

**Study of $B \rightarrow K^* \gamma$ Decays at Belle II and
Low-momentum Electron Identification with its
Silicon Vertex Detector**

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
Subject Board of Physics
for the degree of Doctor of Philosophy

by

Rahul Tiwary

School of Natural Sciences
Tata Institute of Fundamental Research

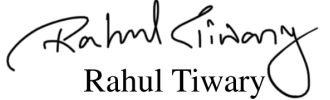
Final Version Submitted in *November, 2024*

Dedicated to my loving family and dear friends, whose never-ending support and endless encouragement have been my strength and inspiration throughout this journey.

DECLARATION

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgment of collaborative research and discussions.

The work was done under the guidance of Prof. Gagan Mohanty, at the Tata Institute of Fundamental Research, Mumbai.


Rahul Tiwary
25/10/2024

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.



Prof. Gagan Mohanty

Date: 25/10/2024

Acknowledgment

I would like to express my deepest gratitude to Prof. Gagan Mohanty, my thesis advisor, for his invaluable guidance, unwavering support, and insightful feedback throughout the research process. His expertise and dedication were crucial in shaping this thesis, and I am sincerely thankful for the enriching experience I gained under his mentorship.

I also extend my heartfelt thanks to Prof. Saurabh Sandilya, my senior collaborator, for his valuable insights and encouragement, which greatly enhanced the scope of my research. I greatly appreciate the spirit of collaboration that defined our work together.

I would also like to extend my heartfelt thanks to my seniors, Mintu Kumar and Soumen Halder. They supported me through various stages of my PhD journey, providing comfort and encouragement during the difficult moments. Soumen, in particular, helped me become a better thinker, and our shared experiences in navigating the challenges of academic life allowed me to grow more independent and find my own path.

I would like to express my gratitude to Prof. Elisa Manoni, Prof. Alexander Glazov, and Dr. Slavomira Stefkova, the Belle II EWP working group convenors, for their invaluable assistance and support throughout my analyses. I wish to convey my sincere thanks to my internal reviewers, Prof. Chunhui Chen, Prof. Mikihiko Nakao, and Prof. Jerome Baudot, for their meticulous review of my internal note and paper. I am also thankful to the publication committee members, Prof. Alan J. Schwartz and Prof. Tom Browder, for refining the paper draft and enhancing its clarity. Additionally, I wish to acknowledge the physics convenors, Prof. James F. Libby and Prof. Diego Tonelli, for their invaluable assistance in guiding me through the stages of the publication process.

Throughout my PhD journey, I had the privilege of assisting and contributing to the summer projects of student interns in various capacities. This experience broadened my perspective on a range of problems and gave me the opportunity to mentor young researchers. I would like to thank each of them – Astha, Rishabh, Sudev, and Sritam – and extend my best wishes to them in their respective PhD journeys.

I wish to acknowledge my fellow DHEP friends: Medha, Ritik, Pruthvi, Soumen, and Mintu. Our time together was marked by enjoyable evenings of poker, drinks, and movie outings, which were a great way to relax and escape the pressures of PhD life. I am deeply grateful for

the support and camaraderie they provided throughout my journey. I also want to recognize Koshvendra, Abhishek, and Ayusa as valuable additions to our group.

I would also like to thank my friends and fellow batch mates from TIFR, the “GG” group: Medha, Monideepa, Ranjan, Shekhar, Shraiyanee, Yash, and Sanjeev. Our occasional gatherings brought many joyful moments and much laughter, whether we were going to the movies or celebrating festivals together. I will always cherish these memories and the camaraderie we shared.

Throughout my PhD journey, I have gained valuable insights from discussions with senior colleagues such as Deepak, Mamta, Hariom, Suryanarayan, Jhansi, Apoorva, Ninad, and Sunil. I am grateful to Deepak for his assistance in helping me adjust to life at TIFR. His support during the early days of my PhD journey was immensely helpful. Additionally, I deeply appreciate the hospitality of Subhashis and Mansi during the challenging COVID-19 period, which provided significant comfort and support in navigating those difficulties.

