Search for the decay $B_s^0 \to \pi^0\pi^0$ at high luminosity electron-positron collider experiment

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

by

Jyotirmoi Borah

DEPARTMENT OF PHYSICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

January 2023


भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

I dedicate my thesis to

“Maa” and “Detamoni”,

“Mahi” and “Moha”,

“Jori Baa”, and
my teachers

Declaration

I declare that the thesis entitled **"Search for the decay $B_s^0 \to \pi^0\pi^0$ at high luminosity electron-positron collider experiment"**, submitted by me to the *Indian Institute of Technology Guwahati*, for the award of the degree of Doctor of Philosophy, is a bonafide work carried out by me under the supervision of **Prof. Bipul Bhuyan**. The content of this thesis, in full or in parts, have not been submitted to any other University or Institute for the award of any degree or diploma.

Signed: Jyotirmoi Borah

Jyotirmoi Borah
Roll No. 176121103
Department of Physics
Indian Institute of Technology Guwahati

Date: 24th January, 2023


भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Certificate

This is to certify that the thesis entitled **"Search for the decay $B_s^0 \to \pi^0\pi^0$ at high luminosity electron-positron collider experiment"**, submitted by Jyotirmoi Borah (176121103), a Research Scholar in the *Department of Physics*, *Indian Institute of Technology Guwahati*, for the award of the degree of Doctor of Philosophy, is a record of an original research work carried out by him under my supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in my opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other Institute or University for the award of any degree or diploma.

Signed: Bipul Bhuyan

Prof. Bipul Bhuyan
Professor
Department of Physics
Indian Institute of Technology Guwahati

Date: 24th January, 2023

Acknowledgements

I am thankful to the people who have made my life better and filled it with hope and delight. I sincerely express my gratitude and thankfulness to Prof. Bipul Bhuyan for accepting the request to be my thesis supervisor and bringing me to this field of experimental high-energy physics. I sincerely thank Prof. Bhuyan for his advice, suggestions, and comments during my doctoral studies. I must mention the freedom and privileges I received during my PhD under his supervision. I will forever be grateful to him for instilling in me the sense of commitment that one needs to have in this profession.

I am deeply thankful towards my institute, the Indian Institute of Technology Guwahati and MHRD, Government of India, for receiving the much-needed support and cooperation for the completion of my research work.

I cannot be more grateful to the physics conveners of the $\Upsilon(5S)$ group at the Belle collaboration, Prof. Kay Kinoshita and Dr. Roman Mizuk, for their several important advice, suggestions, and comments on the analysis. The discussions I had with them during several group meetings shaped the analysis.

I am indebted to my Belle review committee members, Prof. Chunhui Chen, Prof. Martin Sevior, and Dr. Nisar Nellikunnummel, for their thorough review, insightful comments and suggestions for improving the analysis.

I acknowledge the support I received from the Belle and Belle II collaborations during my doctoral studies. This work was only possible with their necessary help and cooperation.

I am thankful to my PhD doctoral committee members, Prof. Poulose Poulose, Dr. Debasish Borah, and Dr. Sayan Chakrabarti for giving their valuable time towards the completion of my academic responsibilities.

I extend my gratitude towards our HoDs of the Physics department, Prof. Poulose Poulose (former), Prof. Subhradip Ghosh (former), and Prof. Perumal Alagarsamy (current) for their support during my PhD tenure.

My mother (“Maa”) has inspired and supported me in my academic journey and life. Her patience, sacrifice, and understanding are beyond my apprehension and experience. I cannot even imagine a life without her advice and guidance. Her untiring efforts in reminding me of my several meetings at Belle and Belle II through phone calls speak volumes about her sincerity. She is also well aware of my analyses, my responsibilities, and my duties at my workplace. I adore you “Maa”. Thank you for everything.

My father ("Detamoni"), who is not with us, has been my ideal person and a strong source of motivation. I feel for our days together, free of fear and anxiety. His farsightedness has always inspired me throughout my childhood and will continue to inspire me in the future. You will always be the person I would love to become.

My aunt ("Mahi") is my second mother. Her concern about my health and well-being exceeds that of any of us in our family. If there is a "child" in my childhood, it is because of "Mahi". I am thankful for your love and affection.

My uncle ("Moha") has been my second father. I would have lost my way had not"Moha" walked with me. His amiable nature and unconditional love have always inspired and held me together. Thank you for giving meaning to our family and my childhood.

My elder sister ("Jori Baa") has always been a constant source of support in my academic journey. Her sincere endeavours for our family's overall well-being are commendable and beyond any acknowledgements. Thank you "Baa" for everything.

My other family members have always been gracious with their blessings. Their support has encouraged me to stay happy and motivated. Thank you all for your love and affection.

My teachers throughout my academic journey have played a prominent role. I cannot forget Indreshwar Terron sir, Bipin Chandra Deka sir, Dr. Manoj Sarmah sir, and Dr. Dilip Sarmah sir, for showing trust and faith in me when I doubted myself. Thank you, dear teachers. You all made my academic life memorable.

My labmates in our HEP lab, Dr. Deepanjali Goswami, Dr. Nitin Yadav, Dr. Kamal Jyoti Nath, Dr. Atanu Nath, Dibya "dada", Maharnab "dada", Devender, Shubhangi, Shivam, Shailesh, and Arkodip have supported me and helped me in several ways beyond academics. I will cherish the moments spent with them forever.

My friends Palash, Nihar, Nilakshi, Jyotishman, Narayani, Pooja, Shilpi, and Taposi have made my life more eventful and enjoyable. Thank you to each one of you for being there for me.

I should acknowledge the support and care I received from Nikumoni. Your presence was a hope for me during my trying times. Thank you for your untiring efforts in making our lives happier and joyful.

I would also like to share my gratitude towards the mess workers, canteen staff at core IV, staff at the physics office, the security guards and cleaners of my department and Lohit hostel for their untiring efforts in making my life a lot easier at IIT Guwahati.

Music has been my solace during times of hope and despair. Zubeen Garg has been the person behind my musical inclination. I often try to emulate his sincerity and dedication towards music in academics. Thank you for all that you have given us.

At the very end, I offer my sincere thankfulness to Nature for giving me all that I need to survive and imbibing the thought that “giving is living”. Thank you. I could not be more grateful for this beautiful life.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

List of Publications

Publication from Thesis

- Search for the decay $B_s^0 \to \pi^0\pi^0$ at Belle, J. Borah, B. Bhuyan *et al.* (Belle Collaboration), Phys. Rev. D **107**, L051101, (2023)
- Search for the decay $B_s^0 \to \pi^0\pi^0$ at Υ(5S) resonance using Belle detector, J. Borah, B. Bhuyan, Springer Proc. Physics 248, 457-462 (2020), Presented at FHEP 2019 (Conference proceedings).

Other publications

1. Search for the decay $B_s^0 \to \eta\eta$, B. Bhuyan, K. J. Nath, J. Borah *et al.* (Belle Collaboration), Phys. Rev. D **105**, 012007 (2022).
2. Measurement of Two-Particle Correlations of Hadrons in e^+e^- Collisions at Belle, Phys. Rev. Lett. **128** (2022) 14, 142005.
3. Search for a light Higgs boson in single-photon decays of Υ(1S) using $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ tagging method, Phys. Rev. Lett. **128** (2022) 8, 081804.
4. B-flavor tagging at Belle II, Eur. Phys. J. C **82** (2022) 4, 283.

Further publications can be accessed through the link iNSPIRE-HEP

Abstract

We report the results of the first search for the decay $B_s^0 \to \pi^0\pi^0$ using 121.4 fb^{-1} of data collected at the $\Upsilon(5S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We observe no signal and set a 90% confidence level upper limit of 7.7×10^{-6} on the $B_s^0 \to \pi^0\pi^0$ decay branching fraction.

Preface

The development of the modern version of the standard model (SM) towards the later part of the 20th century changed the way we describe the fundamental particle interactions of Nature. Its predictions have withstood several confirmations from experiments and continue to provide immense information on particle interactions. However, despite its tremendous successes, it cannot be regarded as a complete theory of particle interactions. At present, it cannot accommodate within its framework some of the most intriguing questions of Nature, such as the observation of the baryon asymmetry, the presence of neutrino masses, the inclusion of gravitational interaction, and the existence of dark matter, to name a few. It, therefore, becomes imperative to look for physics beyond SM. The study of rare decays is one such area where experiments and theories have put forward concerted efforts to look for physics beyond the SM. Several collider experiments have gained insightful information on particle interactions in the past few decades by investigating several rare decays.

This thesis deals with the study of rare decay $B_s^0 \to \pi^0\pi^0$ in the data collected at the Belle experiment corresponding to an integrated luminosity of 121.4 fb^{-1} at $\Upsilon(5S)$ resonance.

Within the SM, the decay $B_s^0 \to \pi^0\pi^0$ proceeds via the W-exchange and "penguin" annihilation topological diagrams. There are theoretical calculations based on the Flavor Diagram Approach, perturbative Quantum Chromodynamics (QCD), and QCD factorization which predict the branching fraction for this decay to be $(0.40 \pm 0.27) \times 10^{-6}$, $(0.28 \pm 0.09) \times 10^{-6}$, and $(0.13 \pm 0.05) \times 10^{-6}$, respectively. The current experimental status on this rare decay channel is an upper limit of 2.1×10^{-4} on the branching fraction at 90% confidence level set by the L3 experiment in 1995.

We perform a blind analysis by developing an analysis strategy on Monte Carlo (MC) samples. We employ several selection criteria to identify and classify the signal from the background. Additionally, neural network algorithms based on multivariate analysis are also used to suppress the dominant continuum background. Probability distribution functions corresponding to the model's variables are designed to extract signal yields in data. Several pseudo-experiments generated from large MC samples are used to look for the stability of the model at the expected sensitivity of the SM. Additionally, validation studies are performed on a similar topology decay channel to verify the analysis procedure. Since, we do not find any significant signal yield in the data, we estimate a 90% confidence level upper limit of 7.7×10^{-6} on the decay branching fraction using a Bayesian approach. Our result provides an improvement of an order of magnitude from the previous result and sets the most stringent upper limit on the branching fraction for this decay channel till date.

We present a chapter wise summary of the thesis in the following paragraphs.

Chapter 1 introduces the theoretical framework of the SM as it is understood today. We elaborate on the SM Lagrangian to understand the kinetic part, which contains the Yukawa interactions associated with the quarks. We briefly describe the origin of the Cabibbo-Kobayashi-Maskawa (CKM) matrix from the Yukawa interactions, which gives rise to the diverse field of flavor physics. We extend our discussions on the unitarity of the CKM matrix and the angles and sides of the unitarity triangle. Towards the later part of the chapter, we present the theoretical motivation for studying rare decays involving weak annihilation amplitudes at collider experiments and their importance in understanding physics beyond SM. We briefly put forward various approaches to study $B \to PP$ decays, where P stand for pseudo-scalar mesons and examine the theoretical branching fraction predictions for the decay, $B_s^0 \to \pi^0\pi^0$.

Chapter 2 describes the KEKB asymmetric-energy e^+e^- collider, the Belle detector, the Belle trigger and data acquisition system, Belle analysis software framework, the physics of the Υ resonances with special reference to the Υ(5S) resonance, and the strategy adopted for this analysis.

Chapter 3 provides an elaborate discussion on the process of event generation and simulation to perform MC studies, the data samples used for this analysis, various skims available at Belle with particular reference to the HadronBJ skim, and about preliminary event selection. The chapter also discusses the various selection criteria developed on MC samples to identify the signal from the background. Sections are devoted to the studies of various topological variables employed in the analysis to suppress the dominant continuum background through neural networks based on multivariate analysis. Further studies on selecting the best B_s^0 candidates are also discussed in detail.

Chapter 4 discusses the parametrization of the variables developed from MC studies to build our model to extract signals from data. We briefly discuss the unbinned extended maximum likelihood procedure, which is eventually used in signal extraction. The chapter also elaborately describes the various studies undertaken to check for the stability of the developed model, its linearity and possible correlations among the model variables.

Chapter 5 discusses in detail the studies performed to validate the analysis strategy adopted for this analysis. The decay channel, $B^0 \to \pi^0\pi^0$ at Υ(4S) resonance, which has similar decay topology and final state particles is chosen as a control sample for this purpose. A section is also devoted to look for data-MC differences in the data sidebands to cross-check the MC simulation of the continuum background.

Chapter 6 reports the first search results for the decay $B_s^0 \to \pi^0\pi^0$ with the Belle data set accumulated at the $\Upsilon(5S)$ resonance with an integrated luminosity of 121.4 fb^{-1}. Several sources of systematic uncertainties are pointed out and summarized in this chapter. Lastly, we briefly discuss the prospects of this decay channel and conclude our work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The beauty of Nature and its manifestations have been intriguing since time immemorial. Scientists have tried to unravel its hidden meaning through their various scientific explorations and inquisitiveness. In this august journey of search and learning, we have been fortunate enough to know some of the inner workings of the beauty of Nature. At the same time, there are mesmerising revelations which still elude scientific explanation.

Our work is another story of quantitative exploration to know our Nature at the scale of minuscules. We forage through various sub-atomic particles and their interactions to ultimately quantify the rate of decay of quark anti-quark pairs to two light-neutral pseudoscalar mesons. In this work, we have used several concepts other than the field of particle physics, such as statistical methods, computer science, and subjects unimaginable to this line of thought, such as biology. This reminds us of the thread of commonality and uniformity that Nature uses to express itself.

We begin by describing our work through one of the successful models of particle physics, the Standard Model, which is the guiding beacon in understanding particle interactions.

1.1 The Standard Model

The Standard Model (SM) is by far the most accurate description of particle interactions, withstanding rigorous experimental tests and numerous validations. At present, it helps us to understand the three fundamental interactions of Nature, namely, the strong, the electromagnetic, and the weak interactions. Consistent efforts from the theoretical side are underway to include the fourth fundamental interaction, the gravitational interaction, within the framework of the SM.

Within the SM, the particles are categorized based on the spin angular momentum. We have, at present, spin half and spin integral particles. The former is called the fermions, and the latter the bosons. In the fermionic type particles, there are three generations of quarks and leptons with their associated anti-particles. While in the bosonic type particles, we have the force carriers, namely gluons, photons, and $W^{\pm}$, Z^0 bosons, for the strong, electromagnetic, and weak interactions, respectively. The SM Higgs boson is a unique particle which provides masses to the fundamental particles. In addition, the fermionic type of particles can be classified into generations based on their effective masses. The first or the lightest generation of the quarks and leptons, barring the neutrino, forms the matter world that we see around us. The fermions of the SM particles are presented in Table 1.1.

TABLE 1.1: Fermionic particle content of the standard model. Here, u, d, c, s, t, b represents the up, down, charm, strange, top, and bottom quarks. While, e, μ, τ, and ν_e, ν_μ, ν_τ, represent the electron, muon, tau, and their corresponding flavors of neutrinos, respectively.

Particle type	Charge	Generations		
		1st	2nd	3rd
Quarks	+2/3	u	c	t
	−2/3	d	s	b
Leptons	−1	e	μ	τ
	0	ν_e	ν_μ	ν_τ

In its modern form, the SM is essentially a gauge theory following the gauge group

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y, \tag{1.1}$$

where, $SU(U)$ stands for special unitary matrix of unit determinant (unitary matrix), the numbers within parenthesis indicates dimensionality of the interaction, and the subscripts c, L, Y, respectively indicates the color, left-handedness, and hypercharge of the SM particles. The $SU(3)_c$ symmetry group is associated with strong interactions where the quarks are treated as color triplets. Quantum Chromodynamics (QCD), which describes the strong interactions, is invariant under $SU(3)_c$ transformation in the color space. The group's generators are the eight gluons which mediate the strong force among the quarks.

The sub-group $SU(2)_L \otimes U(1)_Y$ represents the electroweak interactions, which after spontaneous symmetry breaking (SSB), give rise to the mediating gauge bosons for the weak and electromagnetic interactions: $W^{\pm}$, Z^0, and photons, respectively. The $SU(2)_L$ symmetry group of weak isospin describes the weak interactions among the fermions with its three generators: $W^{\pm}$ and Z^0 bosons, mediating the weak interactions. The left-handed fermions transform as doublets in the weak-isospin space, while the right-handed SM particles under the $SU(2)_L$ group transform as singlets. The unitary group, $U(1)_Y$, represents the electromagnetic interactions upon SSB, where the weak-hypercharge, Y, manifests itself as the force carrier of the electromagnetic interactions, the photons. Therefore, the SM with its associated particles in the gauge theory can be summarized as follows [9]:

1. It follows the gauge group

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y. \tag{1.2}$$

2. Each of the three fermion generations can be represented as

$$Q^I_{Li}(3,2)_{+\frac{1}{6}}, U^I_{Ri}(3,1)_{+\frac{2}{3}}, D^I_{Ri}(3,1)_{-\frac{1}{3}}, L^I_{Li}(1,2)_{-\frac{1}{2}}, E^I_{Ri}(1,1)_{-1}, \tag{1.3}$$

   where, in the term $Q^I_{Li}(3,2)_{+\frac{1}{6}}$: Q_L, i, and I represents left-handed quarks, the three fermion generations ($i = 1,2,3$) and the interaction eigenbasis, respectively. The first and second numbers within the parenthesis specify that the quarks behave as triplets and doublets under $SU(3)$ and $SU(2)$, respectively. At the same time, the numeral subscript represents weak hypercharge under $SU(2) \times U(1)$ symmetry. These notations can be similarly extended for other representations as well, such that, $U_R(D_R)$ stands for right-handed up-type (down-type) quarks which transform as triplets under $SU(3)$ and singlets under $SU(2)$ transformation. $L_L(E_R)$ represents left-handed (right-handed) leptons which transform as singlets under $SU(3)$ and doublets (singlets) under $SU(2)$ transformations.

3. Presence of a scalar field, ϕ to give masses to the gauge bosons and the fermions. The ϕ acquires a vacuum expectation value, v after SSB as

$$\langle\phi\rangle = \langle 0\,|\phi|\,0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{1.4}$$

1.1.1 The SM Lagrangian

The Lagrangian of the SM can be written as [9]

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}, \tag{1.5}$$

where, $\mathcal{L}_{kinetic}$, $\mathcal{L}_{Higgs}$, and $\mathcal{L}_{Yukawa}$ represents the kinetic, Higgs, and Yukawa component of the SM Lagrangian.

The kinetic part of the Lagrangian in Equation 1.5 can be written as

$$\mathcal{L}_{kinetic} = \mathcal{L}_k(Q^I_{Li}) + \mathcal{L}_k(U^I_{Ri}) + \mathcal{L}_k(D^I_{Ri}) + \mathcal{L}_k(L^I_{Li}) + \mathcal{L}_k(E^I_{Ri}). \tag{1.6}$$

We are interested in investigating the quark flavor dynamics within the SM and, therefore, restrict ourselves to quarks and their associated Yukawa interactions. The expansion of the terms is, therefore, done as follows

$$\mathcal{L}_k(Q^I_{Li}) = i\bar{Q}^I_{Li}\gamma_\mu \left[\partial^\mu + \frac{i}{2}g_s G^\mu_a \lambda_a + \frac{i}{2} g W^\mu_b \tau_b + \frac{1}{6} g' B^\mu\right] Q^I_{Li}. \tag{1.7}$$

In the above equation, g_s represents the strong coupling constant, G^μ_a are the eight gluon fields ($a = 1, .., 8$), $\frac{\lambda_a}{2}$ are the generators of $SU(3)_c$, g is the weak coupling constant, W^μ_b are the three weak gauge bosons ($b = 1, 2, 3$), $\frac{\tau_b}{2}$ are the generators of $SU(2)_L, g'$ is the electroweak coupling constant, and B^μ is the hypercharge gauge boson, while the coefficients of B^μ refer to the hypercharges.

The second term of Equation 1.5 can be written as

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi), \tag{1.8}$$

where, $D^\mu\phi = (\partial^\mu + \frac{i}{2}W^\mu_b\tau_b + \frac{i}{2}g'B^\mu)$ and $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$, is a complex scalar field.

The covariant derivative part contains the information of the Higgs interaction with the gauge fields.

The third term of Equation 1.5 is expanded as

$$-\mathcal{L}_{Yukawa} = Y^d_{ij}\bar{Q}^I_{Li}\phi D_{Rj} + Y^u_{ij}\bar{Q}^I_{Li}\phi' U^I_{Rj}, \tag{1.9}$$

where, $\phi' = i\tau_2\phi^*$ and ϕ^* is the complex conjugate of the scalar field, ϕ, $Y_{ij}^{d,u}$ are complex matrices which contain the coupling constants of quarks among different families.

We shall use this information in the subsequent sub-sections to understand the origin of flavor physics within the SM in the quark sector.

1.1.2 Cabibbo-Kobayashi-Maskawa (CKM) matrix : Origin of flavor physics

In this section, we shall briefly discuss the CKM matrix, which gives information about the oscillations of the quarks among its different flavors by connecting the interaction (weak) eigenbasis with the mass eigenbasis. The CKM matrix serves as the origin of the quark flavor physics program within the SM.

The scalar field, ϕ after SSB (in Equation 1.9), acquires a vacuum expectation value, v such that $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \xrightarrow{SSB} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}$, where, h serves the purpose of the Higgs boson.

Concentrating on the first term and omitting the terms containing the Higgs boson in Equation 1.9, we get [10]

$$-\mathcal{L}_{Yukawa} = Y_{ij}^d \bar{Q}_{Li}^I \frac{v}{\sqrt{2}} D_{Rj}^I + \text{otherterms}. \tag{1.10}$$

After evaluating the above expression using quark doublets, we arrive at the following equation

$$-\mathcal{L}_{Yukawa} = \frac{v}{\sqrt{2}} \left[Y_{ij}^d \bar{d}_{Li}^I d_{Rj}^I + Y_{ij}^u \bar{u}_{Li}^I u_{Rj}^I \right] + \text{otherterms} + \text{hc}. \tag{1.11}$$

Considering

$$M_{ij}^d = \frac{v}{\sqrt{2}} Y_{ij}^d, \tag{1.12}$$

$$M_{ij}^u = \frac{v}{\sqrt{2}} Y_{ij}^u. \tag{1.13}$$

The mass matrix after diagonalizing through the biunitary transformation principles, takes the form [11]

$$M^d_{diag} = V^d_L M^d_{ij} V^{d\dagger}_R, \tag{1.14}$$

$$M^u_{diag} = V^u_L M^u_{ij} V^{u\dagger}_R, \tag{1.15}$$

where, V^d_L, V^d_R, V^u_L, and V^u_R are unitary matrices.

Thus, the Yukawa Lagrangian, $\mathcal{L}_{Yukawa}$, can now be represented in the mass eigenbasis. We now have to observe the changes that have taken place due to the transition from the interaction eigenbasis to the mass eigenbasis. Pertaining to this transformation, the kinetic part of the Lagrangian from Equation 1.7 becomes

$$\mathcal{L}_{kinetic} = -\frac{g}{2}\gamma_\mu W^{+\mu} \bar{u}^I_{Li} d^I_{Li} + \text{otherterms}. \tag{1.16}$$

Now, in order that $\bar{u}^I_{Li}$ is changed also to mass eigenbasis, the transformation equation becomes

$$\bar{u}^I_{Li} d^I_{Li} = \bar{u}_{Li} (V^{u\dagger}_L V^d_L)_{ij} d_{Li}, \tag{1.17}$$

where, u_L and d_L are in the mass eigenbasis.

The product of the matrices in Equation 1.17 gives the transition probability of a $d-$type quark to an $u-$type quark. Considering all the elements of such transition leads to the CKM matrix proposed by Kobayashi and Maskawa for three generations of quarks, as shown below

$$V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix}. \tag{1.18}$$

The standard parametrization of the CKM matrix with its four real parameters is shown below

$$V_{\rm CKM} = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix}, \tag{1.19}$$

where, $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ with $(i, j \in [1,3]$ and $i \neq j)$, and δ is the only phase in the CKM matrix which is responsible for the Charge-Parity (CP) violation within the SM.

The Wolfenstein parameterization [12] is another very useful parametrization for the CKM matrix, primarily used in phenomenological studies inspired by experimental observations. It is an approximate parametrization as the expansion parameter is truncated after the third order, which is shown below

$$V_{\rm CKM} = \begin{bmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4), \tag{1.20}$$

where, λ, A, ρ and η are the four parameters unitary to all orders in λ. In this parametrization, η corresponds to the δ phase difference in the standard parametrization. The Wolfenstein parameters are defined accordingly [1], as

$$\lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \;\; \bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \;\; A\lambda^2 = \lambda\left|\frac{V_{cb}}{V_{us}}\right|. \tag{1.21}$$

They are related to the standard parametrization through the following relations

$$\lambda = s_{12}, \qquad A\lambda^3(\rho - i\eta) = s_{13}e^{i\delta}, \tag{1.22}$$

$$\rho = \mathcal{R}_e\left\{\frac{s_{13}e^{i\delta}}{s_{12}s_{23}}\right\}, \qquad \eta = -\mathcal{I}_m\left\{\frac{s_{13}e^{i\delta}}{s_{12}s_{23}}\right\}. \tag{1.23}$$

As per the recent review of particle physics [1], the fit to the Wolfenstein parameters yields the following values

$$\lambda = 0.22500 \pm 0.00067, \qquad A = 0.826^{+0.018}_{-0.015}, \tag{1.24}$$

$$\bar{\rho} = 0.159 \pm 0.010, \qquad \bar{\eta} = 0.348 \pm 0.010. \tag{1.25}$$

The magnitude of the CKM matrix elements, best known today [1], is as follows

$$V_{\text{CKM}} = \begin{bmatrix} 0.97435 \pm 0.00016 & 0.22500 \pm 0.00067 & 0.00369 \pm 0.00011 \\ 0.22486 \pm 0.00067 & 0.973 \pm 0.00016 & 0.04182^{+0.00085}_{-0.00074} \\ 0.00857^{+0.00020}_{-0.00018} & 0.04110^{+0.00083}_{-0.00072} & 0.999118^{+0.000031}_{-0.000036} \end{bmatrix}. \tag{1.26}$$

From the unitarity condition of the CKM matrix, we get the following equations

$$V_{ud}V_{ud}^* + V_{us}V_{us}^* + V_{ub}V_{ub}^* \sim 1, \tag{1.27}$$

$$V_{cd}V_{cd}^* + V_{cs}V_{cs}^* + V_{cb}V_{cb}^* \sim 1, \tag{1.28}$$

$$V_{td}V_{td}^* + V_{ts}V_{ts}^* + V_{tb}V_{tb}^* \sim 1. \tag{1.29}$$

The inferences that can be drawn from the Equations 1.26 - 1.29 are [10]:

- The sum of the squares of the coupling strengths of u to d, s, b is same as the charged coupling of the c and the t quarks: weak universality
- The sum of the squares of the amplitudes(probabilities) add up very close to unity, leaving a very rare chance of existence of a fourth family of quark.

One of the primary purposes of the flavor physics program is to precisely measure the elements of the CKM matrix. Exploiting the following unitary condition

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.30}$$

a unitarity triangle is constructed. The triangle is scaled by the product of the CKM elements, $|V_{cd}V_{cb}^*|$ leading to a scaled-unitarity triangle with the three of its vertices situated at $(0,0)$, $(0,1)$, and $(\bar{\rho},\bar{\eta})$. Figure 1.1 shows the unitarity triangle.

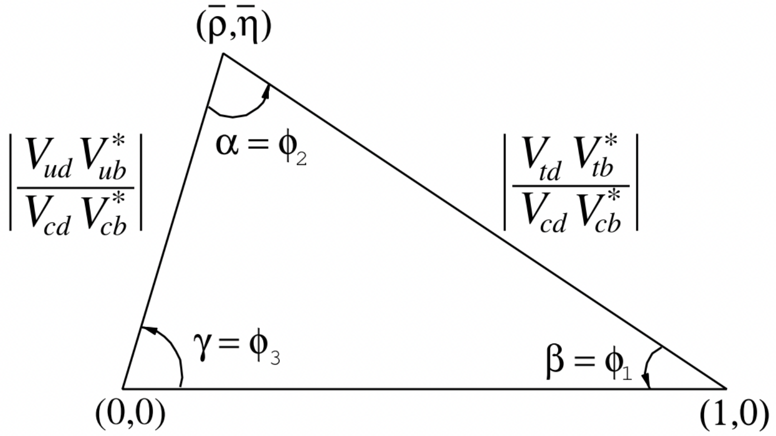

FIGURE 1.1: The unitarity triangle [1].

The sides of the triangle are, therefore, defined as

$$\left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\rho^2+\eta^2}, \qquad \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\rho)^2+\eta^2}. \tag{1.31}$$

and the angles of the triangle are

$$\alpha(\phi_2) \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta(\phi_1) \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma(\phi_3) \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{1.32}$$

Therefore, precise determination of the sides and angles of the unitarity triangle will lead to precise measurements of the CKM matrix elements. Over-constraining the unitarity triangle through precise measurements of its sides and angles is one of the primary goals of the flavor physics program in high-energy physics.

Despite the tremendous success of the SM, it is still regarded as an effective theory of a more general theory at higher energies (the grand unified theory scale $\sim 10^{16}$ GeVs or higher). Currently, the SM cannot accommodate within its theoretical framework some of the most intriguing challenges in the field of particle physics: the existence of dark matter, tiny masses of the neutrinos (in the range of eVs), and the present baryon asymmetry observed in the universe, to name a few. Thus, it becomes imperative to go beyond the SM and look for new physics (NP), which will eventually solve the puzzles.

1.2 Rare decays involving weak annihilation Feynman diagrams

Within the SM, the relative strengths of the quark transitions among different generations can be summarized as [13]:

- Intra-generation transitions: Intra-generation transitions such as, $d \to u$ or $c \to s$ for the different generations are favoured and assigned a strength of $\mathcal{O}(1)$.
- Inter-generation transitions: Transition between third and second (first) are Cabibbo suppressed by $\mathcal{O}(10^{-2})(\mathcal{O}(10^{-3}))$, while transition between second and first generation are Cabibbo suppressed by $\mathcal{O}(10^{-1})$ relative to intra-generation transitions.

Therefore, the transition of a b-quark to an u-quark or d/s-quarks is suppressed within the SM. Further, the transitions $b \to d/s$ cannot occur at the "tree-level", instead can proceed through "one-loop" level due to GIM suppression [14]. These types of transitions within the SM are regarded as rare and provide unique opportunities to search for physics both within and beyond the SM.

The study of heavy-flavoured hadrons decaying to non-leptonic final states provides an essential platform for understanding the interplay between strong and weak interactions. In this regard, rare decays involving weak annihilation amplitudes can be an excellent place to look for disagreement between theoretical predictions and experimental observations. These processes are considered highly suppressed and assumed to be negligible in theoretical calculations. However, the inclusion of rescattering effects into the theoretical framework naturally enhances their contribution [15]. Recently it was observed that the predicted branching fraction for the decay $B_s^0 \to \pi^+\pi^-$, which involves topological annihilation diagrams, was substantially smaller than its measured value by the LHCb experiment [16]. This discrepancy between theoretical prediction and experimental measurement may require some models of strong interaction processes to be revisited [17]. In these aspects, searches for decays involving weak annihilation Feynman diagrams become important and require experimental investigations.

1.3 Significance of the decay $B_s^0 \to \pi^0\pi^0$

The decay $B_s^0 \longrightarrow \pi^0\pi^0$ is a neutral, non-leptonic, charmless, charged current, strangeness non-conserving rare decay mode. The absence of a non-leptonic part makes this decay an important study in understanding the interplay between strong and weak interactions. The absence of charm quarks in the final state hints at a non-perturbative treatment as the final state mesons are well below the $\Lambda_{\rm QCD} \sim 200$ MeV scale and can be a valuable tool to validate most of the non-perturbative methods that exist today.

From a model-building perspective, the presence of charged currents mediated by $W^\pm$ gauge boson can be an important area to search for NP, as the presence of more mediators or mediators of a different kind can eventually affect the branching fractions. A precise measurement of the branching fraction is thus essential to validate different models associated with such processes.

1.4 Approaches to study $B \to PP$ decays

Very often, the study of rare decays is performed model-independently to maximize the conformity of theoretical predictions with experiments. The usual approach in these studies is to express the weak Hamiltonian in terms of the CKM matrix parameters and the coefficients, which consider the short and long-distance effects. In general, the weak Hamiltonian is written as [18]

$$H_w = \frac{G_F}{\sqrt{2}} J_\mu J^\mu, \tag{1.33}$$

where, G_F is the Fermi constant, J_μ is the associated weak current given by

$$J_\mu = \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma_\mu (1-\gamma_5) \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}, \tag{1.34}$$

where, u, c, t are the up, charm, top quark fields and d', s', b' are the down, strange, and bottom quark fields in the weak eigenstate basis, respectively. The CKM matrix connects the weak eigenstate basis to the mass eigenstate basis of d, s, and b.

The expanded Hamiltonian after omitting the space-time components look like

$$
\begin{aligned}
H_w = \frac{G_F}{\sqrt{2}}[&V_{ub}V_{cd}^*(\bar{u}b)(\bar{d}c) + V_{ub}V_{cs}^*(\bar{u}b)(\bar{s}c) \\
&+ V_{ub}V_{ud}^*(\bar{u}b)(\bar{d}u) + V_{ub}V_{us}^*(\bar{u}b)(\bar{s}u) \\
&+ V_{cb}V_{ud}^*(\bar{c}b)(\bar{d}u) + V_{cb}V_{us}^*(\bar{c}b)(\bar{s}u) \\
&+ V_{cb}V_{cs}^*(\bar{c}b)(\bar{s}c) + V_{cb}V_{cd}^*(\bar{c}b)(\bar{d}c). \qquad (1.35)
\end{aligned}
$$

We see from Equation 1.35 that apart from the CKM parameters, the quark fields can also interact among themselves strongly. Thus, it is essential in these types of studies to demarcate the scales at which the different aspects of the strong interaction become essential due to hadronisation to mesons. There are short-distance (high-energy regime) effects and long-distance (low-energy regime) effects. The short-distance effects can be explained perturbatively in QCD, while the long-distance effects are to be solved non-perturbatively. This non-perturbative analysis is complex as it involves a complicated calculation of matrix elements between the initial and the final states.

Under these circumstances, we cannot solely rely on these calculations and should look for other possible theoretical tools to predict branching fractions for such generic decays. It is where the Flavor Diagram Approach (FDA) [19] can be a handy tool. We shall briefly discuss the FDA as one of the tools to predict the branching fraction for this analysis and quote the branching fraction predictions from other approaches such as the perturbative QCD (pQCD) [20], and QCD factorization [21]. The Flavor Diagram Approach uses Flavor SU(3) symmetry and, in general has the following properties [19]:

- Non-perturbative in nature.
- Uses Flavor $SU(3)$ symmetry to relate decay diagrams, both sizes and associated strong phases of the same topology, but differing for the light quarks.
- It can extract matrix elements without relying on any specific models.
- The parameters extracted from this approach have all effects of strong interactions embedded in them, including the long-distance scattering effects.

In the next sub-section, we shall briefly describe how the FDA can be used in predicting the branching fraction for the decay $B_s^0 \to \pi^0\pi^0$.

1.4.1 Topological diagram analysis of $B \to PP$ decays

Within the SM, the decay $B_s^0 \to \pi^0\pi^0$ proceeds via the W-exchange and "penguin" annihilation topological diagrams, as shown in Figure 1.2. In FDA, the transition amplitudes for the heavy meson decays can be categorized based on their validity in explaining experimental data as:

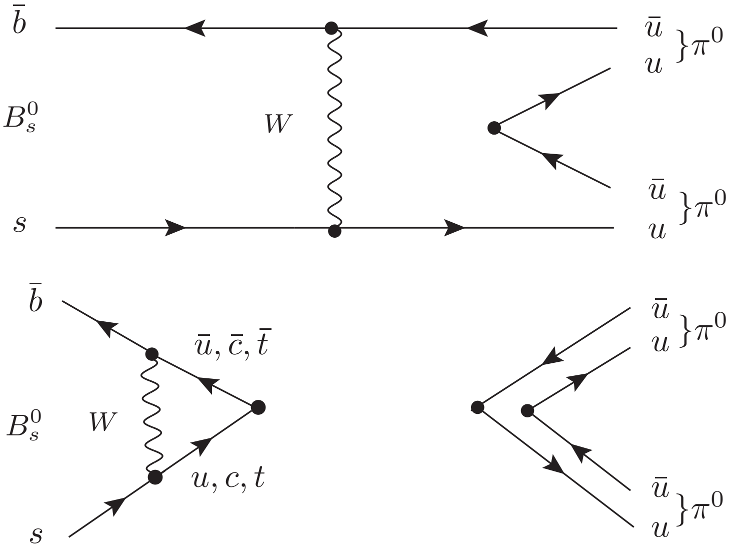

FIGURE 1.2: W-exchange (top) and "penguin" annihilation (bottom) Feynman diagrams for $B_s^0 \to \pi^0\pi^0$.

1. $T \to$ external W emission
2. $C \to$ internal W emission
3. $E \to W$ exchange
4. $P \to QCD$ "penguins"
5. $S \to$ flavor singlet QCD "penguins"
6. $P_{EW} \to$ electroweak "penguin"
7. $PA \to$ "penguin" annihilation

The amplitudes mentioned above can be factorized into a modulus part and its associated strong phase. These theoretical parameters in this approach are derived from experimental data and fitted, which are eventually used in the calculations of the branching fractions and CP asymmetries in $B \to PP$ decay processes, where, P is any pseudoscalar meson.

In our case, as seen from the diagrams, only two amplitudes, E and PA are essential for the analysis. These amplitudes contain CKM matrix parameters and can be expressed as

$$e = Y_{sb}^{u}E = (V_{us}V_{ub}^{*})E, \qquad pa = (Y_{sb}^{u} + Y_{sb}^{c})PA = (V_{us}V_{ub}^{*} + V_{cs}V_{cb}^{*})PA. \tag{1.36}$$

Moreover, as $V_{us} \simeq 0.225$, $V_{ub} \simeq 0.003$, and $V_{cs} \simeq 0.973$, $V_{cb} \simeq 0.042$ [1] it can be seen that

$$|Y_{sb}^{c}| = 0.041 > |Y_{sb}^{u}| = 6.4 \times 10^{-4}. \tag{1.37}$$

Equation 1.37 naively suggests that the contribution from the "penguin" annihilation diagram is more than the W-exchange diagram.

The expression for decay width for any two-body B meson decays can be expressed as

$$\Gamma(B \to M_1M_2) = \frac{k}{8\pi m_B^2}\,|A|^2\,, \tag{1.38}$$

where, m_B is the B meson mass, k denotes the three momentum of the final states mesons, $M_{1,2}$ and A represents the decay amplitude.

In our case, the decay amplitudes correspond to the flavor amplitudes E and PA, while the momentum, k, can be extracted from kinematic relations. Then the flavor amplitude is given as

$$A(B_s^0 \to \pi^0\pi^0) = \frac{1}{\sqrt{2}}(e + 2pa). \tag{1.39}$$

The factor 2 appears due to the transition amplitude PA in Equation 1.36. Then the branching fraction ($\mathcal{B}$) can be calculated from the following expression

$$\mathcal{B}(B_s^0 \longrightarrow \pi^0\pi^0) = (\Gamma(B_s^0 \to \pi^0\pi^0) + \Gamma(\bar{B}_s^0 \to \pi^0\pi^0) \times \tau_{B_s^0}, \tag{1.40}$$

where, $\tau_{B_s^0}$ is the mean lifetime of the B_s^0 meson and $\bar{B}_s^0$ is the conjugate decay mode.

In this approach, the branching fraction becomes a function of the modulus of the flavor amplitudes and the strong phases, which are extracted through a χ^2 fit to data. Proceeding as above, the predicted branching fraction for our decay mode stands at $\sim (0.40 \pm 0.27) \times 10^{-6}$ [19].

We can further convince ourselves using the $SU(2)$ isospin symmetry in $B_s^0 \to \pi^+\pi^-$ decays. The present measured value for this mode is $(7.0 \pm 0.8) \times 10^{-7}$ [16] so, the branching fraction for our decay mode should roughly be in the range of $10^{-6} - 10^{-7}$.

Theoretical calculations based on other approaches such as, perturbative QCD (pQCD) [20], and QCD factorization [21] predict the branching fraction for this decay to be $(0.28 \pm 0.09) \times 10^{-6}$, and $(0.13 \pm 0.05) \times 10^{-6}$, respectively.

The only measurement for this decay was made by the L3 experiment in 1995, which reported an upper limit (UL) of $\mathcal{B} < 2.1 \times 10^{-4}$ at 90% confidence level (CL) using 3 million hadronic Z decays [22].

Summary

In this chapter, we briefly outlined the SM and discussed the origin of the CKM matrix in the quark sector. We also described the significance of the decays involving weak annihilation Feynman diagrams and pointed out the approaches available for studying them. In this context, we briefly described the flavor diagram approach and its prediction on the branching fraction towards the current analysis.

Chapter 2

The Belle experiment

In this chapter, we briefly discuss the Belle experiment emphasizing the collider, the Belle detector, the software framework for physics analyses, the physics of the B_s^0 mesons at $\Upsilon(5\mathrm{S})$ resonance, and the strategy adopted for this analysis.

2.1 History of the Belle experiment

The Belle experiment started as a serious effort from the international high-energy physics community led by Japan to understand and test the CP violation proposed in a model by Kobayashi and Maskawa in 1973 [23]. The construction of the Belle experiment began in 1994 at the High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan. It started taking physics data from 1999 that continued till 2010, and accumulated a data set equivalent to an integrated luminosity of more than 1 ab^{-1}. In 2001, Belle along with BaBar observed for the first time a significant CP violation in the neutral B meson system [24]. This observation led to the Nobel Prize (2008) for the work carried out by Kobayashi and Maskawa, thereby confirming the presence of CP violating sources within the SM.

The Belle experiment comprised of the KEKB collider [25, 26] and the Belle detector [4, 7]. KEKB was an asymmetric-energy electron-positron (e^+e^-) collider, tuned to operate at different centre-of-mass (CM) energies to produce $\Upsilon(n\mathrm{S})$ (*Upsilon*) resonances on threshold, where $n = 1, 2, 3, 4,$ and 5. However, it collected 70% of physics data at $\Upsilon(4\mathrm{S})$ resonance tuning the e^+e^- beam energy to a CM energy of 10.58 GeV primarily due to its broad physics program. It also collected roughly 12% of physics data at $\Upsilon(5\mathrm{S})$ resonance, which remains one of the highest accumulated data sets by any e^+e^- collider experiment

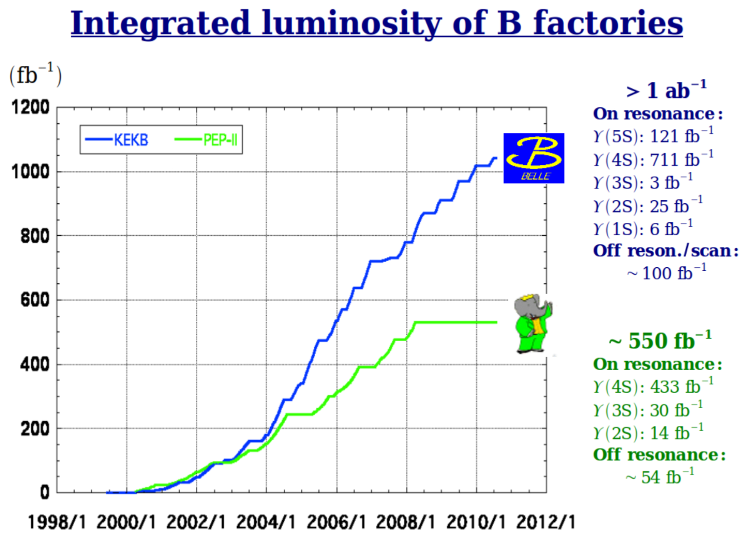

FIGURE 2.1: Data sets accumulated at the Belle experiment over various resonances during its lifetime [2].

to study B_s^0 mesons. Figure 2.1 shows the integrated luminosity collected by the Belle and BaBar experiments over various resonances during its lifetime. The Belle detector was situated at the interaction point of the e^+e^- beam to detect the particles produced during collisions. It had comprehensive physics programs ranging from flavor physics and tau physics to the most exotic dark matter searches. After more than a decade of stopping its physics operation, the data collected by the Belle detector still has physics potential with several ongoing dynamic physics analyses.

2.2 The KEKB collider

The KEKB was an asymmetric-energy e^+e^- collider where the electrons and positrons were accelerated to high energies to produce particles from collisions. Its asymmetric beam energy played a vital role in the first observation of the CP violation in the neutral B mesons. The asymmetry in the beam energy provided the required Lorentz boost factor ($\beta\gamma$) for the separation of the $B\bar{B}$ pairs produced from Υ(4S) on the threshold. Mathematically, the Lorentz boost factor is represented as

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

with, $E_- = 8.0$ GeV (electron beam energy), $E_+ = 3.5$ GeV (positron beam energy), and $\sqrt{s} = 2\sqrt{E_-E_+} = 10.58$ (CM energy).

Figure 2.2 shows a schematic diagram of the KEKB collider. It consisted of two rings encircling each other over a circumference of roughly 3 km. The ring carrying the positrons of energy 3.5 GeV is called the Low Energy Ring (LER), while the ring carrying the electrons of energy 8.0 GeV is called the High Energy Ring (HER). The positrons were produced by the bombardment of an electron beam on a target made of Tungsten. They were separated from a host of other particles and subsequently injected into the LER at the Fuji area of the rings. Simultaneously, electrons were also injected into the HER in the Fuji area. The electrons and the positrons were accelerated to high energies to reach the desired beam energies while maintaining a constant current of 1.4 A and 1.7 A in the HER and LER rings, respectively. The beams during their revolution in the rings gradually lose energy due to synchrotron radiation. Radio-Frequency (RF) cavities were installed at the Oho, Fuji, and Nikko area to re-energize and maintain the stability of the beams. Two types of RF cavities, namely, Accelerators Resonantly coupled with Energy Storage (ARES) and Superconducting Cavities (SCC) were utilized for the LER/HER and HER, respectively. Additionally, wigglers were also installed in the LER ring at Oho and Nikko areas to reduce the longitudinal damping time of the positron beam. Finally, the electron beam in the HER ring is allowed to collide at a crossing angle of ±11 mrad (0.63°) with the positron beam in the LER ring at the interaction region or point (IP) of Tsukuba area. Table 2.1 summarizes the KEKB machine parameters drawing comparison between the designed and the achieved targets.

2.3 The Belle detector

The Belle detector is a general-purpose hermetic detector located at the IP of the e^+e^- beams in the Tsukuba area. It consisted of several sub-detectors arranged in the form of concentric layers for vertex determination, particle identification, and energy measurement. Figures 2.3 and 2.4 depicts the Belle detector and its side-view, respectively. The detector components are enumerated below:

1. Silicon Vertex Detector (SVD)
2. Central Drift Chamber (CDC)
3. Aerogel Cherenkov Counter (ACC)

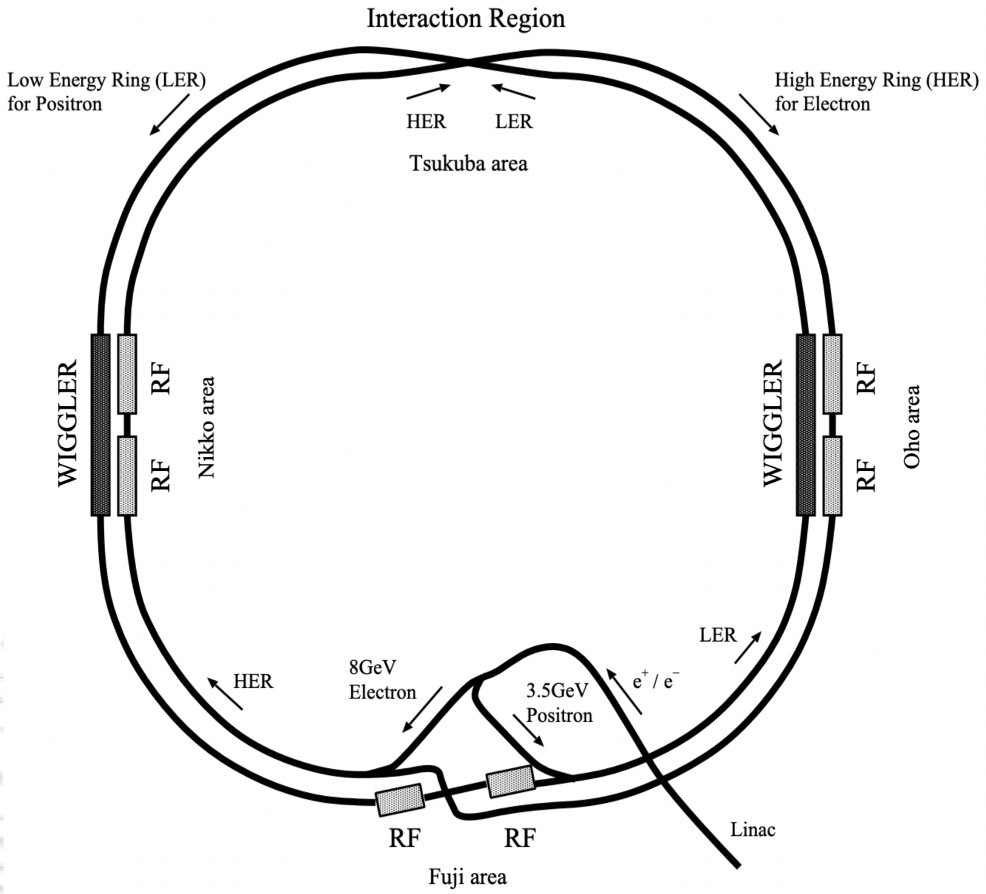

FIGURE 2.2: The KEKB collider [3].

4. Time-of-flight (TOF) counter
5. Electromagnetic Calorimeter (ECL)
6. Long lived Kaon (K_L^0) and muon (KLM) detector
7. Extreme Forward Calorimeter (EFC)

The sub-detectors SVD, CDC, ACC, TOF, and ECL components were placed inside a superconducting solenoid coil, generating a magnetic field of strength 1.5 T. The EFC sub-detector placed close to the beam line was used to monitor the beam and luminosity measurements.

Belle uses the following coordinate system for all of its detector configurations and physics measurements:

- The x-axis is considered to be in the horizontal outward direction of the KEKB ring.

Table 2.1: Machine parameters of the KEKB drawing comparison between designed and achieved targets [7].

	Design	Achieved
Luminosity (10^{33} cm^{-2} s^{-1})	10	5.47
Number of rings	2	-
Number of interaction points	1	-
Circumference (m)	3016	-
Beam energy (GeV)	$8.0(e^-)/3.5(e^+))$	$8.0(e^-)/3.5(e^+))$
Total current per beam (A)	$1.1(e^-)/2.6(e^+))$	$\sim 0.4(e^-)/\sim 0.7(e^+))$
Number of bunches	5120×0.9 (10% gap)	~ 1100
Bunch spacing (m)	0.59	2.36 (regular operation)
Crossing angle (mr)	± 11	± 11
Beam-beam tuneshift		
$e^- \xi_x/\xi_y$	0.039/0.052	0.061/0.035
$e^+ \xi_x/\xi_y$	0.039/0.052	0.073/0.047
Beta-function at the interaction point		
$e^- \beta_x^*/\beta_y^*$ (cm)	33/1	0.63/0.007
$e^+ \beta_x^*/\beta_y^*$ (cm)	33/1	0.59/0.0065
RF frequency (MHz)	508.887	-
Type of cavities	NCC/SCC for e^-; NCC for e^+	-
Detector	Belle	-

NCC and SCC stand for normal conducting and superconducting cavity.

- The y-axis is in the vertical direction from the plane of the KEKB ring.
- The z-axis is along the opposite direction of the positron beam axis.

Following the above configuration, the spherical polar coordinates and the cylindrical coordinates (R, θ, ϕ) and (r, ϕ, z), respectively, are defined as:

- The radial distance, $r = \sqrt{x^2 + y^2}$.
- The polar angle, θ is measured from the z-axis.
- The azimuthal angle, ϕ is measured around the z-axis.

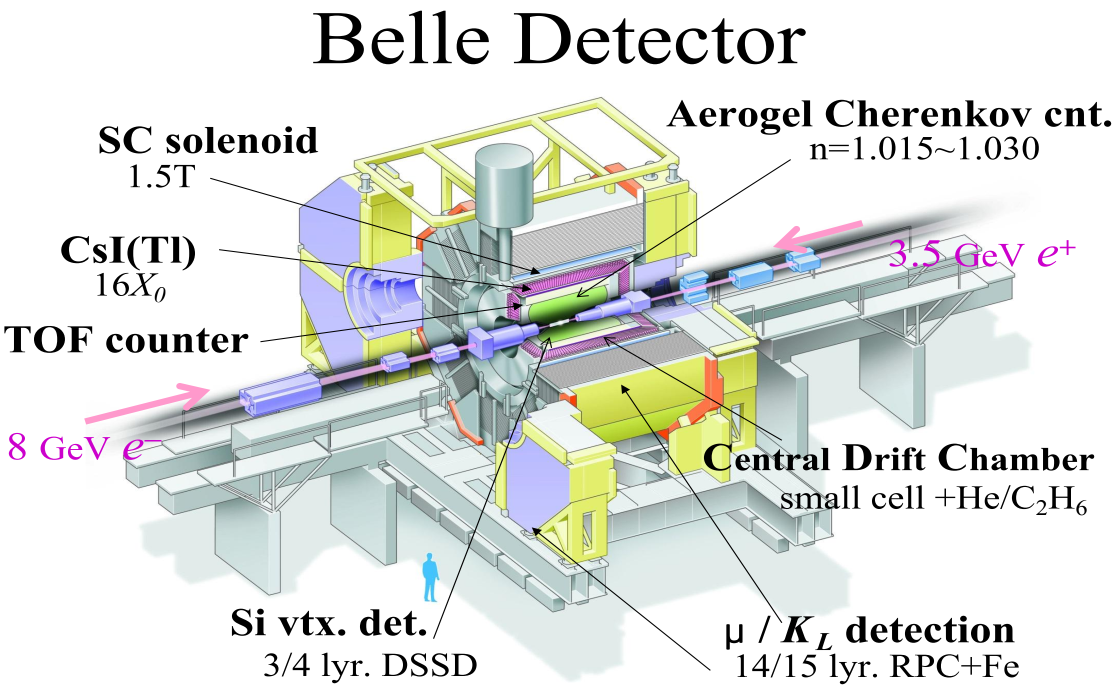

FIGURE 2.3: The Belle detector.

The Belle detector's beam pipe was laid concentrically with the KEKB accelerator's beam pipe. It was a double-walled cylindrical pipe made of Beryllium with the inner gap filled with Helium gas for cooling. The radius of the inner(outer) wall was 20 mm (23 mm) as measured from the IP with a thickness of 0.5 mm for both the walls. The inner walls of the beam pipe were sputtered with gold atoms forming a layer of thickness of 10 μm to prevent synchrotron radiation from entering the other parts of the detector. Figure 2.5 shows the beam pipe of the Belle detector.

A brief description of the sub-detectors are given in the following sub-sections.

2.3.1 Silicon Vertex Detector (SVD)

The primary function of the SVD is to efficiently reconstruct the decay vertices of the B mesons required for CP violation studies. Additionally, information from SVD is also required for vertex reconstruction of the D mesons, τ leptons and to remove backgrounds arising from $e^+e^- \rightarrow q\bar{q}$, where $q = u,\ d,\ s,\ c$. Therefore, placing the SVD as close as possible to the IP is necessary. From 1999 to 2010, two different configurations of SVD (SVD1 and SVD2) were in operation. They are described below:

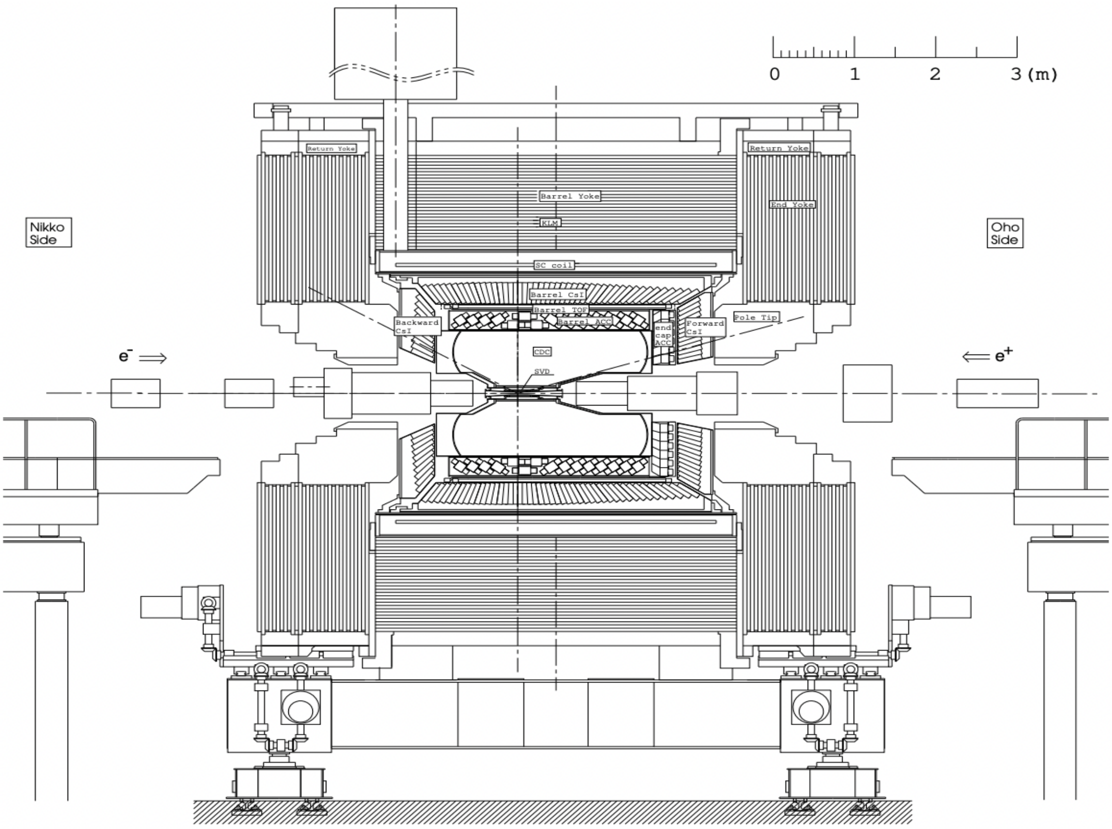

FIGURE 2.4: A cross-sectional view of the Belle detector [3].

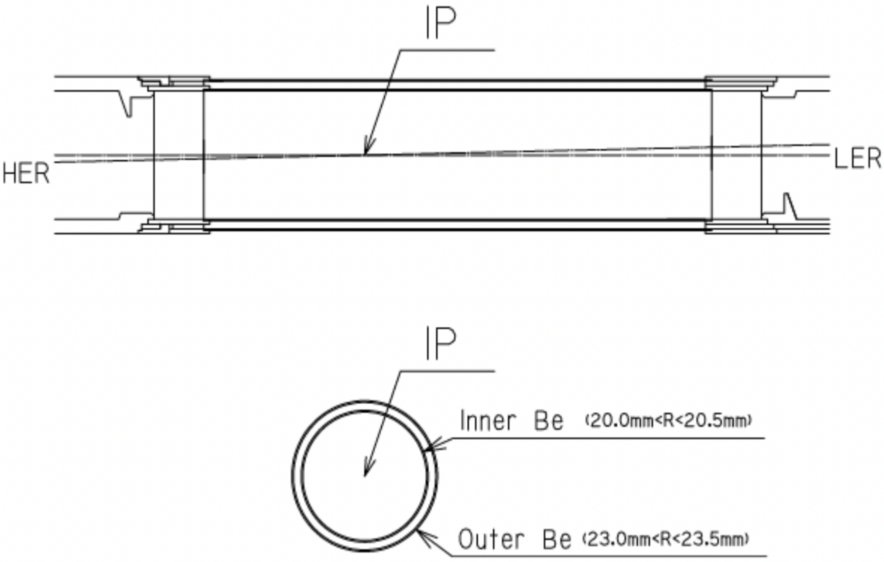

FIGURE 2.5: The beam pipe of the Belle detector [4].

- SVD1 (1999-2003): Figure 2.6 shows the configuration of the SVD1 sub-detector. It comprised of three layers of Double-sided Silicon Strip Detectors (DSSD) arranged concentrically at a distance of 30, 45.5, and 60.5 mm from the IP of the beam pipe. This configuration covered the polar angle, θ, in the range $23^{\circ} < \theta < 139^{\circ}$. There were 8/10/14 independent ladders in each layer. Further, each ladder consisted of 2/3/4 DSSDs for layers 1/2/3, respectively. Therefore, there were, in total, 102 DSSDs in SVD1. Boron-nitride ribs provided the mechanical strength to the DSSDs. A detailed description of SVD1 is discussed in the references [5, 27].

- SVD2 (2003-2010): Figure 2.7 shows the configuration of the SVD2 sub-detector. It comprised of four layers of DSSDs arranged concentrically at a distance of 20, 45.5, 70, and 88 mm from the IP of the beam pipe. The addition of an extra layer and placement of the inner layer closer to the IP, facilitated the reconstruction of charged tracks from SVD hit information alone. The SVD2 covered the polar angle, θ in the range $17^{\circ} < \theta < 150^{\circ}$, which is similar to the geometrical acceptance angle of the CDC. In comparison to SVD1, SVD2 had a better angular coverage. There were 6/12/18/18 independent ladders in each layer. Further, each ladder consisted of 2/3/5/6 DSSDs for layers 1/2/3/4, respectively. To accommodate the inner layer of the SVD2 closer to the IP, the inner radius of the beam pipe was also reduced to 15 mm.

The performance of the SVDs was studied using cosmic rays. The primary parameter to judge their performance was the impact parameter (distance of closest approach of the charged tracks to the IP) resolutions in the $R - \phi$ and z plane. Figure 2.8 shows the performance comparison of SVD1 and SVD2. It was observed that SVD2 performed better than SVD1 after the upgrades.

2.3.2 Central Drift Chamber (CDC)

CDC provided information on the trajectories and momenta of the charged particles (tracks) produced from the e^+e^- collisions in the presence of a magnetic field of strength 1.5 T, generated by a superconducting solenoid. It also determined the energy losses (dE/dx) of the charged particles in the drift chamber medium and provided additional information on the hardware trigger for the detector.

Figure 2.9 shows the cell structure of the CDC. CDC was a cylindrical drift chamber which consisted of layers of wires of two different kinds: sense wires (anode wires) and field wires.

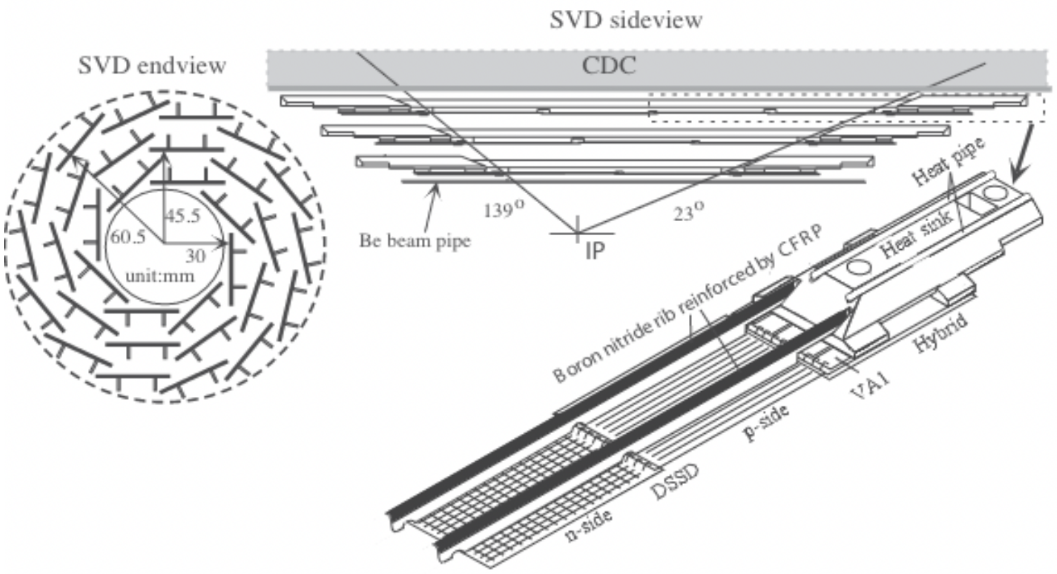

FIGURE 2.6: SVD1 sub-detector [5].

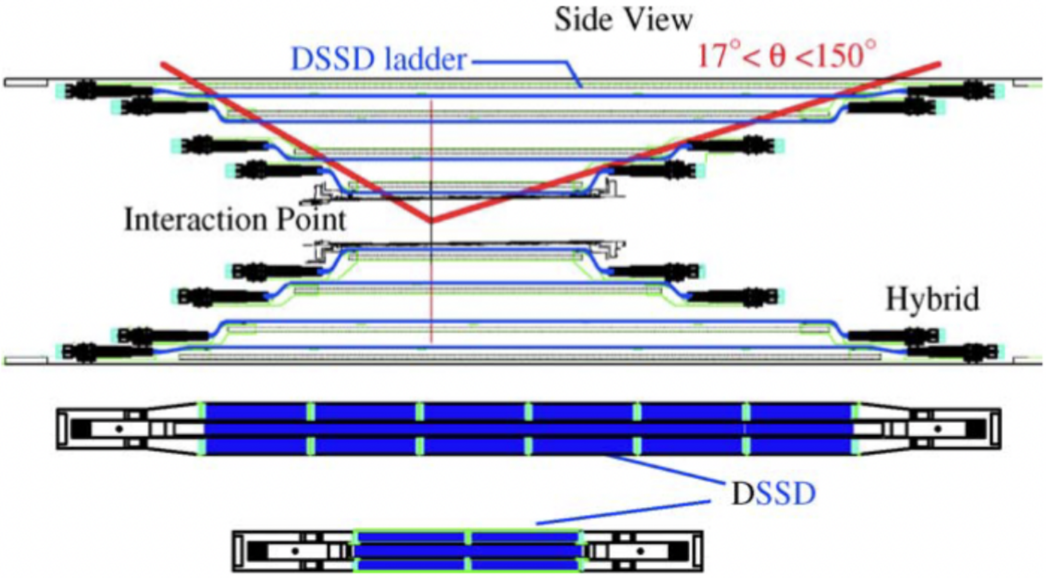

FIGURE 2.7: SVD2 sub-detector [6].

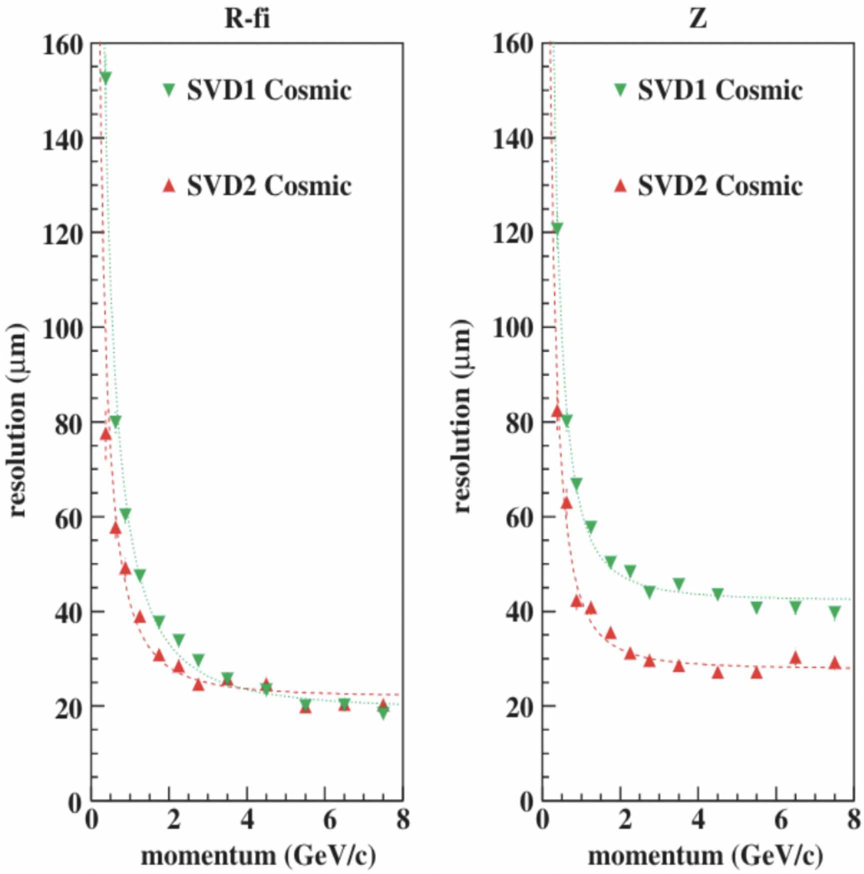

FIGURE 2.8: Comparison of performances between SVD1 and SVD2 in terms of impact parameter resolutions [6].

A positive high voltage was applied to the sense wires while the field wires were kept at the ground potential. The anode wires were arranged in two different configurations: axial and small-angle positions (stereo position). The axially placed (z-direction) anode wires together with the stereo wires provided the 3D information of the charged tracks. In addition to the anode wires, three layers of cathode strips were also installed. These cathode layers helped in the improvement of the precise measurement of the z-direction of the charged tracks. In total, there were 50 layers of anode wires, of which 32 and 18 were placed axially and in stereo position, respectively. Each of these layers were designed to have between three to six axial wires or stereo wires. Figure 2.10 shows the cross-sectional view of the central drift chamber. The drift chamber is asymmetric in the z-direction covering a polar angle, θ in the range $17° < \theta < 150°$.

A precise measurement of the momentum of a charged track requires the position coordinates of the trace of the track under the influence of a strong magnetic field. The wire

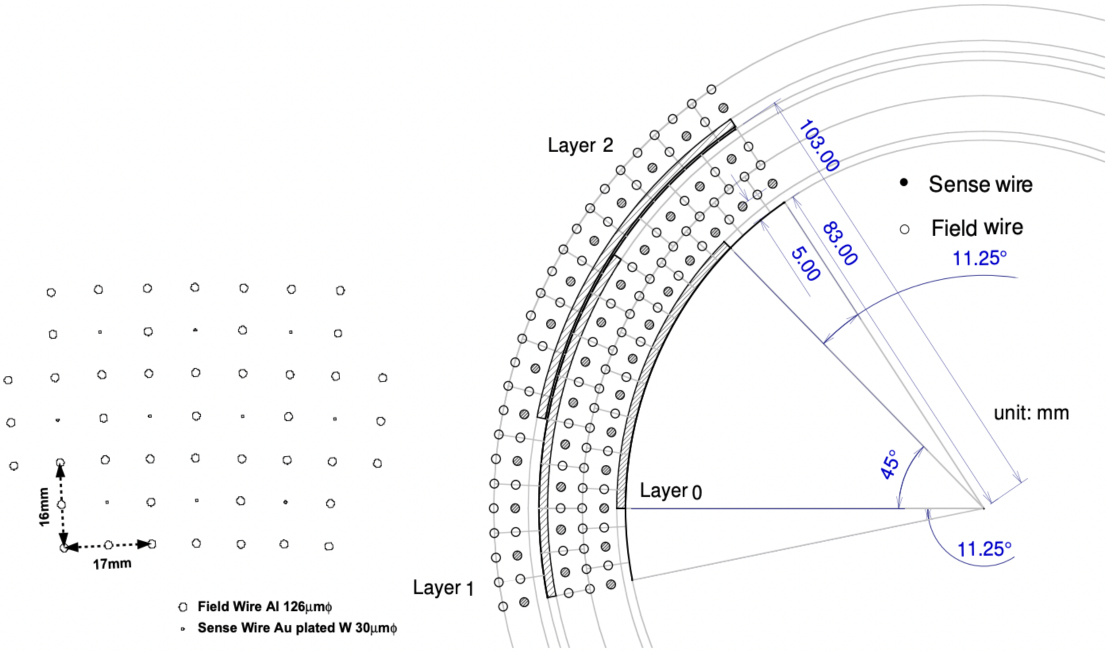

FIGURE 2.9: The cell structure of the CDC [3].

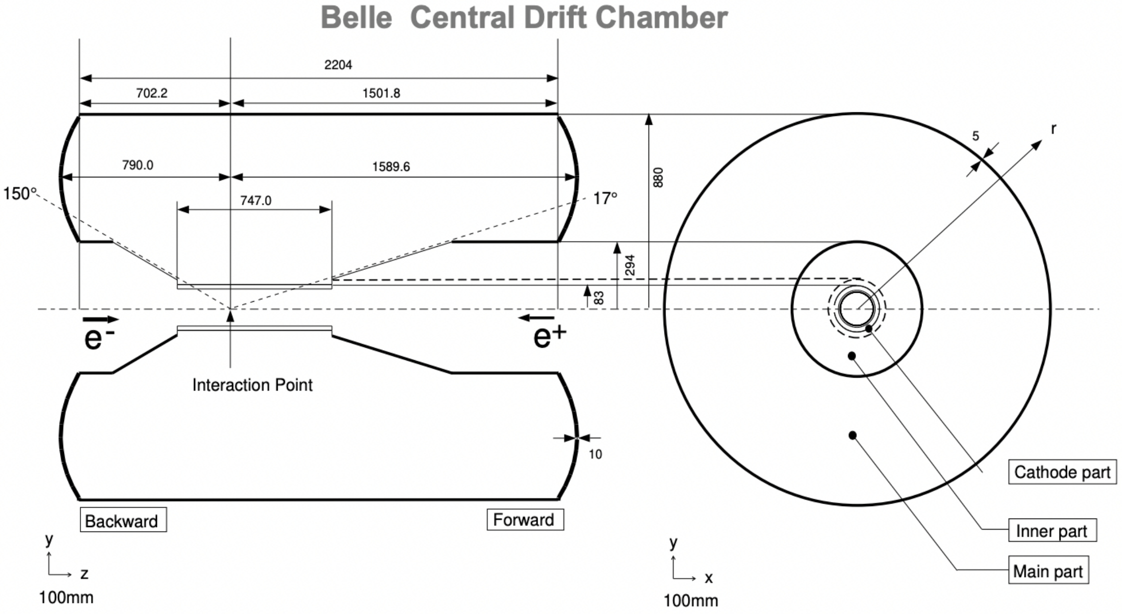

FIGURE 2.10: Cross-sectional view of the CDC [3].

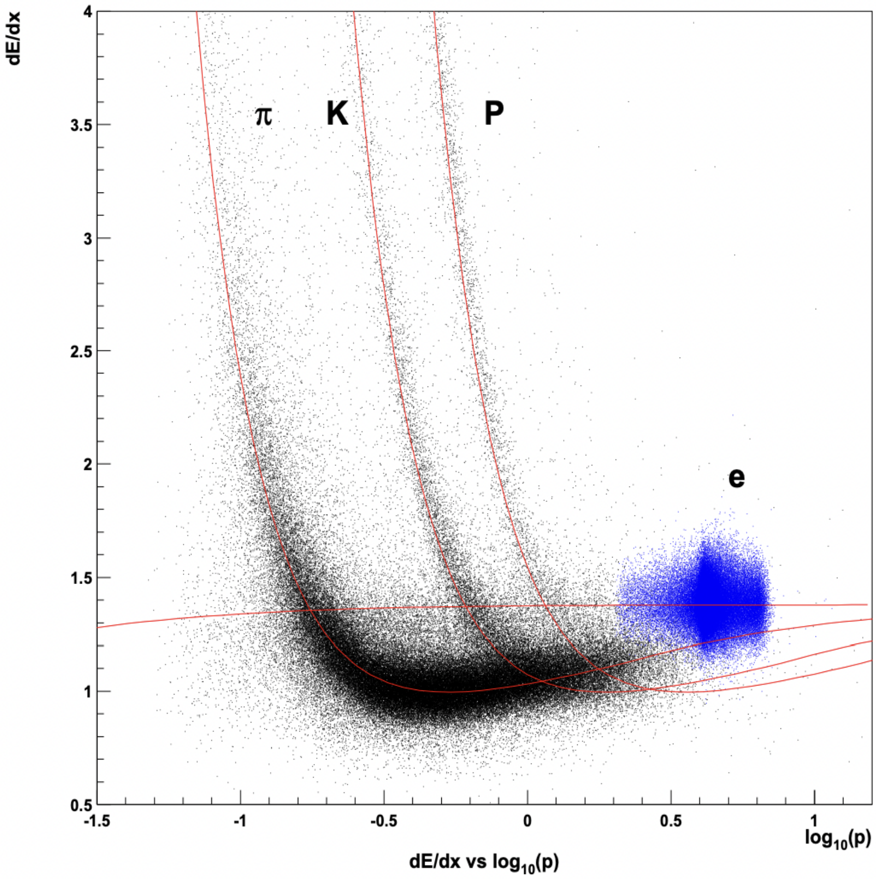

FIGURE 2.11: Variation of dE/dx with momentum of the charged tracks [3].

configuration discussed above, in conjunction with a magnetic field of strength 1.5 T, provided the precise momentum measurement of the tracks. In addition, as mentioned above, the CDC provided valuable inputs on particle identification using the dE/dx information. To extract the energy gradient, dE/dx, CDC was filled with an equal mixture of helium and ethane gases. The gas mixture was chosen to minimize multiple Coulomb scattering of the tracks as it traversed through the medium.

The overall spatial resolution of the CDC detector was determined to be 130 μm. Results from $e^+e^- \to \mu^+\mu^-$ sample estimated the transverse momentum, p_t of the charged tracks in the CDC to be $(1.64 \pm 0.04)\%$ in the momentum range of $4-5$ GeV/c. K_s^0 mass reconstruction from charged pions was used as a calibration sample to determine the low p_t resolution. The low p_t resolution was calculated to be 7.7 MeV/c. The average energy loss, $\langle dE/dx \rangle$ resolution for the charged tracks, was measured to be 5%. Figure 2.11 shows the variation of dE/dx with momenta of the charged tracks. A detailed explanation of the CDC sub-detector and performance studies is described elsewhere (see Section 5 in [4]).

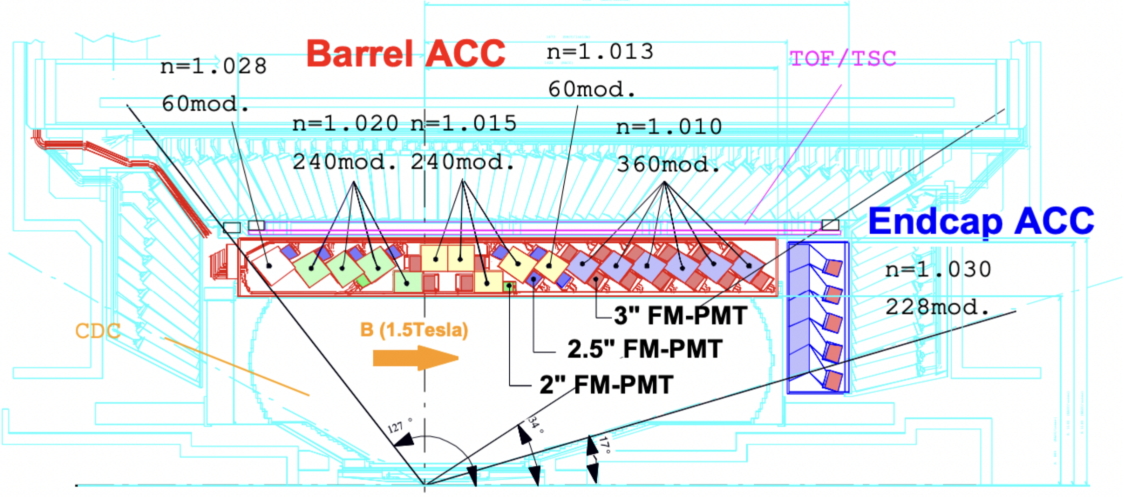

FIGURE 2.12: Cross-sectional view of the ACC sub-detector [3].

2.3.3 Aerogel Cherenkov Counter (ACC)

The primary purpose of the ACC sub-detector was to classify the Kaons (K) and the pions (π). As K and π were two of the most commonly produced final state particles, their identification was of utmost importance for the physics programs at Belle. Particle identification in ACC worked on the principle of detecting Cherenkov radiation from the medium whenever a charged particle traversed through it with speed greater than the velocity of light in the same medium. The choice of the refractive index of the medium is crucial for separating K from π. The aerogel blocks (silica aerogels of hydrophobic nature) at Belle were designed to have a refractive index between 1.01 and 1.03. It ensured that pions could emit the Cherenkov radiation while the Kaons did not; thereby, separating them efficiently.

The aerogel blocks were placed in the barrel and the forward end-cap region of the detector. The barrel region contained 960 modules, divided into 60 cells in the ϕ direction, and the forward end-cap retained 228 modules arranged in 5 concentric layers. The arrangement of the aerogel blocks were such that they projected from the surface towards the IP. In order to detect the Cherenkov radiation, fine-mesh type photo-multiplier tubes (FM-PMT) were designed and operated in presence of the the magnetic field of strength 1.5 T. Figure 2.12 shows the cross-sectional view of the ACC.

The aerogel detector provided particle identification for K/π in the momentum range of $1.2 - 3.5$ GeV/c; thereby, extending the momentum coverage of the detector. A detailed description of the ACC sub-detector is discussed in section 6 of reference [4].

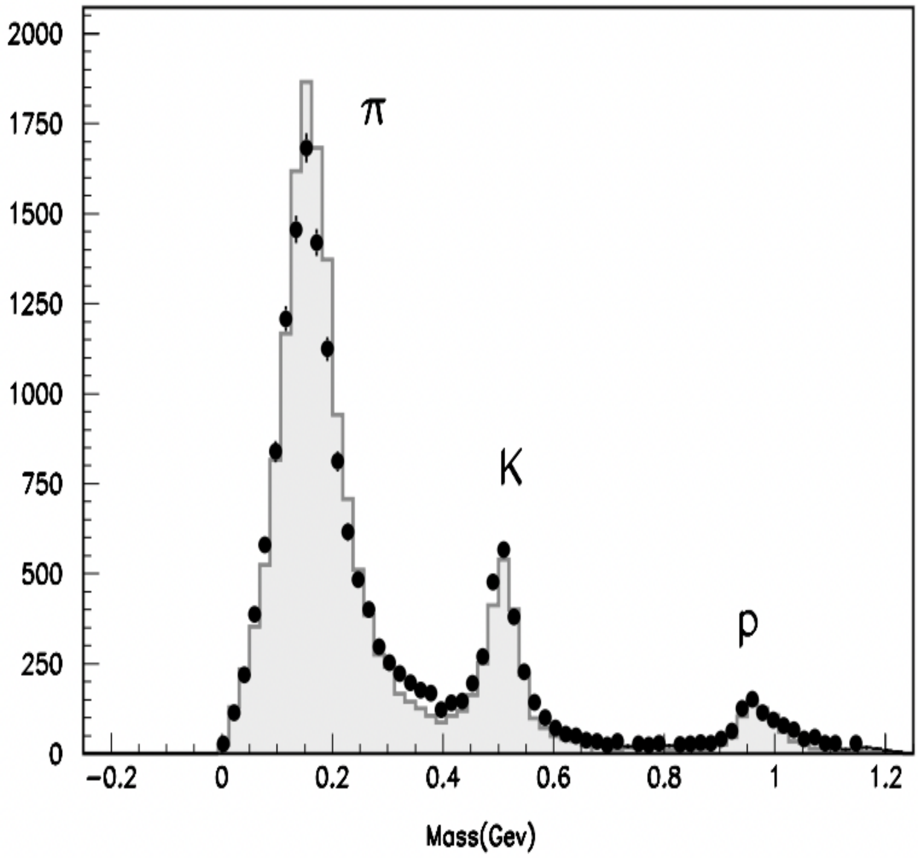

FIGURE 2.13: Distribution of masses of particles using TOF measurements [4].

2.3.4 Time-of-flight counters (TOF)

The TOF counter measured the time-of-flight of the charged particles from the IP and, in conjunction with the momentum information from CDC, provided complete particle identification. It was designed to work efficiently below the momentum range of 1.2 GeV/c. Mathematically, the time-of-flight is calculated using the formula

$$T = \frac{L}{c}\sqrt{1 + \left(\frac{mc}{p}\right)^2},$$

where, L, c, m, and p represent the distance from the IP, the velocity of light, the particle's mass, and the particle's momentum, respectively. The counter measures the time-of-flight and borrows the momentum information from CDC to determine the particle's mass. Fig 2.14 shows the mass distributions from TOF measurements for particle momentum below 1.2 GeV/c.

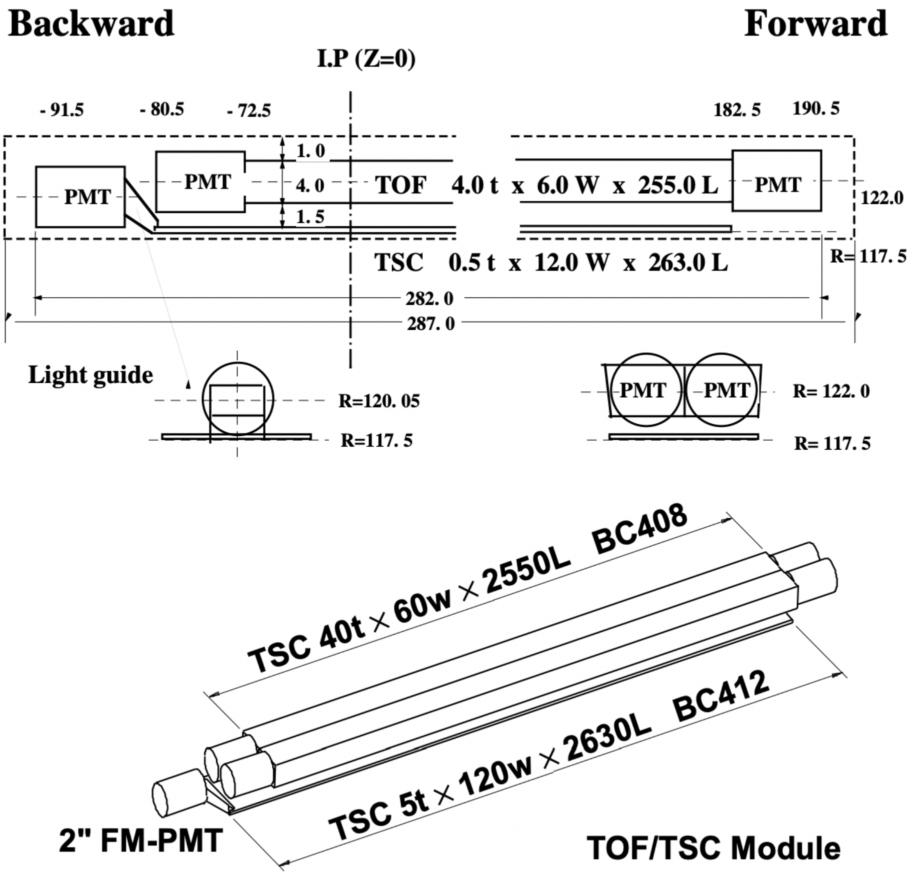

Figure 2.14: TOF sub-detector configuration [3].

Additionally, TOF counter provides information on the fast trigger signals. It was designed to achieve a timing resolution of 100 μs for charged particles below 1.2 GeV/c.

The sub-detector system consisted of two counters: TOF (128 of them) and trigger scintillation counter (TSC, 64 of them). A module was formed from two such TOFs and one TSC to determine the time of flight information of the charged particles from the IP to the point of detection. In total, 64 TOF/TSC modules covered the polar angle, θ in the range of $34° - 120°$ located at a radius of 1.2 m from the IP. The signals from the TOF and TSC were read by FM-PMT. Fig 2.14 shows the configuration of the TOF sub-detector. A detailed description of the TOF sub-detector is described in section 7 of reference [4].

2.3.5 Electromagnetic Calorimeter (ECL)

The ECL sub-detector was an integral part of the Belle detector system developed primarily to measure the energies of photons produced in an event. The photon energies varied over a wide range of values from 0.03 GeV to 4 GeV. The detector was designed to have good performance and high resolution to measure the low and high-energy photons, respectively. It was configured to have a highly segmented array of detector medium to provide good photon separation from high momentum π^0 decays. Additionally, information from ECL on the electron's energy helped in electron identification for several physics analyses.

ECL was designed to have highly segmented CsI(Tl) crystals divided into three regions based on the polar angle, θ: forward end-cap ($12^\circ < \theta < 31.4^\circ$), barrel ($32.2^\circ < \theta < 128.7^\circ$), and backward end-cap ($131.5^\circ < \theta < 157.2^\circ$). The shape of the crystals depended on their angular position; however, on average, they had the following dimensions: the front face pointing towards the interaction point had a smaller area (55×55 mm^2) than the rear face (65×65 mm^2). The size of the crystals were determined based on its ability to contain 80% of the deposited energy by a photon centrally hitting the crystal. Moreover, restrictions on the quality of the crystals were fixed on the condition that non-uniformity in the crystals should not exceed above 9%. The rear face of the crystals were connected with silicon PIN (Positive-Intrinsic-Negative) photodiodes to read the crystals output. In total, there were 8736 CsI(Tl) crystals in the ECL sub-system. Figure 2.15 show the cross-sectional view of the ECL sub-detector.

Additionally, the ECL consisted of trigger cells to provide trigger logic decisions and timing information. Each trigger cell was constituted from $(4 \times 4)16$ ECL cells or clusters. The trigger cell had a finer structure, referred to as trigger sub-cells and comprised of $(4 \times 2)8$ ECL cells. Therefore, each trigger cell contained two sub-cells. Each of the trigger cells operated above a certain threshold of energy recorded by the crystals to provide the trigger logic decisions and the timing information. Further, each of the trigger cells could provide three pairs of time-to-digital converter (TDC) counts corresponding to the rising and falling edges of the signal. To distinguish the signals arriving from the sub-cells in a trigger cell, one of the sub-cells' TDC count was shifted by 1046 counts and had different widths for rising and falling edges. Hence, the timing information from each of the sub-cells became possible. The timing information was crucial to remove the off-time (radiative) Bhabha and $e^+e^- \to \gamma\gamma$ (two photon) events. A detailed description of the implementation of timing criteria to remove the off-time QED backgrounds can be found in the references [28, 29].

As mentioned above, ECL played a seminal role in particle identification (PID), especially electron identification (EID). The following information from the sub-detectors was utilized to identify an electron:

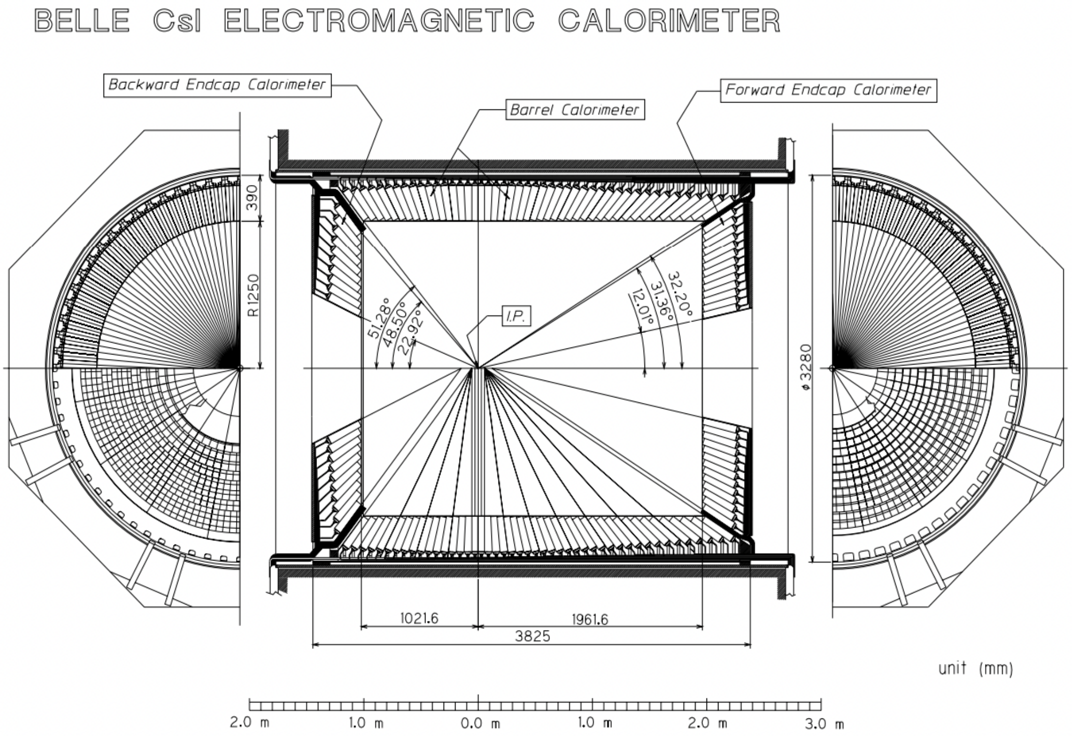

FIGURE 2.15: Cross-sectional view of the ECL sub-detector [4].

- Charged track match: The trace of the charged track in CDC was matched with the energy deposited by the electron in the ECL.
- E/p variable: The ratio of the energy cluster measured by ECL and the track momentum measured by CDC.
- E_9/E_{25} variable: The ratio of shower energy measured by ECL in an array of 3×3 crystals and an array of 5×5 crystals.
- The energy loss (dE/dx) in the drift chamber.
- The response from the ACC.

The energy resolution of the ECL with a threshold enery of 0.5 MeV was calculated using the formula

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

where, the terms in the quadratic sum are contributions from electronic noise, fluctuations due to shower leakage from the crystals, and uncertainty in crystal calibrations. A detailed description of the ECL sub-detector is described elsewhere (see Section 8 in reference [4]).

2.3.6 Long-lived Kaon (K_L^0) and muon detector (KLM)

The KLM sub-detector formed the outermost layer of the Belle detector to detect K_L^0s and muons. It was placed outside the magnetic field and performed particle detection for momentum ranges greater than 500 MeV/c. It consisted of alternating layers of glass-electrode-resistive plate counters (RPC) and iron plates. There were 15 and 14 layers of RPCs in the KLM's barrel and forward/backward end-caps, respectively. At the same time, 14 layers of iron plates were installed in the KLM detector's barrel and forward/backward end-caps. The interaction length played an essential role in detecting K_L^0 mesons. The ECL and the KLM sub-system provided 4.7 interaction lengths to detect the K_L^0 mesons. The interaction of the K_L^0s with the detector medium of the ECL and the RPC/-iron plates produced showers of ionizing particles. The location of the shower production determined the direction of the K_L^0 mesons. The muons are highly penetrating and, based on their range and transverse scattering with the RPCs/iron plates, provided discrimination with other charged hadrons ($\pi^{\pm}$, $K^{\pm}$).

The RPC modules used for the barrel and the forward/backward end-caps had peripheral differences. The barrel RPCs were constituted of superlayers. Each superlayer consisted of two consecutive RPCs placed between strips (determining θ and ϕ) and ground layers. Additionally, the individual RPCs in a superlayer were connected (top and bottom) by copper strips to provide information on particle tracking. The RPCs, in turn, were filled with gas mixtures in between electrodes. Each of these RPCs provided a high degree of resistivity (typically, $> 10^{10}$ Ωcm) for particle location tracking and time. Figures 2.16 and 2.17 show the cross-sectional view of a KLM superlayer and a schematic diagram of the barrel and forward/backward end-caps RPC of the KLM, respectively. A detailed account of the KLM detector is described in section 9 of reference [4].

2.3.7 Extreme Forward Calorimeter (EFC)

The EFC sub-detector provided energy measurement capabilities in the polar regions of the forward ($6.4^{\circ} < \theta < 11.5^{\circ}$) and backward ($163.3^{\circ} < \theta < 171.2^{\circ}$) part of the detector. Additionally, it was used to monitor beam and luminosity parameters for the KEKB collider and the Belle detector, respectively. It also provided tagging information in physics analyses involving two photons. For beam energy between $1-3$ GeV, the energy resolution of the EFC was found to be in the range of $7\% - 10\%$ [30].

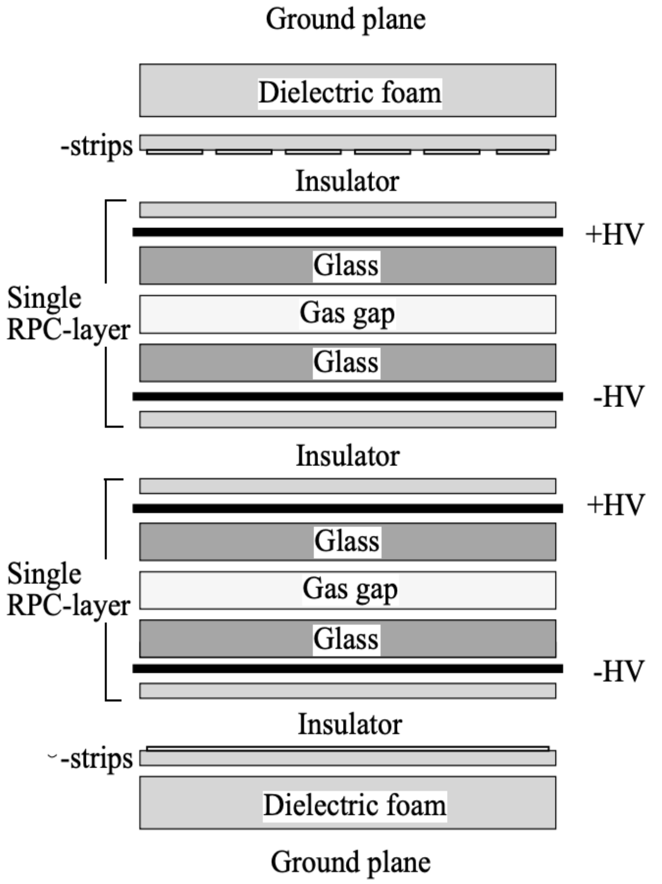

FIGURE 2.16: Cross-sectional view of a KLM superlayer [4].

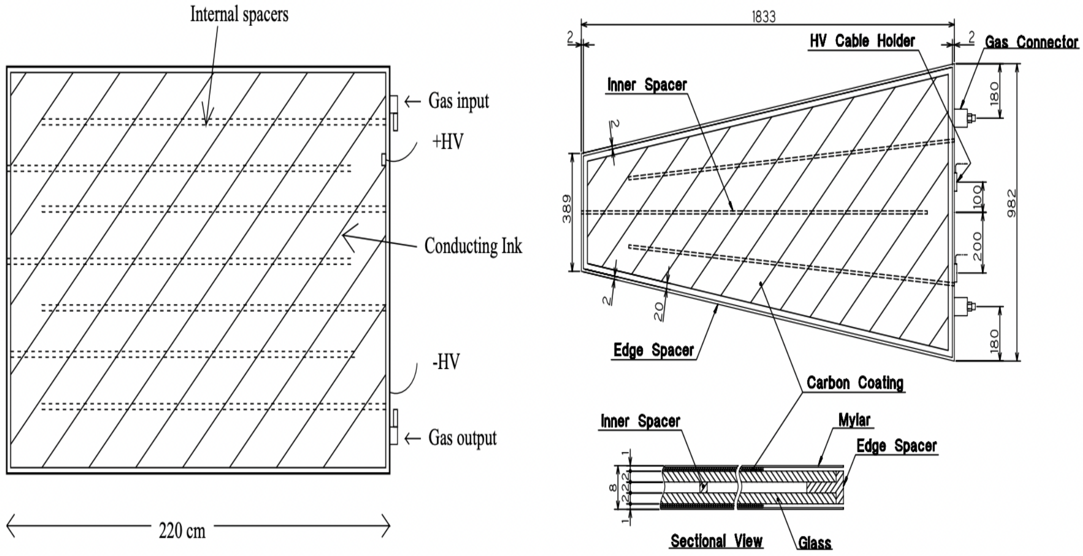

FIGURE 2.17: Schematic diagram of barrel (left) and forward/backward end-caps (right) RPC [4].

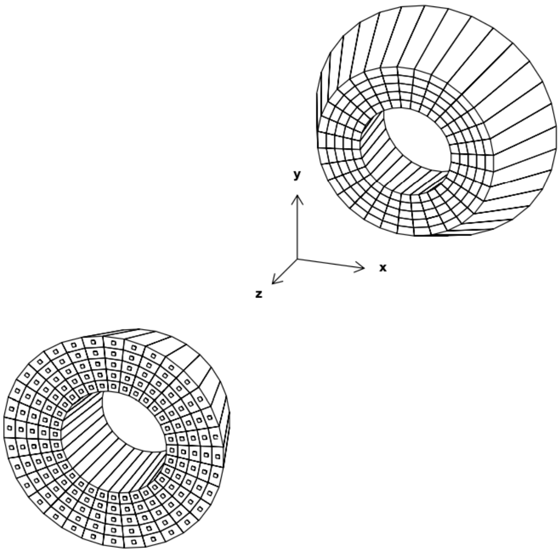

FIGURE 2.18: A 3D view of the EFC sub-detector [4].