My year-long involvement with the TIFR Students’ Society, alongside friends Abhishek and Ritik, provided me with valuable leadership opportunities. I am also grateful to Ruchi and Medha for their support and encouragement in organizing the bi-weekly journal club for DHEP students.

I extend my gratitude to DHEP students-Sagar, Ritik, Mukund, Rishabh, Jasmine, Avik, Lokesh, Atul, and Ashish-for contributing to a vibrant and engaging work environment. I also want to thank Arnab and Abdul for the opportunity to join the DHEP football team, where we proudly won the tournament cup without conceding a single goal!

I want to extend special gratitude to Ms. Medha Chakraborty. She was a true friend, offering invaluable support and countless joyful moments. My time at TIFR would have been incomplete without her and our wide-ranging discussions on everything from physics and good movies to politics and career paths. In my time at TIFR, I was lucky to have several friend circles, and I am proud to say that Medha was a cherished part of all of them.

I extend my heartfelt gratitude to my friends from my undergraduate days at BHU, Varanasi. Special thanks go to Shekhar, my roommate and dear friend, whose support was crucial for my academic success and for securing a PhD position at TIFR. I am profoundly grateful to Shekhar, Saurabh, Prem, Shikhar, Nikhil, Vinay, Ranjiv, Shubham, Shashank, and Subhajit for their

invaluable help and support. Additionally, I would also like to thank my friends Vivekananda, Vivek, Shubham, and Sarvesh, for the joyful memories.

I also wish to acknowledge Dr. Sunil, a PhD student at BHU, when I began my undergraduate studies. His mentorship greatly contributed to my academic growth and personal development.

My heartfelt thanks go to my school friends, Keerti-sundar and Prakash, who have always been just a call away, offering unwavering support despite their busy lives and challenges.

Finally, I am deeply grateful to my father Sri Uday Shankar Tiwary, my mother Smt. Sanju Tiwary, and my brother Avinash Tiwary (Bittu) for their unwavering support and encouragement throughout this journey. My father has been a steadfast presence, always there when I needed him. My mother has continuously believed in me, offering her boundless faith and love. My brother has been my confidant, always ready to listen and provide comfort during challenging times. I am incredibly fortunate to belong to this remarkable family, who are my constant source of inspiration and strength. Their support has been invaluable, and I aspire to make them proud and honor their belief in me.

I want to express my gratitude to everyone I encountered along this journey, who, whether intentionally or not, played a crucial role in helping me complete this work. Although it's impossible to name everyone here, each of you is intricately woven into the fabric of this thesis.

Abstract

We present measurements of the rare radiative decay $B \rightarrow K^*(892)\gamma$ using 365 fb^{-1} of data collected from 2019 to 2022 by the Belle II experiment at the SuperKEKB asymmetric-energy e^+e^- collider. Key measurements include branching fractions, CP and isospin asymmetries for $B \rightarrow K^*(892)\gamma$ decays. Additionally, we discuss the improvement in electron identification performance of the silicon-strip vertex detector of Belle II.

Publications and public documents

1. Measurement of $B \rightarrow K^*(892)\gamma$ decays at Belle II (Belle II Collaboration), I. Adachi et al., [arXiv:2411.10127 \[hep-ex\]](#).
2. Snowmass White Paper: Belle II physics reach and plans for the next decade and beyond (Belle II Collaboration), L. Aggarwal et al., [arXiv:2207.06307 \[hep-ex\]](#).
3. Measurements of the branching fractions for $B \rightarrow K^*\gamma$ decays at Belle II (Belle II Collaboration), F. Abudinén et al., [arXiv:2110.08219 \[hep-ex\]](#).
4. The Design, Construction, Operation and Performance of the Belle II Silicon Vertex Detector. K. Adamczyk et al. 2022 (Belle II SVD Collaboration), [JINST 17 P11042](#).