The crystals of the EFC sub-detector were composed of Bismuth Germanate, $Bi_4Ge_3O_{12}$ (BGO) to handle large radiation fluxes. It was segmented into 5 and 32 in the θ and ϕ directions, respectively, for both the end-caps to provide better position resolutions. The crystals were placed inside a steel case of thickness 1 mm with an inner bore to provide space for the beam pipe. Figure 2.18 refers to three-dimensional view of the EFC sub-detector. A detailed description of the EFC can be found in section 3 of reference [4].

2.4 Trigger

Triggers are selection criteria (logic/conditions) designed to filter events essential to achieve an experiment's goals. Data arriving at the sub-detectors from an e^+e^- collision will be stored by the Data Acquisition System (DAQ) (to be discussed later in the chapter) based on whether the event is accepted by the trigger system or not. At Belle, several physics processes were produced; an example of such processes at $\Upsilon(4S)$ resonance is described in Table 2.2 with their cross-sections and the rate at which the triggers were designed to operate at an instantaneous luminosity of $\mathcal{L}_{\text{inst}} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

The trigger system at Belle [31, 32] was designed to withstand high beam background rates owing to high beam currents. The trigger logic were carefully chosen to limit the background rates within the tolerance range of the DAQ, whose maximum expected rate is $\sim$ 500 Hz. Careful considerations were also placed to increase the efficiency of the

TABLE 2.2: Cross-sections and trigger rates with $\mathcal{L}_{\text{inst}} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$ for physics processes at $\Upsilon(4\text{S})$.

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

*Pre-scaled by a factor 1/100

†with a restricted condition of $p_t \geq 0.3$ GeV/c

events with physics interests. Additionally, the trigger logic was not made too restrictive to increase the efficiencies of various physics processes and acclimatize with the changing accelerator conditions.

Figure 2.19 shows the conceptual design of the Belle trigger system. Belle developed a two-level trigger system to filter out uninteresting events: level-1 (L1) and level-3 (L3) trigger. L1 trigger system was a hardware trigger composed of several sub-triggers from the detector components, as shown in Figure 2.20, while L3 was a software trigger. Additionally, level-0 (L0) and level-4 (L4) triggers were also present; however, they were not a part of the Global Decision Logic (GDL). Level-0 trigger served as an inhibitor for the outputs from the shaping circuits of SVD upon receiving triggering from the TOF sub-detector.

2.4.1 Level-1 (L1) trigger system

The L1 trigger system acted on the information delivered by five sub-detectors: CDC, TOF, ECL, KLM, and EFC. The triggers generated from the sub-detectors were of two types: track triggers and energy triggers. The track triggers were compiled based on the position coordinates, $r - \phi$ and $r - z$, while the energy triggers operated on the basis of total energy deposited and the number of cluster hits.

The CDC sub-detector provided the track trigger signals based on $r-\phi$ and $r-z$ coordinate information. It ensured that the tracks originated from IP and not from background tracks.

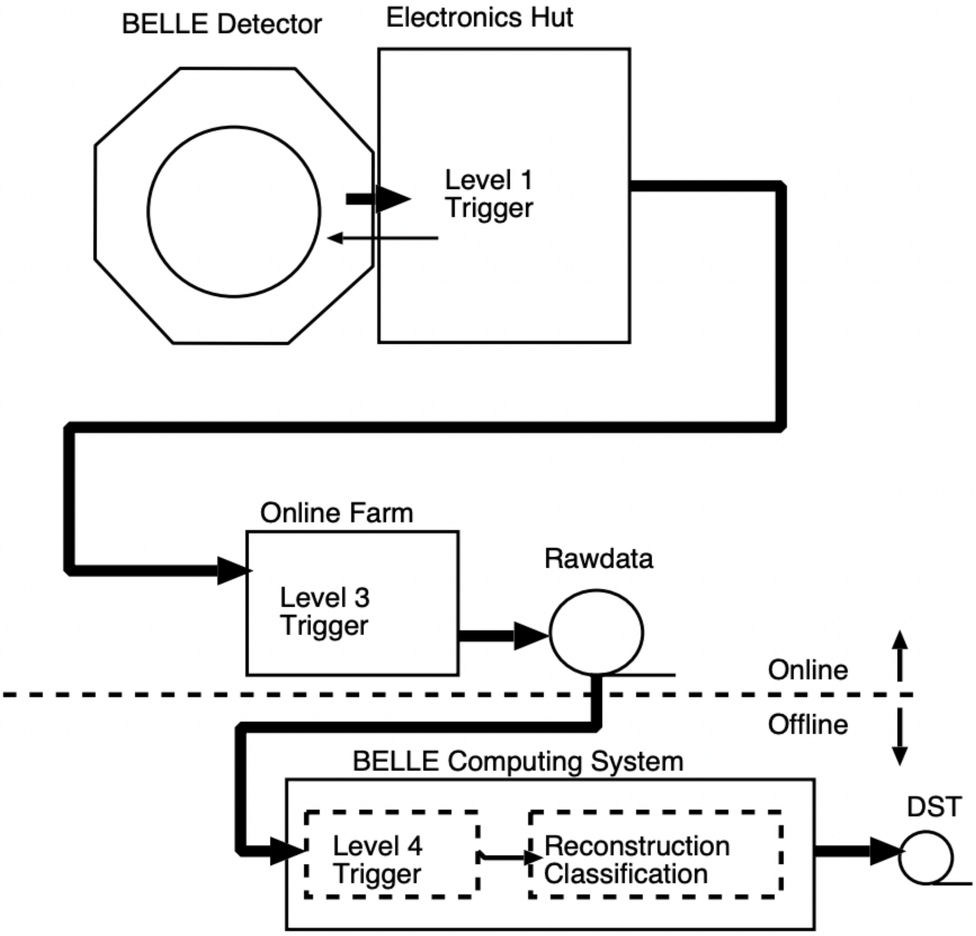

FIGURE 2.19: A conceptual design of the Belle trigger system [4].

The $r - \phi$ information was provided by the axial wires, while the three cathode layers and the stereo wires provided the z information.

The TOF trigger was a crucial component of the L1 trigger system. It delivered signals based on event timing and provided information on the number of hits (multiplicity) and topology. In addition, it provided a prompt trigger signal (L0) to hold the output of the SVDs shaper circuits. The timing signals provided by the TOF were also used in ECL and CDC triggers.

The ECL trigger system provided trigger information based on energy deposition on the ECL trigger cells and the number of cluster hits both for the neutral particles and charged tracks. ECL trigger information from the total energy deposition helped separate hadronic events from other kinds of physics processes produced in the same event. The cluster

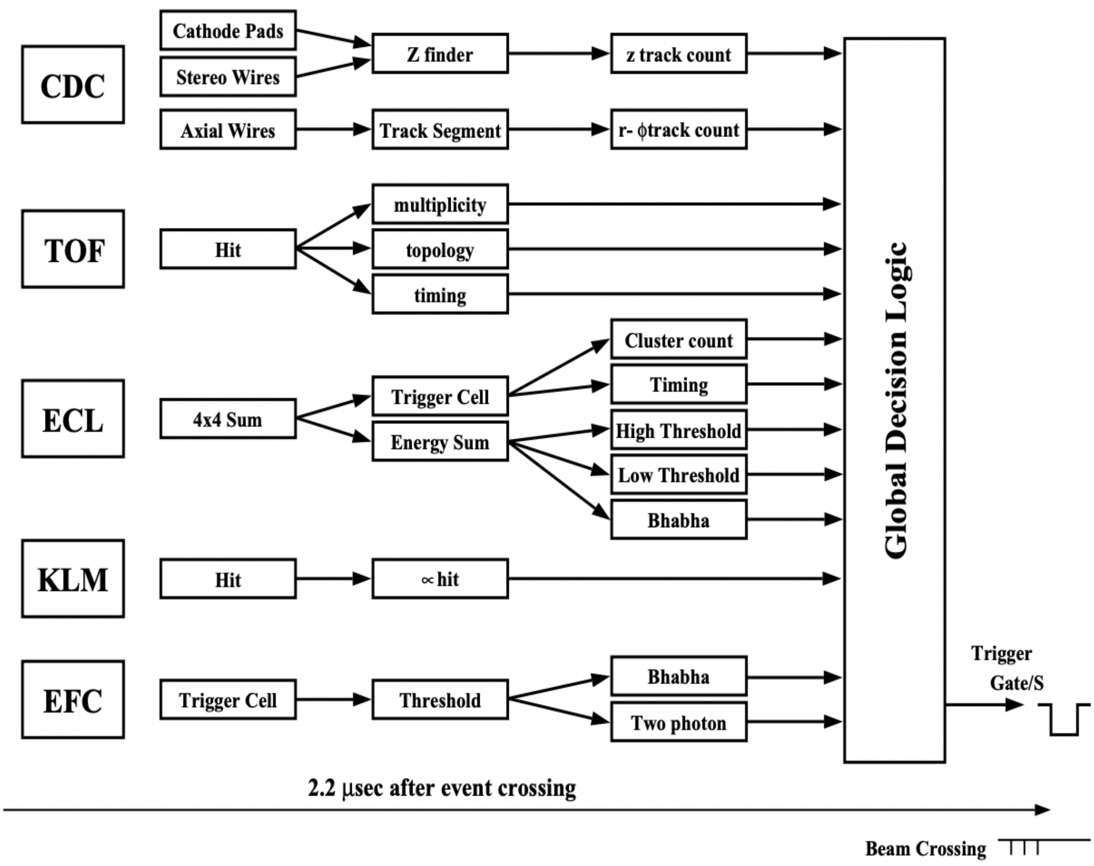

Figure 2.20: A conceptual design of the Belle Level-1 trigger system [4].

counting trigger was crucial to separate hits in the ECL cluster, which were sometimes counted more than once due to a wide spread of the cluster shower.

The KLM trigger system was designed to provide information on events with muons. The trigger information was collected using the barrel, forward and backward end-caps. Trigger efficiency studies were performed using muon pairs with an average efficiency of more than 98%. (see Section 11.4.3 in [4])

The EFC trigger information was required for *Bhabha* and *two-photon* studies. Therefore, the trigger signals from EFC were derived from energy deposition and cluster location in the BGO crystals. To identify a *Bhabha* event, a coincidence of electromagnetic showers from forward or backward end-caps of the ECL was used as a trigger, while for the *two-photon* processes, the signal from a single electromagnetic shower together with track information from CDC or cluster information from ECL was used.

The trigger signals received from the sub-detectors containing information on various parameters such as, z-track count, $r-\phi$ track count, cluster count, timing, high/low threshold, *Bhabha* and *two-photon* events were used as inputs to a central logic system also called as the Global Decision Logic (GDL). GDL processed the information received from the sub-detectors to analyze an event and provided the final triggering signal from the L1 trigger system. The time frame within which the GDL need to take a decision whether to retain an event for further processing was 2.2 μs. The average trigger rate was maintained at around 200 Hz and a cumulative efficiency of more than 99.5% was achieved by the trigger system (see Section 11.7 in [4]).

2.4.2 Data Acquisition System (DAQ)

The DAQ at Belle was designed to acquire and store data after receiving trigger signals from the GDL and detector sub-systems. It was equipped to read data at a rate of roughly 500 Hz with the help of a distributed-parallel system. Figure 2.21 shows the overview of the Belle DAQ. The electrical signals received from the sub-systems (CDC, ACC, TOF, ECL, EFC) were read out using a charge-to-time (Q-to-T) and TDC. While the signals from SVD were read out using a flash analog-to-digital converter (ADC) and a VERSA-Module Euro card (VME) bus, essentially a master-slave computer architecture, the signals from KLM, trigger logic from the sub-systems and the GDL were digitized using the TDC module only. These outputs were combined and sent to an event builder, which converted the sub-detector information to event-based information. A time-division multiplexing (TDM) module and a sequence controller were used to provide the timing information to the VMEs to initiate the readouts from the sub-detectors. An online computer farm received the output from the event builder, which provided additional filtering to the event-based outputs. These filtered data were then transferred to a storage system for further processing.

2.4.3 Level-3 (L3) and Level-4 (L4) trigger system

The L3 was a software trigger and a part of the online computing farm (as shown in Figure 2.19). The L3 trigger contained logic that reduced backgrounds from the data selected by the L1 trigger. It represented the last set of online triggers, which took an important decision to keep data for an event or reject it. With an efficiency of over 90%, the L3 trigger rejected 50% of the events selected by the L1 trigger. The L1 trigger considerably reduced the unwanted load on the storage devices.

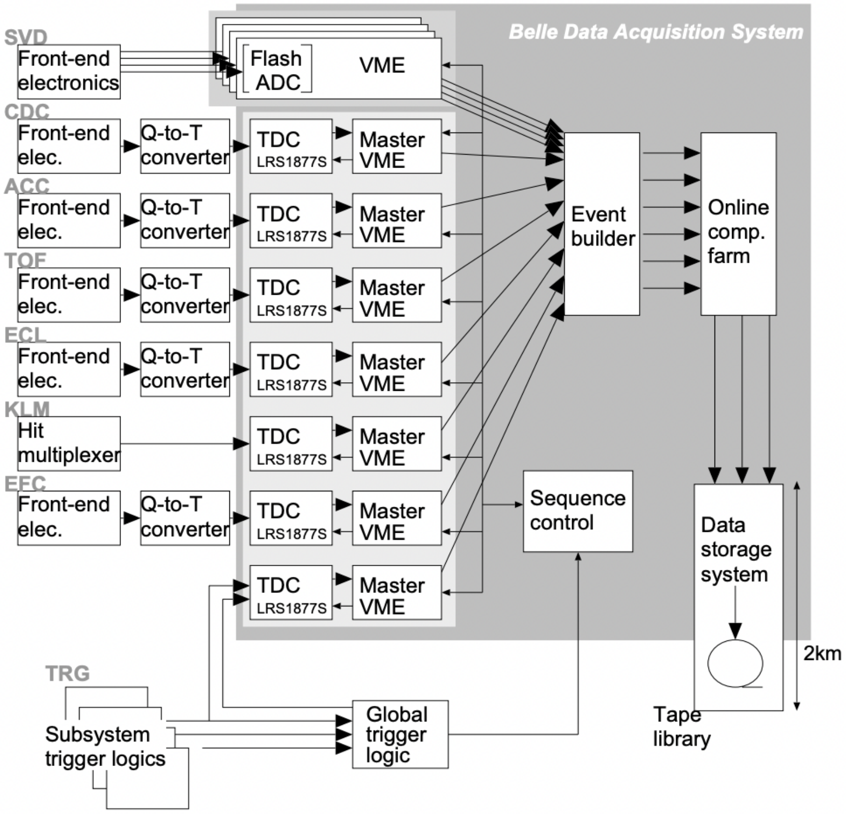

FIGURE 2.21: An overview of the Belle DAQ [4].

The L4 trigger [33] is an offline software trigger which formed a part of the Belle computing system. The role of the L4 trigger was to reduce uninteresting events from the stored data. However, the data lost at this stage were recoverable, unlike the L3 trigger, as L4 filtered out events from stored data. This represented the final trigger before a full event reconstruction. The selection criteria applied in L4 trigger are as follows: Δz resolution of 3 cm, Δr resolution of 1 mm, 80% efficiency for barrel tracks with transverse momentum, $p_t > 300$ MeV/c.

The efficiency of the L4 trigger in selecting the pure hadronic events was close to 100%, while for the rest of the events, the average purity was measured to be around 77%.

2.4.4 Data preparation and software framework

The events (raw data) that passed the L4 offline trigger were reconstructed and stored as Data Summary Tables (DST). The information stored in the DSTs was then converted into physics objects in the form of four vectors $(x, \vec{p})$. As the Belle physics programs were quite broad, event classification and skimming became necessary to provide the analysts with data of their interest. Specific selection criteria were developed, which selected the events of interest from various backgrounds and stored them as skimmed data. Skimming reduced the size of the data and saved time and computing resources (more about skimming criteria are discussed in the next chapter). The data sets are stored for various skims and are referred to as mini-DST (`mdst`), as only the minimum required information is stored. On average, the size of a hadronic event at the `mdst` stage is 40 KB.

The Belle Analysis Software Framework (BASF) is an offline software framework designed to process events and handle reconstruction and analysis algorithms from the users in the form of modules written in `C++`. The event architecture, the format for the input/output (I/O) of the data, and to process the events in parallel, a Symmetric Multiple Processor (SMP) was developed, which was written in `C/C++` and simulated using `GEANT3` [34]. Data transfer among the modules within BASF is performed using `PANTHER`, an event and I/O management package developed by the Belle collaboration. The BASF is the entry point to the Belle experiment for the end users.

2.5 Physics of Υ resonances

In this section, we briefly discuss the physics of the Υ resonances and outline the production of hadronic events at $\Upsilon(5S)$ resonance.

2.5.1 Hadron production in e^+e^- annihilation

The understanding of the production of hadrons in an e^+e^- annihilation through the formation of quark-antiquark ($q\bar{q}$) pairs has been significant in the development of Quantum Electrodynamics (QED) and QCD. The e^+e^- annihilation, which proceeds via a single photon mediator, gives rise to a quark-antiquark pair by coupling to their charges. The quark-antiquark pair later hadronizes to form mesons.

Experimentally measurable quantities such as the cross-sections can give insightful information about the various QCD processes involved in these interactions. However, because of many gluonic interactions, theoretical calculations involving higher-order perturbative terms become complicated. This, in turn, leads to difficulty in the comparison of theoretical results and experimental measurements. As a result, the hadronic cross-section is not directly measured.

On the other hand, the e^+e^- annihilation can also give rise to $\mu^+\mu^-$ pairs with its cross section proportional to that of $e^+e^- \to q\bar{q}$, primarily, because the coupling of photons to the $\mu^+\mu^-$ or to $q\bar{q}$ pair depends only on the charge of the particles involved. As $e^+e^- \to \mu^+\mu^-$ is a well-known process in QED with lesser theoretical uncertainties, the hadronic cross section of $e^+e^- \to q\bar{q}$ is calculated in terms of $e^+e^- \to \mu^+\mu^-$ [35, 36]. Mathematically,

$$R = \frac{\sigma(e^+e^- \to q\bar{q} \to hadrons)}{\sigma(e^+e^- \to \mu^+\mu^-)}, \tag{2.1}$$

where, σ represents the cross-section of the processes involved.

Moreover, the hadronic cross-section, as shown in Equation 2.1, comprises two parts: the resonant and the non-resonant cross sections corresponding to the resonant and non-resonant states. They are briefly described as follows:

1. **Resonant states** : The resonant states are quasi-bound states of $q\bar{q}$ pairs such as $u\bar{u}/d\bar{d}$ (ρ^0, ω), $s\bar{s}$ (ϕ), $c\bar{c}$ (J/Ψ) and $b\bar{b}$ (Υ). They are produced once the quark-antiquark pair reaches the threshold energy of the quarks and protrude from the non-resonant background. At higher CM energies, more than one resonant structure corresponding to a $q\bar{q}$ pair is also observed, as in the case of $\Upsilon(n\mathrm{S})$ systems, where $n = 1, 2, 3$ and so on.

2. **Non-resonant states** : The non-resonant or continuum states are produced by $e^+e^- \to q\bar{q}$, (where $q = u$, d, s, c and b) giving rise to hadrons of the type $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$ and $b\bar{b}$. They are produced once there is available energy for the respective quarks and are different from the resonances in that the quark-antiquark pair do not form quasi-bound states. An important kinematic feature of the continuum events is the formation of two back-to-back jet-like structures in the CM energy frame of the electron and the positron. The jet axis align along the direction of the quark and the antiquark pairs. The hadrons formed from the quark (antiquark) jets carry the momentum of their constituent quarks and thus have directional topology [37, 38].

A feature common to both the resonant and the non-resonant states is the change in the cross-section with CM energy. The total hadronic cross section varies as $1/E_{cm}^2$, where E_{cm} is the CM energy in an e^+e^- annihilation. Thus, the total hadronic cross-section decreases with an increase in the CM energy. Figure 2.22 shows the production cross-section of the Υ system at various CM energy.

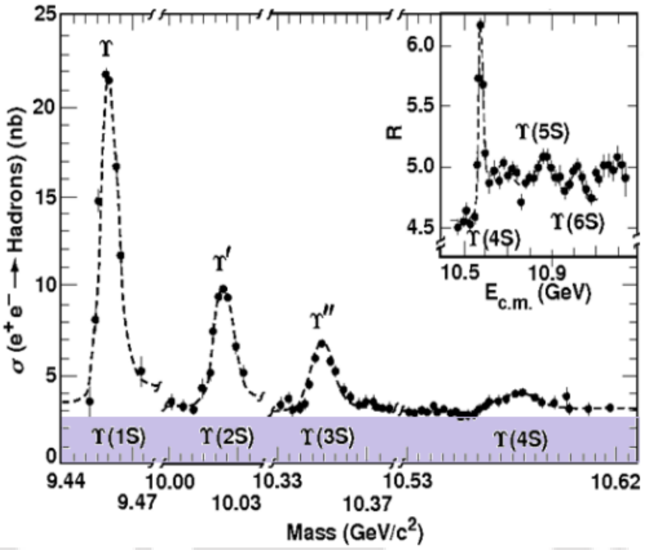

FIGURE 2.22: Production cross sections of Υ systems at various CM energy of e^+e^- collisions. Image courtsey CUSB collaboration

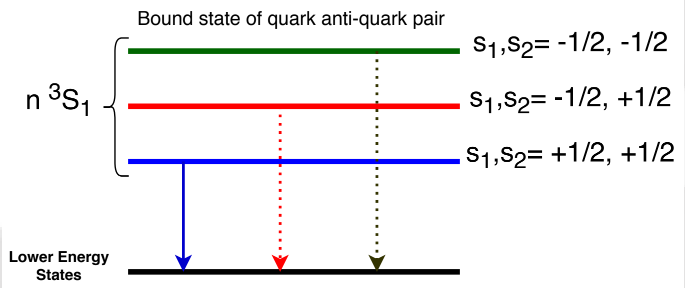

FIGURE 2.23: A diagrammatic representation of the spectroscopic levels of an Υ system.

Among the resonant states, the Υ resonances are important for studying mesons containing b quarks. The $\Upsilon(n\text{S})$ system is composed of $b\bar{b}$ quarks and is present above the CM energy of $\approx$ 9.45 GeVs. The spectroscopic illustration of a quasi-bound $\Upsilon(n\text{S})$ system is shown in Figure 2.23. The low-lying narrow Υ resonances, Υ(1S), Υ(2S) and Υ(3S) do not have sufficient energy to create B meson pairs and annihilate primarily through three gluon production along with competing QED processes as shown in Figure 2.24 [39].

The Υ(4S) resonance has sufficient energy to create a $b\bar{b}$ pair with simultaneous production of the light quarks, u or d. This gives rise to the production of $B\bar{B}$ mesons as pairs at Υ(4S) resonance, as shown in Figure 2.25. At still higher CM energy of $\approx$ 10.86 GeV, the Υ(5S) resonance is formed, which is used to study B_s^0 mesons. A brief description of the production of B_s^0 mesons is discussed in subsection 2.5.2.

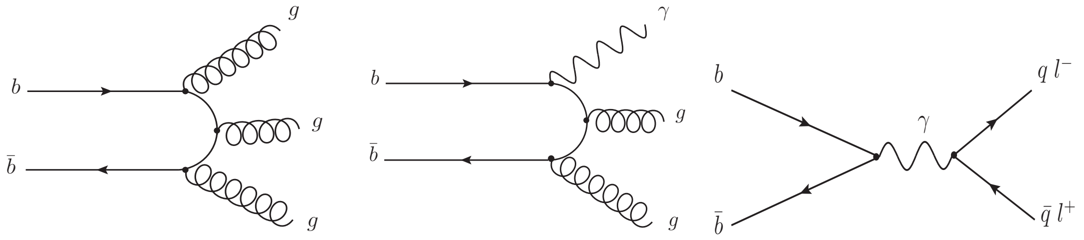

FIGURE 2.24: Various Feynman diagrams depicting the annihilation of the first three resonances, $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$.

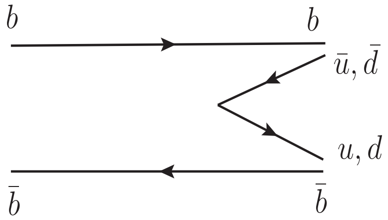

FIGURE 2.25: Feynman diagram for the production of $B\bar{B}$ pairs at $\Upsilon(4S)$ resonance.

2.5.2 Hadronic events near $\Upsilon(5S)$ resonance

The hadronic events near $\Upsilon(5S)$ peak can be classified into three groups, namely, $\Upsilon(5S)$ resonant events, b continuum and u, d, s, c continuum events. The decay products observed near the $\Upsilon(5S)$ peak consists of B_s, B_d and bottomonium systems. As the quark content of the decay products is again b quarks, the source of their parent particle can be anywhere from the $\Upsilon(5S)$ resonant events or from b continuum events. Thus, this distinction cannot be made, and a source of uncertainty prevails. The non-hadronic events near $\Upsilon(5S)$ peak are removed at the triggering stage and adoption of hadronic skim selection criteria (discussed in the next chapter). The hierarchy of hadronic events near $\Upsilon(5S)$ peak is shown in Figure 2.26.

The production cross-section of the $\Upsilon(5S)$ resonance near the e^+e^- CM energy of 10.86 GeV is $\sigma_{b\bar{b}}^{\Upsilon(5S)} = (0.340 \pm 0.016)$ nb [8] and the fraction of $b\bar{b}$ events giving rise to B_s^0 production modes, $B_s^{(*)0}\bar{B}_s^{(*)0}$, is measured to be $f_s = (0.201 \pm 0.031)$ [1]. There are three kinematically allowed modes of production of B_s^0 mesons near the $\Upsilon(5S)$ resonance: $B_s^{*0}\bar{B}_s^{*0}$, $B_s^0\bar{B}_s^{*0}$ or $B_s^{*0}\bar{B}_s^0$, and $B_s^0\bar{B}_s^0$. The production fractions of B_s^0 meson from the former two are $(87.0\pm1.7)\%$ and $(7.3\pm1.4)\%$, respectively [8], while the remaining fraction is from the $B_s^0\bar{B}_s^0$ mode. The B_s^{*0} decays to B_s^0 by radiating a low energy photon which is usually not identified due to its poor reconstruction efficiency. The number of events with $B_s^0\bar{B}_s^0$ is, therefore, estimated to be $N_{B_s^0\bar{B}_s^0} = 121.4\ \text{fb}^{-1} \cdot \sigma_{\text{b}\bar{\text{b}}}^{\Upsilon(5S)} \cdot f_s = (8.30 \pm 1.34) \times 10^6$.

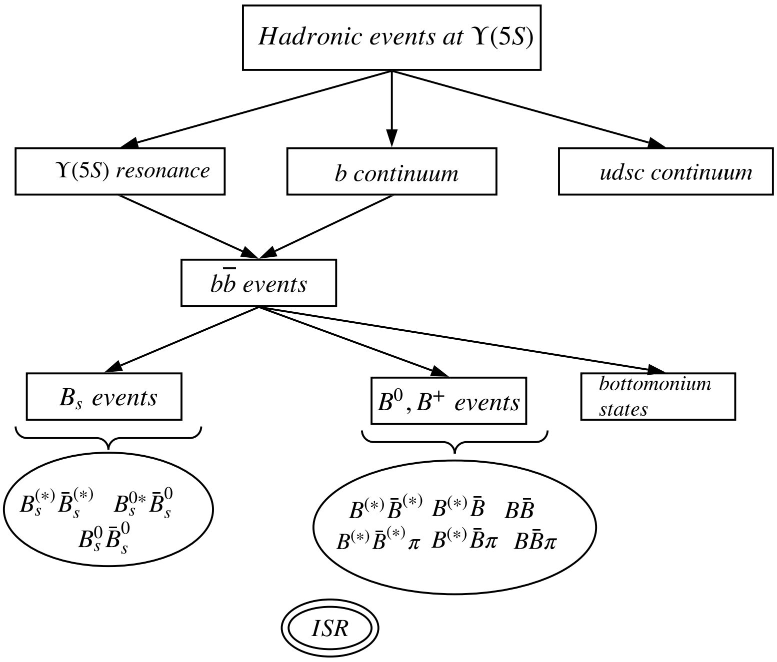

FIGURE 2.26: Hadronic events near Υ(5S) peak.

2.6 Analysis strategy

We briefly outline the strategy adopted for this analysis as follows:

1. To perform blind analysis and improve the reconstruction efficiency, we generate and simulate signal and background Monte Carlo events.
2. We reconstruct the B_s^0 candidates from two neutral pions, each of which, in turn, is reconstructed from two photons. Various pre-selection criteria are applied at this stage to remove the uninteresting events.
3. We then perform signal and background Monte Carlo studies to identify the signal from the background. We employ multivariate analysis to suppress the continuum background ($e^+e^- \to q\bar{q}$, $q = u, d, s, c$) by exploiting the differences in the topological distributions between them and the signal.
4. Probability distribution function parametrizations are performed on the model variables, which are eventually used to extract signal yields in data.

5. Ensemble study based on a large MC sample of pseudo-experiments are used to understand the stability, linearity, and possible correlations in the developed model.

6. We then perform a control sample study using the control mode, $B_d^0 \rightarrow \pi^0\pi^0$ at Υ(4S) to validate the analysis strategy. Further, data-MC comparison studies are performed on data side-bands at Υ(5S) resonance.

7. At the end, we extract the signal from the data and estimate the branching fraction and the upper limit for the decay channel.

A detailed description of the steps are discussed in the subsequent chapters.

Summary

This chapter discussed the history of the Belle experiment, the KEKB collider, and the Belle detector. We briefly explored the trigger, the Belle analysis software framework, and the data acquisition system developed for the Belle experiment. We also discussed the physics of the Υ resonances produced at various CM energies, particularly looking at the Υ(5S) resonance, which is the source for the production of the B_s^0 mesons and concluded the chapter by outlining the strategy adopted for this analysis. The next chapter introduces the event generation and simulation studies to perform the analysis.

Chapter 3

Event generation and Monte Carlo studies

In this chapter, we shall describe the methodology of event generation and its simulation to perform Monte Carlo studies. We discuss the preliminary event selection criteria applied to select the correct event type for this analysis. We shall also discuss the reconstruction of our signal B_s^0 candidates from the neutral pions and apply selection criteria to separate them from the background. We will further study the background using Monte Carlo samples and develop strategies to suppress them.

3.1 Blind analysis

A blind analysis is a technique often adopted in particle physics to leave out the experimenter's biases from the measurement. Experimenter's biases can include one's perception or inclination to apply selection criteria around the signal region (region of interest) or towards theoretical predictions or prior measurements [40].

In our analysis, we have applied the *hidden signal box* type blind analysis technique [40], where the signal region is kept hidden until the selection criteria are finalized. This type of blind analysis technique is particularly helpful for rare processes as we expect to see very few signal events over a large background. We perform our analysis on Monte Carlo data sets (to be discussed in the next section) and then apply the same analysis strategy to data.

3.2 Monte Carlo

The need for Monte Carlo data is twofold. First, to perform blind analysis and second, to reconstruct the signal events to improve the efficiency of our reconstruction algorithm. To perform these steps, we need to have a simulated set of data whose behaviour is similar to real data. These simulated data sets behave like real data and are called Monte Carlo (MC) data sets.

For our analysis, we have simulated signal, as well as background MC events.

3.3 Event generation and detector simulation

In this section, we describe the procedure to generate MC samples and eventually their simulation to model effects similar to data. However, before we proceed, let us first understand the term "event" in the context of this analysis. It is defined as:

- Event: The whole chain starting from $e^+e^- \to \Upsilon(5\mathrm{S}) \to B_s^0\bar{B}_s^0, B_s^0 \to \pi^0\pi^0, \pi^0 \to \gamma\gamma$ and ending at the final state particles (FSP) (here, for example are the γ's) is considered to be an event.

Event generation is an analysis-specific process and, as such, is exclusive to the decay channel studied. There are different event generators available depending on the physics processes involved. Frequently used event generators in the high-energy physics community include QQ [41], JETSET [42], and EvtGen [43]. Belle, for most of the analyses, uses EvtGen because of its high flexibility and easy interfacing with other event generators.

EvtGen is a dynamic package which can connect to other packages for efficient event generation in the form of modules. The events are generated based on decay amplitudes for each node of the decaying tree for the entire decay chain. The advantage of using decay amplitudes is that the algorithm which works on the accept-rejection method need not recreate the entire decay chain for processes where decay amplitudes are not physical or known. Thereby significantly reducing computing time and resources.

Belle uses different packages along with the EvtGen for generating other processes such as `CTOY` for cosmic ray muons (for detector study), `SG` for single tracks, `BHLUMI` for lepton pairs, `TAUOLA` for τ decays, `KK` for fermion pairs and `AAFH` for two-photon production of fermion pairs. The details of these packages can be found in reference [44].

The generated events do not represent the actual physical particles because detector effects are not superimposed on them. For simulating the detector environment, Belle uses `GEANT3`, which is a `FORTRAN`-based package. `GEANT3` is used to incorporate effects on the generated data from different parts of the detector to look real. They are stored in the master database by experiment number and run number. These simulated data sets are then digitized, and additional information such as “truth match” is added for analysis purposes. Truth matching compares the generated data to the reconstructed ones (to be discussed in section 3.7). This information is added as modules (algorithms written for a particular task) to the generated events. In addition, background information for each part of the detector is superimposed to make them closer to reality. Thus, at the end of the simulation, we have a data set that can behave like real data arriving from e^+e^- collisions. A schematic representation of the steps involved in event generation and detector simulation at Belle is shown in Figure 3.1.

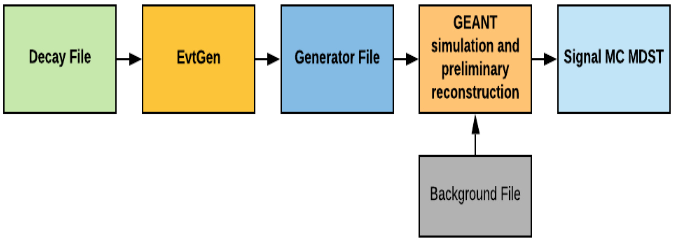

Figure 3.1: Block diagram of the steps involved in event generation and simulation

The first step in event generation is to create a decay file [1], which describes the signal decay chain as well as other generic decay chains (decay chain other than the signal) according to physics models. The decay models that are being used to generate the decay, $B_s^0 \to \pi^0\pi^0$ are summarized in Table 3.1.

In the **PHSP** model, the energy and the momentum of the daughter particles depend on the phase space of the mother particle. No such other constraints are being placed apart from the kinematical ones. The **VSS** model generates decays in which a vector particle disintegrates into two scalar particles. The model utilizes the spin information of the particles for their generation. The **VSP-PWAVE** decay model is used when the initial particles are vector-like while the daughter particles are a scalar and a photon. The

---

[1] The details of the decay file are shown in appendix A

TABLE 3.1: Table depicting the physics models used to generate the various decay process. The production fractions used are taken from reference [8].

Decay process	Physics models
$\Upsilon(5\mathrm{S}) \to \mathrm{B}_\mathrm{s}^{*0}\bar{\mathrm{B}}_\mathrm{s}^{*0}$	N-body phase space (PHSP)
$\Upsilon(5\mathrm{S}) \to \mathrm{B}_\mathrm{s}^{*0}\bar{\mathrm{B}}_\mathrm{s}^{0}$ + cc	N-body phase space (PHSP)
$\Upsilon(5\mathrm{S}) \to \mathrm{B}_\mathrm{s}^{0}\bar{\mathrm{B}}_\mathrm{s}^{0}$	Vector to scalar scalar (VSS)
$B_s^{*0} \to B_s^0\gamma$	Vector to scalar photon P-wave (VSP-PWAVE)
$B_s^0 \to \pi^0\pi^0$	N-body phase space (PHSP)
$\pi^0 \to \gamma\gamma$	N-body phase space (PHSP)

photon propagates in the detector medium in a P-wave state. A detailed description of the models used can be found in the reference [45].

As the expected $\mathcal{B}(B_s^0 \to \pi^0\pi^0)$ is very small, roughly $\sim 10^{-6}$ [19], we have allowed the decay of only one of the $B_s^0(\bar{B}_s^0)$ in a $B_s^0\bar{B}_s^0$ pair to decay to $\pi^0\pi^0$ in an event as otherwise the composite probability to observe the event will be very low, $\sim 10^{-12}$. In an event, the other $\bar{B}_s^0(B_s^0)$ is then allowed to decay generically.

3.4 Data sample

For signal MC study, we have generated 100,000 signal MC events for each of the experiments 43, 53, 67, 69, and 71 dedicated to $\Upsilon(5\mathrm{S})$ resonance which corresponds to an integrated luminosity of 121.4 fb^{-1} data. The generic $\Upsilon(5\mathrm{S})$ MC samples are six times the luminosity of the data at $\Upsilon(5\mathrm{S})$ resonance for each experiment. Here, we have considered all the three possible production channels of B_s^0, namely, $B_s^{*0}\bar{B}_s^{*0}$, $B_s^{*0}\bar{B}_s^0$ + cc (charge-conjugate) and $B_s^0\bar{B}_s^0$. The generated events are run-independent and scaled to the luminosities of their corresponding experiments. In addition, we have generated 100,000 signal MC events separately for the three production channels $B_s^{*0}\bar{B}_s^{*0}$, $B_s^{*0}\bar{B}_s^0$ + cc and $B_s^0\bar{B}^0{}_s$ to determine their one-dimensional (1D) probability distribution functions (PDF) and know the best fit parameter values. We combine the production channels according to the production fractions mentioned in the Table 3.2. We fix the production fractions and use the best fit values (obtained from 1D PDF parametrization) for the channels, $B_s^{*0}\bar{B}_s^0$ + cc and $B_s^0\bar{B}^0{}_s$ while adding them to the dominant production channel, $B_s^{*0}\bar{B}_s^{*0}$.

Table 3.2: Production channels of B_s^0 at $\Upsilon(5S)$ resonance. The production fractions are described in reference [8]

Production channel	Fraction
$B_s^{*0}\bar{B}_s^{*0}$	0.870 ± 0.017
$B_s^{*0}\bar{B}_s^0$ + cc	0.073 ± 0.014
$B_s^0\bar{B}_s^0$	0.057 ± 0.022

For the background MC study, we have used the background MC datasets which is centrally generated and stored at KEKCC (the computing centre at KEK, Japan). There are six streams of continuum, `non-bsbs` and `bsbs` type of backgrounds for each experiment which corresponds to six times the luminosity at $\Upsilon(5S)$ resonance. Section 3.8 discusses these backgrounds in detail.

As recommended, we are using case B data from the SVD2 period for this analysis. Case B data are reprocessed with the latest Belle library `b200901247_0910` as of February 9, 2010.

Table 3.3 summarizes the luminosities of the experiments dedicated to $\Upsilon(5S)$ resonance at Belle.

Table 3.3: Luminosities of experiments at $\Upsilon(5S)$ resonance.

Experiments	Luminosities (fb^{-1})
43	1.857
53	21.513
67	27.222
69	47.830
71	22.938
Total	121.360

3.5 Event skimming

Skimming categorises physics processes for data analysis and rejects uninteresting events to reduce computing time and save resources. It is a crucial step in data storage because if any important physics events get lost during skimming, they cannot be retrieved. Skimming is performed on data, and the corresponding selection criteria are applied to the MC to get the skimmed MC for an analysis.

There are several skims available at Belle for different physics processes. They are of the following types:

1. *HadronBJ*: These skimmed data are stored for studying hadronic events.
2. *TauSkim*: They are used for the study of τ leptons.
3. *LowMult*: These type of skimmed data are used for studying low multiplicity events such as two photon processes.

We shall briefly describe the various selection criteria included in the *HadronBJ* skim, as our decay channel is of hadronic type. A detailed description of the skims is available at [44].

3.5.1 HadronBJ selection criteria

Collision of e^+e^- at the IP leads to the production of many physics processes such as hadronic events, continuum events, Bhabha events, two-photon processes, τ pairs and beam backgrounds. The *HadronBJ* skim retains only the hadronic events by their efficient selection from the rest.

Primarily, the hadronic selection is based on track multiplicity and visible energy.

Hadronic selection criteria based on track multiplicity:

1. Number of charged tracks, $N_{tracks} \geq 3$.
2. Transverse momentum, $p_t > 0.1$ GeV/c. This requirement rejects the other non-hadronic processes and also ensure selection of hadronic events near the IP with the following parameters $|\Delta r| < 2$ cm and $|\Delta z| < 4$ cm.

Hadronic selection criteria based on visible energy:

1. The sum of the energy of the charged tracks and the reconstructed photons, $E^*_{vis}/E_{cm} > 0.20$. The observable denoted by an asterisk are measured in the CM frame.

The beam backgrounds primarily consist of photons from the beam particles' collision with the beam pipe and, therefore, have less energy than the actual photons produced in the collision. The requirement of events coming from near the IP and $E^*_{vis}/E_{cm} > 0.20$ remove most of the beam backgrounds and two-photon processes such as $e^+e^- \rightarrow \gamma\gamma$. Moreover, selection criteria on the event vertex further remove the beam backgrounds. These are the primary vertices from where an event originates. The selection criteria for choosing events from primary vertex are:

1. $|\Delta r| \leq 3.5\ cm$ and $|\Delta z| < 1.5\ cm$.

Now, for the suppression of radiative Bhabha backgrounds and higher multiplicity QED processes (more than two γs in the final state), the following selection criteria are considered:

1. The angle between two or more ECL clusters should be large and lie between $-0.70 < \cos\theta^* < 0.90$.
2. The average energy of each cluster should be less than 1 GeV.
3. The total energy of the ECL cluster, E^*_{sum}, should be less than 80% of the CM energy.

The upper limit on the E^*_{sum} removes the radiative Bhabha and higher multiplicity QED effects. However, we need to put a lower limit of $E^*_{sum} > 0.18\sqrt{s}$ as well to remove τ pair events, beam backgrounds and two-photon processes, which have smaller values of E^*_{sum}. Here $\sqrt{s}$ is the center of mass energy, E_{cm}. The selection criteria mentioned above are highly efficient ($\approx 99\%$) in selecting the hadronic events but at the same time serve to be too restrictive for the continuum events, $e^+e^- \rightarrow q\bar{q}$, $q = u, d, s, c$. Therefore, to increase the continuum events' efficiency, the Heavy Jet Mass (HJM) [46] variable, which is an invariant mass of the τ particles, is introduced as conditional criteria. However, a study on the HJM variable revealed that its shift is deeply correlated to the shift in the visible energy, E^*_{vis}. To suppress these correlations, the ratio, HJM/E^*_{vis} along with the conditional selection criteria on HJM is used to retrieve some of the lost continuum events which do not arise from τ-pair events. These conditions are as follows:

1. $HJM/E^*_{vis} \geq 0.25$ or $HJM > 1.8$ GeV

However, these criteria for hadronic event selection do not have high efficiencies for the inclusive ψ events. So, the J/ψ and $\psi(2S)$ containing events are exclusively added to the *HadronB* events, thus making them the *HadronBJ* skim.

3.6 Preliminary event selection

The successfully simulated events may not have the correct properties in terms of the following: energy deposition in the ECL, the position of the primary vertex, beam energy, smearing of the IP, hadron skimming, L4 trigger (discussed in sub-section 2.4.3), preliminary reconstruction of π^0's and their associated error matrices. Thus, it becomes imperative to correct and properly scale them to reduce the difference between the simulated and the real data. The *fix_mdst* module performs these functionalities, and the surviving events are retained for further analysis. In this respect, the HadronBJ skimming and L4 trigger are an important part of the *fix_mdst* module.

3.6.1 Fix_mdst module

The *fix_mdst* module corrects the differences between the MC and data and rejects events which do not pass the L4 trigger and HadronBJ selection criteria. Some of its modules performing these functions are:

1. `fix_mdst_trk` : Apply various scale errors to charged tracks and extra-smearing to MC tracks.
2. `fix_mdst_ecl` : Energies of particles stored in `mdst_gamma` and `mdst_ecl` and their errors are corrected.
3. `fix_mdst_muid` : Fixes the liklihood ratio of muons, kaons and charged pions by using their probability distribution functions.
4. `fix_mdst_tof` : Fixes the particle's time-of-flight by shifting their value appropriately.
5. `fix_mdst_brecon` : Fixes the relation between photons and the neutral pions.

The *fix_mdst* module is therefore placed prior to the full event reconstruction algorithm to make the necessary corrections. A schematic representation of the *fix_mdst* module is shown in Figure 3.2. Table 3.4 shows the number of signal events which passed the *fix_mdst* module in this analysis.

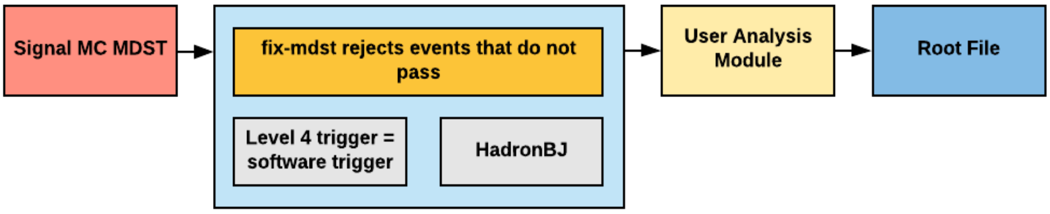

FIGURE 3.2: Block diagram to show the use of *fix_mdst* module.

TABLE 3.4: Preliminary event selection by *fix_mdst* module.

Experiments	Events generated	Events passing *fix_mdst*	Efficiency
43	100000	71704	71%
53	100000	72109	72%
67	100000	72779	72%
69	100000	72559	72%
71	100000	72650	72%

We observe a drop of roughly 28% in the signal efficiency due to the *fix_mdst* module. One of the reasons for this drop in signal efficiency is due to the requirements in the HadronBJ, which rejects events not containing at least three charged tracks.

3.7 Signal Reconstruction

An event is reconstructed from the final state particles following the principle of conservation of four-momentum at each decay vertex using a bottom-up approach. Recreating an event from the final state particles and arriving at an e^+e^- collision is called event reconstruction.

At Belle, only the particles which are stable and have relatively long lifetimes, such as $e^\pm$, $\mu^\pm$, $K^\pm$, $\pi^\pm$, $p^\pm$ and photons (γ), can be detected. For this analysis, the γs are the final state particles used to reconstruct the π^0s and eventually the signal B_s^0.

This type of reconstruction, where all the final state particles can be detected, and the whole decay chain can be recreated using the principle of conservation of four-momentum, is called full hadronic reconstruction or exclusive reconstruction.

Before we proceed, let us define the following terms which will be used extensively throughout this chapter:

- MC truth matched candidates: Candidates which are correctly reconstructed and share their origin to the generator level information (as described in section 3.3) are the MC truth matched candidates.
- Misreconstructed candidates: The candidates which are not correctly reconstructed and are formed due to combinatorial nature of the reconstruction process are termed as misreconstructed or self-cross feed (SCF) candidates.

The following paragraphs discuss in detail the reconstruction of $B_s^0 \to \pi^0\pi^0$ with $\pi^0 \to \gamma\gamma$.

3.7.1 Photon (γ) selection

The photons used for the reconstruction of the π^0's need to qualify the following selection criteria:

- They should not have any associated charged track in the detector's CDC. This ensures that these photons do not originate from any charged particles produced in the detector.
- The photons should have minimum energy of 0.05 GeV in the barrel region and 0.10 GeV in the forward and the backward region of the ECL. A slightly tighter selection criterion on the ECL's forward and backward region rejects high-energetic photons from the beam-induced backgrounds. The forward, barrel, and backward regions of the ECL are considered in the ranges of $12^\circ < \theta < 31.4^\circ$, $32.2^\circ < \theta < 128.7^\circ$, and $131.5^\circ < \theta < 157.2^\circ$, respectively, where θ is the polar angle measured from the e^+ beam direction in the laboratory frame. The photon energy distributions are shown in Figures 3.3 and 3.4.
- To remove the off-time (radiative) Bhabha and $e^+e^- \to \gamma\gamma$ events, a selection criterion based on the photon identification number (ID) is applied only for MC. In the case of data, this criterion is replaced by the timing information of the ECL crystals

TABLE 3.5: Summary of photon selection criteria.

Variables	Selection criteria
Shower Match (no charged track in CDC)	$match == 0$
Photon energy in barrel (forward and backward) ECL	$E_\gamma > 0.05(0.10)$ GeV
Photon ID (MC only)	$\neq 911$

(discussed in section 2.3.5). A detailed description of the selection criterion can be found in [29].

These selection criteria are summarized in Table 3.5.

3.7.2 Neutral pion (π^0) reconstruction

The π^0s are reconstructed from γs which pass the selection criteria mentioned in Table 3.5. These π^0s are then further selected using a set of selection criteria enumerated below:

- The invariant mass of the two photons, $M_{\gamma\gamma}$ is required to be in the range $0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2, which is within 2.36 standard deviations (σ) about the nominal π^0 mass [1]. The σ is determined from the one-dimensional maximum likelihood fit performed on the reconstructed π^0s.

- We perform a mass constraint fit on the π^0's, which pass the above selection criterion to minimize the χ^2 associated with the reconstructed π^0s. The invariant mass distribution of the two photons are shown in Figure 3.5. The χ^2 is defined as follows

$$\chi^2 = \frac{M_{\gamma\gamma} - m_{\pi^0}}{\sigma^2}, \tag{3.1}$$

where, $\sigma^2 = \sigma^2_{\gamma\gamma} + \sigma^2_{\pi^0}$ is the total standard deviation associated with an event and and m_{π^0} is the world average value of π^0 mass as reported in PDG [**?**].

The summary of the selection criteria applied on reconstructed π^0 is shown in Table 3.6.

3.7.3 B_s^0 reconstruction

Candidate B_s^0 are reconstructed from two π^0s which pass the selection criteria summarized in Table 3.6. We introduce kinematic variables to identify the signal B_s^0 from the

TABLE 3.6: Summary of selection criteria applied on the reconstructed π^0's

Variables	Selection criteria
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2

background. The term "signal" henceforth would refer to our decay channel, $B_s^0 \to \pi^0\pi^0$. Reconstruction of the charge conjugate channel, $\bar{B}_s^0 \to \pi^0\pi^0$ is implied throughout the analysis.

3.7.3.1 Invariant mass

The invariant mass variable is the mass of the reconstructed particle obtained from its daughter particles using the Einstein's relativistic energy-momentum relation

$$E = \sqrt{p^2c^2 + m_0^2c^4}, \tag{3.2}$$

where, $\vec{p}$ and m_0 are the three momentum and rest mass of each of the daughter particles.

It is seen that most of the random combinations cannot reproduce the invariant mass of the reconstructed particle and hence can be suppressed by applying a selection criteria on the invariant mass distribution.

3.7.3.2 Beam-constrained-mass and Energy difference

The beam-constrained-mass and energy difference variables are defined by the following relations

$$M_{\rm bc} = \sqrt{(E^*_{\rm beam})^2 - |p^*_{B_s^0}|^2c^2}/c^2, \tag{3.3}$$

$$\Delta E = E^*_{B_s^0} - E^*_{\rm beam}, \tag{3.4}$$

where, $E^*_{\rm beam}$, $E^*_{B_s^0}$, $p^*_{B_s^0}$, and c represents the beam energy, energy and three momentum of the reconstructed B_s^0 candidates in the CM frame, and velocity of light, respectively.

However, the dominant B_s^0 production channel, $B_s^{*0}\bar{B}_s^{*0}$ ($\sim 87\%$) decays to B_s^0 with an accompanying low-energy photon. This photon is not reconstructed because of its poor reconstruction efficiency, which causes the peak of the ΔE variable for the B_s^0 candidates to shift towards more negative ΔE values. Therefore, a modified variable, $\Delta E'$ is used

Table 3.7: Summary of selection criteria applied on the reconstructed B_s^0 candidates.

Variables	Selection criteria
$\Delta E'$	$-0.60 < \Delta E' < 0.15$ GeV
$M_{\rm bc}$	$M_{\rm bc} > 5.30$ GeV/c^2

instead, which is defined as follows

$$\Delta E' = \Delta E + M_{\rm bc}c^2 - m_{B_s^0}c^2, \tag{3.5}$$

where, $m_{B_s^0}$ is the PDG mass of the B_s^0 meson [1].

We select only those B_s^0 candidates which have $\Delta E'$ and $M_{\rm bc}$ values in the range of $[-0.60, 0.15]$ GeV and $[5.300, 5.434]$ GeV/c^2, respectively. These are loose selection criteria chosen to have sufficient MC statistics for parametrizing the signal and background distributions. We expect the $\Delta E'$ variable for the signal MC to peak around zero with an elongated tail, which is observed due to shower energy leakage through the ECL crystals. This effect is dominant for analyses with neutral particles in the final states.

A summary of the selection criteria applied on the reconstructed B_s^0 candidates is shown in Table 3.7.

Figure 3.6 shows the $M_{\rm bc}$ distributions of the two B_s^0 production channels, $B_s^0\bar{B}_s^0$, $B_s^{*0}\bar{B}_s^0$ + cc. Figures 3.7 and 3.8 show the $M_{\rm bc}$, $\Delta E'$ distributions taking into account all the three production channels and the comparison between ΔE and $\Delta E'$ distributions, respectively.

We summarize all the selection criteria applied to select the candidate B_s^0 in Table 3.8.

Table 3.8: Summary of selection criteria

Variables	Selection criteria
Photon ID (MC only)	$\neq 911$
Photon energy, E_γ (GeV)	> 0.05 (0.10) in barrel (forward/backward endcaps) ECL
$M_{\gamma\gamma}$ (GeV/c^2)	$0.118 < M_{\gamma\gamma} < 0.152$
$\Delta E'$ (GeV)	$-0.60 < \Delta E' < 0.15$
$M_{\rm bc}$ (GeV/c^2)	$5.300 < M_{\rm bc} < 5.434$

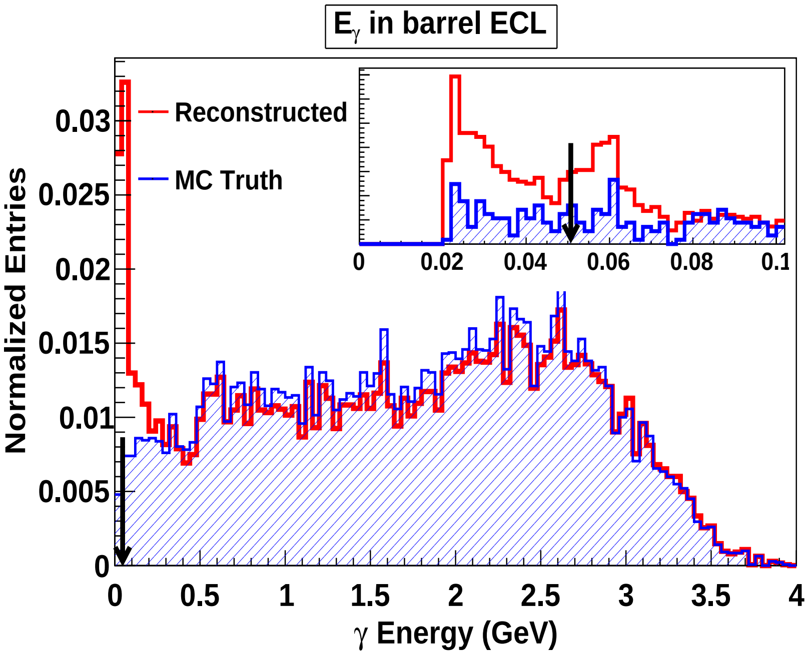

FIGURE 3.3: Photon energy distributions in the barrel regions of ECL. The blue (hatched) distribution represents correctly reconstructed (MC truth matched) candidates while the red (solid-line) distribution is plotted using reconstruction level information. The observation of a peak in the red (solid-line) distribution indicates misreconstruction in the low γ energy regions. The inset figure highlights the low γ energy regions with finer binning. The arrows indicate the positions of the selection criterion.

3.8 Background MC study

At Υ(5S), the following types of backgrounds are produced:

1. $e^+e^- \to q\bar{q}$ (q = u, d, s, c) (continuum: uds + charm)
2. $e^+e^- \to \Upsilon(5S) \to B\bar{B}$ (non-bsbs)
3. $e^+e^- \to \Upsilon(5S) \to B_s^{*0}\bar{B}_s^{*0}$, $B_s^{*0}\bar{B}_s^0$ + cc, $B_s^0\bar{B}_s^0$ (bsbs)

The various hadronic events near Υ(5S) resonance was discussed earlier (see Sub-subsection 2.5.2). These background MCs are produced with six times higher statistics than the expected background at 121.4 fb^{-1} of luminosity. In this section, we discuss their effect on our analysis.

We apply the same set of selection criteria summarized in Table 3.8 on the six streams of backgrounds and obtain the distributions shown in Figures 3.9, 3.10, and 3.11.

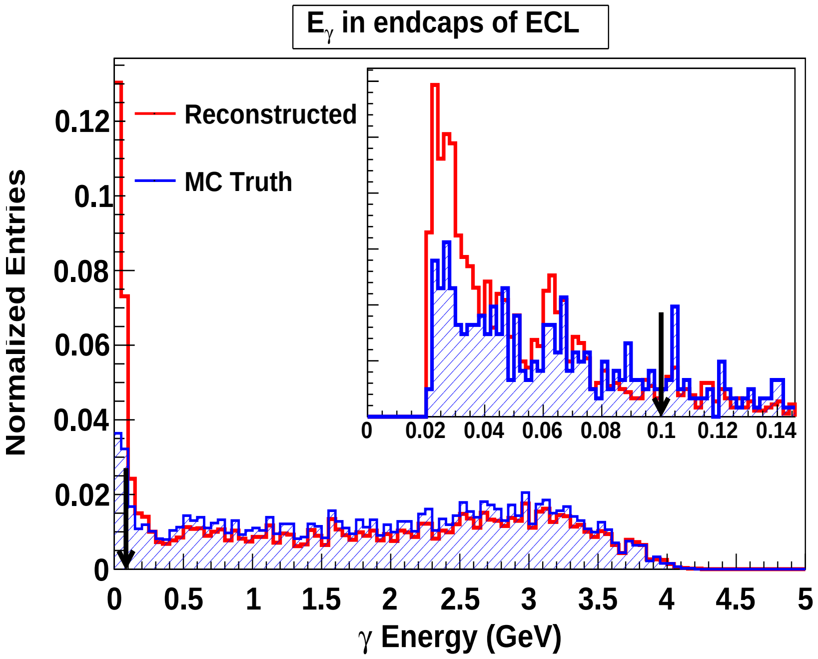

FIGURE 3.4: Photon energy distributions in the endcaps regions of ECL. The blue (hatched) distribution represents correctly reconstructed (MC truth matched) candidates while the red (solid-line) distribution is plotted using reconstruction level information. The observation of a peak in the red (solid-line) distribution indicates misreconstruction in the low γ energy regions. The inset figure highlights the low γ energy regions with finer binning. The arrows indicate the positions of the selection criterion.

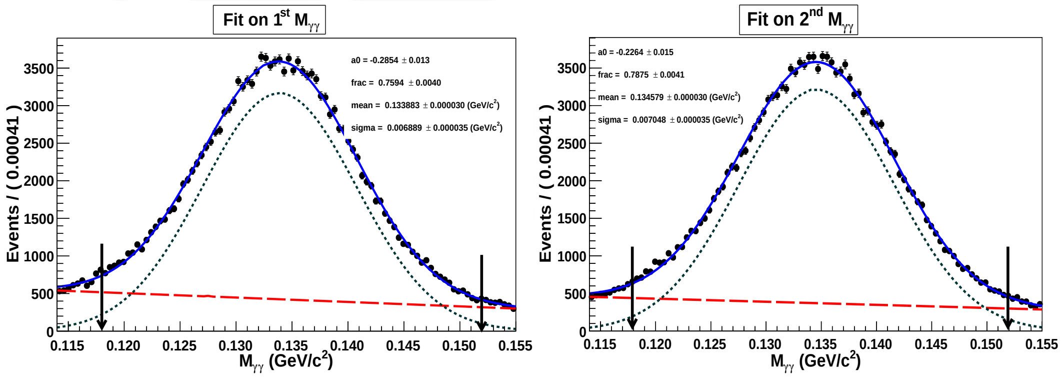

FIGURE 3.5: Fitted distributions of the invariant mass of two photons. The solid, short-dashed, and long-dashed curves indicate the total fit function, a Gaussian distribution, and a first order Chebychev polynomial, respectively. The arrows indicate the invariant mass range of $[0.118, 0.152]$ GeV/c^2.

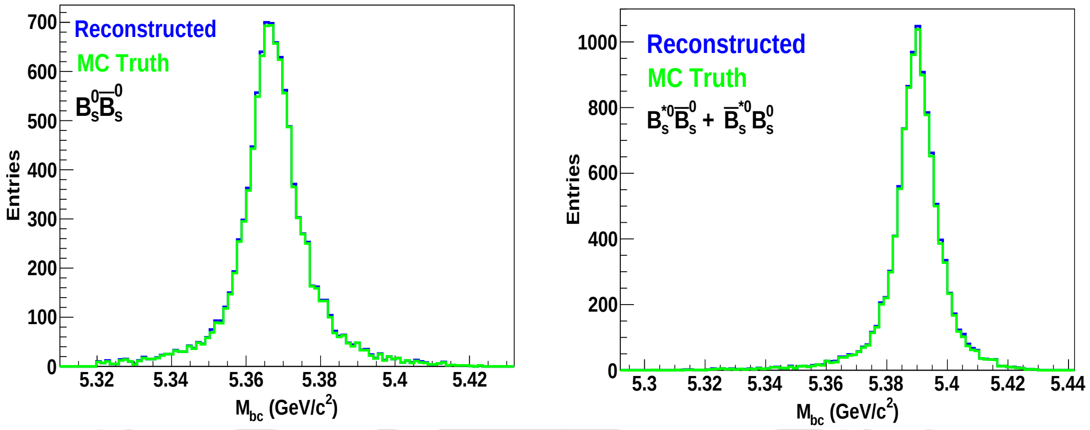

FIGURE 3.6: $M_{\rm bc}$ distributions of B_s^0 candidates from the production channels, $B_s^0\bar{B}_s^0$ (left), $B^{*0}\bar{B}_s^0$ + cc (right), respectively. The green distribution represents correctly reconstructed (MC truth matched) candidates while the blue distribution is plotted using reconstruction level information. Of all the signal MC candidates, 98.2% of them are MC truth matched.

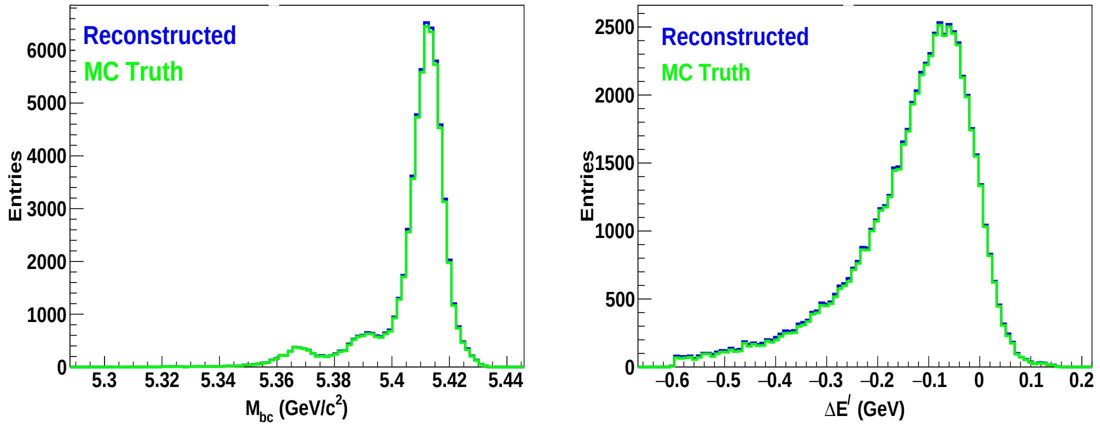

FIGURE 3.7: $M_{\rm bc}$ (left) and $\Delta E'$ (right) distributions of B_s^0 candidates from the production channels, $B_s^0\bar{B}_s^0$, $B^{*0}\bar{B}_s^0$ + cc, $B_s^{*0}\bar{B}_s^{*0}$. The highest peak in the M_{bc} distribution represents the dominant production channel, $B_s^{*0}\bar{B}_s^{*0}$. The green distribution represents correctly reconstructed (MC truth matched) candidates while the blue distribution is plotted using reconstruction level information. Of all the signal MC candidates, 98.2% of them are MC truth matched.

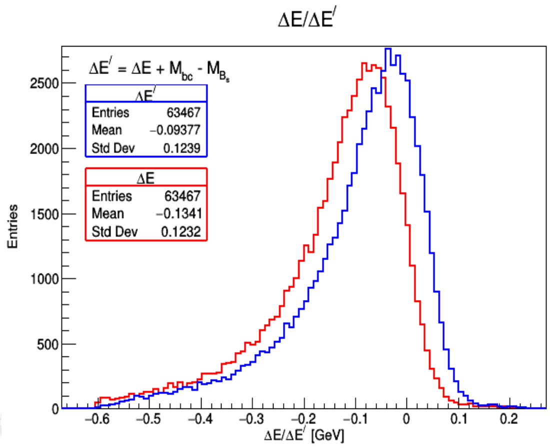

FIGURE 3.8: ΔE and $\Delta E'$ distributions of B_s^0 candidates. The red histogram represents ΔE distribution while blue histogram represents the $\Delta E'$ distribution. Even though the resolution remains the same, the peak in $\Delta E'$ distribution shifts towards zero.

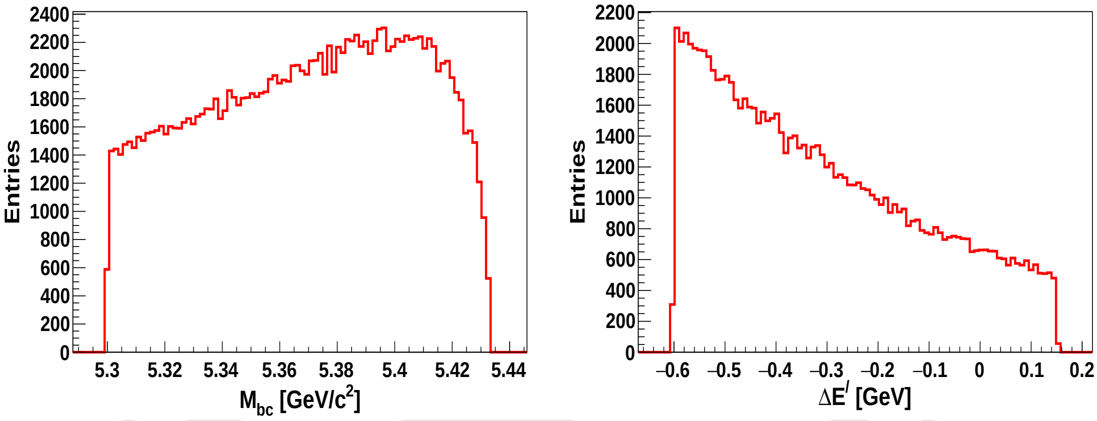

FIGURE 3.9: M_{bc} (left) and $\Delta E'$ (right) distributions of B_s^0 candidates for uds background

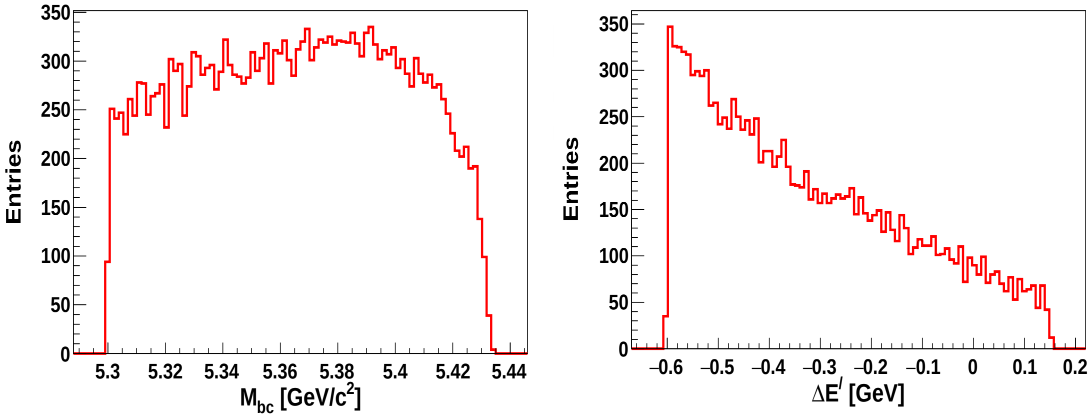

FIGURE 3.10: M_{bc} (left) and $\Delta E'$ (right) distributions of B_s^0 candidates for charm background

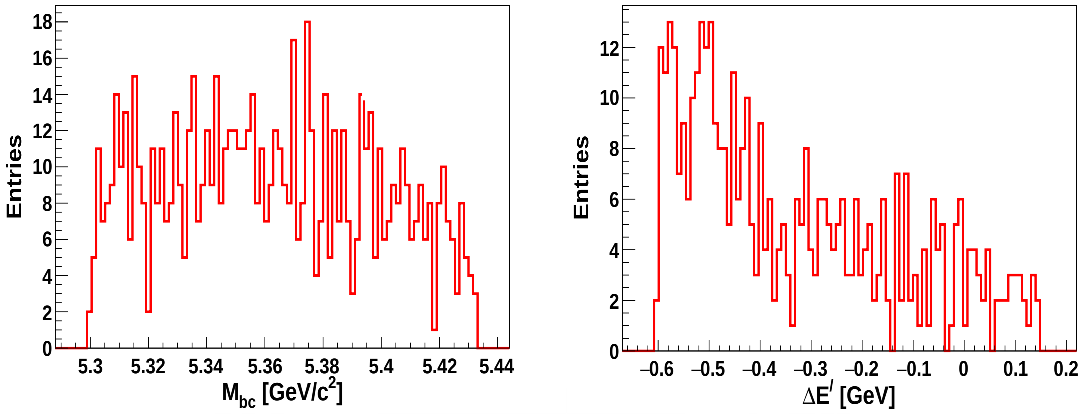

FIGURE 3.11: $M_{\rm bc}$ (left) and $\Delta E'$ (right) distributions of B_s^0 candidates for non-bsbs background

TABLE 3.9: Cut-flow table. The figures in the parenthesis indicates percentage loss. The backgrounds are considered for six streams.

Variables	Selection Criteria	MC Truth	Signal	uds	charm	bsbs	nonbsbs	Background
$M_{\rm bc}$	$M_{\rm bc} > 5.30$ GeV/c^2	147751	157279	1452869	239717	49	1757	1694392
$\Delta E'$	$-0.60 < \Delta E' < 0.15$ GeV	146274 (1%)	154134 (2%)	894766 (38%)	140889 (41%)	1 (83%)	770 (56%)	1036433 (39%)
E_γ	$E_\gamma > 0.05(0.1)$ (GeV) Barrel(Fwd/Bwd)	140716 (3.8%)	144732 (6.1%)	831577 (7%)	131537 (6.6%)	1 (12.5%)	705 (8.4%)	963826 (7%)
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2	131007 (6.9%)	135470 (6.4%)	771407 (7.2%)	120707 (8.2%)	0 (28%)	653 (7.3%)	892772 (7.3%)

We do not find any surviving bsbs backgrounds after applying the selection criteria summarized in Table 3.8. In Table 3.9, we summarize the cut-flow table where the selection criteria are applied sequentially.

From the cut-flow table, we find that the continuum background is dominant over other backgrounds, and hence we need to suppress them using topologically discriminating variables.

3.8.1 Continuum suppression

Continuum suppression is a technique of differentiating between the signal and the continuum background by utilising the information based on their topology. The variables which distinguish these differences in their topology are the topologically discriminating variables. They are of different types, such as event shape variables, thrust angle, and flight angle of the B_s^0 mesons, to name a few. We discuss the ones which are used in this analysis.

3.8.1.1 Event shape variables

In the CM frame of an event, the three momenta of the B_s^0 candidates are very low compared to its invariant mass and hence can be assumed to be at rest. Therefore, their decay product distribution would be isotropic about the beam axis. On the other hand, the continuum events produced directly from an e^+e^- collision distribute in the form of two back-to-back jets in their CM frame. This difference in their event shapes distinguishes the signal from the continuum background. Figure 3.12 shows the event shape distributions for continuum background and signal. Fox and Wolfram first studied these event shape variables in the year 1978 and they now known as the Fox-Wolfram moments [38]. Mathematically, they are described as

$$H_l = \sum_{i,j}^{N} \frac{|\vec{P}_i||\vec{P}_j|}{s} P_l(\cos\theta_{i,j}), \tag{3.6}$$

where, N is the total number of final state particles, $\vec{P}_i$ and $\vec{P}_j$ are the three momentum of the i^{th} and j^{th} final state particles, s is the square of the CM energy used as a normalization constant, $P_l(\cos\theta_{i,j})$ are associated Legendre's polynomial, θ_{ij} represents the angle between the i^{th} and j^{th} final state particle.

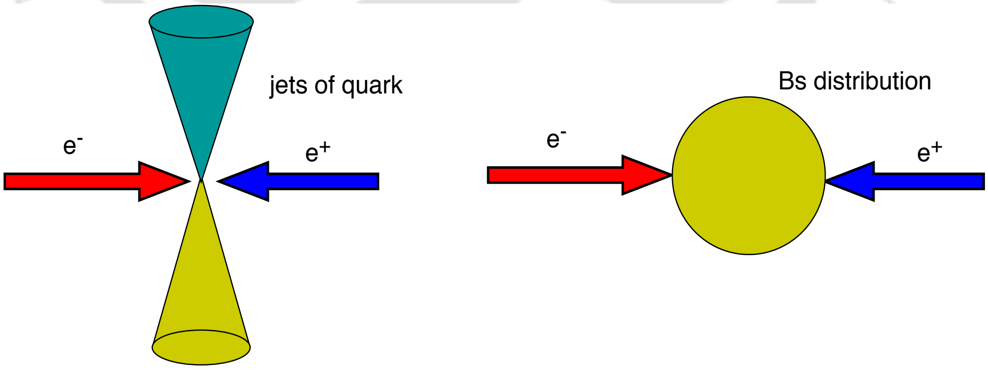

FIGURE 3.12: Distribution of continuum background (left) and signal events (right).

These variables may have non-linear dependencies, making a simple rectangular cut-based selection criterion less efficient. It is observed that these variables tend to have more substantial suppression power when combined and a singular selection criterion applied to them. Thus, we need the best combination among these variables to suppress the continuum backgrounds efficiently. This search for the best combination is done by forming a discriminant, often called Fischer discriminant at Belle, after the scientist Fischer, who first used this method. The Fischer discriminant consists of the Fox-Wolfram moments and an additional variable, the transverse momentum of the final state particle introduced by Belle, for better discrimination. Mathematically, it is represented as

$$KSFW = \sum_{l=0}^{4} R_l^{so} + \sum_{l=0}^{4} R_l^{oo} + \gamma \sum_{n=1}^{N_t} |(P_t)_n|, \tag{3.7}$$

where, R_l^{so}, R_l^{oo} are the modified Fox-Wolfram moments, P_t is the transverse momentum of each particle, γ is a free parameter, and N_t is the total number of particles. These modified Fox-Wolfram moments are also known as KSFW (Kakuno-Super-Fox-Wolfram) moments. A detailed discussion on the KSFW moments is described in references [44, 47, 48, 49]. We briefly define the modified Fox Wolfram moments below

R_l^{so} and R_l^{oo}

(I) R_l^{so}

$$R_l^{so} = \frac{\alpha_{cl} H_{cl}^{so} + \alpha_{nl} H_{nl}^{so} + \alpha_{ml} H_{ml}^{so}}{E^*_{\text{beam}} - \Delta E'}, \tag{3.8}$$

where, all the particles and their momenta are divided into three categories:

1. Charged particles (c) → $\alpha_{cl} H_{cl}^{so}$
2. Neutral particles (n) → $\alpha_{nl} H_{nl}^{so}$
3. Missing particles (m) → $\alpha_{ml} H_{ml}^{so}$

For odd l

$$H_{nl}^{so} = H_{ml}^{so} = 0, \tag{3.9}$$

and

$$H_{cl}^{so} = \sum_i \sum_{jx} Q_i Q_{jx} |\vec{p}_{jx}| P_l(\cos\theta_{i,jx}), \tag{3.10}$$

where, i iterates over B daughters and jx runs over rest-of-events (ROE) (all other particles not associated with the signal B_s^0) in the category $x (x = c, n, m)$, Q_i and Q_{jx} are the charges of particle i and jx, respectively, p_{jx} is the momentum of particle jx, and $P_l(\cos\theta_{i,jx})$ is the l-th order Legendre polynomial of the cosine of the angle between particles i and jx.

For even l

$$H_{xl}^{so} = \sum_i \sum_{jx} |\vec{p}_{jx}| P_l(\cos\theta_{i,jx}). \tag{3.11}$$

There are two odd parameters and nine (3 × 3) even parameters, totalling to 11 free parameters from this expression.

(II) R_l^{oo}

For odd l

$$R_l^{oo} = \sum_j \sum_k \beta_l Q_j Q_k |\vec{p}_j||\vec{p}_k| P_l(\cos\theta_{j,k}), \tag{3.12}$$

where, both j and k runs over ROE, β represents Fischer coefficients, and other variables are the same as described earlier.

For even l

$$R_l^{oo} = \sum_j \sum_k \beta_l |\vec{p}_j||\vec{p}_k| P_l(\cos\theta_{j,k}). \tag{3.13}$$

Therefore, the total number of moments from R_l^{oo} is, 5.

These KSFW moments are then used as inputs to the neural network for signal and background classification.

3.8.1.2 Thrust

The thrust of an event containing N final state particles with momenta $\vec{p}_i (i = 1, 2, 3, ..., N)$ is along the axis, which has the maximum total projection of the moment. The thrust axis, $\vec{T}$, is a unit vector along this axis. Mathematically, the thrust, $|\vec{T}|$, is defined as

$$|\vec{T}| = \frac{\sum_{n=1}^{\infty} |\vec{T} \cdot \vec{p}_i|}{\sum_{i=1}^{N} |\vec{p}_i|}. \tag{3.14}$$

The angle between the thrust axis of the reconstructed B_s^0 candidates and the thrust axis of the ROE candidates in an event has very strong discriminating power. Thus, the cosine of the thrust angle, $\cos\theta_T$, is used to differentiate between the signal and background events.

Figures 3.13 show one of the event shape variables (R_2^{oo}) and the thrust angle of the B_s^0 mesons for both signal and continuum background, respectively. We have used 16 KSFW moments and the thrust angle of the B_s^0 mesons as inputs to the Neural Network (NN) for

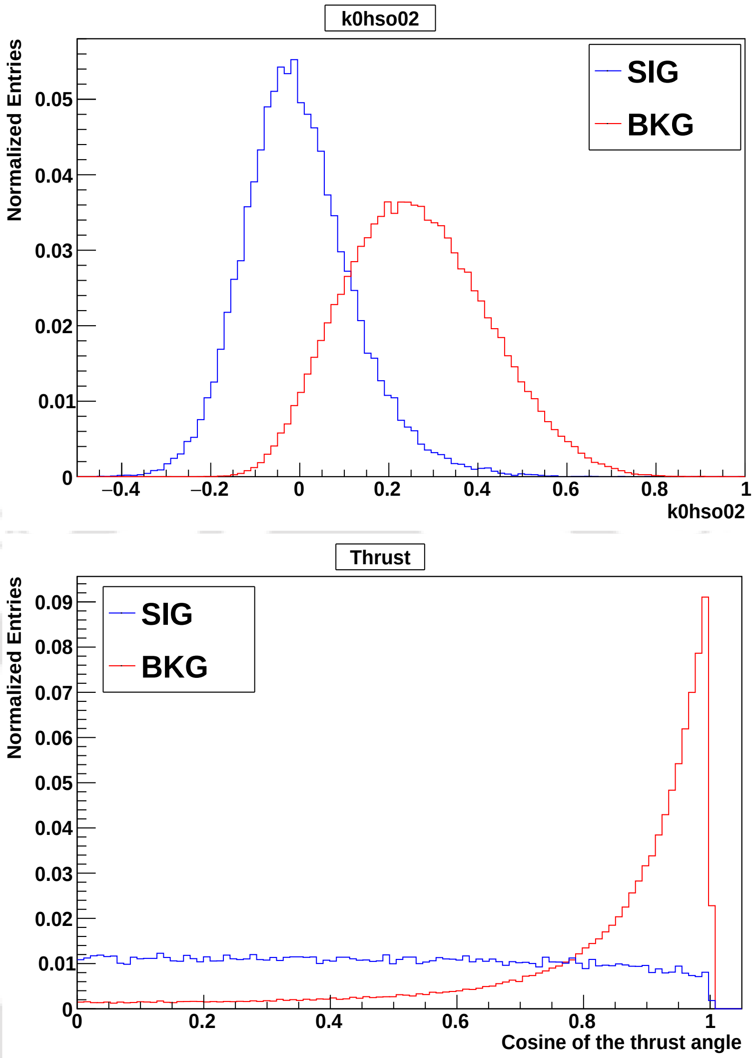

FIGURE 3.13: Plots of topologically discriminating variables, R_2^{so}(top) and thrust angle (bottom). The blue and red distributions represent correctly reconstructed (MC truth) and background MC candidates, respectively.

suppressing the dominant continuum background. The distributions of the remaining event shape variables are shown in the Figures B.1, B.2, and B.3 in appendix B.1. Moreover, we do not find any significant correlations of the topological variables with the fit variables, $M_{\rm bc}$ and $\Delta E'$. Table B.1 in appendix B.1 shows the correlation coefficients.

3.8.1.3 Neural Network

Neural networks are advanced techniques applied in solving multivariate analysis problems that efficiently handle the correlations among the input variables. The technique draws cues from the complex working of the neurons in the human brain. Mathematically, the

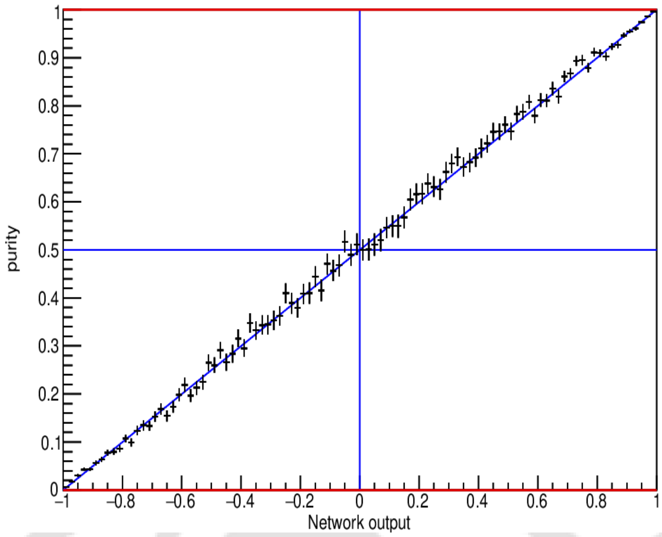

FIGURE 3.14: Distribution of Neural Network output versus purity. The MC data points with error bars represent the training sample and the solid blue line (inclined at 45°) indicates the expected trajectory for an ideal training.

neural network is a layered framework of a function defined as

$$x_j^n = g(t)(\sum_k w_{jk}^n \cdot x_k^{n-1} + \mu_j^n), \tag{3.15}$$

where, $g(t)$ is a sigmoid function which activates a node of a layer and μ_j^n is a constant which determines the threshold of activation. The superscript n denotes the n^{th} layer and the output of the node j of the n^{th} layer is the outcome of the weighted sum of all the nodes in the $(n-1)^{th}$ layer. In our analysis, we have used the NeuroBayes package [50] to classify the signal and the background. The NeuroBayes package contains two stages of implementation, namely, NeuroBayes Teacher and NeuroBayes Expert. A random sample of MC (Signal MC + Background MC) amounting to $2/3^{rd}$ of the total MC data sample is first trained to learn from the intricacies present in the MC and the correlation among the input variables.. This learned information about the network topology is then displayed to the user in the form of plots such as purity versus network output and signal versus background classification to check network information during learning. This information is then transferred to the NeuroBayes expert to classify the MC sample's signal and background. For running the NeuroBayes expert algorithm, the remaining $1/3^{rd}$ MC is used as the validation sample to know whether the network has learned properly. Only after consistent results between the training and the expert may we conclude that there is no significant over-training in the neural network. Figures 3.14 and 3.15 show the purity vs network output and training/validation, respectively.

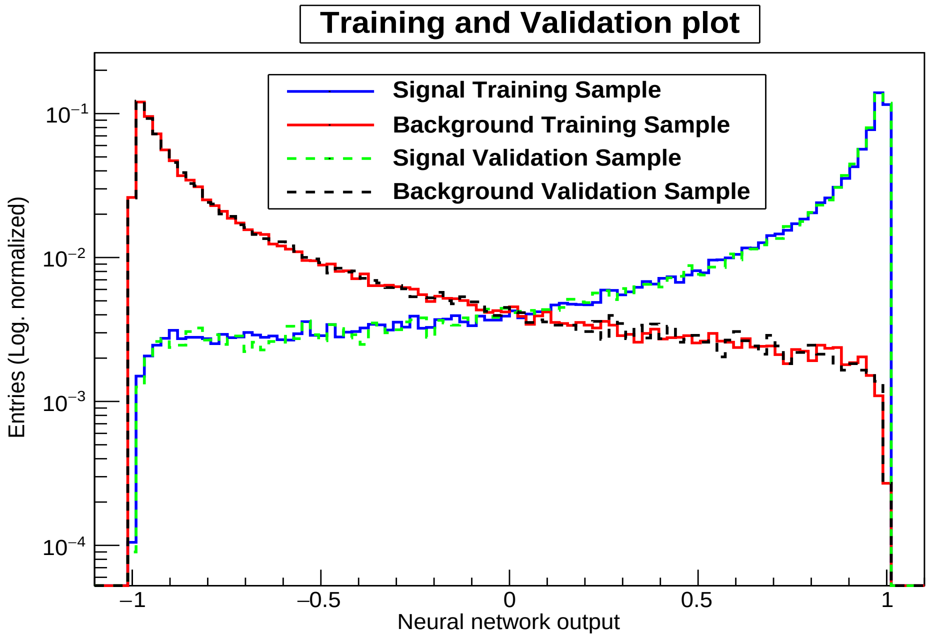

FIGURE 3.15: Signal and background MC training and validation plot.

The training and validation distributions are superimposed to determine the over-training of the NN. We see good agreement between the two distributions in Figure 3.15, and hence conclude that no considerable over-training is observed in our neural network.

Figure 3.16 shows the neural network output for the target MC. The distributions for the signal and the continuum background are pushed to two opposite poles (signal peaks near $+1$ and background peaks near -1). The non-linearities and the correlations among the seventeen input variables (sixteen KSFW moments and $\cos\theta_T$) are handled by the neural network and, in the end, give a single output which can easily discriminate between the signal and the continuum background.

We apply a selection criterion on the neural network output at 0.90, which removes 99% of the continuum background and 53% of the signal as seen in Table 3.10. We obtain a signal MC purity of approximately 92% with this choice of NN selection criterion. This selection criterion on the NN output is not optimised but chosen to remove the dominant continuum backgrounds to reach the SM sensitivity using the Belle data set. Figure 3.17 shows Punzi's figure-of-merit (FOM) [51] for reference. The Punzi's FOM is defined as

$$\text{FOM} = \frac{\epsilon(t)}{a/2 + \sqrt{B}},$$

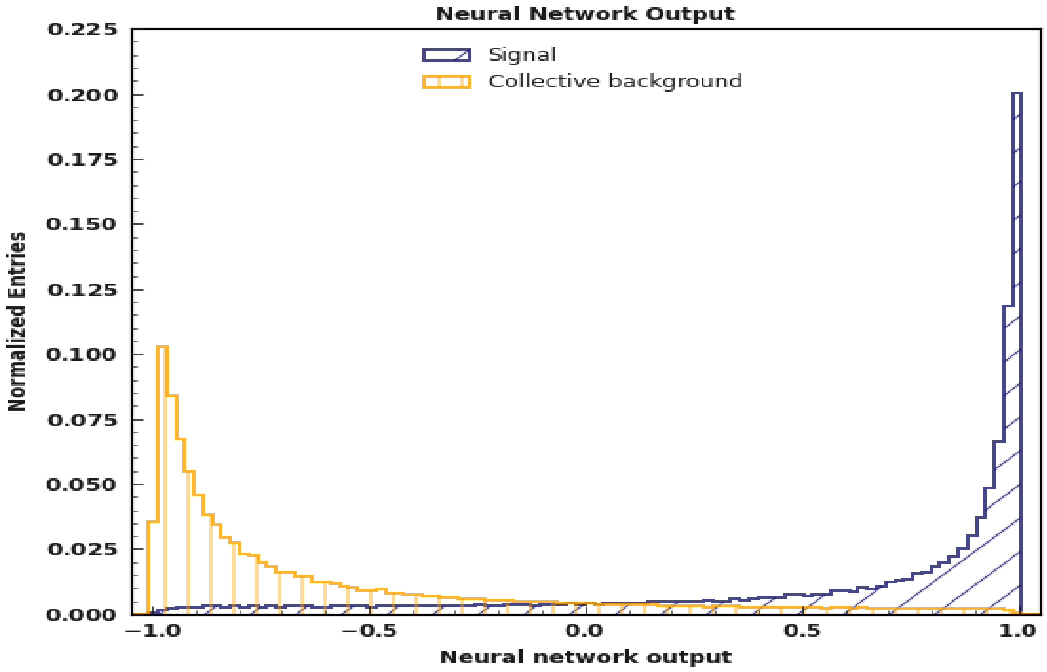

FIGURE 3.16: Neural network output of signal and background MC. The blue (slanted-hatched) and yellow (vertically-hatched) distributions refer to the output from the neural network for signal and background MC, respectively.

TABLE 3.10: Cut-flow table. The figures in the parenthesis indicates percentage loss. The backgrounds are considered for six streams.

Variables	Selection Criteria	MC Truth	Signal	uds	charm	bsbs	nonbsbs	Background
$M_{\rm bc}$	$M_{\rm bc} > 5.30$ GeV/c^2	147751	157279	1452869	239717	49	1757	1694392
$\Delta E'$	$-0.60 < \Delta E' < 0.15$ GeV	146274 (1%)	154134 (2%)	894766 (38%)	140889 (41%)	1 (83%)	770 (56%)	1036433 (39%)
E_γ	$E_\gamma > 0.05(0.1)$ GeV Barrel(Fwd/Bwd)	140716 (3.8%)	144732 (6.1%)	831577 (7%)	131537 (6.6%)	1 (12.5%)	705 (8.4%)	963826 (7%)
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2	131007 (6.9%)	135470 (6.4%)	771407 (7.2%)	120707 (8.2%)	0 (28%)	653 (7.3%)	892772 (7.3%)
NN Output	$\mathcal{C}_{\rm NN} > 0.90$	62359 (52%)	63477 (53%)	4439 (99%)	723 (99%)	–	0 (100%)	5169 (99%)

where, $\epsilon(t)$ is the signal efficiency and a is function of the selection criteria, t, a is the required level of significance, and B is the number of background events. In the present scenario, we require the level of significance to be three standard deviations.

The signal MC purity is defined as follows

$$\text{Purity} = \frac{\text{Number of remaining signal MC after NN selection criterion}}{\text{Number of remaining signal and background MC after NN selection criteria}}. \tag{3.16}$$

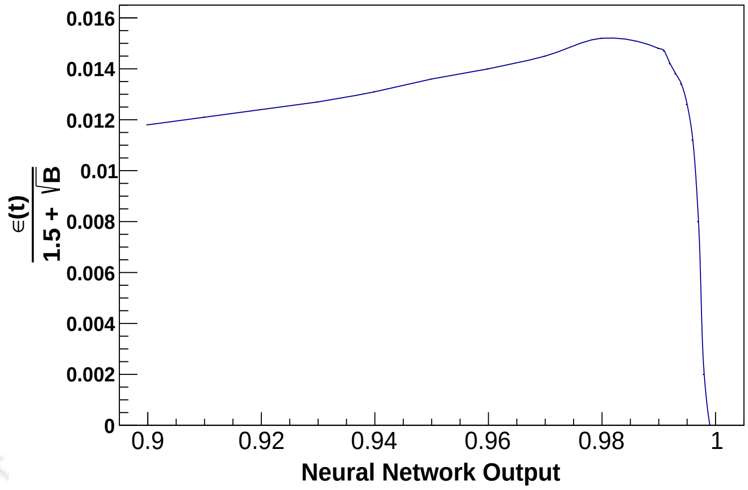

FIGURE 3.17: Distribution of Punzi's FOM for the neural network output variable. The maximum FOM occurs at NN = 0.985 (shown for reference).

TABLE 3.11: Values for NN output

$\mathcal{C}_{\rm NN(min)}$	$\mathcal{C}_{\rm NN(max)}$
0.90	0.998

3.8.2 Modified neural network output

The output of the NN is transformed to another variable, $C'_{\rm NN}$, to facilitate the data model. It is defined as

$$\mathcal{C}'_{\rm NN} = \log\left(\frac{\mathcal{C}_{\rm NN} - \mathcal{C}_{\rm NN(min)}}{\mathcal{C}_{\rm NN(max)} - \mathcal{C}_{\rm NN}}\right), \tag{3.17}$$

where, $\mathcal{C}_{\rm NN}$ is the neural network output, $\mathcal{C}_{\rm NN(min)}$ is the selection criterion applied on the NN output, $\mathcal{C}_{\rm NN(max)}$ is the maximum value of the $\mathcal{C}_{\rm NN}$ obtained from the NN distribution. The values of these variables are summarized in Table 3.11. Figure 3.18 shows the distribution of the modified neural network output.

3.9 Rare B_s^0 background MC study

In this section, we perform rare B_s^0 background MC study, which may behave very much like the signal events. We perform this study in the following steps:

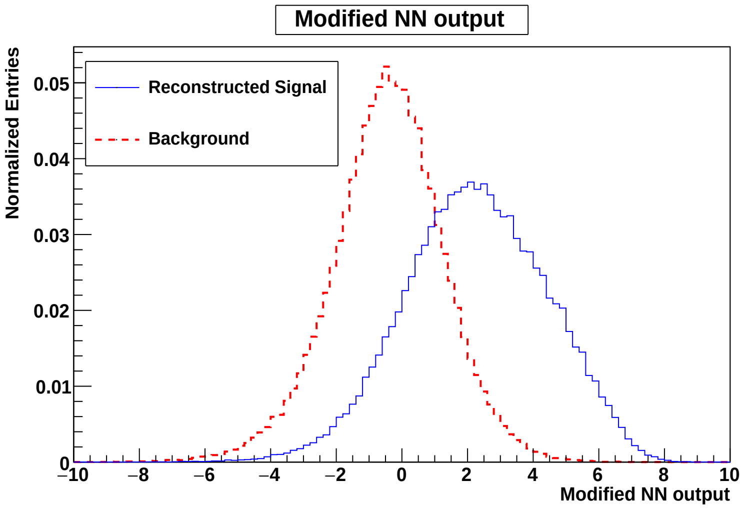

FIGURE 3.18: Modified neural network output, $\mathcal{C}'_{\text{NN}}$. The blue and red distribution represents signal MC and continuum MC background, respectively.

1. We compare the decay channels which are considered for **bsbs** backgrounds (in DECAY.DEC file) with the charmless (rare) decay channels listed for B_s^0 in PDG [1]. DECAY.DEC file consists of all the decay channels to be produced centrally at KEKCC for background MC study. The generated background events are stored in streams, where each stream corresponding to all experiments is of luminosity equivalent to real data collected at that resonance.

2. From step 1, we find that out of 32 B_s^0 charmless (rare) decay channels considered in PDG, 12 of them are not considered in the DECAY.DEC file. The decays $B_s^0 \to \rho^+\rho^-$ and $B_s^0 \to \eta\pi^0$ are also not reported in the PDG. We, therefore, use theoretical branching fractions for their generation. We also regenerate the decay channel $B_s^0 \to \bar{K}_s^0\pi^0$ according to theoretically predicted branching fraction. Moreover, the decay channel $B_s^0 \to \bar{K}_S K_S$ has an updated value of branching fraction as reported in the PDG. Thus, we also regenerate this decay channel with the latest value. Therefore, we consider $15(12+3)$ rare decay channels, which may behave as a potential peaking background for this analysis. At the same time, the remaining $(32-15=)17$ decay channels are already present in the DECAY.DEC file. As we did not find any surviving **bsbs** backgrounds for the remaining decay channels after applying all the selection criteria, they are not of much concern for this analysis.

   Therefore, it remains to see whether these 15 decay channels have any significant contribution as a background in this analysis. Table 3.12 shows the list of these 15 decay channels.

3. We generated 100,000 MC events for these decays for all the four experiments (experiment 53, 67, 69, and 71) at $\Upsilon(5S)$. We did not consider experiment 43 as its luminosity (1.857 fb^{-1}) is twenty five times less than the highest luminosity experiment (69 with $\mathcal{L} = 47(\text{fb})^{-1}$).

4. After applying all the selection criteria, we find the only remaining backgrounds appear from the decay channels, $B_s^0 \to \rho^+\rho^-$ and $B_s^0 \to K_S\pi^0$. Table 3.13 summarizes the backgrounds that have passed the selection criteria. The expected number of events at 121.4 fb^{-1} are obtained using the formula

$$N_{\text{expected}} = 2 \times \mathcal{B} \times N_{B_s\bar{B}_s} \times f, \tag{3.18}$$

where, $\mathcal{B}$ is the branching fraction, total number of $B_s\bar{B}_s$ pairs, $N_{B_s\bar{B}_s} = 8.3 \times 10^6$, and f is the fraction of events which have survived the selection criteria. The predicted $\mathcal{B}$ for $B_s^0 \to \rho^+\rho^-$ and $B_s^0 \to \bar{K}_S\pi^0$ are taken from references [20] and [19], respectively.

5. The expected number of rare B_s^0 background events at $\Upsilon(5S)$ is 0.0024(total), which is four hundred times less than the expected number of signal events, which is unity. Therefore, we do not expect these remaining backgrounds to hinder signal extraction significantly.

3.10 Best candidate selection (BCS)

The theoretically predicted branching fraction for this decay is very rare $\mathcal{O}(10^{-6})$ [19]. So, we do not expect multiple B_s^0 candidates per event in our analysis. However, after applying all the selection criteria, we still find 10.3% of multiple candidates. Therefore, we select the best B_s^0 among the multiple candidates by applying the following selection criterion

$$\chi^2_{min}(B_s^0) = \chi^2_{min}(\pi_1^0) + \chi^2_{min}(\pi_2^0), \tag{3.19}$$

where, $\chi^2(\pi^0)$ is defined as follows

$$\chi^2(\pi^0) = \frac{(M_{\gamma\gamma} - m_{\pi^0})^2}{\sigma^2}, \tag{3.20}$$

TABLE 3.12: List of decay channels considered for rare MC B_s background study.

Serial No.	Decay channel	DECAY.DEC $\mathcal{B}(10^{-5})$	PDG $\mathcal{B}(10^{-5})$	Remarks $\mathcal{B}(10^{-5})$
1	$B_s \to \rho^+\rho-$	–	–	0.1(Theory)
2	$B_s \to \eta\pi^0$	–	–	0.012(Theory)
3	$B_s \to \bar{K}^0_S\pi^0$	0.02	–	0.23(Theory)
4	$B_s \to \bar{K}^0_S K^0_S$	0.79	1.76	Re-generated
5	$B_s \to K^*(892)^+K^-$	–	1.9	–
6	$B_s \to K^*(892)^+\pi^-$	–	0.29	–
7	$B_s \to K^*_2(1430)^+K^-$	–	1.0	–
8	$B_s \to K^*_0(1430)^+K^-$	–	3.1	–
9	$B_s \to K^*_0(1430)^+\bar{K}^0$	–	3.3	–
10	$B_s \to K^{*0}_0(892)\bar{K}^0$	–	2.0	–
11	$B_s \to K^0_S\bar{K}^{*0}(892)$	–	1.6	–
12	$B_s \to K^{*0}(892)\bar{K}^{*0}(892)$	–	1.11	–
13	$B_s \to K^0_S\pi^+\pi^-$	–	0.95	–
14	$B_s \to K^0_S K^+\pi^-$	–	0.84	–
15	$B_s \to K^0_S K^+K-$	–	0.13	–

where, $M_{\gamma\gamma}$, m_{π^0}, and σ represent the invariant mass of the reconstructed π^0's, mass of π^0 from PDG [1], and the standard deviation (added in quadrature) associated with the measurements $M_{\gamma\gamma}$ and m_{π^0}, respectively.