Belle II Internal Note

1. Measurements of observables for $B \rightarrow K^*\gamma$ decays at Belle II.
R. Tiwary, S. Halder, S. Sandilya, G. Mohanty.
BELLE2-NOTE-PH-2022-047.
2. Study of $B \rightarrow K^*\gamma$ decays at Belle II.
R. Tiwary, R. Mehta, R. Dhamija, S. Halder, S. Sandilya, G. Mohanty.
BELLE2-NOTE-PH-2021-008.
3. Low-momentum lepton identification using SVD.
R. Tiwary, G. Mohanty.
BELLE2-NOTE-TE-2020-034.

Conference Proceedings

1. Proceedings of 22nd Conference on Flavor Physics and CP Violation (FPCP 2024) on “Rare B decays at e^+e^- colliders (Belle and Belle II Collaboration)”, R. Tiwary, *EPJ Web of Conferences*, vol. 312.
2. Proceedings of 12th Workshop on the CKM Unitarity Triangle (CKM 2023) on “Radiative B decays at Belle and Belle II (Belle and Belle II Collaborations)”, R. Tiwary, [arXiv:2404.01957 \[hep-ex\]](https://arxiv.org/abs/2404.01957).
3. Proceedings of the XXV DAE-BRNS High Energy Physics Symposium (DAE-HEP 2022) on “Probing $B \rightarrow K^*\gamma$ Decays at Belle II”, R. Tiwary, S. Halder, S. Sandilya, G. Mohanty *Springer Proceedings in Physics*, vol 304. Springer, Singapore.
4. Proceedings of the XXIV DAE-BRNS High Energy Physics Symposium (DAE-HEP 2020) on “Lepton Identification using the Belle II Silicon-strip Vertex Detector”, R. Tiwary, G. Mohanty, *Springer Proceedings in Physics*, vol 277. Springer, Singapore.

Contents

Abstract	vii
1 Introduction	3
2 Theory	9
2.1 Building the theory	9
2.1.1 Symmetries	10
2.1.2 Particles or fields	11
2.2 The Standard Model Lagrangian	11
2.3 Electroweak unification	14
2.4 The CKM matrix	17
2.4.1 Unitary triangles and Jarlskog invariant	19
2.4.2 Flavor-changing neutral current	21
2.5 CP violation	22
2.6 Effective field theory	22
2.6.1 Building a generic EFT	23
2.6.2 Renormalization group and EFT	24
2.7 $B \rightarrow K^*(892)\gamma$ decays	26
2.7.1 Effective Hamiltonian	27
2.7.2 Observables	29
3 Experimental Setup	33
3.1 SuperKEKB accelerator	34
3.2 Belle II Detector	35
3.2.1 Pixel Detector	37

3.2.2	Silicon Vertex Detector	39
3.2.3	Central Drift Chamber	40
3.2.4	Time of Propagation Subdetector	42
3.2.5	Aerogel Ring Imaging Cherenkov Subdetector	44
3.2.6	Electromagnetic Calorimeter	45
3.2.7	Solenoid Magnet	46
3.2.8	KLong and Muon Subdetector	46
3.2.9	TDAQ System	48
4	Event reconstruction	51
4.1	Analysis strategy	51
4.1.1	Blind analysis methodology	51
4.2	Data and simulated samples	52
4.3	B meson reconstruction	53
4.3.1	M_{bc} and ΔE	53
4.3.2	Modification for M_{bc}	55
4.3.3	Vertex fit	56
4.3.4	Rest-of-event	57
4.3.5	Selection optimization strategy	57
4.4	K^+ and π^+ selection	58
4.5	γ selection	59
4.5.1	Photon timing selection	59
4.5.2	Selection on shower shape	60
4.5.3	Veto photons from π^0 and η decays	60
4.5.4	Impact of ECL region on ΔE Resolution	61
4.6	K_S^0 selection	63
4.7	π^0 selection	63
5	Background study	65
5.1	Continuum suppression	66
5.1.1	Input variables	66
5.1.2	Hyperparameter Optimization	69
5.1.3	Overtraining check	69