Selecting the minimum $\chi^2(\pi^0)$ for the two π^0's enables us to select the best B^0_s among the multiple candidates. Our BCS criterion is 56% efficient in selecting the correctly reconstructed (MC truth matched) B^0_s candidates. The definition of BCS efficiency is as follows

$$\text{BCS efficiency (MCtruth)} = \frac{\text{Remaining MC truth candidates after BCS}}{\text{Number of multiple MC truth candidates before BCS}}. \tag{3.21}$$

Figure 3.19 shows the distribution of the multiplicity of B^0_s candidates.

TABLE 3.13: Table summarizing the scaled rare B_s^0 backgrounds at $\Upsilon(5S)$ resonance for the analysis, $B_s^0 \to \pi^0\pi^0$.

Expt. No.	Luminosity fb^{-1}	Events generated	Events passed	Decay channel	$\mathcal{B}$ $(10)^{-5}$	Expected events at 121.4 fb^{-1}
53	21.513	100000	2	$B_s \to \rho^+\rho^-$	0.1	0.000032
			1	$B_s \to \bar{K}_S^0\pi^0$	0.23	0.00038
67	27.222	100000	2	$B_s \to \rho^+\rho^-$	0.1	0.000032
			1	$B_s \to \bar{K}_S^0\pi^0$	0.23	0.00038
69	47.830	100000	5	$B_s \to \rho^+\rho^-$	0.1	0.000083
			3	$B_s \to \bar{K}_S^0\pi^0$	0.23	0.0011
71	22.938	100000	2	$B_s \to \rho^+\rho^-$	0.1	0.000032
			1	$B_s \to \bar{K}_S^0\pi^0$	0.23	0.00038

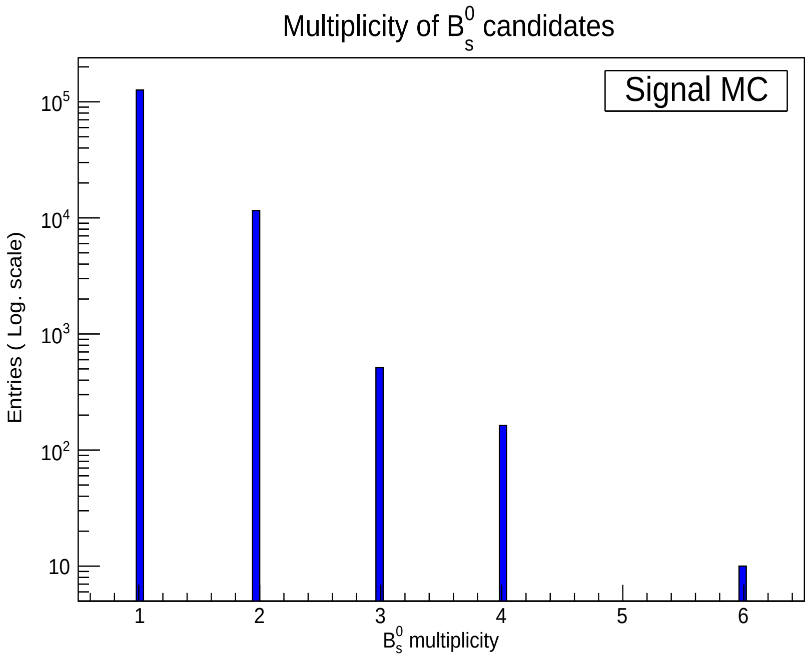

FIGURE 3.19: Distribution of multiplicity of B_s^0 candidates.

Summary

In this chapter, we discussed the procedure for event generation and simulation. We have also seen the preliminary event selection requirements used to select our event types for this analysis. We also discussed the selection criteria applied to identify the signals from the backgrounds. Background MC studies reveal the dominance of continuum background over the other types of backgrounds. We employed multivariate analysis using the neural network algorithm to suppress the continuum background. We also performed rare B_s^0 MC background studies and found their contribution to be negligible. The best candidate selection criterion is adopted to choose one from several B_s^0 candidates based on the sum of the minimum $\chi^2(\pi^0)$ of the two π^0 candidates.

Chapter 4

PDF parametrization and ensemble study

The previous chapter discussed the signal and background MC study in detail. This chapter explains the process of parametrization of the signal and the background MC using the variables M_{bc}, $\Delta E'$, and $\mathcal{C}'_{\text{NN}}$ and eventually developing a model to extract signal yields in data. We also perform "ensemble" studies to estimate the biases associated with the fit model and check its linearity and correlations among the variables.

4.1 Need for parametrization

The idea behind parametrizing a data sample is to convert its discreteness to a continuous distribution. The advantage of a continuous distribution is the ease of predicting the probability of occurrence of the event at any point within the distribution. This continuous distribution is a probability distribution function (PDF) and has a universal appeal because of its versatility in comparing results with similar experiments.

4.2 Tools involved in parametrization

Parametrization or PDF modelling is performed using `ROOFIT` [52] package, which nowadays comes integrated with the `ROOT` [53] analysis framework for hassle-free use. It involves primarily of the steps enumerated below:

1. Building the PDF
2. Constructing the likelihood function
3. Minimizing the log-likelihood function
4. Estimating parameter values
5. Calculating the errors on the parameters

The steps $2, 3, 4$, and 5 are jointly referred to as fitting data to a PDF. The following paragraphs briefly describe the above-mentioned steps.

4.2.1 Building the PDF

The `ROOFIT` package comes prepared with a few essential PDFs. There is also an option to build PDFs according to a user's need in case the desired PDF is not there using `EXPR` and `CEXPR` interfaces which are inbuilt into it. The most frequently used PDFs are listed below:

- `RooGaussian(x,mean,sigma)`: Gaussian PDF
- `RooExponential(x,alpha)`: Exponential PDF
- `RooPolynomial(x,a0,a1,a2,...)`: Polynomial PDF
- `RooChebychev(x,t0,t1,t2,...)`: Chebychev polynomial PDF
- `RooBreitWigner(x,mean,width)`: Non-relativistic Breit-Wigner PDF
- `RooLandau(x,mean,width)`: Landau PDF
- `RooCBShape(x,x0,sigma,alpha,n)`: Crystal Ball PDF
- `RooPoisson(x,`μ`)`: Poisson PDF
- `RooArgusBS(x,m,c)`: ARGUS phase space PDF
- `RooHistPdf`: Histogram-based shape PDF
- `RooKeysPdf`: Kernel estimation PDF

The PDF is built by declaring a variable with a user-defined range. The parameters of the PDF are then set accordingly with appropriate ranges. Sometimes, it may so happen that a single PDF does not fit the data accurately. In those situations, adding two or more PDFs has to be done to be consistent with the data. There is also a provision for adding two PDFs in `ROOFIT` through the use of `RooAddPdf` interface. As any PDFs generated are normalized probability density functions, the two added PDFs would also be normalized PDFs ensured by the relation

$$S(x) = f \cdot F(x) + (1 - f) \cdot G(x), \tag{4.1}$$

where, $S(x)$ is the total PDF, $F(x)$ and $G(x)$ are the component PDFs, and f is the weight of the first PDF $(F(x))$.

4.2.2 Constructing the likelihood function

The likelihood function is one of the most frequently used estimator methods for finding out the best fit values of the parameters of an observed sample of data. It is a good estimator because of its consistency, minimum bias, minimum variance bound, and efficiency. It is defined as the probability density of an observed data sample given a set of parameters. Mathematically,

$$\mathcal{L}(x_1, x_2, ..., x_N; \theta_1, \theta_2, ..., \theta_N) = |\frac{\partial P}{\partial x_i}|_{(\theta_1,\theta_2,...,\theta_N)}, \tag{4.2}$$

where, $i = 1, 2, ..., N$ and $P(x_1, x_2, ..., x_N)$ is the probability density with N unknown parameters, $\theta_1, \theta_2, ..., \theta_N$.

4.2.3 Minimizing the log-likelihood function

The likelihood function is maximized to estimate the best fit values of the parameters of an observed data sample. This procedure is known as the maximum likelihood method. From a numerical point of view, minimizing the function $-ln\mathcal{L}$ is more convenient than maximizing the likelihood function. Mathematically,

$$-\frac{\partial ln\mathcal{L}(x_1, x_2, ..., x_n; \theta_1, \theta_2, ..., \theta_n)}{\partial \theta_i} = 0, \tag{4.3}$$

where, $i = 1, 2, ..., n$.

In `ROOFIT`, the `RooMinimizer` interface class minimizes the likelihood functions using the `MINUIT` package. `MINUIT` on the other hand uses two algorithm for completing this task: the `MIGRAD` and the `HESSE` algorithms. `MIGRAD` algorithm tries to find out the minimum negative logarithm of the likelihood function by repeatedly calculating the gradient against the parameter values.

4.2.4 Estimating the parameter values

The parameter values which are obtained at the end of the minimization technique by `MIGRAD` algorithm are the considered to be the best fit parameter values for the PDF.

4.2.5 Calculating the errors on the parameters

The `HESSE` algorithm, based on an extended secant method, evaluates the error associated with each parameter by calculating the double derivative of the logarithm of the likelihood function. Therefore, at the end of a successful hunt for the best fit values of the parameters against a set of data, the fit of the PDF to the data sample is said to be completed.

4.3 Unbinned extended maximum likelihood

Sub-sections 4.2.2 and 4.2.3 describe the likelihood and minimization of the likelihood, respectively, in general. This subsection will describe them in the context of unbinned and extended approaches.

Let us consider a random variable x which is distributed according to the PDFs defined as, $S(x;\vec{\theta})$ and $B(x;\vec{\phi})$ with p and q unknown parameters, $\vec{\theta} = \theta_1, \theta_2, ..., \theta_p$ and $\vec{\phi} = \phi_1, \phi_2, ..., \phi_q$, respectively. We make N observations such that $\vec{x} = x_1, x_2, ..., x_N$.

We note that x is not segregated into certain class intervals or binned for any of the observations. This approach of not constraining data to bins is called an unbinned approach. Unbinned data analysis is particularly helpful for studying rare processes as the observable does not change across any bin width preventing any loss of information.

We can combine the PDFs to construct a singular model, $F(x;\vec{\theta},\vec{\phi})$ of the form

$$F(x;\vec{\theta},\vec{\phi}) = N_{\mathrm{S}}S(x;\vec{\theta}) + N_{\mathrm{B}}B(x;\vec{\phi}), \tag{4.4}$$

where, $N = N_\mathrm{S} + N_\mathrm{B}$. Normalizing Equation 4.4 with respect to N gives

$$F'(x;\vec{\theta},\vec{\phi}) = \frac{N_\mathrm{S}}{N_\mathrm{S}+N_\mathrm{B}} S(x;\vec{\theta}) + \frac{N_\mathrm{B}}{N_\mathrm{S}+N_\mathrm{B}} B(x;\vec{\phi}). \tag{4.5}$$

Here, we can relate $S(x;\vec{\theta})$ and $B(x;\vec{\phi})$ to signal and background PDFs, respectively. Correspondingly, N_S and N_B can be related to the number of signal and background events in data.

It is important to note that unlike a standard PDF normalized to unity, in Equation 4.5 the normalization is with respect to the total number of observations.

In addition, if N itself be treated as a random variable which distributes in accordance with a Poisson's distribution such as

$$\mathcal{N}(N_\mathrm{expected};N_\mathrm{observed}) = \frac{N_\mathrm{observed}^{N_\mathrm{expected}} e^{-N_\mathrm{expected}}}{N_\mathrm{expected}!}, \tag{4.6}$$

where, $N_\mathrm{expected} = N$. The likelihood function then takes the form

$$\mathcal{L}_\mathrm{e}(x_i;\vec{\theta},\vec{\phi},N_\mathrm{observed}) = \mathcal{N}(N_\mathrm{expected};N_\mathrm{observed}) \prod_{i=1}^{i=N} F'(x;\vec{\theta},\vec{\phi}), \tag{4.7}$$

$$\mathcal{L}_\mathrm{e}(x_i;\vec{\theta},\vec{\phi},N_\mathrm{observed}) = \frac{N_\mathrm{observed}^{N_\mathrm{expected}} e^{-N_\mathrm{expected}}}{N_\mathrm{expected}!} \prod_{i=1}^{i=N} F'(x;\vec{\theta},\vec{\phi}). \tag{4.8}$$

Equation 4.7 is called the unbinned extended likelihood. The inclusion of the number of observations to the likelihood function is the "extension" to the standard likelihood method. This enables the observation of signal and background yields in data. The minimization of the unbinned extended likelihood follows from Equation 4.3 to estimate the parameter values.

We use 3D ($1 \times 1 \times 1$) unbinned extended maximum likelihood fit to extract signal yield in real data. Therefore, we need to ensure that the fit variables, M_bc, $\Delta E'$ and $\mathcal{C}'_\mathrm{NN}$, are not correlated significantly. Otherwise, the correlation will reduce the signal identification capability of the fit variables.

4.4 Candidate and signal region

The candidate and the signal regions are defined as:

TABLE 4.1: Table showing candidate and signal regions. Here, 'all' indicates all the production channels.

Region	Production channel	M_{bc} (GeV/c^2)	$\Delta E'$ (GeV)	$\mathcal{C}'_{\mathrm{NN}}$
Candidate	all	$5.30 - 5.434$	$-0.6 - 0.15$	$-6.0 - 8.0$
Signal	$B_s^0 \bar{B}_s^0$	$5.345 - 5.380$	$-0.31 - 0.14$	$-3.901 - 7.451$
	$B_s^0 \bar{B}_s^{*0} + cc$	$5.380 - 5.395$	$-0.31 - 0.14$	$-3.901 - 7.451$
	$B_s^{*0} \bar{B}_s^{*0}$	$5.395 - 5.434$	$-0.31 - 0.14$	$-3.901 - 7.451$
	all	$5.345 - 5.434$	$-0.31 - 0.14$	$-3.901 - 7.451$

- Candidate region: This includes the region of the variables (M_{bc}, $\Delta E'$, $\mathcal{C}'_{\mathrm{NN}}$) over which we perform the PDF parametrization.
- Signal region: The signal region is chosen within three standard deviations of the best fit parameter values of the variable in concern. The best fit parameter values are obtained after 1D PDF parametrization.

The candidate and the signal regions are shown in Table 4.1.

4.4.1 Correlation coefficients among the fit variables

The linear correlation coefficient, r_{pq} between two variables, p and q for a number of observations is defined as

$$r_{pq} = \frac{a\sum_{k=1}^{a} p_k q_k - \sum_{k=1}^{a} p_k \sum_{k=1}^{a} q_k}{\sqrt{a\sum_{k=1}^{a} p_k^2 - (\sum_{k=1}^{a} p_i)^2}\sqrt{a\sum_{k=1}^{a} q_k^2 - (\sum_{k=1}^{a} q_k)^2}}. \tag{4.9}$$

The two variables are strongly correlated when r_{pq} approaches either $+1$ or -1 and are said to be independent of each other for $r_{pq} \to 0$. For this analysis, we assume any correlation below 20% or $r < 0.20$ to be negligible. The threshold on the correlation coefficient for assuming correlations not to be of significant concern is analysis dependent. It has to be adequately investigated and studied. Therefore, we perform a correlation study (to be discussed in Sub-section 4.7.3) to verify our assumption that the correlations among the fit variables can be safely neglected.

Table 4.2 and 4.3 summarizes the linear correlation coefficients among the fit variables for MC signal and continuum MC background, respectively. The corresponding scatter plots are shown in Figures 4.1 and 4.2, respectively.

Table 4.2: Linear correlation coefficients among the fit variables from signal MC study. Here, 'all' indicates all the production channels. The candidate and signal regions are defined in Table 4.1

Variables	Production channel	Candidate region	Signal region
$M_{\rm bc} : \Delta E'$	all	$+0.095$	-0.006
	$B_s^0 \bar{B}_s^0$	–	-0.077
	$B_s^0 \bar{B}_s^{*0} + cc$	–	-0.078
	$B_s^{*0} \bar{B}_s^{*0}$	–	-0.048
$\Delta E' : \mathcal{C}'_{\rm NN}$	all	$+0.043$	$+0.028$
	$B_s^0 \bar{B}_s^0$	–	$+0.035$
	$B_s^0 \bar{B}_s^{*0} + cc$	–	$+0.020$
	$B_s^{*0} \bar{B}_s^{*0}$	–	$+0.021$
$M_{\rm bc} : \mathcal{C}'_{\rm NN}$	all	$+0.003$	-0.0002
	$B_s^0 \bar{B}_s^0$	–	$+0.006$
	$B_s^0 \bar{B}_s^{*0} + cc$	–	$+0.009$
	$B_s^{*0} \bar{B}_s^{*0}$	–	-0.005

As we do not find any significant correlation among the fit variables, we proceed with 1D PDF parametrizations for the three fit variables, $M_{\rm bc}$, $\Delta E'$ and $\mathcal{C}'_{\rm NN}$.

4.4.2 1D PDF modelling

The 1D PDFs are modelled from a large sample of signal and background MC statistics. The signal $M_{\rm bc}$ PDF consists of three PDFs corresponding to the three B_s^0 production channels. Each of them is again separately modelled from large MC samples. They are then combined according to their production fractions [8] to produce the final signal PDF for the $M_{\rm bc}$ variable. The PDF used for parametrizing $B_s^0 \bar{B}_s^0$ is a sum of two Gaussian functions while each of $B_s^0 \bar{B}_s^{*0}$ or $B^{*0} \bar{B}_s^0$, and $B_s^{*0} \bar{B}_s^{*0}$, are parametrized using a sum of a Gaussian function and an empirical PDF shape introduced by Crystal Ball [54]. The signal $\Delta E'$ variable is modelled using the Crystal Ball function, which is modified for this analysis to include the asymmetric nature of the distribution about the mean position. The output from the NN is parametrized using a Gaussian and an asymmetric (bifurcated) Gaussian PDF for the signal $\mathcal{C}'_{\rm NN}$ variable. The continuum background distribution of the $M_{\rm bc}$

TABLE 4.3: Linear correlation coefficients among the fit variables from background MC study. Here, 'all' indicates all the production channels. The candidate and signal regions are defined in Table 4.1

Variables	Production channel	Candidate region	Signal region
$M_{\rm bc} : \Delta E'$	all	+0.091	−0.011
	$B_s^0\bar{B}_s^0$	–	−0.070
	$B_s^0\bar{B}_s^{*0} + cc$	–	−0.011
	$B_s^{*0}\bar{B}_s^{*0}$	–	−0.026
$\Delta E' : \mathcal{C}'_{\rm NN}$	all	+0.029	+0.011
	$B_s^0\bar{B}_s^0$	–	−0.038
	$B_s^0\bar{B}_s^{*0} + cc$	–	+0.009
	$B_s^{*0}\bar{B}_s^{*0}$	–	+0.031
$M_{\rm bc} : \mathcal{C}'_{\rm NN}$	all	−0.020	+0.026
	$B_s^0\bar{B}_s^0$	–	−0.051
	$B_s^0\bar{B}_s^{*0} + cc$	–	+0.06
	$B_s^{*0}\bar{B}_s^{*0}$	–	−0.002

variable is modelled through an empirically determined parametrized background shape introduced by ARGUS [55]. The continuum background is parametrized using a first order Chebychev polynomial and a sum of two Gaussian distributions for the $\Delta E'$ and $\mathcal{C}'_{\rm NN}$ variables, respectively. Table 4.4 and 4.5 summarizes the various PDF models that are being used for parametrizing the signal and background MC for the fit variables, $M_{\rm bc}$, $\Delta E'$ and $\mathcal{C}'_{\rm NN}$.

The formulae of CB, ARGUS and ACB are defined in appendix C.1. The 1D PDF parametrization of the fit variables for signal and background MC are shown in Figures 4.3, 4.4, 4.5, and 4.6.

4.5 Strategy for signal extraction

This section describes the strategy we follow to extract the signal in data. They are enumerated below:

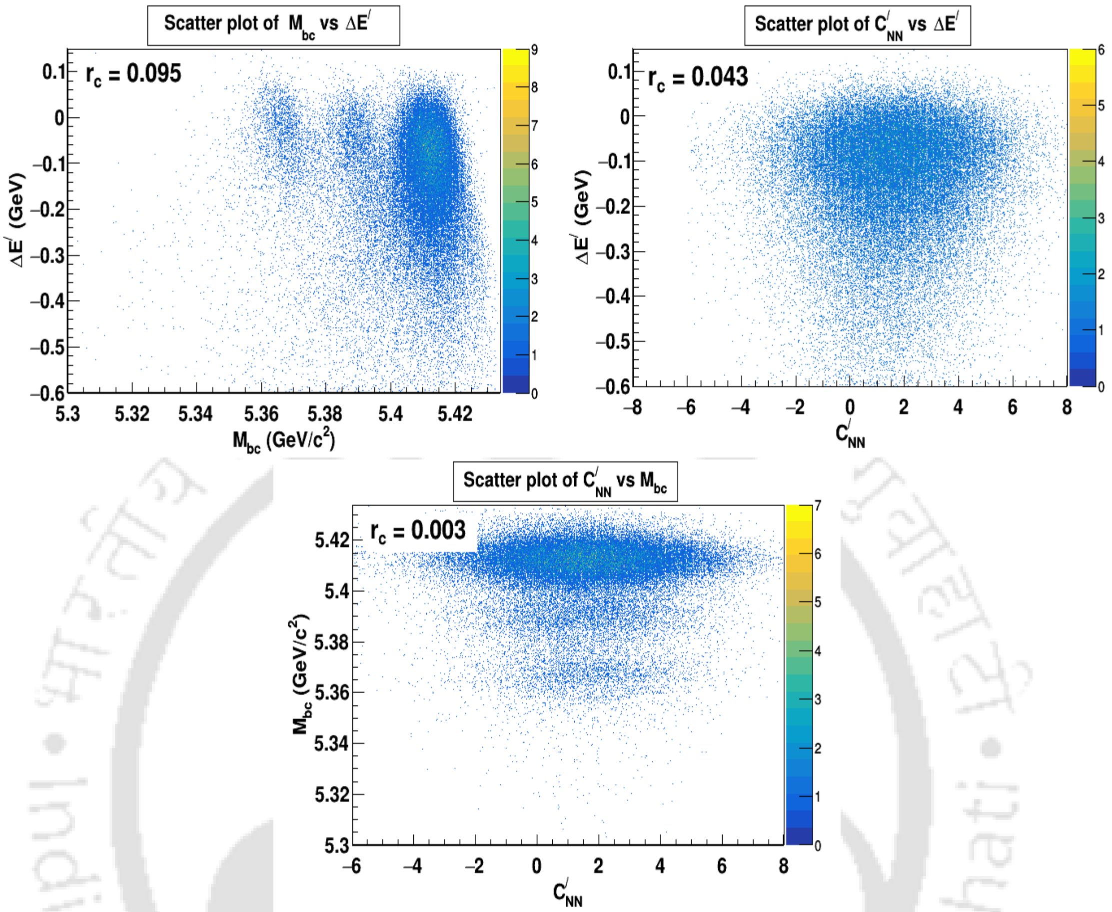

FIGURE 4.1: Scatter plots among the fit variables for signal MC, $M_{\rm bc}$: $\Delta E'$ (top-left), $\mathcal{C}'_{\rm NN}$: $\Delta E'$ (top-right), and $\mathcal{C}'_{\rm NN}$: $M_{\rm bc}$ (bottom). Inset variable, r_c represents the linear correlation coefficient.

TABLE 4.4: PDF models for signal MC. Here, G, CB, ACB, and BG signify Gaussian, Crystal Ball, Asymmetric Crystal Ball, and Bifurcated Gaussian, respectively.

Variables	PDF model	Parameters	Remarks
$M_{\rm bc}$	G + G ($B_s^0\bar{B}_s^0$)	μ, σ_1, σ_2	Common mean
	G + CB ($B_s^{*0}\bar{B}_s^0$ + cc)	μ, σ(G), σ(CB), n, α	Common mean
	G + CB ($B_s^{*0}\bar{B}_s^{*0}$)	μ, σ(G); μ, σ (CB), n, α	–
$\Delta E'$	ACB	μ, σ_L, σ_R, n, α	–
$\mathcal{C}'_{\rm NN}$	G + BG	μ, σ (G) ; μ, σ_L, σ_R (BG)	–

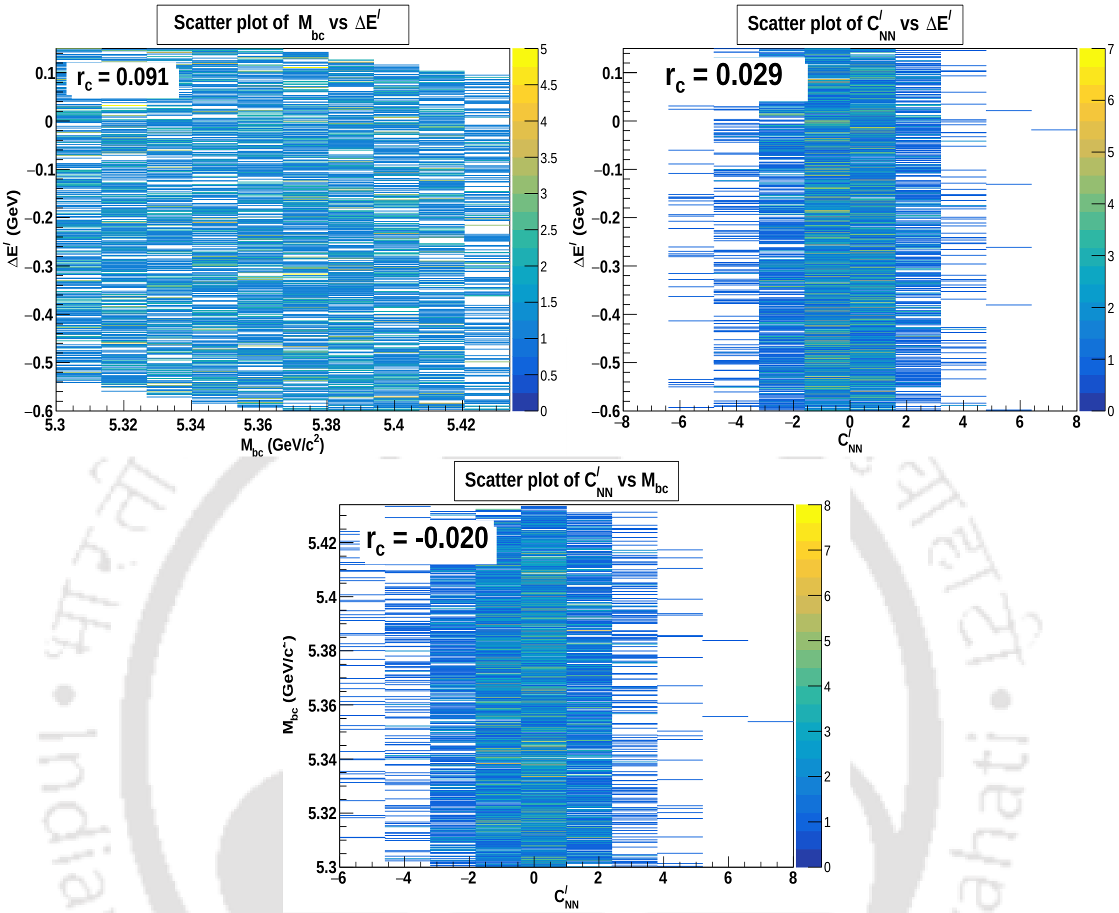

FIGURE 4.2: Scatter plots among the fit variables for background MC, $M_{\rm bc}$: $\Delta E'$ (top-left), $\mathcal{C}'_{\rm NN}$: $\Delta E'$ (top-right), and $\mathcal{C}'_{\rm NN}$: $M_{\rm bc}$ (bottom). Inset variable, r_c represents the linear correlation coefficient.

TABLE 4.5: PDF models for background MC. Here, CP, and G signify Chebychev Polynomial, and Gaussian.

Variables	PDF model	Parameters	Remarks
$M_{\rm bc}$	ARGUS (all three production channels)	c, End point	End point fixed at 5.434 GeV/c^2
$\Delta E'$	First order CP	a	–
$\mathcal{C}'_{\rm NN}$	G + G	μ, σ (G) ; μ, σ (G)	–

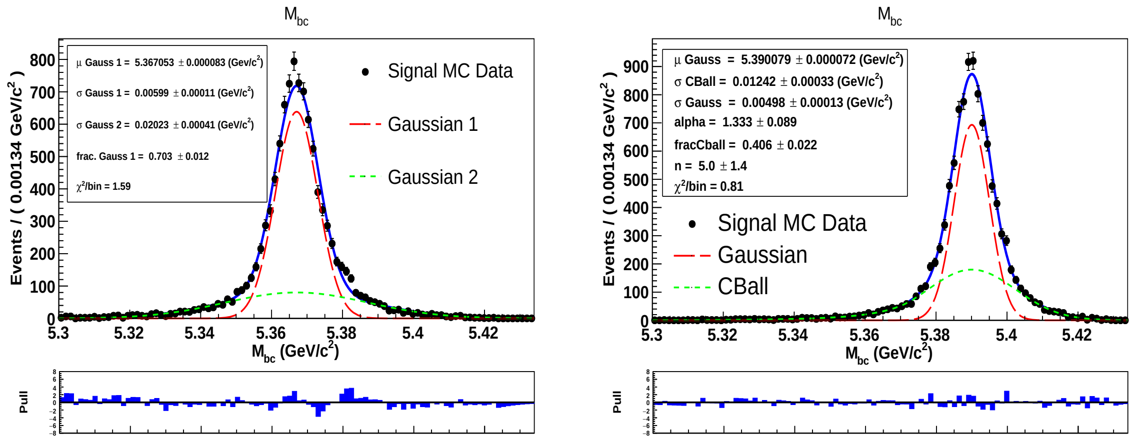

FIGURE 4.3: 1D PDF parametrizations of $M_{\rm bc}$ distributions for the production channels, $B_s^0\bar{B}_s^0$ (left) and $B_s^{*0}\bar{B}_s^0$ + cc (right) for signal MC.

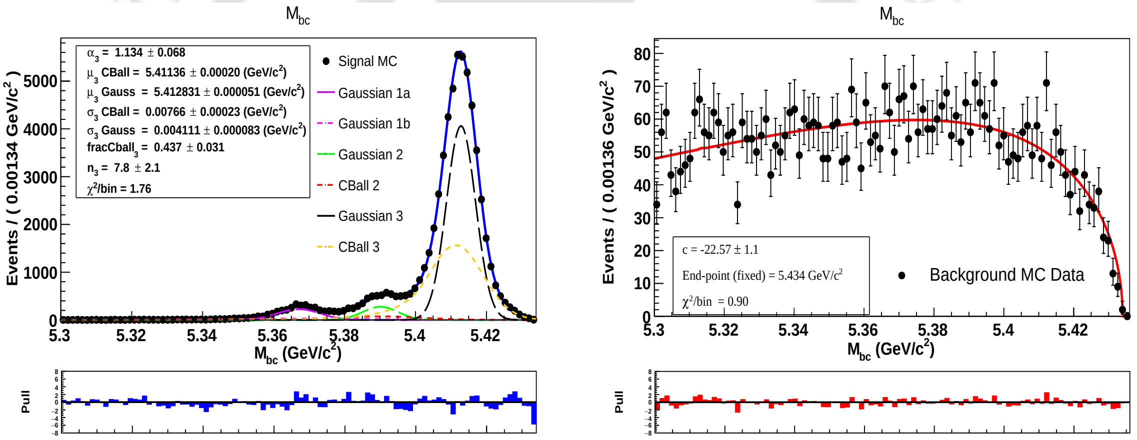

FIGURE 4.4: 1D PDF parametrizations of $M_{\rm bc}$ distributions for signal (all three production channels) (left) and background (right) MC.

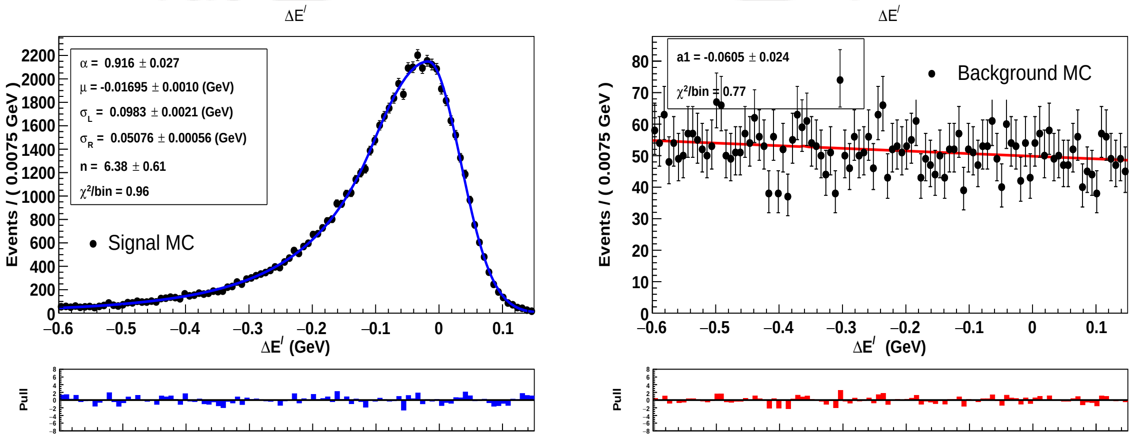

FIGURE 4.5: 1D PDF parametrizations of $\Delta E'$ distributions for signal (left) and background (right) MC.

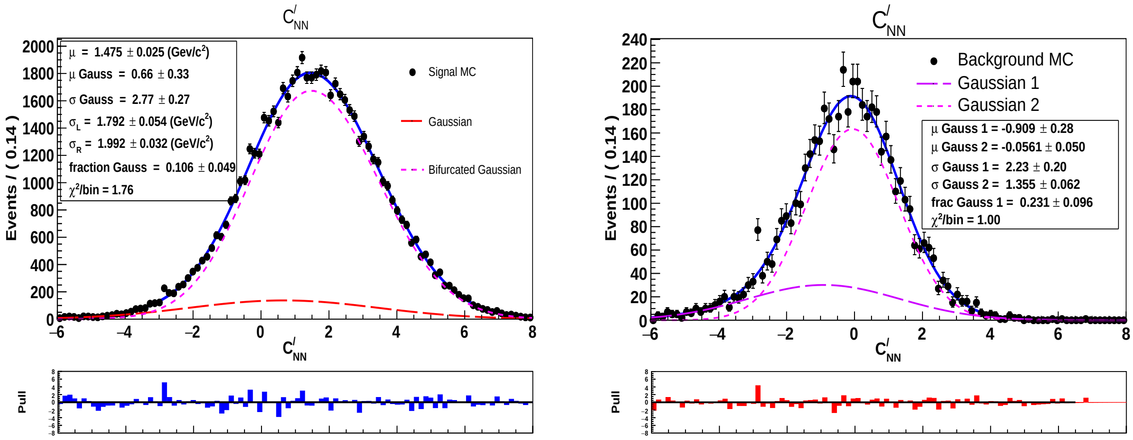

FIGURE 4.6: 1D PDF parametrizations of $\mathcal{C}'_{\rm NN}$ distributions for signal (left) and background (right) MC.

1. **Build the fit model**: We use a 3D($\Delta E' \times M_{\rm bc} \times \mathcal{C}'_{\rm NN}$) unbinned extended maximum likelihood fit method to extract the signal and background yields. The signal and the background yields are floated. The likelihood function of the fit model is mathematically represented as

$$\mathcal{L}_{\rm fit} = e^{-\sum\limits_j n_j} \prod_i^N \left(\sum_j n_j P_j(M^i_{\rm bc}, \Delta E'^i, \mathcal{C}'_{\rm NN}{}^i) \right), \tag{4.10}$$

   where, $P_j(M^i_{\rm bc}, \Delta E'^i, \mathcal{C}'_{\rm NN}{}^i)$ is the PDF of the signal or background component (specified by index j), n_j is the yield of this component, and N is the total number of events in the sample.

   Further, we do not observe any significant correlation among the fit variables and therefore, the three-dimensional PDF P_j factorizes as

$$P_j \equiv P_j(M_{\rm bc}) P_j(\Delta E') P_j(\mathcal{C}'_{\rm NN}). \tag{4.11}$$

2. **Fixed and Floated parameters** : To extract the signal and background yields from the three-dimensional fit model, we fix all the signal PDF parameters obtained from signal MC study and float all the background PDF parameters except the ARGUS end-point. In addition, we also float the signal and the background yields. Tables 4.6 and 4.7 summarize the signal and background PDF parameters obtained from 1D PDF parametrization.

TABLE 4.6: Summary of Signal MC PDF parameters which will be used for signal extraction. Here, G, CB, ACB, and BG signify Gaussian, Crystal Ball, Asymmetric Crystal Ball, and Bifurcated Gaussian PDFs, respectively. Mean, standard deviation, and tail parameters of CB are represented by μ, σ, $\sigma_{L,R}$, and α, n, respectively. Further, 'all' represents all the production channels taken together.

Sl. No	Variable	Channel	PDF	Parameters	Parameter Values	Remarks
1	M_{bc} (GeV/c^2)	$B_s^0\bar{B}_s^0$	G1	μ	5.367053	Fixed
2				σ	0.00599	Fixed
3				fraction	0.703	Fixed
4			G2	σ	0.02023	Fixed
5		$B_s^0\bar{B}_s^0$		fraction	0.057	Fixed
6	M_{bc} (GeV/c^2)	$B_s^0\bar{B}_s^{*0}+\mathrm{cc}$	G	μ	5.390079	Fixed
7				σ	0.00498	Fixed
8			CB	σ	0.01242	Fixed
9				α	1.333	Fixed
10				n	5.0	Fixed
11				fraction	0.406	Fixed
12		$B_s^0\bar{B}_s^{*0}+cc$		fraction	0.073	Fixed
13	M_{bc} (GeV/c^2)	$B_s^{*0}\bar{B}_s^{*0}$	G	μ	5.41283	Fixed
14				σ	0.00411	Fixed
15			CB	μ	5.41136	Fixed
16				σ	0.00766	Fixed
17				α	1.134	Fixed
18				n	7.8	Fixed
19				fraction	0.437	Fixed
20		$B_s^{*0}\bar{B}_s^{*0}$		fraction	0.87	Fixed
21	$\Delta E'$ (GeV)	all	ACB	μ	-0.01695	Fixed
22				σ_L	0.0983	Fixed
23				σ_R	0.0507	Fixed
24				α	0.916	Fixed
25				n	6.38	Fixed
26	$\mathcal{C}'_{\mathrm{NN}}$	all	BG	μ	1.475	Fixed
27				σ_L	1.792	Fixed
28				σ_R	1.992	Fixed
29			G	μ	0.66	Fixed
30				σ	2.77	Fixed
31				fraction	0.106	Fixed

Table 4.7: Summary of background MC PDF parameters. Here, ARGUS, CP and G signify ARGUS function, Chebychev polynomial amd Gaussian PDFs, respectively. Mean, standard deviation and coefficient of first order CB are represented by μ, σ, and a_0, respectively. Further, ‘all’ represents all the production channels taken together.

Sl. No	Variable	Channel	PDF	Parameters	Remarks
1	$M_{\rm bc}$ (GeV/c^2)	all	ARGUS	m (End point)	Fixed (5.434)
2				c (Curvature)	Floated
3	$\Delta E'$ (GeV)	all	First order CP	a_0	Floated
4	$\mathcal{C}'_{\rm NN}$	all	G	μ	Floated
5				σ	Floated
6		all	G	μ	Floated
7				σ	Floated

4.6 Results from signal and background MC study

In this section, we report the results obtained from signal and background MC study.

1. **Signal reconstruction efficiency** : After applying all the selection criteria we are left with 63477 signal B_s^0 out of 500000 generated signal events. We, therefore, have a signal reconstruction efficiency of 12.69% (considering all the production channels), which is 36% more than the L3 experiment [22]. The break up of the reconstruction efficiencies considering the production channels are shown in Table 4.8.

2. **Best candidate selection** : From signal MC study, we observe that we have 10.3% of multiple candidates in our analysis. We, therefore, select the best B_s^0 candidates by requiring the sum of the two χ^2 obtained after a mass constrained fit on the π^0s to be minimum. This selection criteria is 56% efficient in selecting the correct B_s^0 among the multiple candidates in an event.

3. **Signal misreconstruction fraction or self cross-feed (SCF)**: We find 1.7% and 1.2% of SCF events in the fitting and the signal region, respectively. They are small in quantity and hence not parametrized separately. The SCF events are those events which are not correctly reconstructed, that is, the MC truth matched information are not available for these candidates. The distributions of the SCF fractions are shown in Figures D.1, D.2, and D.3 in appendix D.1.

4. **Signal and background MC events in the candidate and the signal region**: We have seen from Table 3.10 that the dominant background in our analysis are the continuum. We expect 862 ± 30 backgrounds in the candidate region and 328 ± 18 of them in the signal region. Table 4.9 shows the expected signal and background yield from MC study in real data. The expected signal yield is calculated considering the predicted $\mathcal{B}(B_s^0 \to \pi^0\pi^0)$ of $(0.40 \pm 0.27) \times 10^{-6}$ using the flavor diagram approach. The expected signal yield is calculated using the formula

$$N^{\rm sig}_{\rm expected} = 2 \times N_{B_s^0\bar{B}_s^0} \times \mathcal{B}(B_s^0 \to \pi^0\pi^0) \times \epsilon^{\rm sig}_{\rm rec} \times \prod_{i=1}^{2} \mathcal{B}(\pi^0 \to \gamma\gamma) \tag{4.12}$$

   where, $N_{B_s^0\bar{B}_s^0} = (8.30 \pm 1.34) \times 10^6$ is the number of $B_s^0\bar{B}_s^0$ pairs at $\Upsilon(5S)$, $\epsilon^{\rm sig}_{\rm rec}$ is the signal reconstruction efficiency and $\mathcal{B}(\pi^0 \to \gamma\gamma) = (98.823 \pm 0.034)\%$ [1]. The factor 2 accounts for the two $B_s^0(\bar{B}_s^0)$ mesons produced in an event.

5. **Correlation among the fit variables**: The correlations among the fit variables are not very significant. Therefore, each of the fit variables can be assumed to be independent and can be parametrized separately.

6. **3D unbinned extended maximum likelihood fit**: We plan to perform a 3D unbinned extended maximum likelihood fit to extract the signal yields in data, where the yields of the signal and the backgrounds will be floated.

4.7 Ensemble study

The 3D function (described in Equation 4.11) developed in the preceding sections for signal extraction is modelled from large MC samples. However, we do not expect to observe these many signal events in data (Table 4.9 tabulates the expected signal yields). We thus need to verify the stability of the model at the expected sensitivity of the SM. We also look for correspondence between the input signal events and the output signal yields obtained from the 3D function. In addition, we perform correlation studies among the fit variables ($M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$) at the expected sensitivity of the SM. This section is devoted to these studies for assigning systematic uncertainties to the model. Henceforth, "fit" will mean the 3D function developed using an unbinned extended maximum likelihood method (discussed in Sub-section 4.3). Therefore, reference to any bias in the model will be reported as "fit bias".

Table 4.8: Summarizing the signal reconstruction efficiencies. The signal and candidate regions are defined in Table 4.1

Production channel	Production fraction (%)	Region	Signal reconstruction efficiency ($\epsilon_{\rm rec}^{\rm sig}$)
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$ $+\ B_s^0\bar{B}_s^0$	100%	Candidate	12.69%
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$ $+\ B_s^0\bar{B}_s^0$	100%	Signal	11.67%
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$	$\sim 94.3\%$	Signal	10.99%
$B_s^{*0}\bar{B}_s^{*0}$	$\sim 87\%$	Signal	10.21%
$B_s^0\bar{B}_s^{*0} + cc$	$\sim 7.3\%$	Signal	0.78%
$B_s^0\bar{B}_s^0$	$\sim 5.7\%$	Signal	0.68%

Table 4.9: Expected signal and background yield from MC study. The signal and candidate regions are defined in Table 4.1. The background yields have been scaled to real data luminosity of 121.4 fb^{-1}.

Production channel	Production fraction (%)	Region	$\epsilon_{\rm rec}^{\rm sig}$	Signal Yield	Background Yield
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$ $+\ B_s^0\bar{B}_s^0$	100%	Candidate	12.69%	0.78 (1 ± 1)	862 ± 30
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$ $+\ B_s^0\bar{B}_s^0$	100%	Signal	11.67%	0.81 (1 ± 1)	328 ± 18
$B_s^{*0}\bar{B}_s^{*0}$ $+\ B_s^0\bar{B}_s^{*0} + cc$	$\sim 94.3\%$	Signal	10.99%	0.75 (1 ± 1)	185 ± 14
$B_s^{*0}\bar{B}_s^{*0}$	$\sim 87\%$	Signal	10.21%	0.65(1 ± 1)	121 ± 11
$B_s^0\bar{B}_s^{*0} + cc$	$\sim 7.3\%$	Signal	0.78%	0.05(0)	65 ± 8
$B_s^0\bar{B}_s^0$	$\sim 5.7\%$	Signal	0.68%	0.04(0)	142 ± 12

We define a few terminologies used throughout the chapter for convenience of the readers, which are discussed below:

- Ensemble: An ensemble is a set of several similar pseudo-experiments which impersonate data. They are a crucial part of the validation process and helps to gauge the functioning of the developed model.
- Pull: Any random variable, say, x, when is observed to occur by Poissonian (or Binomial or Gaussian) distribution with an expected mean of μ for many trials, it distributes itself as a Gaussian distribution with a mean, μ and standard deviation, σ as a consequence of the central limit theorem. The weighted deviation observed in the distribution from the expected mean, μ, is defined as the pull of the distribution. Mathematically,

  $$\text{Pull} = \frac{x-\mu}{\sigma}. \tag{4.13}$$

  In the context of the analysis, the random variable will refer to the signal yields from the fit function, $N^{\text{sig}}_{\text{yield}}$, the expected mean will be the input signal events to the fit function, $N^{\text{sig}}_{\text{gen}}$ and σ refers to the uncertainty on the signal yields. Therefore, the expression for pull takes the form

  $$\text{Pull} = \frac{N^{\text{sig}}_{\text{yield}} - N^{\text{sig}}_{\text{gen}}}{\sigma}. \tag{4.14}$$

  An ideal pull distribution will have a mean value of zero and a unit standard deviation.

4.7.1 Fit bias

We create ensembles from MC samples to represent the data. The model developed from the signal and the background PDFs are repeatedly fitted to this generated sample of MC to look for differences in the yield from the expected number of signal events during each iteration. The differences are reported as systematic uncertainties. A detailed description of the ensemble test is as follows:

1. **Extraction of signal and background events from MC samples** : Signal MC are generated and simulated based on the physics processes described in Section 3.3 using EvtGen and `GEANT3`, respectively. Once reconstructed, the selection criteria

mentioned in Table 3.10 are applied. A random sampling based on a Poissonian distribution is performed on the pool of surviving candidates to extract the expected number of signal events in data (Table 4.9 tabulates the expected signal yields). Random sampling chooses the events independent of each other and reduce selection biases. For this analysis, we select the following set of expected signal events to perform the ensemble test, (1, 2, 3, 5, 7, 10, 20) . This range is chosen based on the maximum expected signal yield with a signal reconstruction efficiency of 12.69% for this analysis.

In contrast, background MC events are extracted at random using a Poissonian distribution based on their expected yields in data (Table 4.9 tabulates the expected background yields) from the background PDFs (Table 4.5 refers to the background PDFs).

While extracting the signal and background MC events, we fix all the signal parameters along with the ARGUS end-point of the background PDF and float all other parameters of the background PDFs to reduce systematic uncertainties.

This process of random extraction of signal events after its generation and simulation using MC will be termed as “Gsim (**G**eneration and **sim**ulation)” throughout the text to distinguish it from Toy MC procedure (to be discussed in later paragraphs).

2. **Iterating 3D unbinned extended maximum likelihood procedure on the extracted events**: We perform a fit based on the developed 3D model on the extracted MC events to calculate the signal yields. The fit is performed for several iterations (1000 in our case) to access the systematic effects associated with the model. The differences in the signal yields from the input signal events follow a Gaussian distribution. The mean of this distribution gives the average difference in yields, while the standard deviation gives the systematic uncertainty. The distribution is normalized with respect to the standard deviation such that the mean is around zero and the standard deviation is unity. We calculate the pull distributions for each of the expected signal yields to estimate the systematic uncertainty associated with the model.

3. **Assigning systematic uncertainty to the model (fit bias):** The average deviation of the zeroth order polynomial fit to the mean of the pull values for various expected signal yields from the no bias condition (zero mean and unit standard deviation) is recognized as a fit bias.

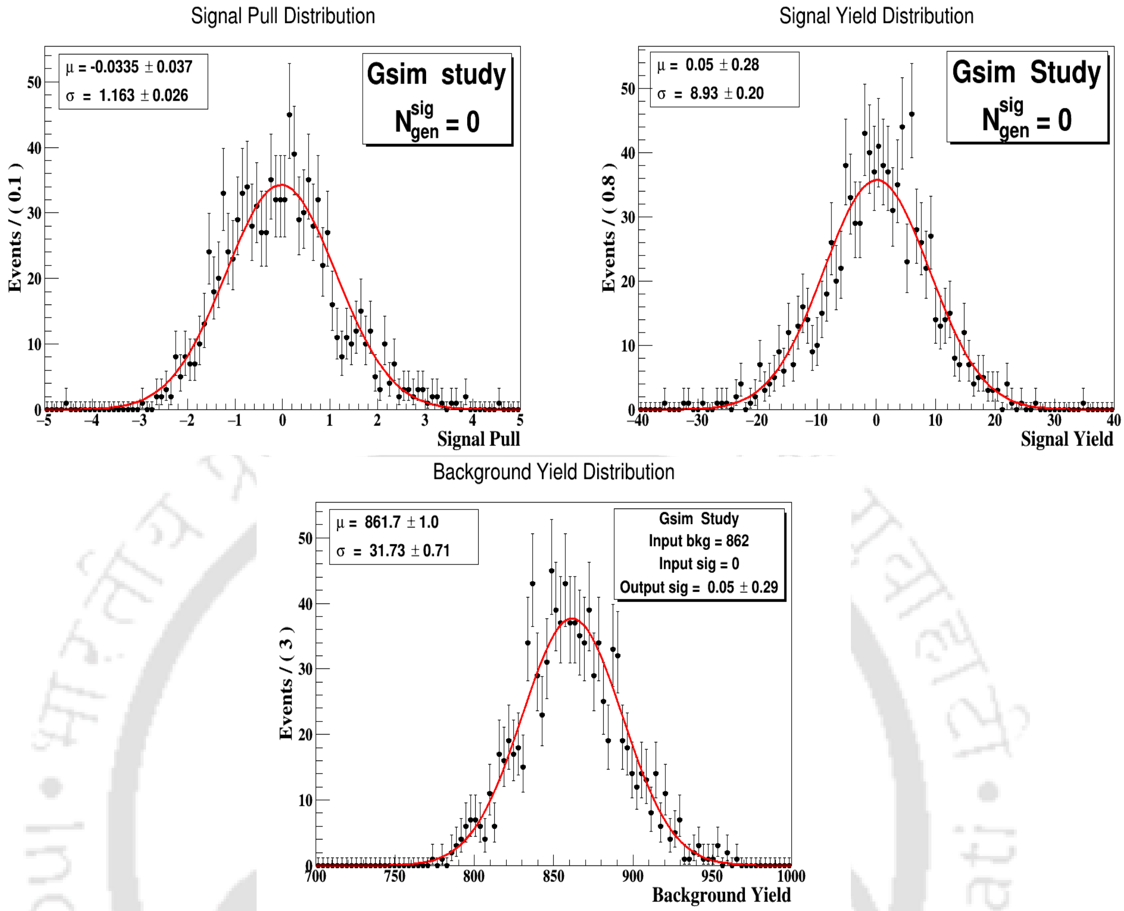

FIGURE 4.7: Signal pull (top-left), signal yield (top-right) and background yield (bottom) distribution for $N^{sig}_{gen} = 0$ (i.e., background only)

Figure 4.7 shows the pull, yields of signal MC, and yields of background MC from the fit, for $N^{sig}_{gen} = 0$ (i.e., background only), where, N^{sig}_{gen} is the input number of signal events given to the fit.

The signal pull distribution parameters (mean, μ and standard deviation, σ) agree with a standard normal distribution. The signal yield from the fit ($N^{sig}_{yield} = 0.050 \pm 0.28$) agrees with the input signal events ($N^{sig}_{gen} = 0$) within the MC statistical uncertainty. We also do not find any significant difference between the background yield from the fit (861.7 ± 1.0) and the input background (862).

The pull and signal yield distributions for $N^{sig}_{gen} = 1, 2, 3, 5, 7, 10,$ and 20 are shown in appendix E.0.1. Figures 4.8 and 4.9 show the average pull distributions for the cases when we consider $N^{sig}_{gen} = 0, 1, 2, 3, 5, 7, 10,$ and 20 and $N^{sig}_{gen} = 0, 1, 2, 3, 5,$ and 7, respectively.

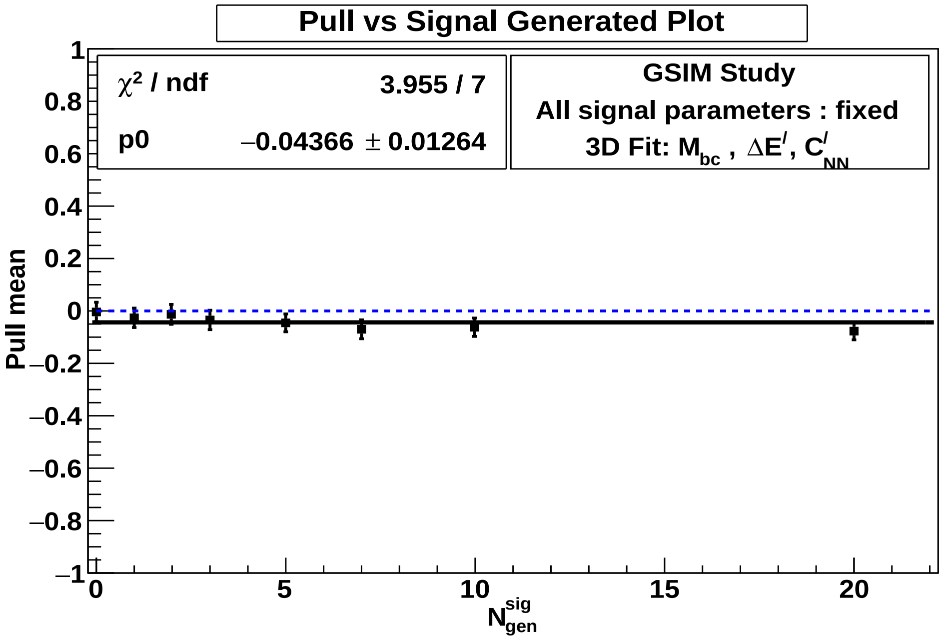

FIGURE 4.8: Average pull distribution for $N^{\mathrm{sig}}_{\mathrm{gen}} = 0, 1, 2, 3, 5, 7, 10, 20$. The data points with error bars, dashed-blue, and solid-black line represent signal MC, the no-bias condition, and zeroth order polynomial fit, respectively.

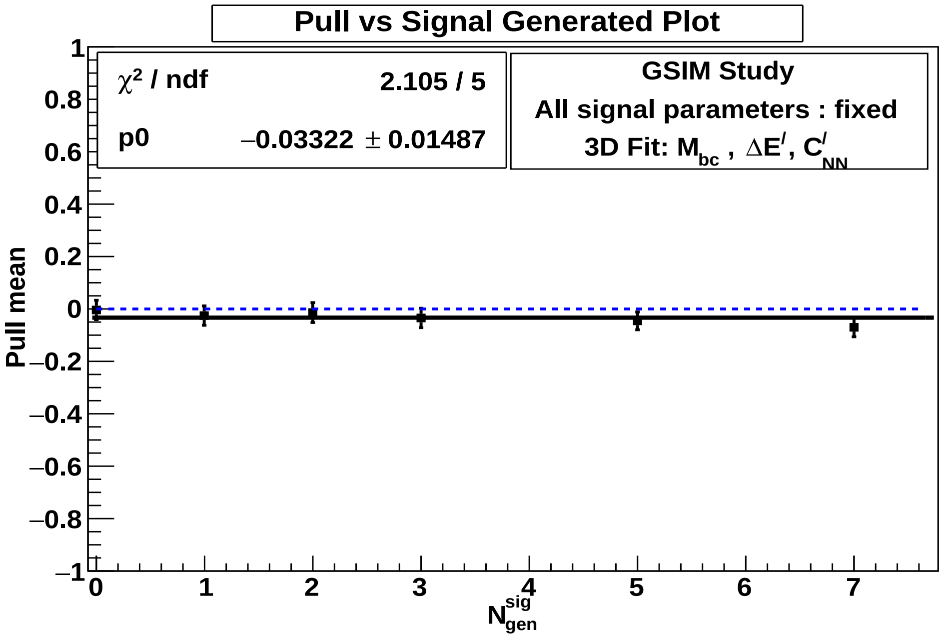

FIGURE 4.9: Average pull distribution for $N^{\mathrm{sig}}_{\mathrm{gen}} = 0, 1, 2, 3, 5, 7$. The data points with error bars, dashed-blue, and solid-black line represent signal MC, the no-bias condition, and zeroth order polynomial fit, respectively.

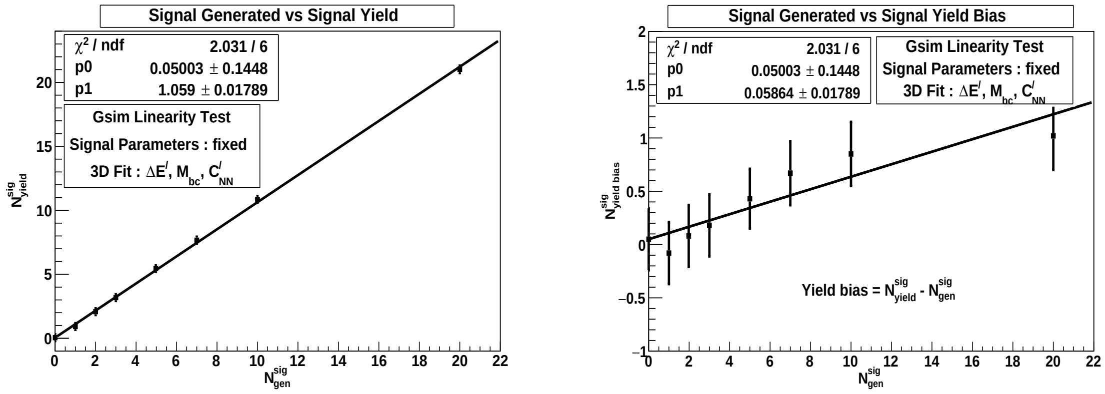

FIGURE 4.10: Plots of $N^{\rm sig}_{\rm gen}$ versus $N^{\rm sig}_{\rm yield}$ (left) and $N^{\rm sig}_{\rm gen}$ versus $N^{\rm sig}_{\rm yield\ bias}$ (right). The left plot gives the information on the linearity and the second shows the bias in the yields obtained from the fit model. The signal MC points and the first order polynomial fit are shown as black dots with error bars and solid-black line, respectively.

We assign the deviation from zero (as indicated by zeroth order polynomial fit parameter, p0) observed in the average pull distribution plots (Figures 4.8 and 4.9) as a source of systematic uncertainty coming from the fit model (systematic uncertainties will be discussed in the subsequent chapters). Since, we do not expect more than one signal event in data, the average pull distribution of -3.3% (as shown in Figure 4.9) will be a sufficient estimate for this bias.

4.7.2 Linearity test

The linearity test checks the correspondence between the output signal yield and the input signal events. If the fit is unbiased, we expect this correspondence to be linear. The bias in the linearity is defined as

$$N^{\rm sig}_{\rm yield\ bias} = N^{\rm sig}_{\rm yield} - N^{\rm sig}_{\rm gen}. \tag{4.15}$$

From Figure 4.10, we infer that the fit model is linear. This is verified by comparing the first order polynomial fit parameters $(\text{slope}, \text{p1} = 1.059 \pm 0.018$ and $\text{intercept}, \text{p0} = 0.050$ $\pm\ 0.144)$ with the expected values $(\text{slope}, \text{p1} = 1$ and $\text{intercept}, \text{p0} = 0)$ in the first plot. The second plot verifies this result as the intercept of the bias(p0) matches the linearity plot's intercept. We, therefore, do not assign any systematic uncertainty due to linearity in the fit model.

TABLE 4.10: Average linear correlation coefficients among the fit variables in signal MC and (signal + background) MC obtained from Gsim study.

Variables	Signal MC	(signal + background) MC
$\Delta E'$ vs. $M_{\rm bc}$	-4.7%	-0.02%
$\Delta E'$ vs. $\mathcal{C}'_{NN}$	$+0.7\%$	$+0.05\%$
$M_{\rm bc}$ vs. $\mathcal{C}'_{\rm NN}$	$+0.16\%$	$+0.06\%$

4.7.3 Correlation studies

The correlation coefficients among the fit variables $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$ for signal and background MC are defined in Tables 4.2 and 4.3, respectively. As stated earlier, we do not observe any significant correlation among the fit variables. However, we perform a correlation test using the Gsim study to reaffirm our observation.

For correlation studies, we extract the signal and the background events as discussed in the section 4.7.1 and iterate over them to calculate the linear correlation coefficients between any two fit variables at a time. We then fit the obtained distribution with a zeroth order polynomial to assign the correlation coefficient among the fit variables. Figures 4.11 and 4.12 show the correlation coefficients among the fit variables for $\mathrm{N}^{\rm sig}_{\rm gen} = 0, 1, 2, 3, 5, 7, 10, 20$ considering signal MC and (signal + background) MC in the signal region (for signal region see Table 4.1).

As seen from the values of p0 in Table 4.10 obtained from Figures 4.11 and 4.12, the correlations among the fit variables are not very significant.

We summarize our Gsim study in the following points:

1. We observe an average fit bias of -3.3% in the fit model. We will assign this fit bias as a source of systematic uncertainty.
2. We find the fit model to be linear. So, we do not assign any systematic uncertainty from the fit's linearity.
3. We do not find any significant correlation among the fit variables. Therefore, the PDFs parametrizing the fit variables can be assumed to be independent and can be factorized.

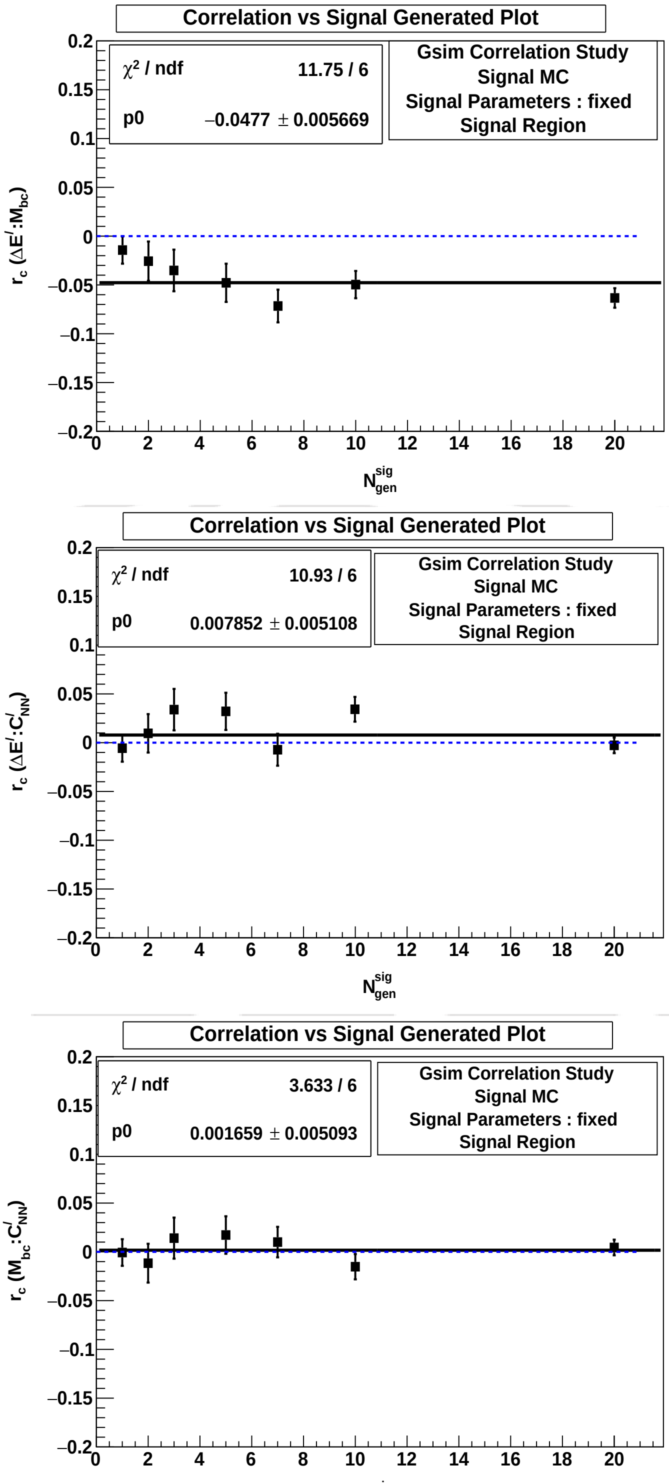

FIGURE 4.11: $N_{\rm gen}^{\rm sig}$ versus correlation coefficient plots between the fit variables, $\Delta E'$ vs. $M_{\rm bc}$ (top), $\Delta E'$ vs. $\mathcal{C}'_{\rm NN}$ (middle) and $M_{\rm bc}$ vs. $\mathcal{C}'_{\rm NN}$ (bottom) in signal MC from Gsim studies. The data points with error bars, dashed-blue, and solid-black line represent signal MC, no correlation condition (correlation coefficient $= 0$), and zeroth-order polynomial fit, respectively.

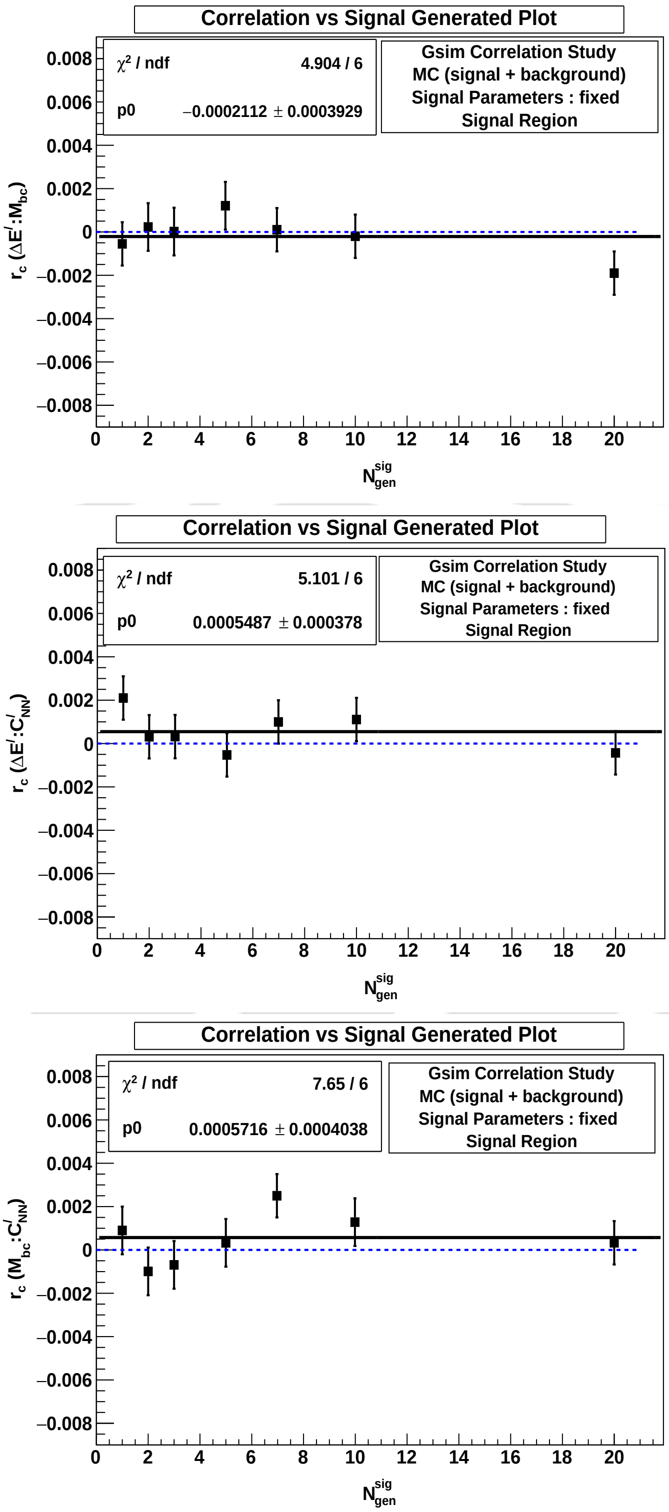

FIGURE 4.12: Correlation vs. $N_{\rm gen}^{\rm sig}$ plots between the fit variables, $\Delta E'$ vs. $M_{\rm bc}$ (top), $\Delta E'$ vs. $\mathcal{C}'_{\rm NN}$ (middle) and $M_{\rm bc}$ vs. $\mathcal{C}'_{\rm NN}$ (bottom) in (signal + background) MC from Gsim studides. The data points with error bars, dashed-blue, and solid-black line represent (signal + background) MC, no correlation condition (correlation coefficient = 0), and zeroth-order polynomial fit, respectively.

4.7.4 Toy MC study

Toy MC is another ensemble study class where the correlations among the fit variables are not considered. Therefore, any bias from such a study appears to solely originate from the PDF parametrization of the fit variables. To perform the toy MC study, we adopt the procedure enumerated below:

1. We extract signal and background MC events from the signal and background PDFs, respectively, by incorporating Poissonian fluctuations.
2. At the end of this procedure, we have MC events resembling data barring the correlations among the fit variables.
3. In this way, we have produced 1000 pseudo-experiments for each of the expected signal yields in data, namely, 0 (background only), 1, 2, 3, 5, 7, 10, and 20.
4. For the pull distribution, we expect the mean to be zero and the standard deviation to be one, ideally. Any deviation from these expected values show bias from the fit model.
5. For the yield distribution, we expect the yield from the fit to match the input signal events within MC statistical uncertainty. Any larger deviation will have to be corrected.

Figures in 4.13 are important, as we do not expect more than one (Table 4.9 tabulates the expected signal yields) signal event from this analysis. The signal pull distribution parameters (mean, μ and standard deviation, σ) agrees with a standard normal distribution. The signal yield from the fit model ($\mathrm{N}^{\mathrm{sig}}_{\mathrm{yield}} = -0.071 \pm 0.029$) agrees with the input signal events ($\mathrm{N}^{\mathrm{sig}}_{\mathrm{gen}} = 0$) within the MC statistical uncertainty. We also do not find any significant difference between the background yield from the fit (862.15 ± 0.98) and the input background of 862.

The pull and signal yield distributions for $\mathrm{N}^{\mathrm{sig}}_{\mathrm{gen}} = 1, 2, 3, 5, 7, 10, \text{and } 20$ are shown in appendix E.0.2. Figures 4.14 and 4.15 show the average pull distribution and linearity of the fit model for $\mathrm{N}^{\mathrm{sig}}_{\mathrm{gen}} = 0, 1, 2, 3, 5, 7, 10, \text{and } 20$.

We, therefore, conclude the following from Toy MC study:

1. From signal pull , signal and background MC yield distributions for the background-only case, we find that they agree with the expected results. The expected result for

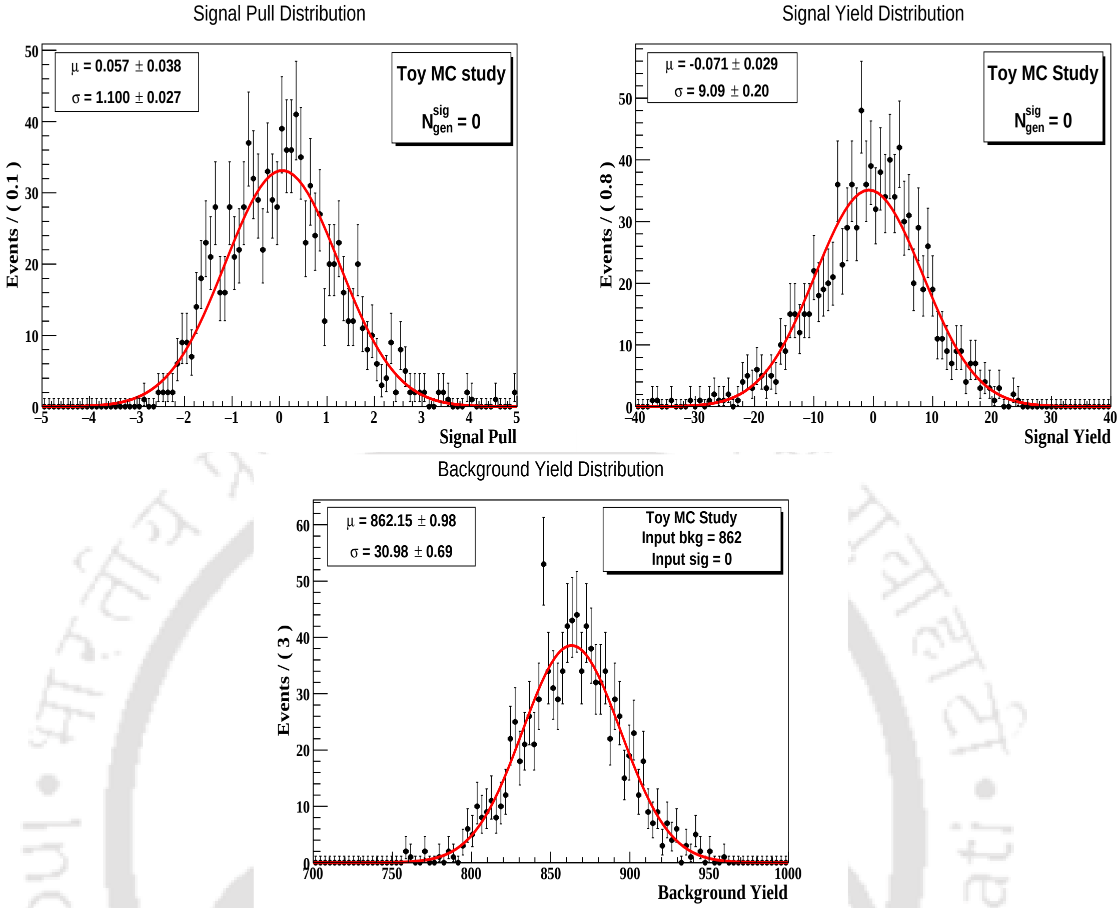

FIGURE 4.13: Signal pull (top-left), signal yield (top-right) and background yield (bottom) distribution for $N^{sig}_{gen} = 0$ (i.e., background only) from Toy MC study.

the background-only case (no signal) is that the fit function should return a signal yield close to zero within MC statistical uncertainty.

2. From Figure 4.14, we see that the pull distribution from Toy MC is consistent with the Gsim studies.

3. We also observe from Figure 4.15 that the fit is linear.

Even though we calculate the fit bias and linearity from the toy MC study, we only assign systematic uncertainties associated with the fit bias from the Gsim study. This is because the Gsim ensemble study closely resembles data compared to the toy MC study owing to considering correlations among the fit variables, if any.

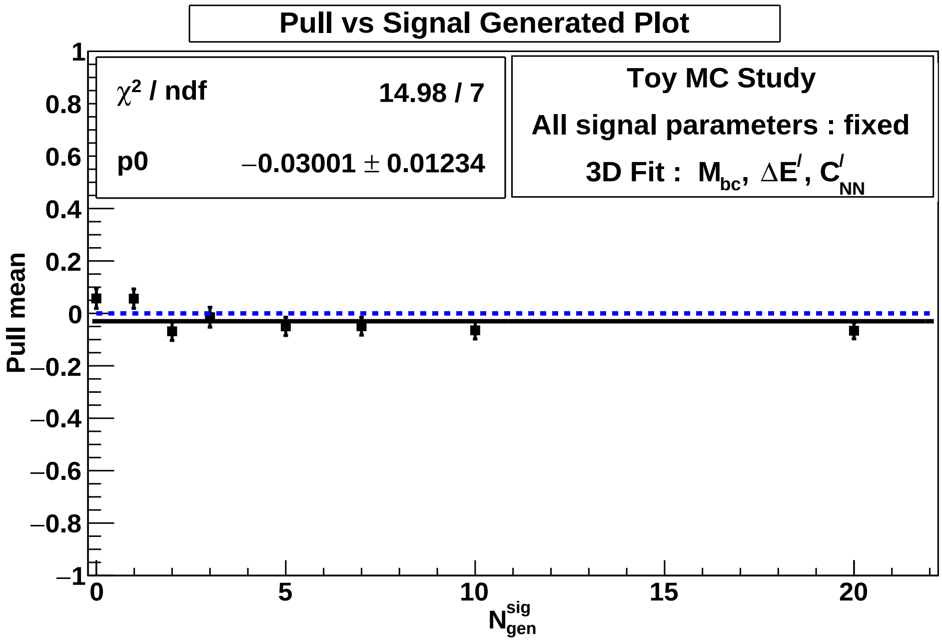

FIGURE 4.14: Average pull distribution from toy MC study for $N^{sig}_{gen} = 0, 1, 2, 3, 5, 7, 10, 20$. The data points with error bars, dashed-blue, and solid-black line represent signal MC, the no-bias condition, and zeroth order polynomial fit, respectively.

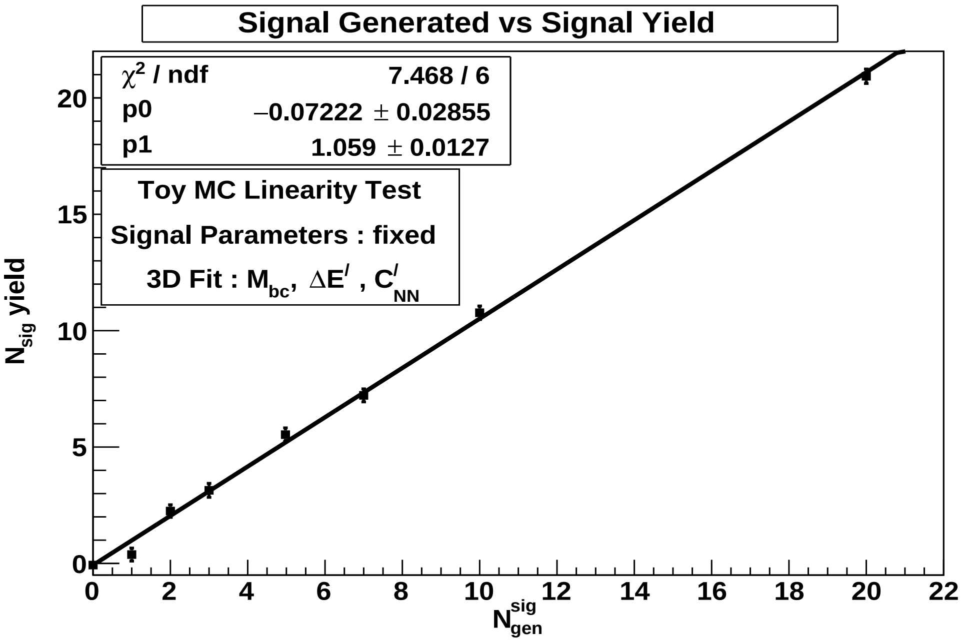

FIGURE 4.15: Linearity test from toy MC study for $N^{sig}_{gen} = 0, 1, 2, 3, 5, 7, 10, 20$. The signal MC points and the first order polynomial fit are shown as black dots with error bars and solid-black line, respectively.

Summary

In this chapter we discussed the PDF parametrizations of the fit variables: $M_{\rm bc}$, $\Delta E'$, $\mathcal{C}'_{\rm NN}$. We found the correlations among the fit variables negligible and opted for factorizing the 3D model. We described the unbinned extended maximum likelihood fit procedure, which will be used to extract signals in data. We summarized the results from signal and background MC studies. We also highlighted our study on the fit bias, linearity, and the correlations among the fit variables to assign systematic uncertainty due to the fit model. While doing so, we adopted two procedures: Gsim and toy MC. We prefer to keep the former for assigning any systematic uncertainty on the fit model due to its better resemblance with data. From the Gsim study, we obtained -3.3% of fit bias and found the fit model to be linear at the expected sensitivity of the SM. Further conducting the correlation studies does not reveal any significant correlation among the fit variables. In the next chapter, we will perform validation studies to cross-check the analysis procedure.

Chapter 5

Validation study

This chapter discusses in detail the methods adopted to validate the analysis procedure. We perform the validation through two methodical studies: Control sample and data side-bands. The control sample study substantiates the various selection criteria applied in this analysis. In contrast, the data side-band study looks for data-MC differences in the side-bands of the fit variables: $M_{\rm bc}$, $\Delta E'$, $\mathcal{C}'_{\rm NN}$.

We will refer to the analysis, $B_s^0 \to \pi^0\pi^0$ throughout this chapter as the "target" analysis to differentiate it from the "control" channel.

5.1 Control sample study

To validate the MC studies, we perform a control sample study on a decay channel with the closest decay topology to the target sample with similar final state particles. For this purpose, we use $B^0 \to \pi^0\pi^0$ at $\Upsilon(4S)$ resonance as our control sample. We refer to Table 5.1 for the current status of $B^0 \to \pi^0\pi^0$ decay, performed previously by Belle [56] and information on the MC studies. The description of the selection criteria applied to the variables is similar to the target analysis discussed in section 3.7.

The data set used for the control sample study is enumerated below:

1. Two silicon vertex detectors (SVD) were used for data accumulation at Υ(4S) resonance at Belle for different periods. Moreover, two different data/MC reprocessing (case A and case B) are also available. We use the recommended case B (SVD2) reprocessing data and MC for the control sample. The case B (SVD2) set corresponds

to an integrated luminosity of 693 fb^{-1} of data collected at the $\Upsilon(4\text{S})$ resonance amounting to 619×10^6 $B\bar{B}$.

2. We have generated 100,000 signal MC events for the highest luminosity experiment 55 for the control mode signal MC study at $\Upsilon(4\text{S})$ resonance. These events are scaled to their corresponding luminosity.

3. We use six streams of continuum background, ten streams of generic $B\bar{B}$ background ($e^+e^- \to \Upsilon(4S) \to B^0\bar{B}^0/B^+B^-$) and fifty streams of rare B background ($e^+e^- \to \Upsilon(4S) \to B^0\bar{B}^0/B^+B^-$, where the B's make $b \to u$ transition) for some of the highest luminosity experiments at $\Upsilon(4\text{S})$, such as experiments 37, 41, 43, and 55.

Table 5.1: Table summarizing some of the important information on the control sample, $B^0 \to \pi^0\pi^0$, from the previous Belle analysis.

Data type	Properties	Remarks
Data	$B^0 \to \pi^0\pi^0$	Control sample
	Data set at $\Upsilon(4\text{S})$	693 fb^{-1}
	Measured $\mathcal{B}$	$(1.31 \pm 0.19(\text{stat.}) \pm 0.19(\text{syst.})) \times 10^{-6}$ (Belle [56])
		$(1.59 \pm 0.26) \times 10^{-6}$ (PDG [1])
MC	M_{bc}	> 5.26 GeV/c^2
	ΔE	$-0.3 < \Delta E < 0.2$ GeV
	E_γ	$> 0.05(0.10)$ GeV for barrel (Fwd/Bwd) ECL
	$M_{\gamma\gamma}$	$0.115 < M_{\gamma\gamma} < 0.152$ GeV/c^2
	T_c (Neural network output)	> -0.3
	Reconstruction efficiency	$\sim 22\%$

5.1.1 Signal and background MC study

Signal and background studies are performed using MC samples shown in Table 5.2. The experiments used in the MC studies are some of the highest luminosity experiments at $\Upsilon(4\text{S})$ resonance. Table 5.3 shows the luminosities of experiments at $\Upsilon(4\text{S})$ resonance. We adopt the procedures discussed in section 3.3 for generating and simulating the MC samples.

The control channel's signal reconstruction procedure is similar to our target sample's reconstruction discussed in section 3.7. To have a one-to-one correspondence between the target and the control sample, we apply similar selection criteria to the control sample

TABLE 5.2: MC samples used for signal and background studies. The term "generic" and "rare" refers to $B\bar{B}$ and $b \to u$ type of backgrounds, respectively. The term "mixed" refers to both B^+B^- and $B^0\bar{B}^0$ type of backgrounds.

MC sample	Experiment	Remarks
Signal	55	$100,000$ (scaled to experiment 55 luminosity)
$e^+e^- \to q\bar{q}, (q \to u,\ d,\ s\ c)$	55	$6\times$ real data
$\Upsilon(4S) \to B^+B^-$ (charged generic)	37, 41, 43, 55	$10\times$ data
$\Upsilon(4S) \to B^0\bar{B}^0$ (rare generic)	37, 41, 43, 55	$10\times$ data
$b \to u$ (charged rare)	37, 41, 43, 55	$50\times$ data
$b \to u$ (mixed rare)	37, 41, 43, 55	$50\times$ data

TABLE 5.3: Some of the highest luminosity experiments at Υ(4S) resonance.

Experiment	Luminosity (fb^{-1})
37	62
41	60
43	57
55	80
	$\sum = 259$

(selection criteria for target sample is tabulated in Table 3.8). We summarize the selection criteria for the control channel in Table 5.4.

At Υ(4S) resonance, ΔE and $\Delta E'$ are to a great extent identical as shown in Figure 5.3. However, to maintain consistency with the target sample, we use $\Delta E'$, where $\Delta E' = \Delta E + M_{\rm bc}c^2 - m_{B^0}c^2$, as one of our fit variables. The mass of the B^0 meson is referred from PDG [1]. We have used sixteen KSFW moments and the $\cos\theta_T$ as inputs to the neural network. Figures 5.1 and 5.2 show the purity versus Neural Netowork output and training and validation plots, respectively.

After applying all the selection criteria tabulated in Table 5.4, we note the following points:

1. We find 9% of multiple candidates after applying all the selection criteria. To select the best B^0 candidate, we use the BCS criterion similar to the target sample. We

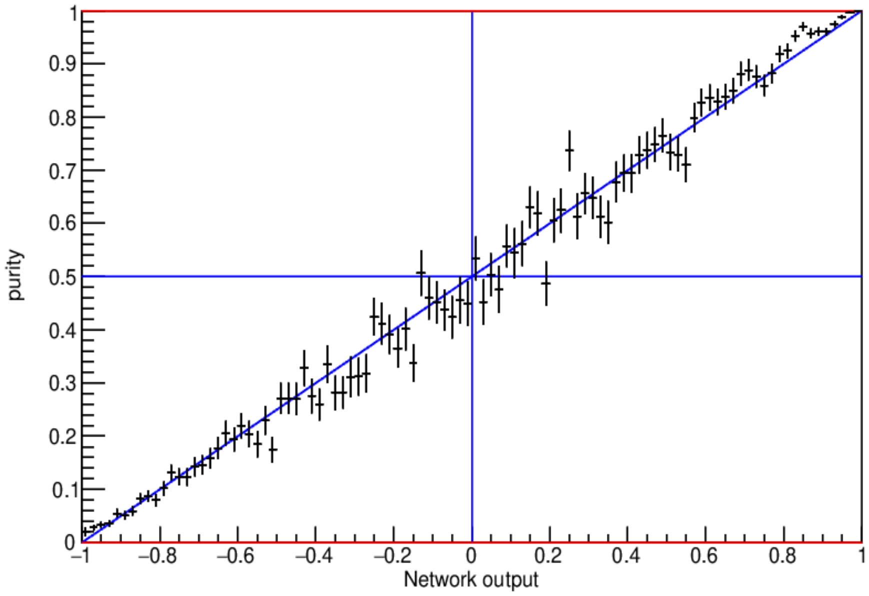

FIGURE 5.1: Distribution of Neural Network output versus purity for the control channel, $B^0 \to \pi^0\pi^0$. The MC data points with error bars represent the training sample and the solid blue line (inclined at 45°) indicates the expected trajectory for an ideal training.

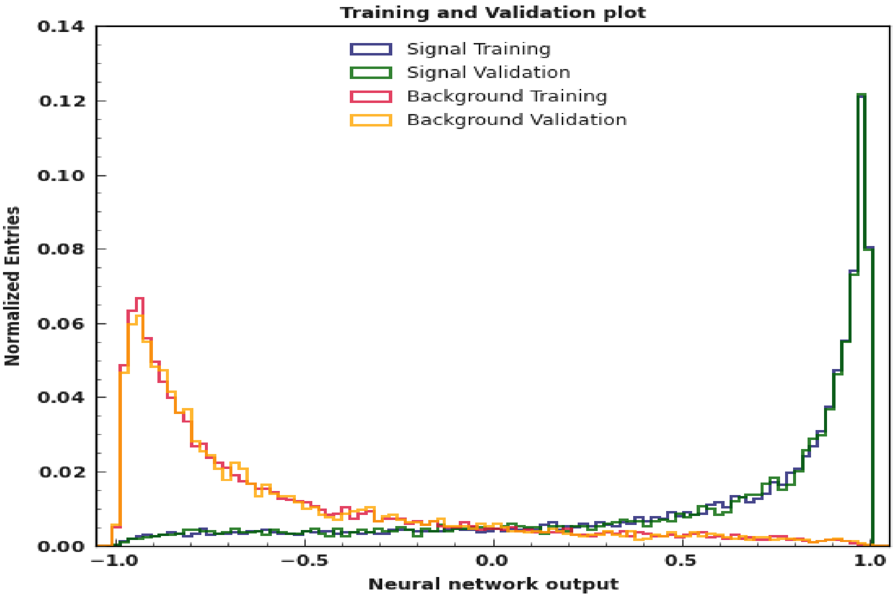

FIGURE 5.2: Signal and background MC training and validation plot for the control channel, $B^0 \to \pi^0\pi^0$. The signal training and validation distributions are represented by blue and green colors, while, background training and validation distributions are shown with red and yellow colors.

TABLE 5.4: List of selection criteria applied to the control sample.

Variables	Selection Criteria
Photon ID	$\neq 911$ (MC only)
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2
E_γ	0.050 (0.100) GeV barrel (end-caps) ECL
$\Delta E'$	$-0.6 < \Delta E' < 0.20$ GeV
$M_{\rm bc}$	$M_{\rm bc} > 5.22$ GeV/c^2
$\mathcal{C}'_{\rm NN}$	$\mathcal{C}'_{\rm NN}$ transformed after applying $\mathcal{C}_{\rm NN} > 0.90$

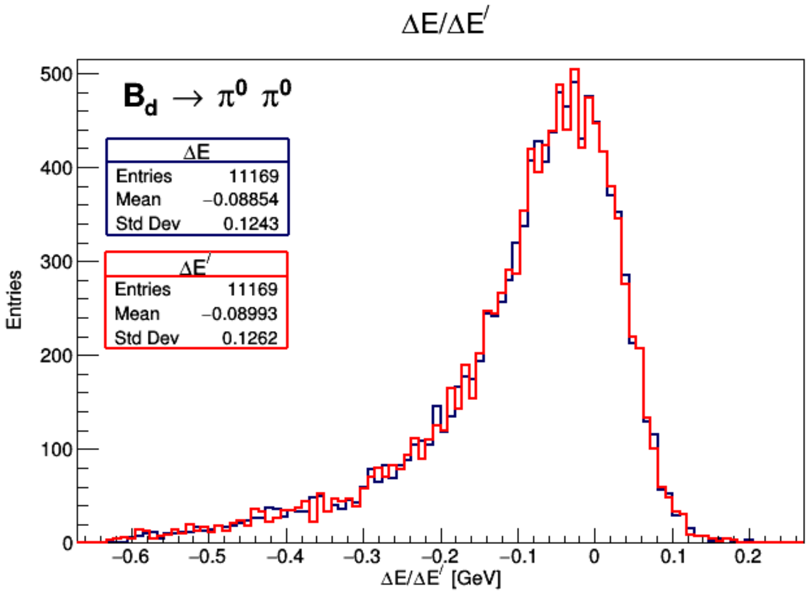

FIGURE 5.3: ΔE and $\Delta E'$ distributions of signal MC. The blue and red distributions represent ΔE and $\Delta E'$, respectively.

re-write it here for convenience.

$$\chi^2_{min}(B^0) = \chi^2_{min}(\pi^0_1) + \chi^2_{min}(\pi^0_2), \tag{5.1}$$

where, $\chi^2(\pi^0)$ is obtained from the mass-constrained fit on the π^0s (as discussed in Equation 3.20), and, therefore, $\chi^2_{min}(\pi^0)$ represents the minimum χ^2 value. The BCS criterion is 57% efficient in selecting the correctly reconstructed (MC truth) B^0 candidates.

2. The continuum is the dominant background, and we expect to observe 2798 from case B (SVD2) data.

TABLE 5.5: Table showing candidate and signal regions in the control sample.

Region	$M_{\rm bc}$ (GeV/c^2)	$\Delta E'$ (GeV)	$\mathcal{C}'_{\rm NN}$
Candidate	$[5.22, 5.30]$	$[-0.6, 0.20]$	$[-6.0, 6.5]$
Signal	$[5.27, 5.29]$	$[-0.23, 0.15]$	$[-3.1, 4.61]$

3. We do not find any significant contamination from generic $B\bar{B}$ background. Table 5.6 shows the surviving generic $B\bar{B}$ backgrounds in the candidate and the signal region. The candidate and the signal regions are defined in Table 5.5. On projecting the generic $B\bar{B}$ backgrounds, for case B (SVD2) data amounting to 619×10^6 $B\bar{B}$ events, we expect eighty-six of them in the candidate region and not more than ten in the signal region. We, therefore, allow the dominant continuum background PDFs to account for them.

4. We also do not find any significant contamination from rare $B\bar{B}$ ($b \to u$ type) background from MC studies. Table 5.7 shows the surviving rare $B\bar{B}$ background in the candidate and signal regions. On projecting these backgrounds, for 619×10^6 $B\bar{B}$ events, we expect thirty-six of them in the candidate region and not more than five in the signal region. They mostly come from $B^\pm \to \rho^\pm(\pi^\pm\pi^0)\pi^0$ (92%) and $B^0 \to K_s^0(\pi^0\pi^0)\pi^0$ (5%) when one of the charged/neutral pion is lost in an event. Taking into account the rarity of these events, they are separately parametrized..

5. The misreconstruction (candidates whose generator-level information are not found) in the candidate and the signal region is found to be 1% and 0.8% from signal MC study, respectively. Since the fraction of misreconstructed events are not significant, we do not parametrize them separately.

6. From MC study, we find the signal reconstruction efficiency to be 11.16%.

5.1.2 1D PDF parametrization

Prior to performing 1D PDF parametrization, we calculate the correlation coefficients among the fit variables, $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$. Table 5.8 shows the linear correlation coefficients among the fit variables for signal and background (continuum + generic $B\bar{B}$ + rare $B\bar{B}$) MC in the signal region. The correlation plots are shown in Figures 5.4 and 5.5.

As the fit variables are not significantly correlated, we proceed to 1D PDF parametrization for the control sample. Table 5.9 show the PDF models used to parametrize signal and

Table 5.6: Table to summarize the surviving generic $B\bar{B}$ backgrounds after applying all the selection criteria. The numbers within parenthesis in column 'Streams' indicate a single stream. The numbers within parenthesis in column 'Events' indicate the expected number of backgrounds at the corresponding luminosity.

Expt. no.	Luminosity (fb^{-1})	Background	Streams	Events	Events in signal region
37	62	charged	10 (1)	32 (3)	0
		mixed		73 (7)	0
41	60	charged	10 (1)	18 (2)	0
		mixed		44 (4)	0
43	57	charged	10 (1)	27 (3)	0
		mixed		62 (6)	0
55	80	charged	10 (1)	31 (3)	0
		mixed		84 (8)	2

Table 5.7: Table to summarize the surviving rare $B\bar{B}$ backgrounds after applying all the selection criteria. The numbers within parenthesis in column 'Streams' indicate a single stream. The numbers within parenthesis in column 'Events' indicate the expected number of backgrounds at the corresponding luminosity.

Expt. no.	Luminosity (fb^{-1})	rare B background	Streams	Events	Events in signal region
37	62	charged	50 (1)	34 (1)	0
		mixed		175 (4)	1
41	60	charged	50 (1)	84 (2)	0
		mixed		120 (2)	0
43	57	charged	50 (1)	62 (1)	0
		mixed		143 (3)	0
55	80	charged	50 (1)	78(2)	0
		mixed		132 (3)	1

Table 5.8: Linear correlation coefficients among the fit variables obtained from MC studies in the signal region.

Variables	Signal correlation coefficients	Background correlation coefficients
$M_{\text{bc}} : \Delta E'$	-0.0582	$+0.0009$
$\mathcal{C}'_{\text{NN}} : \Delta E'$	$+0.0074$	$+0.0316$
$\mathcal{C}'_{\text{NN}} : M_{\text{bc}}$	-0.0053	-0.0271

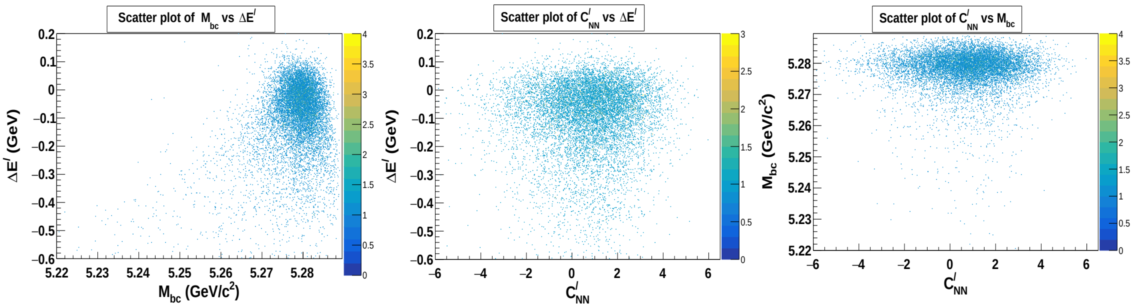

FIGURE 5.4: Scatter plots among the fit variables, M_{bc} vs. $\Delta E'$ (left), $\mathcal{C}'_{\text{NN}}$ vs. $\Delta E'$ (middle), and $\mathcal{C}'_{\text{NN}}$ vs. M_{bc} (right) for signal MC.

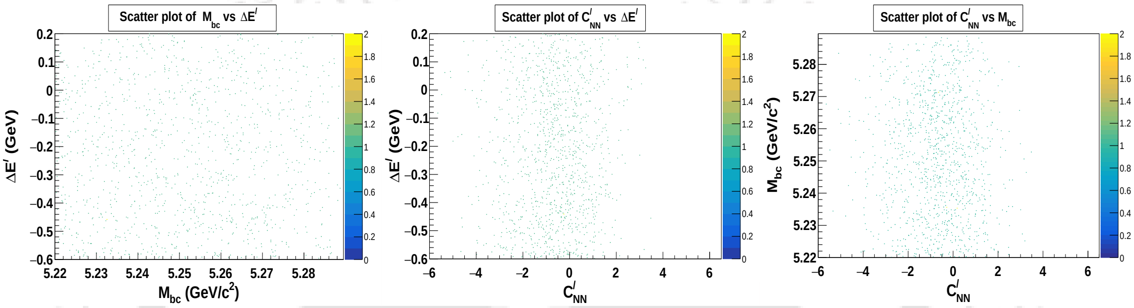

FIGURE 5.5: Scatter plots among the fit variables, M_{bc} vs. $\Delta E'$ (left), $\mathcal{C}'_{\text{NN}}$ vs. $\Delta E'$ (middle), and $\mathcal{C}'_{\text{NN}}$ vs. M_{bc} (right) for background MC.

TABLE 5.9: List of PDFs used for parametrizing the signal and background MC. Here, G, CB, ACB, CP and AG signify Gaussian, Crystal Ball, Asymmetric Crystal Ball, Chebychev Polynomial, and Bifurcated Gaussian, respectively.

Variables	Signal MC PDFs	Continuum MC PDFs	Rare $B\bar{B}$ PDFs
M_{bc}	G + CB	Argus function	CB
$\Delta E'$	ACB	Second order CP	CB
$\mathcal{C}'_{NN}$	G + BG	G + G	G + G

background MC. The PDF models used for parametrizing the control sample are almost similar to those used for the target sample.

Figure 5.6 shows the 1D PDF parametrizations of the fit variables, M_{bc}, $\Delta E'$, and $\mathcal{C}'_{\text{NN}}$, for signal MC. Figures 5.7, 5.8, and 5.9 show 1D PDF parametrizations of the fit variables, M_{bc}, $\Delta E'$, and $\mathcal{C}'_{\text{NN}}$, for background MC respectively.