5.1.4	CSBDT performance	71
5.2	Best Candidate Selection	72
5.3	$B\bar{B}$ Background	74
5.4	Impact of Helicity variable	75
6	Measurement of observables	79
6.1	Observables	79
6.2	Fit strategy	80
6.2.1	$B \rightarrow K^*\gamma$ signal	81
6.2.2	Continuum background	83
6.2.3	$B\bar{B}$ background	83
6.3	Fit to MC events	89
7	Systematic Uncertainty	93
7.1	Double-ratio method	94
7.2	External parameters	94
7.3	B counting	95
7.4	Tracking	95
7.5	Charged hadron identification	96
7.6	Photon selection	96
7.7	K_S^0 reconstruction	96
7.8	π^0 selection	97
7.9	χ^2 selection	97
7.10	π^0/η Veto	98
7.11	Continuum suppression BDT	98
7.12	Fit bias	99
7.13	Limited MC statistics	100
7.14	Fit model	101
7.14.1	Signal PDF shape	101
7.14.2	KDE Model for $B\bar{B}$	102
7.15	Best Candidate Selection	103
7.16	Detector asymmetry	103
8	Results	107

8.1	Background Modeling	107
8.1.1	Off-resonance data	107
8.1.2	On-resonance Sideband	108
8.2	Unblinding Strategy	108
8.3	Results	109
9	Lepton identification with Belle II SVD	113
9.1	Introduction	113
9.2	Reconstruction and extraction of likelihood	114
9.3	Performance	114
9.4	Tests for muons	116
10	Summary	117
10.1	Discussion	117
10.2	Future Work	118
A	Statistical tools	121
A.1	Likelihood discriminator	121
A.1.1	Particle identification	122
A.1.2	Maximum-likelihood fit	123
A.2	Background subtraction	123
A.3	Multivariate methods	124
A.4	Variance for observables	124
A.5	Combining measurements	126
B	Variables for continuum suppression	129
C	Plots for various cross-checks	135
D	Plots for fit bias study	139

List of Figures

2.1	Elementary particles	12
2.2	$b \rightarrow s\gamma$ loop diagram	26
3.1	A schematic diagram of the SuperKEKB accelerator	35
3.2	Pictorial presentation of the difference between head-on collision scheme (top) and nano-beam scheme (bottom).	36
3.3	Beam profile of nano-beam scheme at SuperKEKB (blue) and KEKB (black) is portrayed	37
3.4	Belle II detector	38
3.5	Working principles of TOP and ARICH subdetectors are depicted in the top and bottom plots, respectively.	43
4.1	Schematic of a $B \rightarrow K^*\gamma$ event	52
4.2	A schematic representation of (a) M_{bc} and (b) ΔE for $B \rightarrow K^*\gamma$ decays. The plots are prepared using MC sample.	55
4.3	(a) $M_{bc} - \Delta E$ and (b) modified $M_{bc} - \Delta E$ correlations for correctly reconstructed $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ signal events. The plots are prepared using MC sample.	56
4.4	Distributions of (a) <code>clusterZernikeMVA</code> and (b) FOM for <code>clusterZernikeMVA</code> to select photons. The dashed blue line shows the cut applied on the variable, and the arrow shows the region of interest selected by the cut.	61
4.5	Distributions of (a) π^0 and (b) η probability for photons, and the FOM for (c) π^0 and (d) η probability. Here, the dashed blue line shows the veto criterion, and the arrow shows the region of interest retained by the veto.	62
4.6	Distributions of ΔE for (a) $B^0 \rightarrow K^{*0}\gamma$ and (b) $B^+ \rightarrow K^{*+}\gamma$ modes.	62