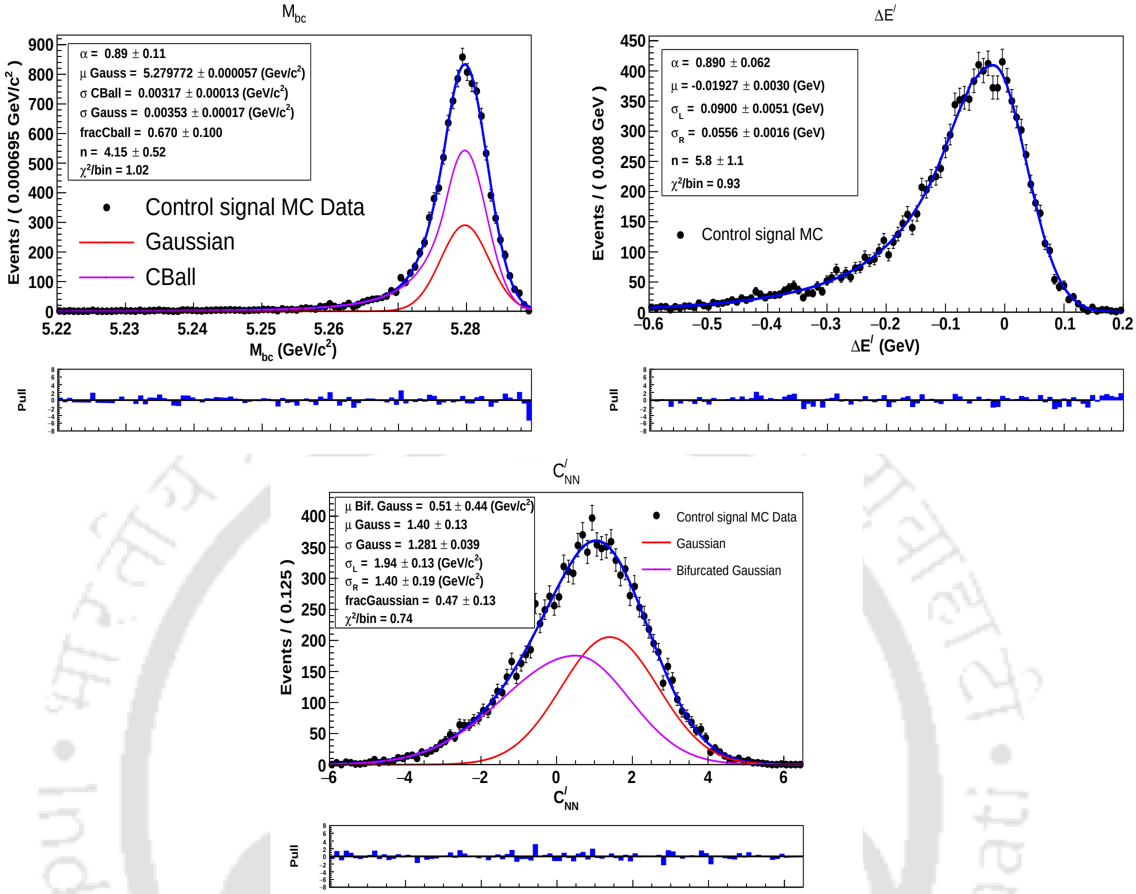

FIGURE 5.6: 1D PDF parametrizations of $M_{\rm bc}$ (top-left) and $\Delta E'$ (top-right), and $\mathcal{C}'_{\rm NN}$ (bottom) distributions for signal MC.

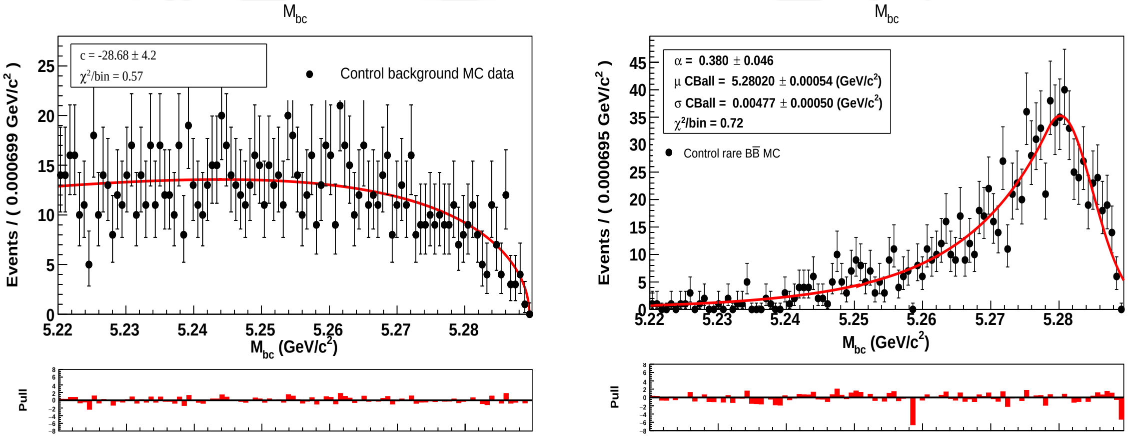

FIGURE 5.7: 1D PDF parametrizations of $M_{\rm bc}$ distributions for continuum background (left) and rare $B\bar{B}$ background (right) MC.

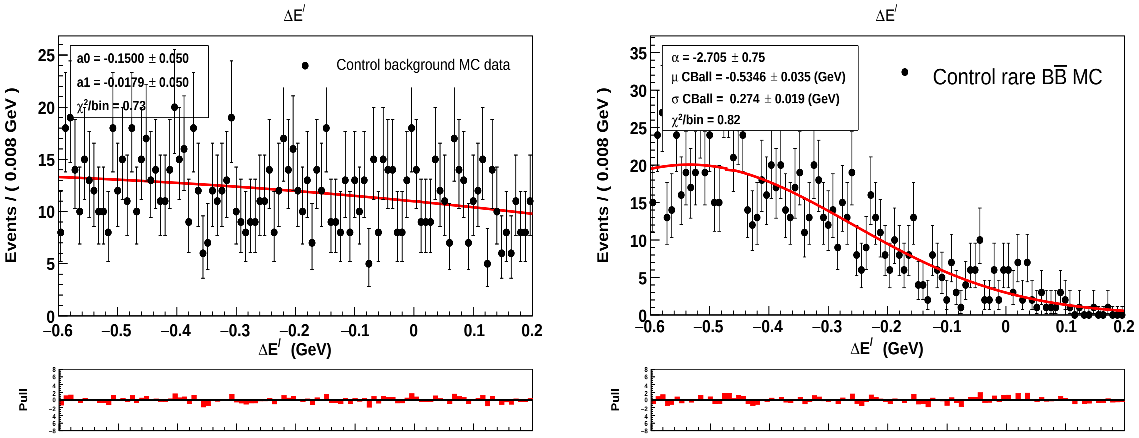

FIGURE 5.8: 1D PDF parametrizations of $\Delta E'$ distributions for continuum background (left) and rare $B\bar{B}$ background (right) MC.

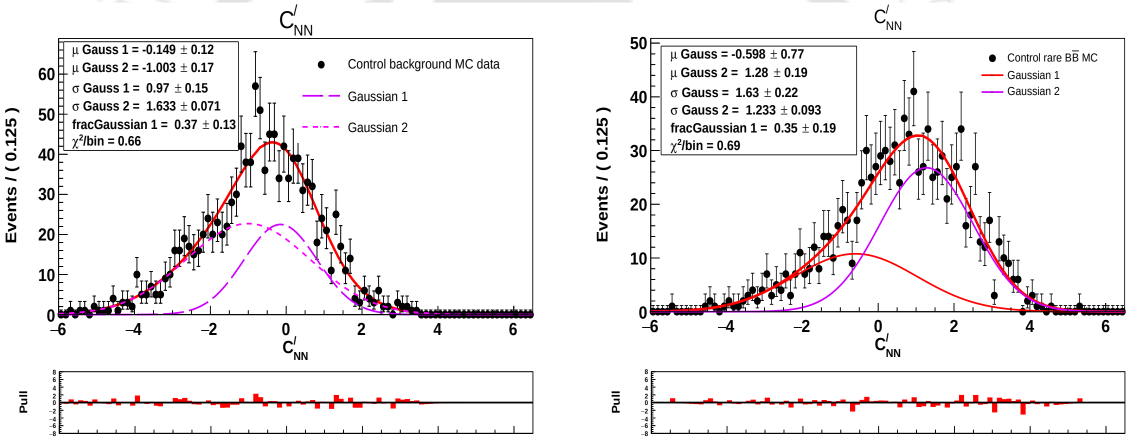

FIGURE 5.9: 1D PDF parametrizations of $\mathcal{C}'_{\rm NN}$ distributions for continuum background (left) and rare $B\bar{B}$ background (right) MC.

5.1.3 Data fit results for the control sample, $B^0 \to \pi^0\pi^0$ at $\Upsilon(4S)$ resonance

We perform 3D unbinned extended maximum likelihood fit (discussed in sub-section 4.3) to extract signals from data, just like we use it in our target sample. We perform the signal extraction procedure for two cases:

1. Case I: Comparison by applying selection criteria similar to the previous Belle analysis, which is another way of verifying our analysis procedure.
2. Case II: Apply our selection criteria, calculate the branching fraction and compare with the previous Belle analysis [56] and the PDG average [1].

Table 5.10: List of selection criteria applied to the control sample in data for case I. Here, we use ΔE instead of $\Delta E'$ so as to maintain conformity with the previous Belle analysis.

Variables	Selection Criteria
ECL timing	(9000 - 11000) ns
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2
E_γ	0.050 (0.100) GeV barrel (Fwd/Bwd) ECL
$\mathbf{\Delta E}$	$\mathbf{-0.3 < \Delta E < 0.20}$ **GeV**
$\mathbf{M_{bc}}$	$\mathbf{M_{bc} > 5.26}$ **GeV**/c^2
$\mathcal{C}'_{\mathrm{NN}}$	$\mathcal{C}'_{\mathrm{NN}} > \mathbf{-0.3}$

Table 5.11: Comparison of reconstruction efficiency from this study and previous Belle analysis MC studies.

Variable	This MC study	Previous Belle analysis MC study
Reconstruction efficiency, ϵ	21.10%	$\sim$ 22%

5.1.3.1 Case I: Comparison by applying selection criteria similar to the previous Belle analysis

Table 5.10, summarizes the selection criteria applied in case of the control sample in data for case I.

The bold-faced figures in Table 5.10 are the selection criteria applied in the previous Belle analysis. However, they did not consider the modified NN output, $\mathcal{C}'_{\mathrm{NN}}$; instead used the NN output obtained directly from the classifier. We, on the other hand, have used $\mathcal{C}'_{\mathrm{NN}}$ with the selection criterion greater than -0.3 applied by the previous Belle analysis. Table 5.11 shows the comparison of signal reconstruction efficiency from this study and previous Belle MC studies.

Figure 5.10 show the data signal enhanced projection plots. We have used case B (SVD2) data for producing these plots.

From the data fit, we obtain a signal yield of 164 ± 27 with a reconstruction efficiency of 21.10%. We calculate the branching fraction using the formula

$$\mathcal{B} = \frac{N_{\mathrm{fit}}^{\mathrm{sig}}}{N_{B\bar{B}} \times \epsilon \times \prod_i \mathcal{B}_i},$$

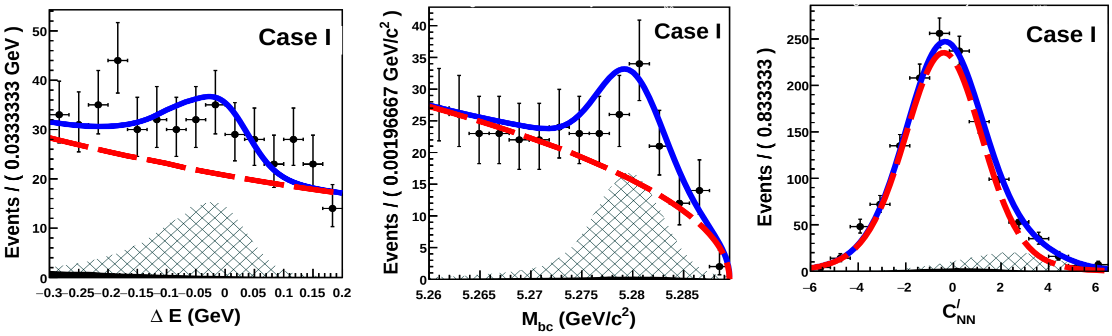

FIGURE 5.10: Signal enhanced projection plots of ΔE (left), $M_{\rm bc}$ (middle), and $C'_{\rm NN}$ (right) for case I. Each of the plots are generated by applying the signal region selection criteria for the other variables. The signal regions for the three variables are as follows, $-0.15 < \Delta E < 0.05$ (GeV), $5.27 < M_{\rm bc} < 5.285$ (GeV/c^2), and $2.2 < C'_{\rm NN} < 6.5$. The solid (blue), dashed (red), hatched, and filled (black) distributions represent the total fit model, continuum background, signal, and rare $B\bar{B}$ backgrounds, respectively.

TABLE 5.12: Comparison of branching fractions with the previous Belle result and PDG value.

Analysis	Branching Fractions
Our result	$(1.28 \pm 0.21(\text{stat.})) \times 10^{-6}$
Previous Belle result	$(1.31 \pm 0.19(\text{stat.})) \times 10^{-6}$
PDG value	$(1.59 \pm 0.26) \times 10^{-6}$

where, $N_{fit}^{sig} = 164 \pm 27$, $N_{B\bar{B}} = 619 \times 10^6$ (SVD2 data), $\epsilon = 21.10\%$, and $\mathcal{B}_i = (98.823 \pm 0.034)\%$, is the branching fraction of, $\pi^0 \to \gamma\gamma$ [1].

Table 5.12 shows the branching fractions from the comparison study (Case I) with the previous Belle analysis and PDG value. It is observed that our result on the branching fraction agrees within 1 standard deviation (σ) of the previous Belle result [56] and 1.6σ of the PDG value [1].

5.1.3.2 Case II : Comparison with previous Belle analysis by applying selection criteria similar to the target analysis

Table 5.13 shows the selection criteria applied in this case. After applying all the selection criteria, we have a signal reconstruction efficiency of 11.16%. We notice a drop in the signal reconstruction efficiency from MC studies due to a tighter selection criterion of greater than 0.90 on $\mathcal{C}'_{\rm NN}$.

Table 5.13: List of selection criteria applied to the control sample in data.

Variables	Selection Criteria
ECL timing	(9000 - 11000) ns
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ (GeV/c^2)
E_γ	0.050 (0.100) GeV barrel (Fwd/Bwd) ECL
$\Delta E'$	$-0.6 < \Delta E' < 0.20$ (GeV)
$M_{\rm bc}$	$M_{\rm bc} > 5.22$ GeV/c^2
$\mathcal{C}'_{\rm NN}$	$\mathcal{C}'_{\rm NN} > 0.90$

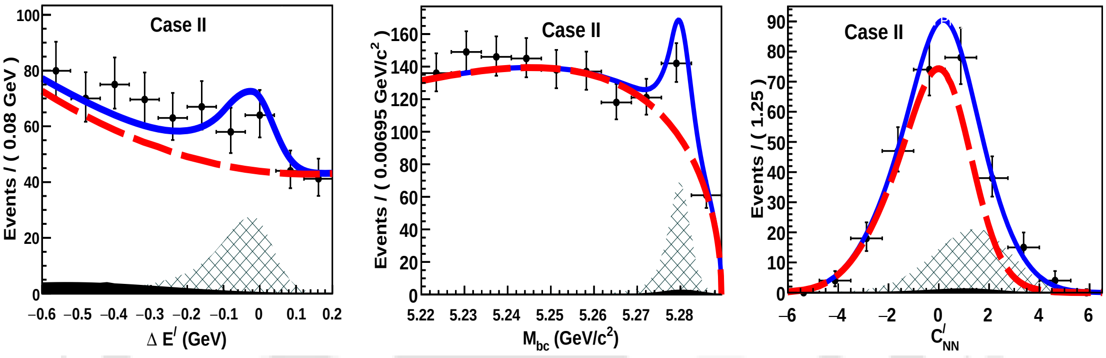

Figure 5.11: Signal enhanced projection plots of $\Delta E'$ (left), M_{bc} (middle), and $\mathcal{C}'_{\rm NN}$ (right) for case II. Each of the plots are generated by applying the signal region selection criteria for the other variables. The signal regions for the three variables are as follows, $-0.23 < \Delta E' < 0.16$ (GeV), $5.27 < M_{\rm bc} < 5.29$ (GeV/c^2), and $-6.0 < \mathcal{C}'_{\rm NN} < 6.5$. The solid (blue), dashed (red), hatched, and filled (black) distributions represent the total fit model, continuum background, signal, and rare $B\bar{B}$ backgrounds, respectively.

Figure 5.11 shows the signal enhanced projections of the fit variables ($\Delta E'$, $M_{\rm bc}$, $\mathcal{C}'_{\rm NN}$), respectively.

After a 3D unbinned extended maximum likelihood fit to data, we obtain a signal yield of 80 ± 14 with a signal reconstruction efficiency of 11.16%. Table 5.14 shows the branching fraction for case II in comparison with the previous Belle analysis and PDG value.

From Table 5.14, we see that the branching fraction agrees within 1σ of the previous Belle result [56] and 1.6σ of the PDG value [1].

From the control sample study, we conclude that the analysis procedure adopted for the target decay channel, $B_s^0 \to \pi^0\pi^0$ is well placed and understood.

Table 5.14: Comparison of branching fractions (case II) with the previous Belle result and PDG value.

Analysis	Branching Fractions
Our result	$(1.18 \pm 0.21(\text{stat.})) \times 10^{-6}$
Previous Belle result	$(1.31 \pm 0.19(\text{stat.})) \times 10^{-6}$
PDG value	$(1.59 \pm 0.26) \times 10^{-6}$

Table 5.15: Data side-band region.

Fit variable	Side-band region
$\Delta E'$	$-0.6 < \Delta E' < -0.37$ GeV
M_{bc}	$5.30 < M_{\text{bc}} < 5.33$ GeV/c^2

Table 5.16: Data-MC correlation coefficients in side-band. Here we compare data side-band with background MC as the analysis is continuum background dominated.

Variables	Data	Background MC
$\Delta E' : M_{\text{bc}}$	+0.016	−0.064
$\Delta E' : C'_{\text{NN}}$	+0.069	+0.01
$M_{\text{bc}} : C'_{\text{NN}}$	+0.015	−0.003

5.2 Data-MC difference in side-bands for $B_s^0 \to \pi^0\pi^0$ in data at $\Upsilon(5S)$ resonance

In this section, we look for data-MC differences in the fit variables ($\Delta E'$, M_{bc}, $\mathcal{C}'_{\text{NN}}$) and $\chi^2_{\pi^0}$ (defined in Equation 3.1). As continuum background dominates our target channel, a data-MC side-band compatibility check is necessary to validate the background PDFs developed using MC studies. We have performed a 3D unbinned extended maximum likelihood fit on data side-band region for experiment 69 (highest luminosity experiment at $\Upsilon(5S)$ with $\mathcal{L} \sim 47\ \text{fb}^{-1}$), which corresponds to roughly 40% of the total data collected at $\Upsilon(5S)$ resonance. Table 5.15 defines the data side-band region, determined using the variables $\Delta E'$ and M_{bc}.

Figures 5.12, 5.13, 5.14, and 5.15 refer to the data-MC comparison for $\Delta E'$, M_{bc}, C'_{NN}, and $\chi^2_{\pi^0}$ variable, respectively. The data-MC differences in the correlation coefficients are tabulated in 5.16.

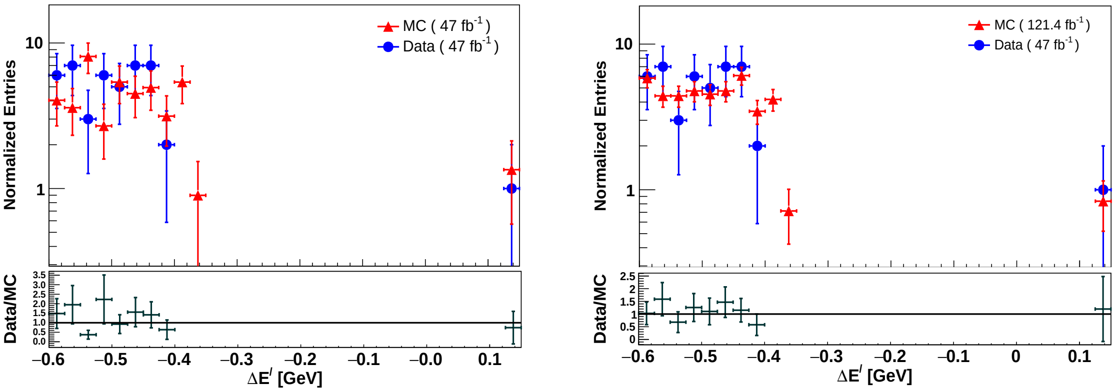

FIGURE 5.12: Data-MC difference plots of $\Delta E'$ variable when MC is considered for $\mathcal{L} = 47$ fb^{-1}(left) and $\mathcal{L} = 121.4$ fb^{-1} (right). The red (filled-triangular) and blue (filled-circular) points with error bars represent MC and data, respectively.

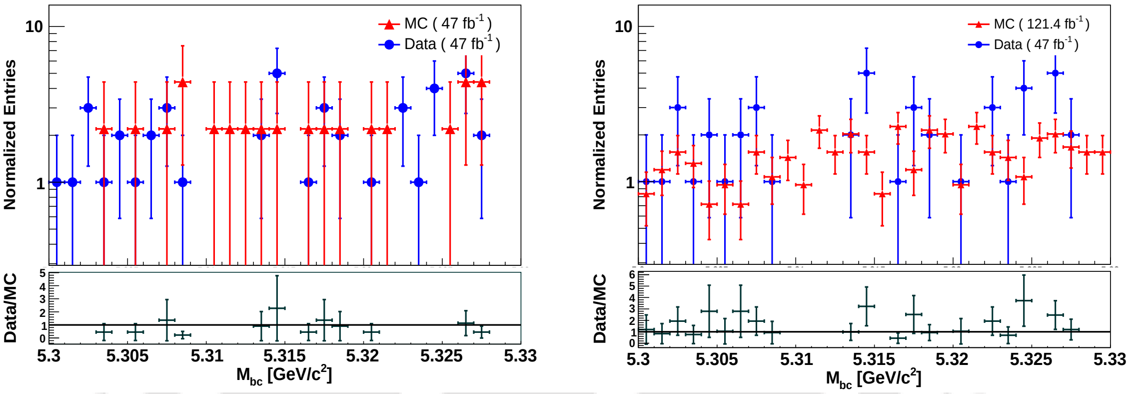

FIGURE 5.13: Data-MC difference plots of M_{bc} variable when MC is considered for $\mathcal{L} = 47$ fb^{-1}(left) and $\mathcal{L} = 121.4$ fb^{-1} (right). The red (filled-triangular) and blue (filled-circular) points with error bars represent MC and data, respectively.

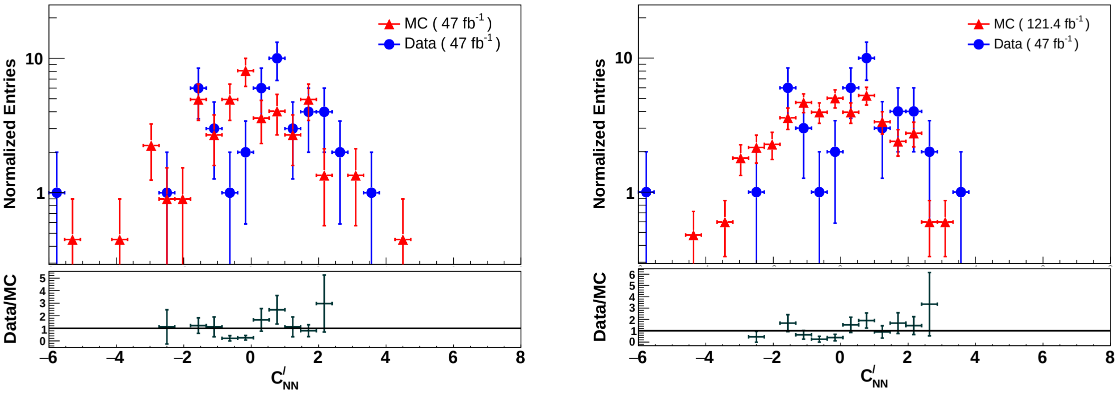

FIGURE 5.14: Data-MC difference plots of C'_{NN} variable when MC is considered for $\mathcal{L} = 47$ fb^{-1}(left) and $\mathcal{L} = 121.4$ fb^{-1} (right). The red (filled-triangular) and blue (filled-circular) points with error bars represent MC and data, respectively.

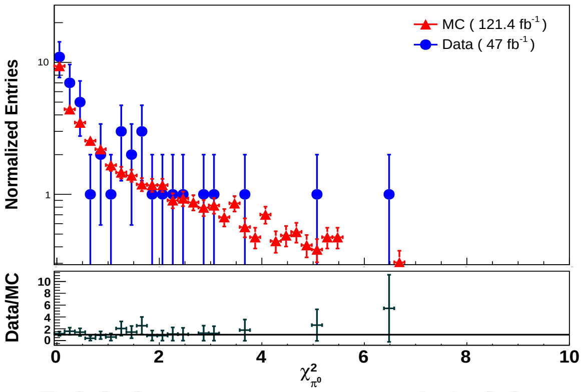

FIGURE 5.15: Data-MC difference plot of $\chi^2_{\pi^0}$ variable when MC is considered for $\mathcal{L} =$ 121.4 fb^{-1}. The red (filled-triangular) and blue (filled-circular) points with error bars represent MC and data, respectively.

From data-MC comparison distributions, we do not observe any serious differences. However, we note that the data distributions are statistically limited because of considering only a portion of the whole data sample.

Summary

In this chapter, we discussed two validation methods, that we adopted to cross-check our analysis procedure and PDF parametrization. The first validation method, the control sample study was performed with a decay channel, $B^0 \to \pi^0\pi^0$ at $\Upsilon(4S)$ resonance, having a similar topology and final state particles to the target sample, $B^0_s \to \pi^0\pi^0$. The measured branching fraction was found to be within 1σ of the previous Belle result and 1.6σ of the PDG value. The results from the control sample was, therefore, found to be consistent with recent measurements. The second validation method was performed using a data-MC side-band comparison with 40% of the data at $\Upsilon(5S)$ resonance for the target channel. We found the data-MC side-band comparison to be consistent.

Chapter 6

Results and outlook

The previous chapters discussed the procedures we adopted for the analysis. We started with event generation and simulation of MC samples to develop the selection criteria. We performed signal and background MC studies to identify the signal and suppress the continuum background. To extract signal yields in data, we developed a fit model based on a 3D unbinned extended maximum likelihood fit and validated the model with a similar topology control sample, $B^0 \to \pi^0\pi^0$ at $\Upsilon(4S)$ and with data-MC side-band comparison. We found our fit model to be stable and linear from the ensemble study performed over large MC samples. From MC studies, we conclude that our analysis procedures are well-placed, and we understand our fit model within the systematic uncertainties associated with the analysis (to be discussed in this chapter). We, therefore, use the developed model to search for the rare decay $B_s^0 \to \pi^0\pi^0$ with the data sets accumulated at $\Upsilon(5S)$ resonance corresponding to an integrated luminosity of 121.4 fb^{-1}.

6.1 Data results for $B_s^0 \to \pi^0\pi^0$

In this section, we report the results obtained from data, for all the experiments (43, 53, 67, 69, and 71) at $\Upsilon(5S)$ resonance which corresponds to an integrated luminosity of 121.4 fb^{-1}. We briefly recall the signal extraction procedure as follows:

- We have used a 3D unbinned extended maximum likelihood fit with the variables, $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$.

TABLE 6.1: Summary of selection criteria for $B_s^0 \to \pi^0\pi^0$ in data.

Variables	Selection criteria
Photon energy, E_γ	$> 0.05(0.10)$ GeV in barrel (Fwd/Bwd) ECL
ECL timing criterion	$9000 - 11000$ ns
$M_{\gamma\gamma}$	$0.118 < M_{\gamma\gamma} < 0.152$ GeV/c^2
$\Delta E'$	$-0.60 < \Delta E' < 0.15$ GeV
$M_{\rm bc}$	$5.300 < M_{\rm bc} < 5.434$ GeV/c^2
$\mathcal{C}'_{\rm NN}$	$\mathcal{C}'_{\rm NN} > 0.90$

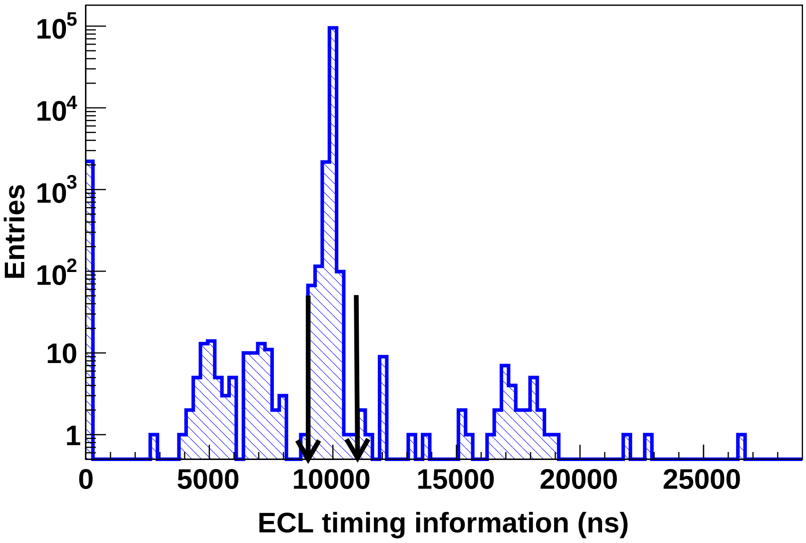

FIGURE 6.1: Distribution of ECL timing information in data. The arrows indicate the region of the selected events.

- We have fixed the signal parameters and background ARGUS end-point from MC study. In contrast, we float the signal and background yields along with the other background PDF parameters.

- Table 6.1 summarizes the selection criteria for this analysis in data.

Figure 6.1 shows the ECL timing information distribution in data.

The fit projections obtained from a 3D unbinned extended maximum likelihood fit in the signal regions are shown in Figure 6.2 for the fit variables, $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$. We obtain

5.7 ± 5.8 signal events and 989 ± 32 continuum background events in our fit to the data. The data-MC comparison of yields are shown in Table 6.2.

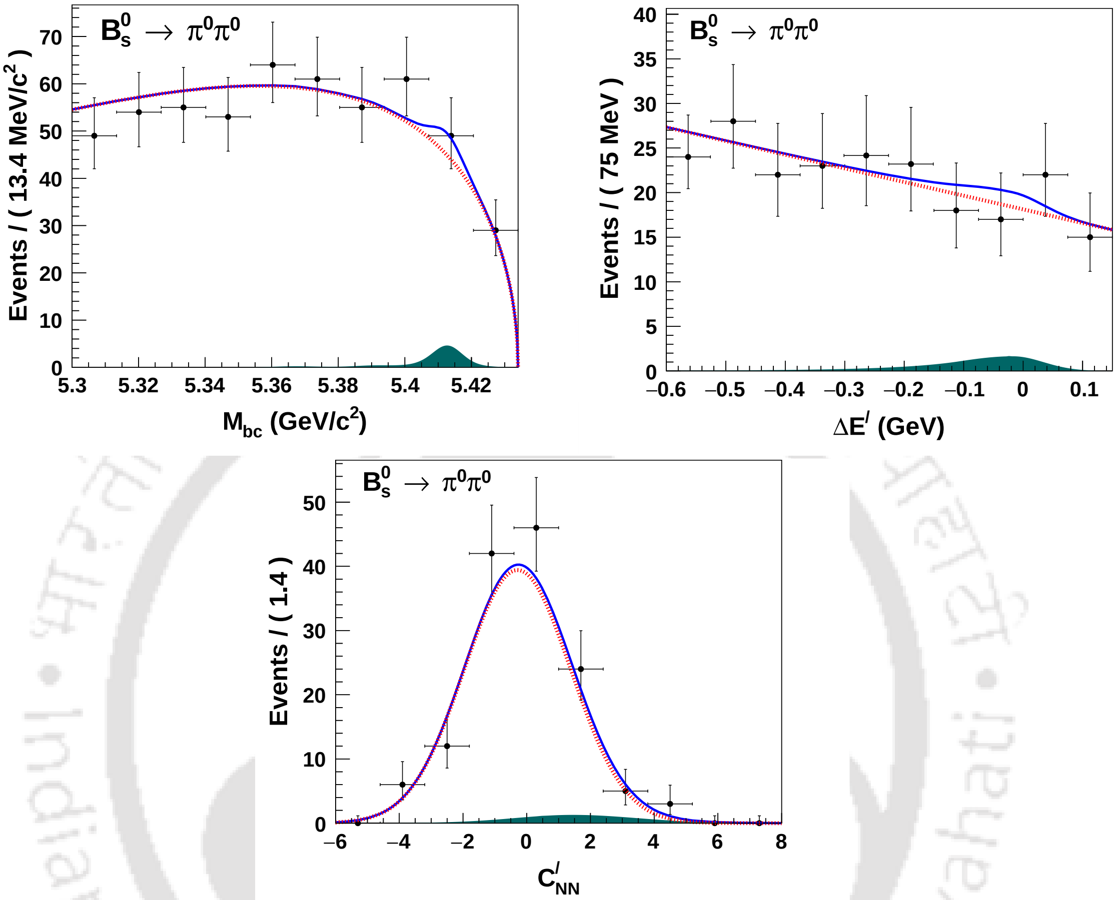

FIGURE 6.2: Signal enhanced projections of $M_{\rm bc}$ (top-left), $\Delta E'$ (top-right), and $\mathcal{C}'_{\rm NN}$ (bottom) for the analysis, $B_s^0 \to \pi^0\pi^0$. Each plot is generated by applying the signal region selection criteria on the two variables other than the plotted variable. The signal regions for the three variables are as follows, $5.395\ {\rm GeV}/c^2 < M_{\rm bc} < 5.434\ {\rm GeV}/c^2$, $-0.310\ {\rm GeV} < \Delta E' < 0.140\ {\rm GeV}$, and $-3.901 < \mathcal{C}'_{\rm NN} < 7.451$. The dark-filled, red (dotted), and blue (solid) color distributions represent the signal, continuum NN background and total fit function, respectively. Points with error bars represent data. The peak in the $M_{\rm bc}$ distribution is due to the dominant B_s^0 production channel, $B_s^{*0}\bar{B}_s^{*0}(87.0\%)$. The other two production channels, $B_s^{*0}\bar{B}_s^0(7.3\%)$ and $B_s^0\bar{B}_s^0(5.7\%)$ are present, but suppressed in the plot.

6.1.1 Estimation of systematic uncertainties

In this subsection, we consider all the possible systematic uncertainties that may have affected our analysis. They are enumerated below:

TABLE 6.2: Table comparing the signal and background yields from MC and data. The candidate and signal regions are defined in Table 4.1

Type	MC	Data	Region
Signal	1	5.7 ± 5.8	Candidate
Background	862	989 ± 32	Candidate
(continuum)			
	327	338 ± 19	Signal

1. **PDF parametrization** : We assign systematic uncertainty on the signal yield by varying all the fixed parameters by ±1 standard deviation (σ) about the fixed values obtained from the MC study. We consider this systematic uncertainty to arise because of the PDF parametrization. We obtain a fractional systematic uncertainty of $^{+3.5}_{-5.2}\%$. Table F.1 in appendix F.1 shows the detailed calculation of the bias arising due to PDF parametrization.

2. **Fit bias** : Uncertainties may also arise due to the fit model we chose to extract the signal yield in data. This study was performed using ensemble tests in MC (reported in section 4.7.1). We assign a systematic uncertainty of -3.3% to our fit model.

3. π^0 **selection efficiency** : It is estimated from the decay of $\tau^- \to \pi^-\pi^0\nu_\tau$ [57]. We assign a systematic uncertainty of 2.2% per π^0. Therefore, for the two π^0s, a total systematic uncertainty of 4.4% is assigned, considering the fact that the two π^0s are correlated.

4. **Uncertainty on** $\mathcal{B}(\pi^0 \to \gamma\gamma)$: The branching fraction measurement of $\pi^0 \to \gamma\gamma$ is $(98.823 \pm 0.034)\%$ [1]. We assign a fractional systematic uncertainty of 0.03% on the branching fraction of $\pi^0 \to \gamma\gamma$.

5. **Uncertainty on signal reconstruction efficiency,** $\epsilon_{\rm rec}$: The uncertainty on $\epsilon_{\rm rec}$ is calculated using the following formula, which obeys a binomial distribution

$$\Delta\epsilon = \sqrt{\frac{\epsilon_{\rm rec}(1-\epsilon_{\rm rec})}{\rm N}}, \tag{6.1}$$

where, N is the number of generated signal MC events. In our case, $\epsilon_{\rm rec} = 12.69\%$ and $\rm N = 500000$. This gives, $\epsilon_{\rm rec} = (12.69 \pm 0.05)\%$. We then assign a fractional systematic uncertainty of 0.39% on $\epsilon_{\rm rec}$.

6. **Uncertainty on the neural network selection criterion** : We assign systematic uncertainty due to the efficiency of our neural network selection criterion. We calculate this uncertainty using the control sample ($B^0 \to \pi^0\pi^0$) with a parameter, R, defined as

$$R = \frac{\epsilon_{\rm Data}}{\epsilon_{\rm MC}}, \tag{6.2}$$

where,

$$\epsilon_{\rm Data/MC} = \frac{\text{Entries remaining with } \mathcal{C}'_{\rm NN} \text{ selection criterion}}{\text{Entries remaining without } \mathcal{C}'_{\rm NN} \text{ selection criterion}}. \tag{6.3}$$

Table 6.3 shows the R-value obtained using this procedure. The fractional uncertainty in R is assigned as a source of systematic uncertainty due to the efficiency of the neural network selection criteria.

Table 6.3: Table to estimate the systematic uncertainty associated with the choice of the selection criterion on the neural network output.

Type	Selection Criterion	Yield	$\epsilon_{\rm Data}/\epsilon_{\rm MC}$	R
Data	Without $\mathcal{C}'_{\rm NN}$ cut	294020 ± 542	0.0101 ± 0.0001	1.01 ± 0.03
	With $\mathcal{C}'_{\rm NN} > 0.90$	2980 ± 55		
MC	Without $\mathcal{C}'_{\rm NN}$ cut	286531 ± 531	0.0100 ± 0.0001	
	With $\mathcal{C}'_{\rm NN} > 0.90$	2890 ± 54		

From the value of R, as shown in Table 6.3, we assign a fractional uncertainty of 3%.

7. **Uncertainty on production cross-section of $b\bar{b}$** : The production cross-section of $b\bar{b}$ at $\Upsilon(5S)$ is measured to be, (0.340 ± 0.016) nb [8]. We assign 4.7% as a source of systematic uncertainty due to this production cross-section.

8. **Uncertainty due to production of $B_s^{(*0)}\bar{B}_s^{(*0)}$ from $b\bar{b}$, f_s** : The fraction with which $B_s^{(*0)}\bar{B}_s^{(*0)}$ is produced from $b\bar{b}$ is, 0.201 ± 0.031 [1]. We, therefore, assign a fractional systematic uncertainty of 15.4% on f_s.

TABLE 6.4: Systematic uncertainties associated with the decay channel, $B_s^0 \to \pi^0\pi^0$. The lower part part of the table summarizes the systematic uncertainties which are external to the analysis.

Type	Systematic uncertainties	Fractional uncertainties (%)
Internal	PDF parametrization	$^{+3.5}_{-5.2}$
	Fit bias	-3.3
	Reconstruction efficiency, $\epsilon_{\rm rec}^{\rm sig}$	± 0.4
	$\mathcal{C}'_{\rm NN}$ requirement	± 3.0
	Fractions of $B_s^{(*0)}\bar{B}_s^{(*0)}$	$^{+5.2}_{-3.5}$
	Uncertainties added in quadrature	$^{+7.0}_{-7.7}$
External	π^0 efficiency	± 4.4
	$\mathcal{B}(\pi^0 \to \gamma\gamma)$	± 0.03
	$b\bar{b}$ cross-section, $\sigma_{b\bar{b}}$	± 4.7
	f_s	± 15.4
	Uncertainties added in quadrature	± 16.7
Total	Uncertainties added in quadrature	$\mathbf{^{+18.1}_{-18.4}}$

9. **Uncertainty due to production of B_s^0 from $B_s^{(*0)}\bar{B}_s^{(*0)}$** : The $B_s^{(*0)}\bar{B}_s^{(*0)}$ pairs include $B_s^{*0}\bar{B}_s^{*0}$, $B_s^0\bar{B}_s^{*0}$ + $B_s^{*0}\bar{B}_s^0$ ($B_s^0\bar{B}_s^{*0}$), and $B_s^0\bar{B}_s^0$, the fraction of the first two being $f_{B_s^{*0}\bar{B}_s^{*0}} = (87.0 \pm 1.7)\%$ and $f_{B_s^0\bar{B}_s^{*0}} = (7.3 \pm 1.4)\%$, respectively [8]. We vary these fractions by $\pm 1\sigma$ about their fixed values to see the changes in the signal yield. The deviation in the yield is assigned as a source of systematic uncertainty. We obtain a fractional uncertainty of $^{+5.2}_{-3.2}\%$.

Table 6.4 summarizes the systematic uncertainties associated with the analysis.

6.1.2 Branching fraction and significance level

From the data fit, we obtain a signal yield of 5.7 ± 5.8 (shown in Table 6.2) with a reconstruction efficiency of $(12.69 \pm 0.05)\%$. We calculate the branching fraction using the formula

$$\mathcal{B} = \frac{N^{\text{sig}}_{\text{fit}}}{2 \times N_{B^0_s \bar{B}^0_s} \times \epsilon \times \prod_i^2 \mathcal{B}_i}, \tag{6.4}$$

where, $N^{\text{sig}}_{\text{fit}} = 5.7 \pm 5.8$, $N_{B^0_s \bar{B}^0_s} = (8.30 \pm 1.34) \times 10^6$, is the number of $B^0_s \bar{B}^0_s$ pairs, $\epsilon = (12.69 \pm 0.05)\%$ is the signal reconstruction efficiency, and $\mathcal{B}_i = (98.823 \pm 0.034)\%$, is the branching fraction of, $\pi^0 \rightarrow \gamma\gamma$.

The value of the branching fraction obtained using Equation 6.4 is 2.8×10^{-6}.

The statistical uncertainty associated with the $\mathcal{B}$ is calculated using the formula,

$$\Delta\mathcal{B} = \frac{\Delta N^{\text{fit}}_{\text{sig}}}{N^{\text{fit}}_{\text{sig}}} \times \mathcal{B}, \tag{6.5}$$

where, $\Delta N^{\text{fit}}_{\text{sig}} = 5.8$ and $N^{\text{fit}}_{\text{sig}} = 5.7$, is the uncertainty associated with the signal yield and the signal yield obtained from the 3D fit, respectively.

From Equation 6.5, the statistical uncertainty on the branching fraction is found to be 2.8×10^{-6}.

The total systematic uncertainty is estimated to be $^{+18.1}_{-18.4}\%$ for this analysis, as reported in Table 6.4. This corresponds to a systematic uncertainty of 0.5×10^{-6} on the branching fraction.

Therefore, the branching fraction calculated, can be reported as

$$\mathcal{B}(B^0_s \rightarrow \pi^0\pi^0) = (2.8 \pm 2.8\ (\text{stat.}) \pm 0.5\ (\text{syst.})) \times 10^{-6}, \tag{6.6}$$

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

The significance of the result is calculated using $\sqrt{-2ln\lambda} = \sqrt{ln(\mathcal{L}_0) - ln(\mathcal{L}_{max})}$, where, $\mathcal{L}_0$ and $\mathcal{L}_{max}$ are the plain likelihoods considering background only (no signal) hypothesis and maximum value obtained from the 3D fit, respectively [58]. The significance of the result is found to be 0.1σ. Since, the significance is below 3σ, we estimate the upper limit on the branching fraction using the Bayesian approach.

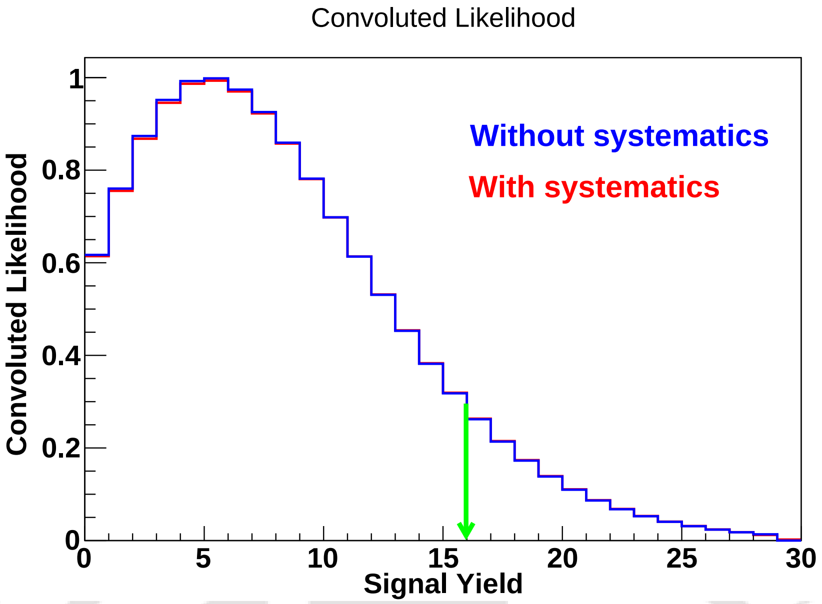

FIGURE 6.3: Plot of convoluted likelihood obtained from 3D unbinned extended maximum likelihood fit to data corresponding to an integrated luminosity of 121.4 fb^{-1} at $\Upsilon(5S)$ resonance. The green arrow represents the area of the distribution integrated upto 90% and points to corresponding signal yield ($N^{\text{sig}}_{90\%\ \text{CL}} = 16$).

6.1.3 Estimation of upper limit on the branching fraction

The process of obtaining the upper limit (UL) is enumerated below (following reference [58]):

1. Construct the profile likelihood, $-2ln\lambda = ln(\mathcal{L}_0) - ln(\mathcal{L}_{max})$ where, $\mathcal{L}_0$ and $\mathcal{L}_{max}$ are the plain likelihoods considering background only (no signal) hypothesis and maximum value obtained from the 3D fit, respectively.
2. Convolute the plain likelihood, λ with a Gaussian distribution of width equal to the total systematic uncertainty estimated from the analysis. We find the total systematic uncertainty associated with the procedures adopted for signal extraction to be $^{+18.1}_{-18.4}\%$.
3. Integrate the convoluted plain likelihood upto 90% of its area to estimate the 90% confidence level (CL) [1] UL. Figure 6.3 shows the convoluted likelihood obtained from data.

---

[1] As we use Bayesian method, this is "credible interval" but we use "confidence level" here following common convention.

The formula to obtain the UL on the branching fraction is defined as

$$\mathcal{B} = \frac{N^{\rm sig}_{90\%\ \rm CL}}{2 \times N_{B^0_s \bar{B}^0_s} \times \epsilon \times \prod_i^2 \mathcal{B}_i}, \tag{6.7}$$

where, $N^{sig}_{90\%\ \rm CL} = 16$ (with systematics, position of green line in Figure 6.3) is the signal yield corresponding to 90% CL, $N_{B^0_s \bar{B}^0_s} = (8.30 \pm 1.34) \times 10^6$, $\epsilon = (12.69 \pm 0.05)\%$, and $\mathcal{B}_i = (98.823 \pm 0.034)\%$ [1], is the branching fraction of, $\pi^0 \to \gamma\gamma$.

We obtain a UL on the BF of 7.3×10^{-6} at 90% CL (without systematics) and 7.7×10^{-6} at 90% CL (with systematics). This is an improvement by a magnitude of twenty seven from the previous upper limit set by L3 experiment ($< 2.1 \times 10^{-4}$ at 90% CL) [22].

6.1.4 Summary on the results from data

We summarize our data results briefly as follows:

1. We have used all the experiments dedicated to $\Upsilon(5S)$ resonance to look at the results for this analysis.
2. We do not find any significant deviation between the MC results and data.
3. The branching fraction obtained from this analysis is $\mathbf{(2.8 \pm 2.8\ (stat.) \pm 0.5\ (syst.)) \times 10^{-6}}$.
4. In absence of any significant signal yield, we use the Bayesian approach to estimate the upper limit on the branching fraction.
5. The upper limit on the branching fraction at **90%** confidence level is estimated to be $\mathbf{7.7 \times 10^{-6}}$. This is an improvement by a magnitude of twenty seven from the previous upper limit set by the L3 experiment ($\mathbf{< 2.1 \times 10^{-4}}$ at **90%** CL) [22] in 1995.

6.2 Outlook

B^0_s physics studies involving weak-annihilation amplitudes are gradually gaining momentum with experimental interest both from LHCb and Belle. As discussed in chapter 1, they are highly suppressed within the SM and are assumed to have negligible contributions to theoretical calculations. However, as we gradually proceed towards an era of precision, where experimental uncertainties compete with theoretical ones, the inclusion of higher-order terms becomes imperative. The decays involving weak-annihilation amplitudes provide ample scope to compare the theoretical predictions and experimental measurements and also provide the opportunity to search for physics beyond the SM. However, progress in this front is limited due to the unavailability of high luminosity data. Nevertheless, with LHCb upgrade for its run 3 and Belle II estimated to collect ~ 5 ab^{-1} of data at $\Upsilon(5S)$ will provide a large number of data sets to measure these decays.

Our analysis, so far, puts the most stringent UL on the branching fraction for the decay $B_s^0 \to \pi^0\pi^0$ to date. We note that the predicted values of the branching fractions from FDA, pQCD and QCD approaches (discussed in chapter 1) are still lower than our UL estimation. This, therefore, remains a potential discovery channel for the Belle II experiment with a dedicated run at the $\Upsilon(5S)$ resonance.