5.1	Pie chart illustrating the approximate relative contribution of various events originating from e^+e^- collisions at $\sqrt{s} = 10.58$ GeV.	66
5.2	Schematic of $e^+e^- \rightarrow q\bar{q}$ process with jet-like topology (left), and $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events with spherical topology (right). Credit: Markus Rörken.	67
5.3	ROC curves for $B \rightarrow K^*\gamma$ modes.	70
5.4	KS test results for $B \rightarrow K^*\gamma$ modes.	71
5.5	Plots of FOM for CSBDT of $B \rightarrow K^*\gamma$ modes.	72
5.6	Plots for candidate multiplicity of $B \rightarrow K^*\gamma$ modes obtained from simulation.	73
5.7	Distributions of ΔE and M_{bc} for $B \rightarrow K^*\gamma$ modes after applying all selection cuts.	74
5.8	Distribution of ΔE for $B\bar{B}$ background events, the number in parentheses is the number of MC events selected for the specific decay.	76
6.1	Fit to ΔE and M_{bc} distributions for correctly reconstructed $B \rightarrow K^*\gamma$ signal events from the MC sample after applying all selection criteria.	85
6.2	Fit to ΔE and M_{bc} distributions for self-crossfeed events from the MC sample after applying all selection criteria.	86
6.3	Fit to ΔE and M_{bc} distributions for continuum background events from the MC sample after applying all selection criteria.	87
6.4	Fit to ΔE and M_{bc} distributions for $B\bar{B}$ background events from the MC sample after applying all selection criteria.	88
6.5	Fit projections for $M_{bc} - \Delta E$ distributions of $B^0 \rightarrow K^{*0}\gamma$ modes. The fit is performed on an MC dataset of around 4 times the data luminosity. The fit projections in (e-f) include both B^0 and \bar{B}^0 events.	89
6.6	Fit projections for $M_{bc} - \Delta E$ distributions of $B^+ \rightarrow K^{*+}\gamma$ modes. The fit is performed on an MC sample of around 4 times the data luminosity.	90
6.7	Comparison between $sPlot$ background-subtracted distribution of $\cos \theta_{hel}$ in the MC events (black points with error bars) with that obtained from correctly reconstructed signal events (red histograms). The plots are obtained from a fit performed on an MC sample of around 4 times the data luminosity.	92
8.1	Fit projections for $M_{bc} - \Delta E$ distributions of $B^0 \rightarrow K^{*0}\gamma$ modes. The fit is performed on the full dataset, equivalent to 365 fb^{-1} . The fit results for (e-f) include both B^0 and \bar{B}^0 events	110

8.2	Fit projections for $M_{bc} - \Delta E$ distributions of $B^+ \rightarrow K^{*+}\gamma$ modes. The fit is performed on the full dataset, equivalent to 365 fb^{-1}	111
9.1	Generator-level p_T distributions for low-momentum electrons coming from (a) $B \rightarrow K^*\tau\tau$ and (b) $\tau \rightarrow eee$ decays. The dashed vertical lines indicate the minimum p_T value required to reach the ECL.	114
9.2	Electron efficiency vs. (a) $K \rightarrow e$ and (b) $\pi \rightarrow e$ misidentification probability for different criteria on the total PID likelihood.	115
9.3	SVD-based (a) simulated dE/dx vs. p distributions for leptons and hadrons and (b) separation metric s_{ij} as a function of p . The dE/dx distributions for charged particles are obtained in bins of momentum of $100 \text{ MeV}/c$ width.	116
B.1	Variable importance	129
B.2	Distribution of CSBDT input variables for $B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ modes	130
B.3	Distribution of CSBDT input variables for $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ modes	131
B.4	Distribution of CSBDT input variables for $B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ modes	132
B.5	Distribution of CSBDT input variables for $B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ modes	133
C.1	Comparison between $s\mathcal{P}$ lot background subtracted distribution of $\cos \theta$ helicity variable (blue data points with error) for data ($\int \mathcal{L} = 365 \text{ fb}^{-1}$) with that obtained from correctly reconstructed signal events (red solid line) of MC sample.	135
C.2	Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes in the off-resonance sideband. The MC histogram is scaled down to the number of events in the data histogram.	136
C.3	Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes in the on-resonance sideband. Here we blind the signal region with a ΔE selection of $\Delta E < -0.15 0.07 < \Delta E$. The MC histogram is scaled down to the number of events in the data histogram.	137
C.4	Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes. Here we blind the signal region with $M_{bc} < 5.27 \text{ GeV}/c^2$ selection. The MC histogram is scaled down to the number of events in the data histogram.	138
D.1	Pull distributions and scatter plots for \mathcal{A}_{CP} and \mathcal{B} for the fit model of $B^0 \rightarrow K^{*0}\gamma$ modes.	140

D.2 Pull distributions and scatter plots for fit parameters \mathcal{A}_{CP} and \mathcal{B} for the fit model of $B^+ \rightarrow K^{*+}\gamma$ modes.	141
---	-----

List of Tables

2.1	Summary of measurements performed for the observables of $B \rightarrow K^*\gamma$ decays. The first uncertainty is statistical, and the second is systematic.	31
4.1	Cross-sections of various physics processes at the $\Upsilon(4S)$ resonance.	53
4.2	Correlation for $M_{bc} - \Delta E$ calculated for correctly reconstructed signal events.	56
4.3	Summary of preselection criteria applied on B meson.	58
4.4	Summary of preselection criteria for charged tracks.	59
4.5	Summary of selection criteria for primary photons coming from B meson.	63
4.6	Additional selection criteria for the K_S^0 candidates	63
4.7	Selection criteria of π^0	64
5.1	Hyperparameters of the CSBDT training.	70
5.2	CSBDT cut, signal loss, and background rejection for $B \rightarrow K^*\gamma$ modes.	72
5.3	Results for best candidate selection.	73
6.1	Summary of the fit model.	84
6.2	Status of parameters of the fit model for M_{bc}	84
6.3	Status of parameters of the fit model for ΔE variable.	84
6.4	Fit results for an MC equivalent to $\mathcal{L} = 1.4 \text{ ab}^{-1}$	91
7.1	Data-MC correction factors and assorted systematics due to charged hadron identification. (Here, NA denotes not applicable.)	96
7.2	Data-MC correction factors and systematics for the MVA selection.	99
7.3	Systematics due to fit bias.	100
7.4	Systematic due to limited signal MC statistics.	101
7.5	Systematic due to fixed PDF shape parameters.	102

7.6	Systematic due to parametrization of $B\bar{B}$ background using KDE PDF.	102
7.7	Systematic due to the best candidate selection.	103
7.8	Systematics due to detector asymmetry.	104
7.9	Systematic uncertainties (%) on the branching fraction and Δ_{0+}	105
7.10	Systematic uncertainties (%) on the \mathcal{A}_{CP} measurement.	105
8.1	Comparison of statistical uncertainties (%) between data and simulation for various observables.	108
8.2	Fit results for the partial dataset of 89 fb^{-1} luminosity. The quoted uncertainties are statistical only.	109
8.3	Fit results for full dataset equivalent to 365 fb^{-1} . The first uncertainty is statistical, the second is systematic, and the third uncertainty on Δ_{0+} is due to f^{\pm}/f^{00}	109
8.4	Comparison of results with PDG world averages. The first uncertainty is statistical, the second is systematic, and the third uncertainty on Δ_{0+} is due to f^{\pm}/f^{00}	112

Chapter 1

Introduction

“It’s like trying to decide the rules of a game by being mere observers.”

– Richard P. Feynman

Since time immemorial, humanity has always been fascinated by nature’s mysteries. Thinkers have spent endless hours considering questions like how the cosmos came to be. How does it function? Who or what could have made it? What if there was no creator at all?

Various cultures have devised their own explanations for the universe’s origin, intricacies, and possible fate. Ancient Indian thinkers made significant efforts to investigate these puzzles. The hymns they created reflect their interests and observations. In Indian philosophy, the Nāsadīya Sūkta, or “Hymn of Creation”, concludes with an open-ended question:

“Whence all creation had its origin,
the creator, whether he fashioned it or whether he did not,
the creator, who surveys it all from highest heaven,
he knows, or maybe even he does not know.”

The present view is that matter is made up of some rudimentary building blocks, which get increasingly elementary as we delve deeper. The overall assumption is that if we keep looking, we will eventually reach a point where the laws of nature cannot become any more fundamental. Ordinary matter, at energies in the keV range, consists of atoms that may be classified into approximately 100 different types. For the past century, the method for probing

farther up the energy scale, or, more accurately, further down the length scale, has stayed practically unchanged. This procedure involves accelerating and colliding the test objects.