Appendix A

Decay file

A.1 Decay File to generate $B_s^0 \to \pi^0\pi^0$

```
  Alias MyB_s*0 B_s*0
  Alias Myanti-B_s*0 anti-B_s*0
  Alias MyB_s0 B_s0
  Alias Myanti-B_s0 anti-B_s0
  Alias Mypi0 pi0

# The production fractions are normalized to unity

  Decay     Upsilon(5S)
  0.435     MyB_s*0  anti-B_s*0              PHSP ;
  0.435     B_s*0    Myanti-B_s*0            PHSP ;
  0.01825   MyB_s*0  anti-B_s0               PHSP ;
  0.01825   B_s*0    Myanti-B_s0             PHSP ;
  0.01825   MyB_s0   anti-B_s*0              PHSP ;
  0.01825   B_s0     Myanti-B_s*0            PHSP ;
  0.0285    MyB_s0   anti-B_s0               VSS  ;
  0.0285    B_s0     Myanti-B_s0             VSS  ;
  EndDecay

  Decay     MyB_s*0
  1.0       MyB_s0       gamma      VSP_PWAVE ;
  EndDecay

  Decay     Myanti-B_s*0
  1.0       Myanti-B_s0  gamma      VSP_PWAVE ;
  EndDecay

  Decay     MyB_s0
  1.0       Mypi0        Mypi0      PHSP ;
  EndDecay
```

```
Decay      Myanti-B_s0
1.0        Mypi0          Mypi0       PHSP ;
EndDecay

Decay      Mypi0
1.0        gamma          gamma       PHSP ;
EndDecay

End
```

Appendix B

Neural Network

B.1 Event shape distributions

The distributions of the event shape variables are shown below in Figures B.1, B.2 and B.3. The blue and red distributions represent correctly reconstructed MC signal and continuum backgrounds, respectively.

The correlations of the event shape distributions and the thrust variable with the fit variables, M_{bc} and $\Delta E'$ for signal MC are shown in Table B.1. Table B.1 suggests that the correlations are not significant.

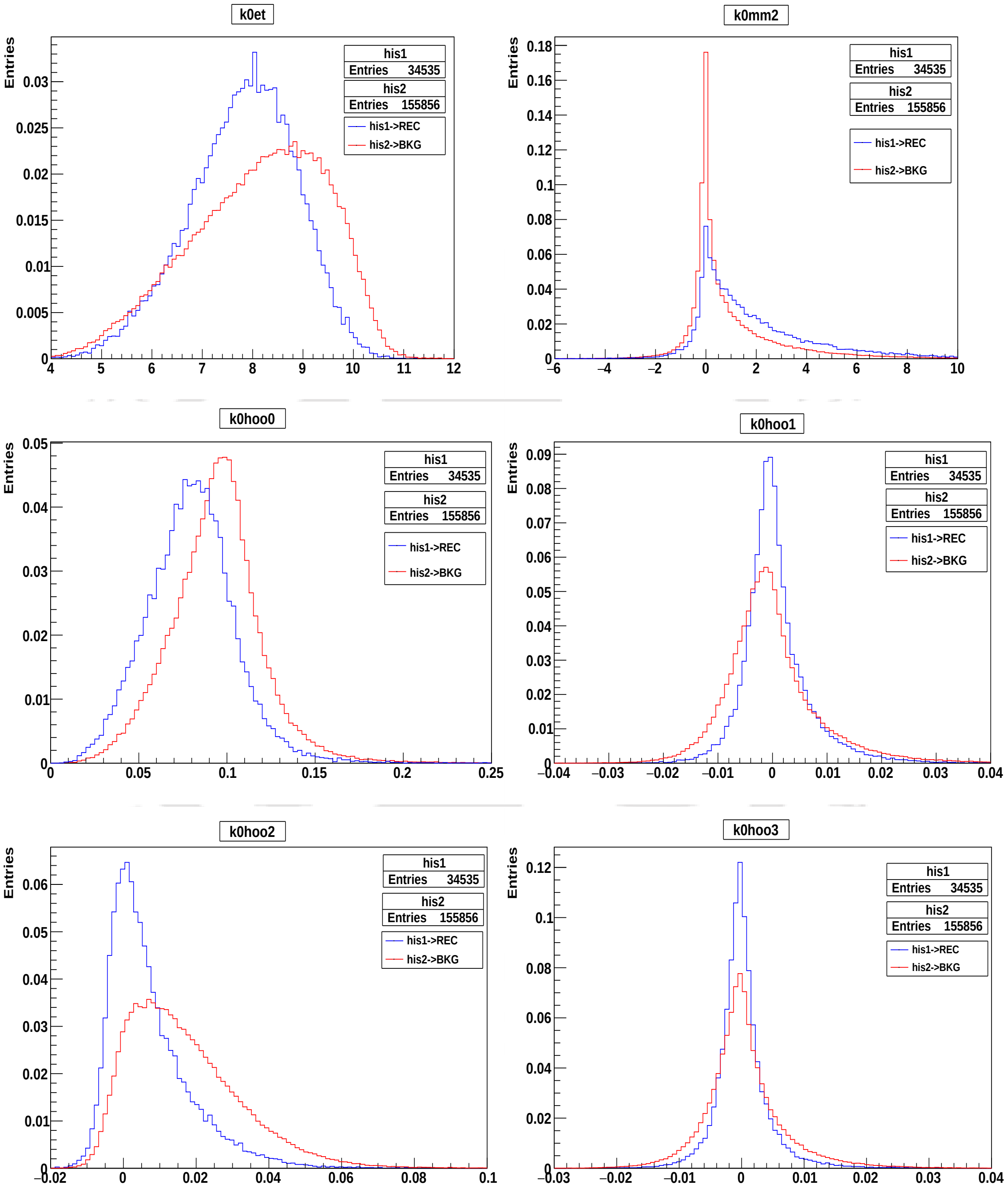

FIGURE B.1: Event shape distributions.

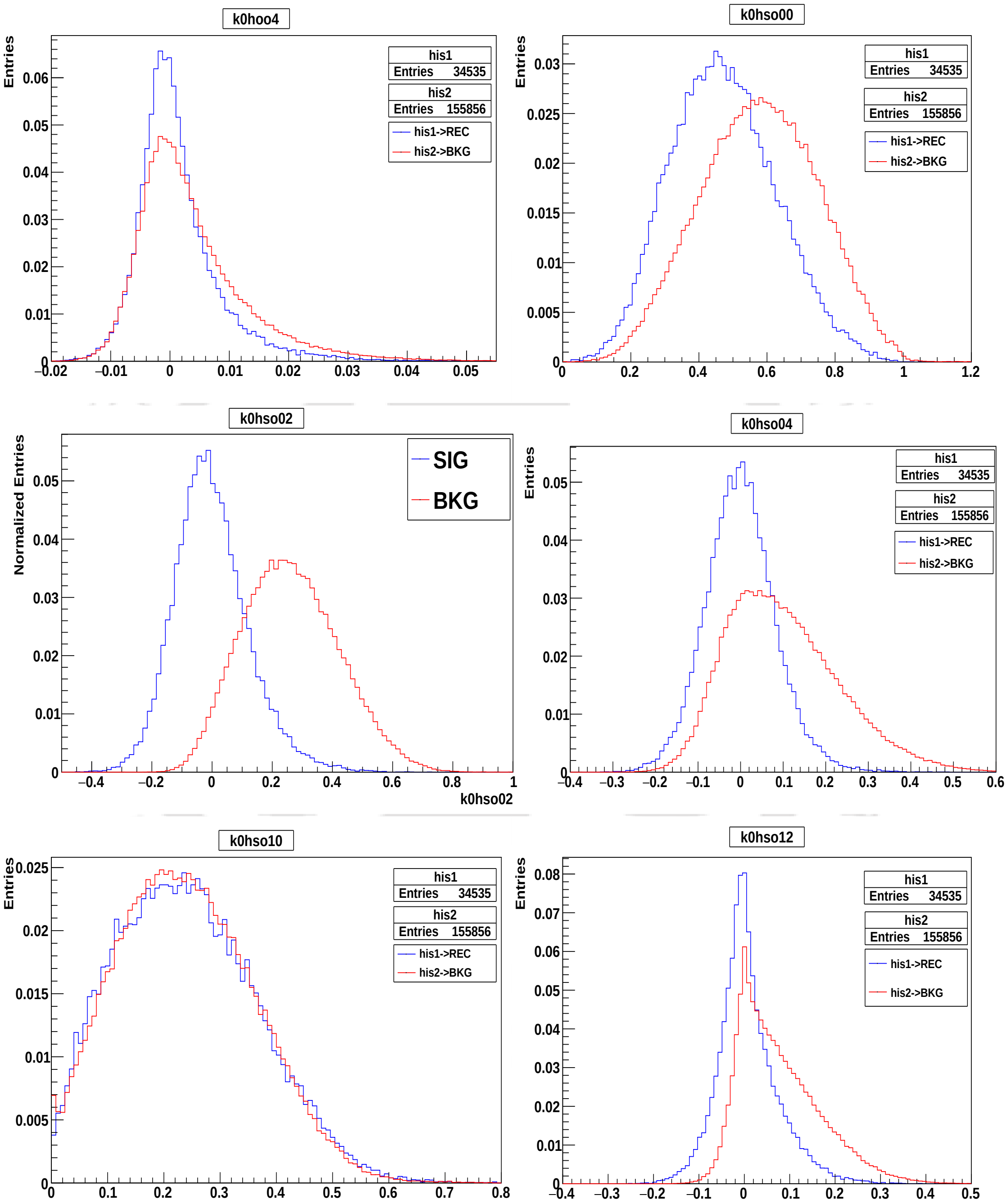

FIGURE B.2: Event shape distributions.

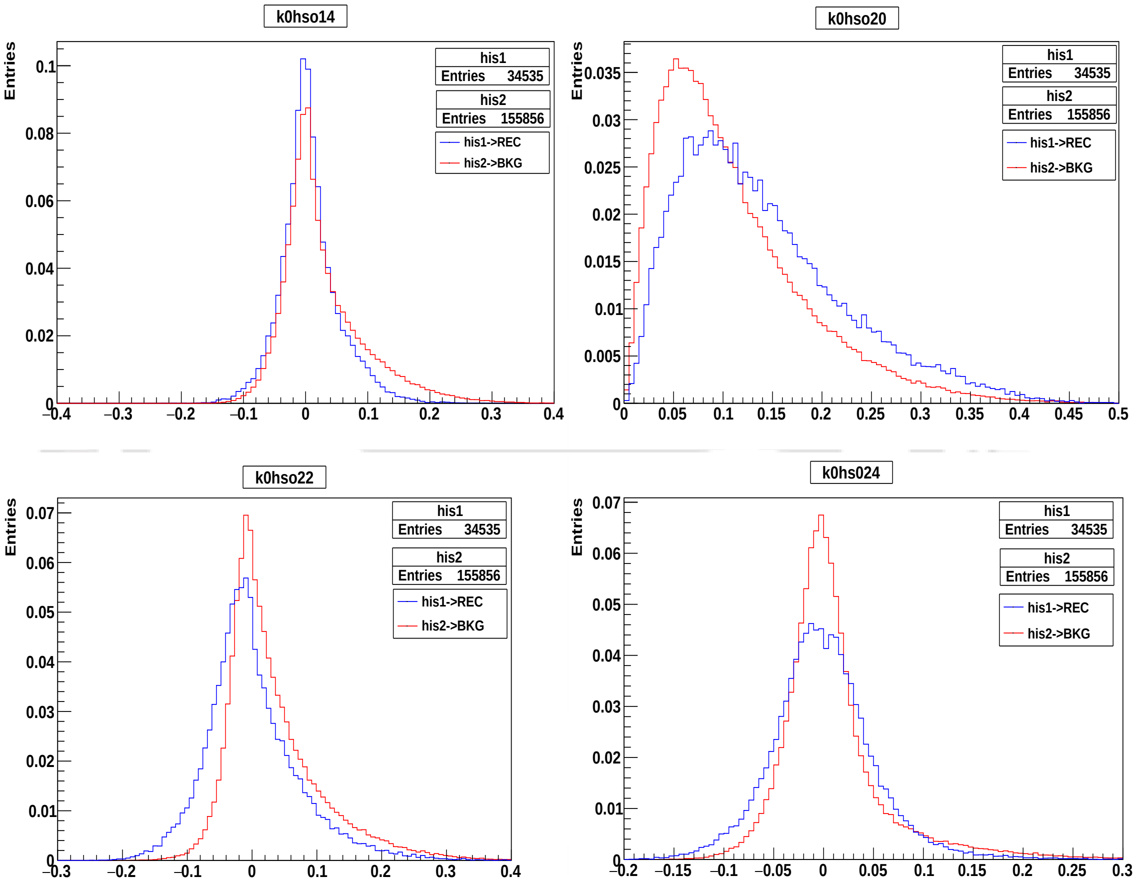

FIGURE B.3: Event shape distributions.

Table B.1: Linear correlation coefficients of the event shape variables and the thrust variable with the fit variables, $M_{\rm bc}$ and $\Delta E'$ for signal MC.

Variable	Correlation coefficients	
	$M_{\rm bc}$	$\Delta E'$
k0et	0.0150	0.0816
k0mm2	-0.0048	-0.0997
k0hoo0	-0.0065	0.0790
k0hoo1	0.003	0.005
k0hoo2	0.0144	0.0237
k0hoo3	-0.0022	0.0008
k0hoo4	0.0071	0.0067
k0hso00	0.0123	0.0692
k0hso02	-0.0033	0.0019
k0hso04	-0.0076	-0.0115
k0hso10	-0.0083	-0.0075
k0hso12	-0.0385	-0.0469
k0hso14	-0.0452	-0.0622
k0hso20	-0.0196	0.0171
kohso22	-0.0426	-0.0536
kohso24	-0.0051	-0.0038
$\cos\theta_T$	-0.0112	-0.0147

Appendix C

Formulae of PDFs

C.1 Formulae of PDFs

The following formulae are used for Crystal Ball, ARGUS, and Asymmetric Crystal Ball PDFs.

- The Crystal Ball function is algorithmically defined as follows,

```
Double_t RooCBShape::evaluate() const {
Double_t t = (m-m0)/sigma;
if (alpha < 0) t = -t;

Double_t absAlpha = fabs((Double_t)alpha);

if (t >= -absAlpha) {
return exp(-0.5*t*t);
}
else {
Double_t a =  TMath::Power(n/absAlpha,n)*exp(-0.5*absAlpha*absAlpha);
Double_t b= n/absAlpha - absAlpha;

return a/TMath::Power(b - t, n);
}
}
```

where, m and m_0 are the fit variable and initial value of the fit variable, $sigma$ is the standard deviation, $alpha$ and n are the parameters which determine the tail of the Gaussian. The algorithm is written in accordance with the RooCBShape function definition found in the following link : `https://root.cern.ch/doc/master/RooCBShape_8cxx_source.html`

- The ARGUS function is algorithmically defined as follows,

```
Double_t RooArgusBG::evaluate() const {
Double_t t= m/m0;
if(t >= 1) return 0;

Double_t u= 1 - t*t;
return m*TMath::Power(u,p)*exp(c*u) ;
}
```

where, m and m_0 are the mass parameter and initial value of the mass parameter, c is the slope and p is the power whose value is fixed to 0.5. The documentation of the algorithm can be found at : `https://root.cern.ch/doc/master/RooArgusBG_8cxx_source.html`.

- The asymmetric Crystal Ball function is algorithmically defined as,

```
Double_t RooAsymCBShape::evaluate() const
{
Double_t t = (x - mean) / (x < mean ? sigmaL : sigmaR);

if (alpha < 0) t = -t;

Double_t absAlpha = fabs((Double_t)alpha);

if (t >= -absAlpha) {
return exp(-0.5*t*t);
}
else {
Double_t a =  TMath::Power(n/absAlpha,n)*exp(-0.5*absAlpha*absAlpha);
Double_t b= n/absAlpha - absAlpha;

return a/TMath::Power(b - t, n);
}
}
```

where, x is the variable name, $mean$ is the mean value of the variable, $sigmaL$ and $sigmaR$ are the left and right standard deviations, $alpha$ and n define the tail of the distribution. This is a privately defined function.

Appendix D

Self cross-feed signal MC distributions

D.1 Distributions of self cross-feed (SCF) signal MC candidates

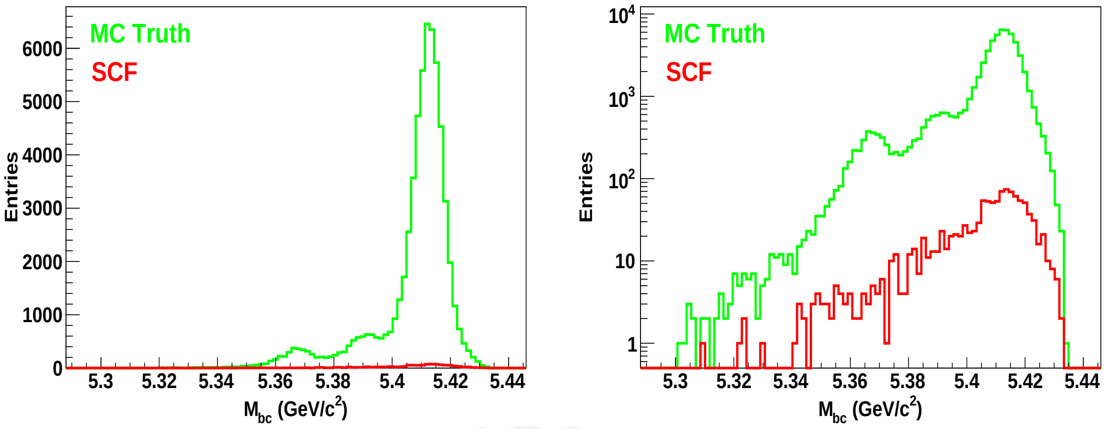

FIGURE D.1: Without normalization (left) and with logarithmic scale (right) distributions of $M_{\rm bc}$ variable. The green and red distributions represent correctly reconstructed (MC truth) and SCF signal MC candidates .

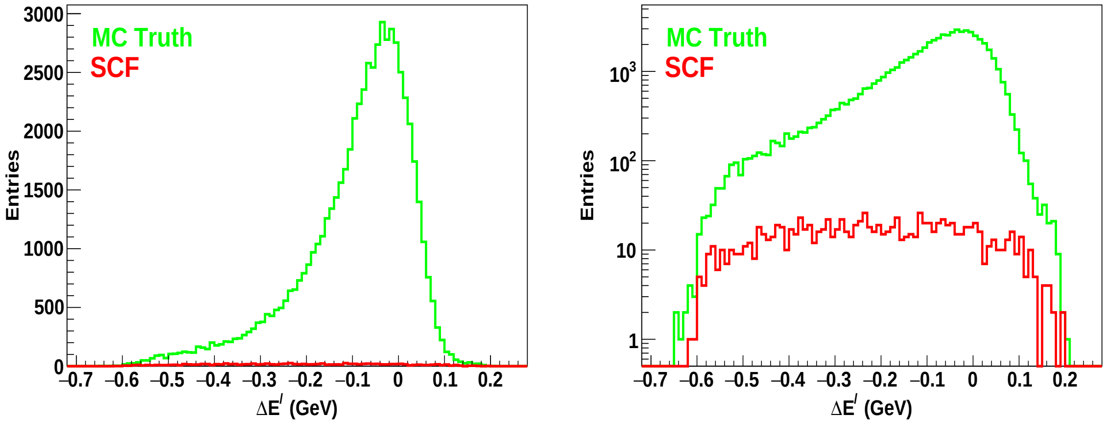

FIGURE D.2: Without normalization (left) and with logarithmic scale (right) distributions of $\Delta E'$ variable .The green and red distributions represent correctly reconstructed (MC truth) and SCF signal MC candidates.

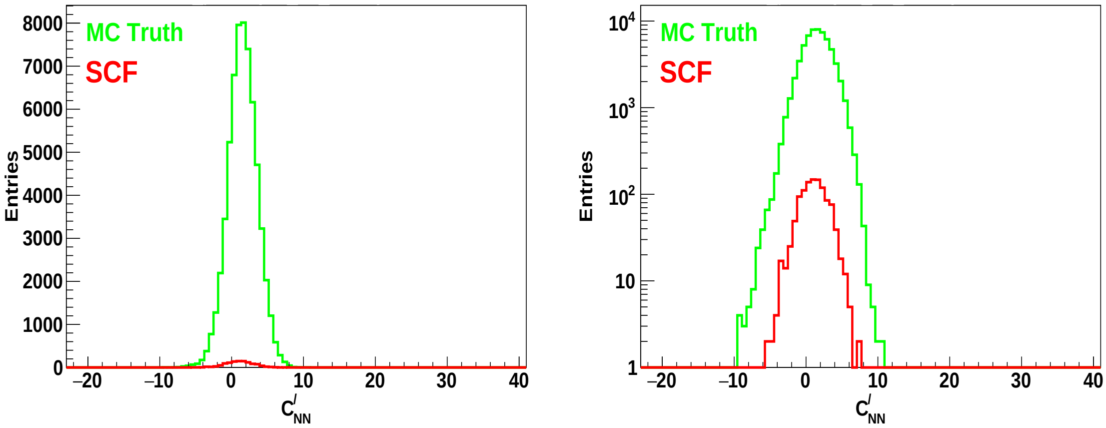

FIGURE D.3: Without normalization (left) and with logarithmic scale (right) distributions of $\mathcal{C}'_{\rm NN}$ variable .The green and red distributions represent correctly reconstructed (MC truth) and SCF signal MC candidates.

Appendix E

Fit bias, linearity, and correlation studies

E.0.1 Gsim study

Figures E.1 to E.7 show the pull and yield distributions for $N_{\rm gen}^{\rm sig} = 1, 2, 3, 5, 7, 10,$ and 20.

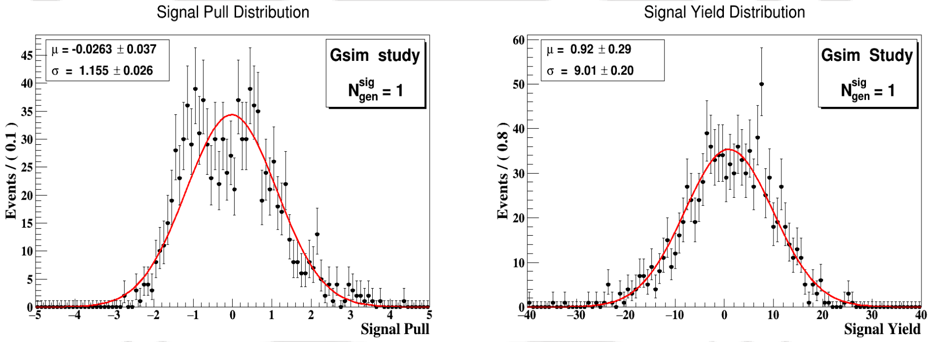

FIGURE E.1: Signal pull (left) and signal yield (right) distribution for $N_{\rm gen}^{\rm sig} = 1$

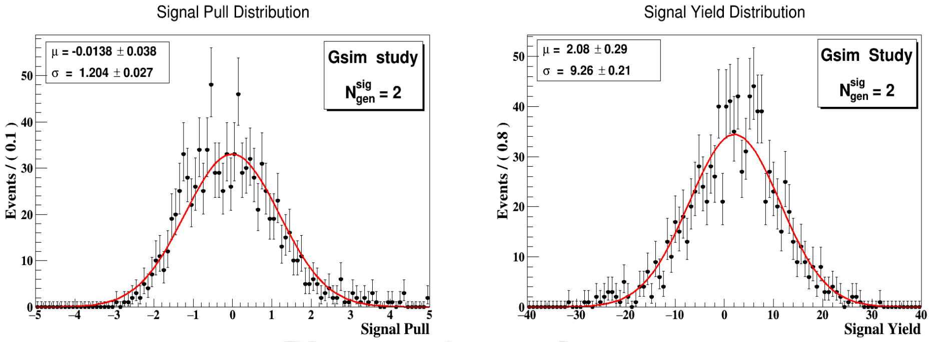

FIGURE E.2: Signal pull (left) and signal yield (right) distribution for $N^{\mathrm{sig}}_{\mathrm{gen}} = 2$

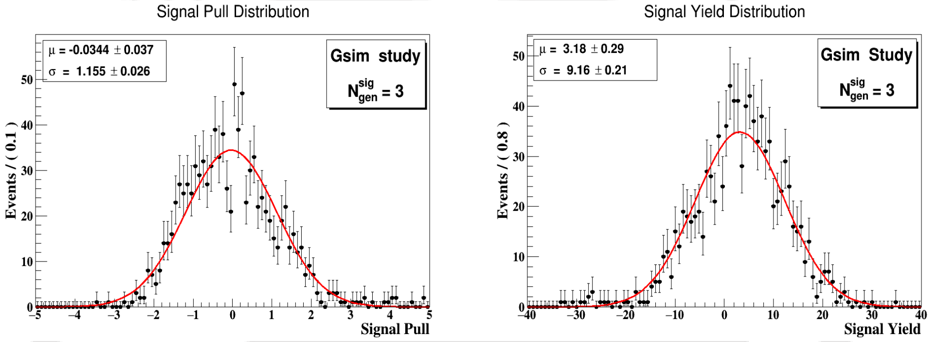

FIGURE E.3: Signal pull (left) and signal yield (right) distribution for $N^{\mathrm{sig}}_{\mathrm{gen}} = 3$

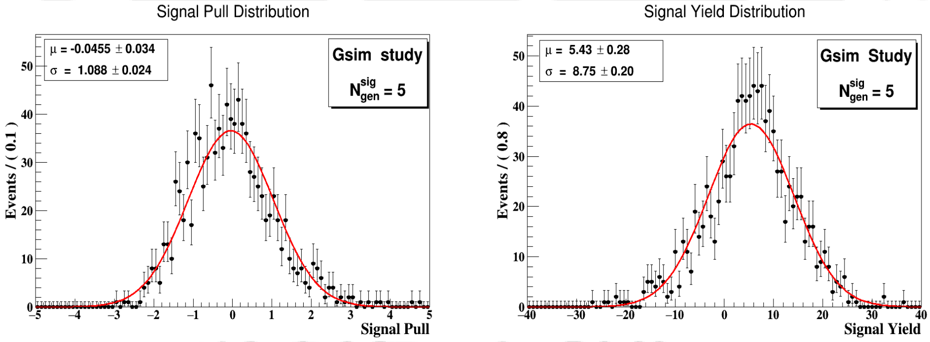

FIGURE E.4: Signal pull (left) and signal yield (right) distribution for $N^{\mathrm{sig}}_{\mathrm{gen}} = 5$

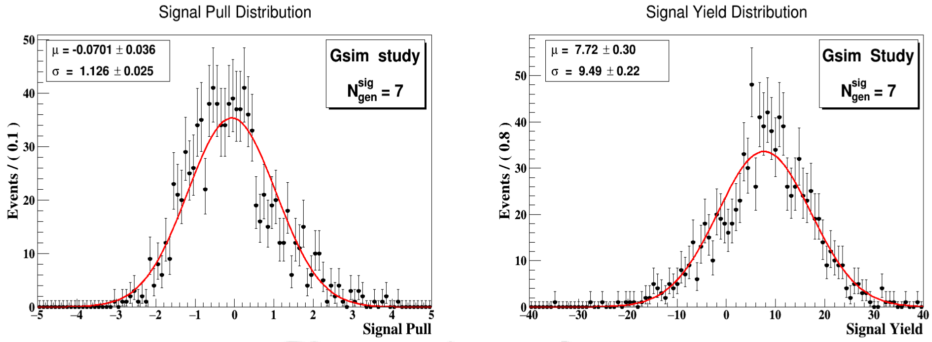

FIGURE E.5: Signal pull (left) and signal yield (right) distribution for $\mathrm{N}_{\mathrm{gen}}^{\mathrm{sig}} = 7$

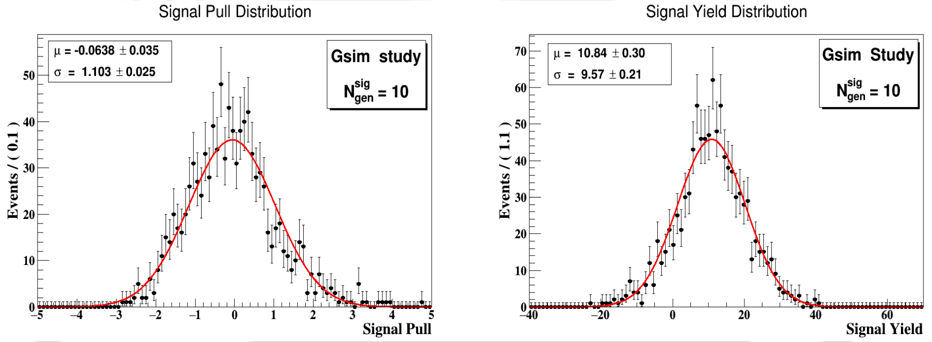

FIGURE E.6: Signal pull (left) and signal yield (right) distribution for $\mathrm{N}_{\mathrm{gen}}^{\mathrm{sig}} = 10$

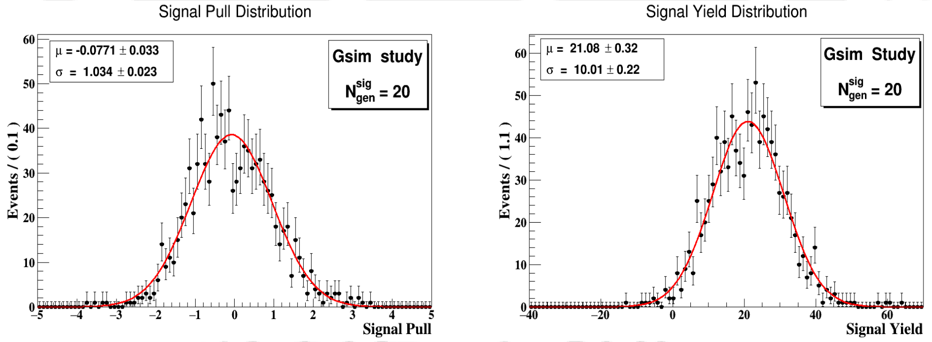

FIGURE E.7: Signal pull (left) and signal yield (right) distribution for $\mathrm{N}_{\mathrm{gen}}^{\mathrm{sig}} = 20$

E.0.2 Toy MC study

Figures E.8 to E.14 show the pull and yield distributions for $N_{\mathrm{gen}}^{\mathrm{sig}} = 1, 2, 3, 5, 7, 10,$ and 20.

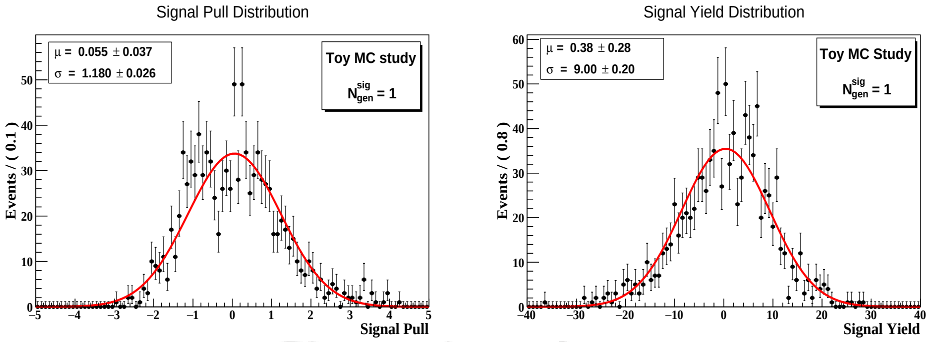

FIGURE E.8: Signal pull (left) and signal yield (right) distribution for $N_{\rm gen}^{\rm sig} = 1$

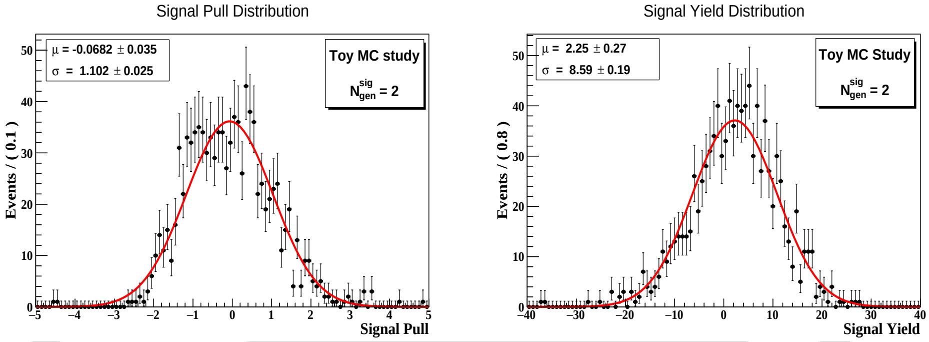

FIGURE E.9: Signal pull (left) and signal yield (right) distribution for $N_{\rm gen}^{\rm sig} = 2$

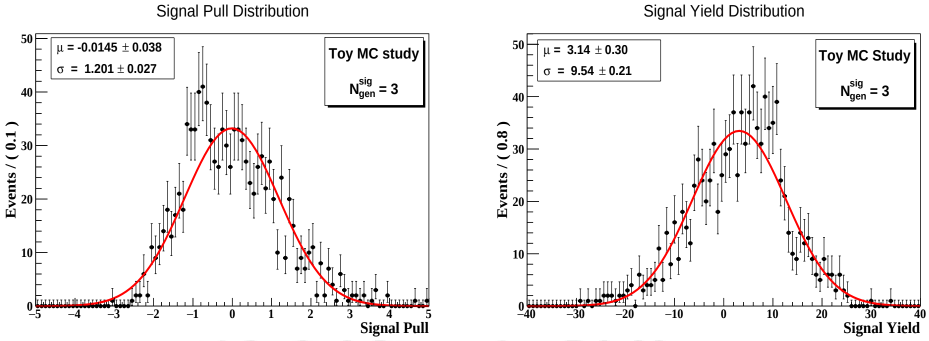

FIGURE E.10: Signal pull (left) and signal yield (right) distribution for $N_{\rm gen}^{\rm sig} = 3$

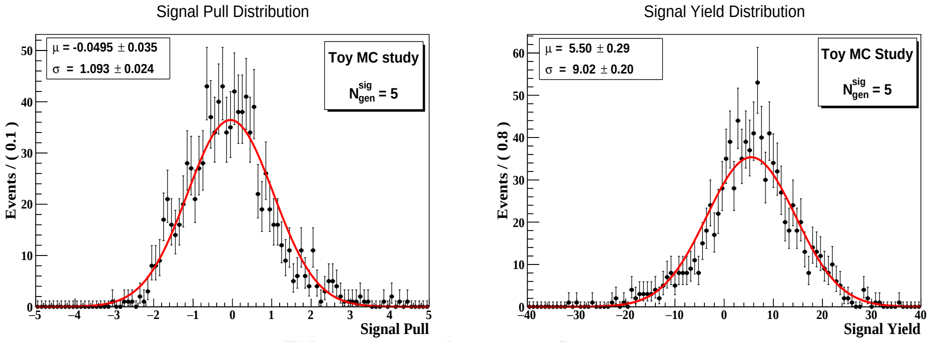

FIGURE E.11: Signal pull (left) and signal yield (right) distribution for $N_{\mathrm{gen}}^{\mathrm{sig}} = 5$

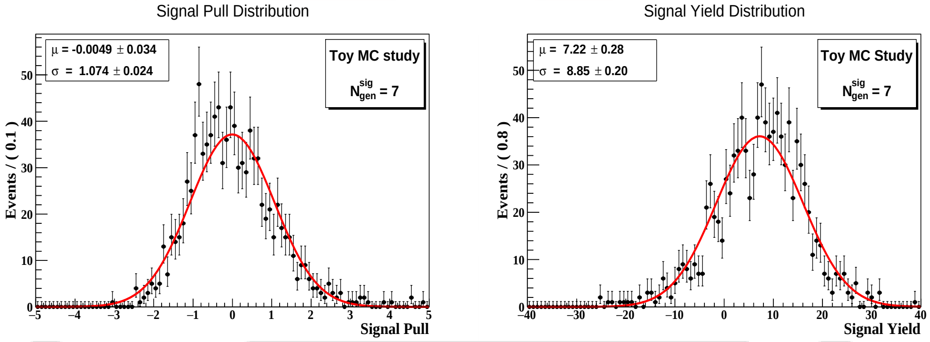

FIGURE E.12: Signal pull (left) and signal yield (right) distribution for $N_{\mathrm{gen}}^{\mathrm{sig}} = 7$

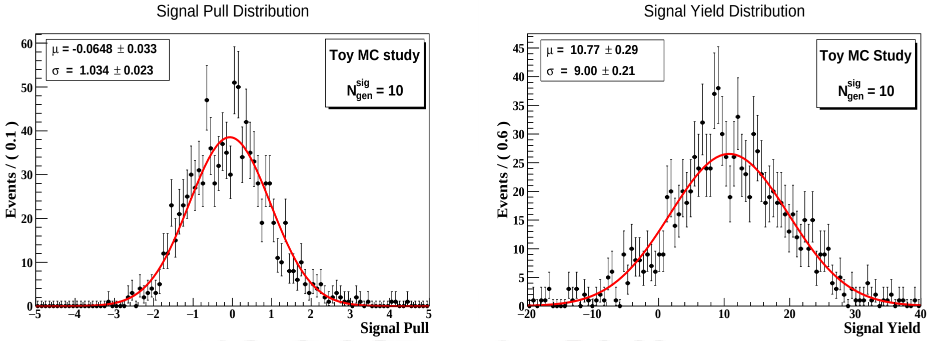

FIGURE E.13: Signal pull (left) and signal yield (right) distribution for $N_{\mathrm{gen}}^{\mathrm{sig}} = 10$

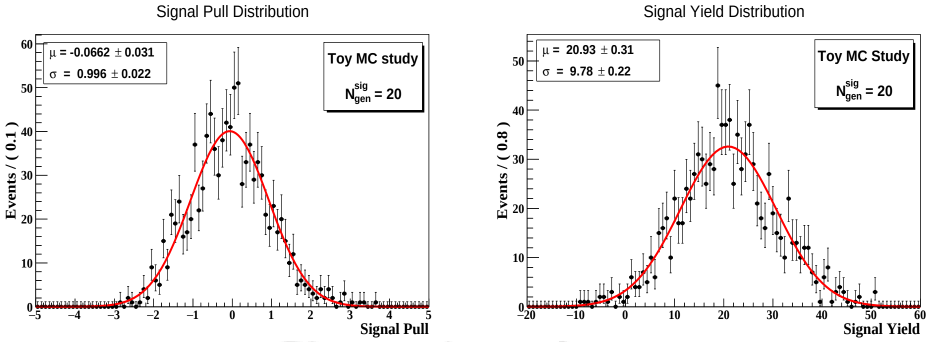

FIGURE E.14: Signal pull (left) and signal yield (right) distribution for $N_{\text{gen}}^{\text{sig}} = 20$

Appendix F

Systematic uncertainties

F.1 Systematic uncertainties

Table F.1 shows the systematic uncertainty arising due to the PDF parametrization in real data.

Table F.1: Table showing the systematic uncertainties associated with the PDF parametrization. It is obtained by varying each of the fixed parameter values by 1σ.

Sl. No	Variable	Channel	PDF	Parameters	Parameter Values	Yield (-1σ)	Yield	Yield ($+1\sigma$)
1	$\mathrm{M_{bc}}$ (GeV/c^2)	$B_s^0\bar{B}_s^0$	G1	μ	5.367053 ± 0.000083	5.7	5.7	5.7
2				σ	0.00599 ± 0.00011	5.7	5.7	5.7
3				fraction	0.703 ± 0.012	5.7	5.7	5.7
4			G2	σ	0.02023 ± 0.00041	5.7	5.7	5.7
5		$B_s^0\bar{B}_s^0$		fraction	0.057 ± 0.022	5.6	5.7	5.7
	Uncertainty						$^{+0.0}_{-0.1}$	
6	$\mathrm{M_{bc}}$ (GeV/c^2)	$B_s^0\bar{B}_s^{*0}+\mathrm{cc}$	G	μ	5.390079 ± 0.000072	5.7	5.7	5.7
7				σ	0.00498 ± 0.00013	5.7	5.7	5.7
8			CB	σ	0.01242 ± 0.00033	5.7	5.7	5.7
9				α	1.333 ± 0.089	5.7	5.7	5.7
10				n	5.0 ± 1.4	5.7	5.7	5.7
11				fraction	0.406 ± 0.022	5.7	5.7	5.7
12		$B_s^0\bar{B}_s^{*0}+cc$		fraction	0.073 ± 0.014	5.6	5.7	5.8
	Uncertainty						$^{+0.1}_{-0.1}$	
13	$\mathrm{M_{bc}}$ (GeV/c^2)	$B_s^{*0}\bar{B}_s^{*0}$	G	μ	5.41283 ± 0.000051	5.6	5.7	5.8
14				σ	0.00411 ± 0.000083	5.6	5.7	5.7
15			CB	μ	5.41136 ± 0.00020	5.7	5.7	5.7
16				σ	0.00766 ± 0.00023	5.6	5.7	5.7
17				α	1.134 ± 0.068	5.7	5.7	5.6
18				n	7.8 ± 2.1	5.7	5.7	5.7
19				fraction	0.437 ± 0.031	5.6	5.7	5.8
20		$B_s^{*0}\bar{B}_s^{*0}$		fraction	0.87 ± 0.017	5.6	5.7	5.4
	Uncertainty						$^{+0.4}_{-0.2}$	
21		all (background)	ARGUS	m (end-point)	5.434 ± 0.000095	5.8	5.7	5.8
	Uncertainty						$^{+0.1}_{-0.1}$	
22	$\Delta E'$ (GeV)	all	ACB	μ	-0.01695 ± 0.0010	5.6	5.7	5.7
23				σ_L	0.0983 ± 0.0021	5.6	5.7	5.7
24				σ_R	0.0507 ± 0.00056	5.6	5.7	5.7
25				α	0.916 ± 0.027	5.7	5.7	5.7
26				n	6.38 ± 0.61	5.7	5.7	5.7
	Uncertainty						$^{+0.0}_{-0.2}$	
27	$\mathcal{C}'_{\mathrm{NN}}$	all	BG	μ	1.475 ± 0.025	5.7	5.7	5.6
28				σ_L	1.792 ± 0.054	5.6	5.7	5.8
29				σ_R	1.992 ± 0.032	5.7	5.7	5.7
30			G	μ	0.66 ± 0.33	5.7	5.7	5.6
31				σ	2.77 ± 0.27	5.7	5.7	5.7
32				fraction	0.106 ± 0.049	5.6	5.7	5.8
	Uncertainty						$^{+0.2}_{-0.1}$	
	Total Uncertainty						$\mathbf{^{+0.5}_{-0.3}}$	

Bibliography

[1] R. L. Workman *et al.*, "Particle Data Group," *Prog. Theor. Exp. Phys.* ***083C01***, (2022).

[2] Belle public webpage : `https://belle.kek.jp/`.

[3] S. Nishida, "Radiative B meson decays into $K\pi\gamma$ and $K\pi\pi\gamma$ final states," PhD thesis, University of Kyoto, (2003).

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

[5] Y. Ushiroda *et al.* (Belle Collaboration), "Belle silicon vertex detectors," *Nucl. Instr. and. Meth. A* ***511***, *6*, (2002).

[6] Z. Natkaniec *et al.* (Belle Collaboration), "Status of the Belle silicon vertex detector," *Nucl. Instr. and. Meth. A* ***560***, *1*, (2006).

[7] J. Brodzicka *et al.* (Belle Collaboration), "Physics achievements from the Belle experiment," *PTEP* **2012**, *04D001*, (2012).

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

[9] Y. Grossman, "Introduction to flavor physics," *arXiv:1006.3534*, (2010).

[10] P. Kooijman and N. Tuning, *Lectures on CP Violation.* PhD thesis, 2015.

[11] S. M. T. Morii, C.S. Lim, *The Physics of the Standard Model and Beyond.* World Scientific.

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

[13] R. Fleischer, “Flavor Physics and CP violation: Expecting the LHC,” *arXiv:0802.2882v1*, (2008).

[14] S. L. Glashow, J. Iliopoulos, L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry,” *Phys. Rev. D **2**, 1285*, (1970).

[15] M. Gronau, D. London, and J. L. Rosner, “Rescattering contributions to rare B-meson decays,” *Phys. Rev. D* **87**, *036008*, (2013).

[16] R. Aaij *et al.*, “Observation of the annihilation decay mode $B^0 \to K^+K^-$,” *Phys. Rev. Lett.* **118**, *081801*, (2017).

[17] Q. Chang, J. Sun, Y. Yang, and X. Li, “A combined fit on the annihilation corrections in $B_{u,d,s} \to PP$ decays within QCDF,” *Phys. Lett. B* **740**, *56*, (2015).

[18] A. S. M. Kaur, R. Dhir and R. C. Verma, “Topological diagram analysis of bottom meson decays emitting two pseudoscalar mesons,” *Phys. Part. Nuclei Lett. **12***, (2015).

[19] H. Cheng, C. Chiang, A. Kuo, “Updating $B \to PP, VP$ decays in the framework of flavor symmetry,” *Phys. Rev. D **91**, 014011*, (2015).

[20] A. Ali, G. Kramer, “Charmless non-leptonic B_s^0 decays to PP, PV and VV final states in pQCD method,” *Phys. Rev. D **76**, 074018*, (2007).

[21] M. Beneke and M. Neubert, “QCD factorization for $B \to PP$ and $B \to PV$ decays,” *Nucl. Phys. B* **675**, *333-415*, (2003).

[22] M. Acciarri *et al.*, “Search for neutral charmless B decays at LEP,” *Phys. Lett. B **363**, 127*, (1995).

[23] M. Kobayashi, T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction,” *Prog. Theor. Phys **49**, 652*, (1973).

[24] K. Abe *et al.*, “Observation of Large CP Violation in the Neutral B Meson System,” *Phys. Rev. Lett. **88**, 091802*, (2001).

[25] S. Kurokawa and E. Kikutani, “Overview of the KEKB accelerators,” *Nucl. Instr. and. Meth. A **499**, 1*, (2003), and other papers included in this volume.

[26] T. Abe *et al.*, “Achievements of KEKB,” *Prog. Theor. Exp. Phys. **2013**, 03A001*, (2003), and references therein.

[27] A. Abashian *et al.*, “The Belle detector,” *NIM A: Accelerators, Spectrometers, Detectors and Associated Equipment*, (2002).

[28] I. Nakamura, “Implementation of timing information of ECL clusters,” *Belle Internal Note*, BN 774.

[29] S. Villa, “Search for the $B^0 \to \gamma\gamma$ decay in the presence of off-time QED background,” *Belle Internal Note*, BN 811.

[30] M. Z. Wang *et al.* (Belle Collaboration), “Beam test of the BELLE extreme forward calorimeter at KEK,” *Nucl. Instr. and. Meth. A **455**, 319*, (2000).

[31] Y. Ushiroda, A. Mohapatra, H. Sakamoto *et al.*, “Developement of the central trigger system for the BELLE detector at the KEK B-factory,” *Nucl. Instrum. Meth. Phys. Res. A **438**, 460*, (1999).

[32] T. Zeigler, R. Abe, T. Abe *et al.*, “The Belle trigger system with new silicon vertex detector SVD2,” *IEEE Trans. Nucl. Sci. **51**, 1852*, (2004).

[33] K. Hanagaki, M. Hazumi, and H. Kakuno, “Belle internal note,” BN 299.

[34] GEANT Detector Description and Simulation Tool, CERN Program Library Long Writeup W5013.

[35] Martin L. Pearl, “The Total Cross-Section in e^+e^- Annihilation and the New Particles,” *Comments Nucl. Part. Phys.*, vol. 7, (1977).

[36] R. Giles *et al.* (CLEO Collaboration), “The Total Cross-section for Electron - Positron Annihilation Into Hadron Final States in the Υ Energy Region,” *Phys. Rev. D **29***, p. 1285, (1984).

[37] S. Behrends *et al.* (CLEO Collaboration), “Inclusive Hadron Production in Upsilon Decays and in Nonresonant electron-Positron Annihilation at 10.49-GeV,” *Phys. Rev. D **31***, p. 2161, (1985).

[38] G. C. Fox and S. Wolfram, “Observables for the analysis of event shapes in e^+e^- annihilation and other processes,” *Phys. Rev. Lett. **41**, 1581*, (1978).

[39] P. Franzini and J. Lee-Franzini , “Upsilon Resonances,” *Annu. Rev. Nucl. Part. Sci.*, vol. 33, (1983).

[40] A. Roodman, “Blind Analysis in Particle Physics,” *Proceedings of the PHYSTAT 2003, SLAC eConf C030908*, p. 166, (2003). `http://www.slac.stanford.edu/econf/C030908`.

[41] Documentation available at : `http://www.lns.cornell.edu/public/CLEO/soft/QQ`.

[42] T. Sjöstrand, “High-energy-physics event generation with pythia 5.7 and jetset 7.4,” *Computer Physics Communications, 82 74*, (1994).

[43] D. Lange, “The EvtGen particle decay simulation package,” *NIM A* ***462***, (2001).

[44] A. J. Bevan, B. Golob, T. Mannel, S. Prell, and B. D. Yabsley, “The Physics of the B Factories,” *The European Physical Journal* ***C***, (2014).

[45] A. Ryd, D. Lange, N. Kuznetsova, S. Versille, M. Rotondo, D. P. Kirkby, F. K. Wuerthwein, and A. Ishikawa, “EvtGen: A Monte Carlo Generator for B-Physics,” (2005).

[46] B. Casey, “Measurement of the number of $B\bar{B}$ events in experiment 5 data and the $B\bar{B}$ cross-section at KEKB,” *Belle Internal Note* , BN 296.

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

[48] K. A. *et al.* (Belle Collaboration), “Measurement of Branching Fractions for $B \to \pi\pi$, $K\pi$, and KK decays ,” *Phys. Rev. Lett.* **87***, 101801*, (2001).

[49] K. A. *et al.* (Belle Collaboration), “A measurement of the branching fraction for the inclusive $B \to X_s\gamma$ decays with the Belle detector ,” *Phys. Lett. B* **511***, 151*, (2001).

[50] U. K. M. Feindt, “The neurobayes neural network package,” *NIM A* ***559***, (2006).

[51] G. Punzi, “Sensitivity of searches for new signals and its optimization,” *arXiv:physics/0308063*, (2003).

[52] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” *arXiv:physics/0306116v1 [physics.data-an]*, (2003).

[53] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework,” *Nucl. Instrum. Meth. A* ***389****, 81*, (1997).

[54] *T. Skwarnicki (Crystal Ball Collaboration).* PhD thesis, Institute for Nuclear Physics, Krakow, 1986.

[55] H. A. *et al.* (ARGUS Collaboration), “Search for $b \to u$ decays,” *Phys. Lett. B* **241**, *278*, (1990).

[56] T. Julius *et al.* (Belle Collaboration), “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. D* ***96****, 032007*, (2017).

[57] S. Ryu, "Study of π^0 efficiency using $\tau^- \to \pi^-\pi^0\nu_\tau$," *Belle Internal Note*, BN 1224.

[58] Luca Lista, *Statistical Methods for Data Analysis in Particle Physics.* Springer, (2016).