Rutherford conducted the first experiment to shed light on atomic characteristics [1]. The experimental design was novel, and the outcomes were equally intriguing. In a fixed-target experiment, he directed alpha particles towards a gold foil. The distribution of alpha particles after scattering helped him and his team members comprehend the structure of the atom, which revealed that its vast majority is empty and includes a positively charged core known as the nucleus.

Physicists today explore nature's mysteries through high-energy collider experiments employing specialized detectors. The field of physics that deals with the fundamental particles and their interactions is nowadays referred to as particle physics [2–4]. The theoretical framework used to describe particles and interactions is grounded in a quantum field theory [5–8]. Decades of collaborative efforts by physicists and engineers have culminated in a comprehensive phenomenological description of the universe at its most fundamental level within this framework. Known as the Standard Model (SM) [9–19], it stands as the most precise and extensively validated framework to understand the fundamental workings of the universe.

The SM is a theoretical framework based on a set of elementary particles and their interactions through fundamental forces. The elementary particles come in two classes, quarks [20] and leptons; they constitute the fermionic matter fields. The quarks combine to form composite objects, called hadrons [21; 22]. The hadrons are classified into two kinds, mesons which contain quark and antiquark, and baryons which contain three quarks. Experimentally, other composite objects have also been discovered that can be combinations of four (tetraquarks) or five quarks (pentaquarks).

The fundamental forces are described with a specific mathematical structure called a gauge group [3; 23–28]; the SM gauge group is $SU(3)_c \times SU(2)_L \times U(1)_Y$. The $SU(2)_L \times U(1)_Y$ part addresses electromagnetic and weak interactions, while the $SU(3)_c$ part deals with the strong interaction. This gauge group has 12 generators, identified as bosons or force carriers: 8 gluons g for the strong interaction, three weak bosons W^\pm and Z^0 that mediate the weak interactions, and the photon γ , which mediates the electromagnetic interaction. The gauge symmetries require massless gauge bosons, but in nature, the weak bosons are found to be massive. The gauge group $SU(2)_L \times U(1)_Y$ undergoes spontaneous symmetry breaking to

furnish the massive weak bosons by introducing an $SU(2)$ complex doublet scalar field Φ , known as the Higgs field. The mass terms for the fermionic matter fields are introduced through Yukawa couplings to Φ .

The SM fermions exist in a number of copies of the same gauge representation, each referred to as a generation or family. The individual fermions are then identified using quantum numbers, known as “flavor” in the literature. The unbroken gauge group $SU(2)_L \times U(1)_Y$ comprises four types of fermions, each coming in three generations: up type quarks (u, c, t), down type quarks (d, s, b), charged leptons (e, μ, τ), and neutrinos (ν_e, ν_μ, ν_τ). The term “flavor physics” [29–31] refers to interactions that identify different flavors. Gauge interactions, defined as those connected to unbroken symmetries and mediated by massless gauge bosons, do not distinguish between flavors and are hence not part of flavor physics. Thus, within the SM, flavor physics refers to the weak and Yukawa interactions.

Flavor physics is interesting, on the one hand, as a tool for discovery and, on the other hand, because of intrinsic puzzling features. Several significant, open questions in flavor physics remain, and finding answers to them is highly desirable. Below, we highlight three such questions:

- Gauge interactions in the SM make no distinction between fermions from different generations. All charged leptons carry the same electric charge, and all quarks have the same color charge. The primary distinction between generations is their mass. Our knowledge of why generations exist, why there are three, and why fermion masses and mixing angles have a distinct hierarchy is limited.
- The SM meets the Sakharov criteria for baryogenesis [32], including baryon-number violations at high temperatures, CP violations [33–35] from complex couplings in the quark sector, and non-equilibrium events during cosmic phase changes. However, it cannot fully explain the observed matter-antimatter imbalance, suggesting the need for contributions from novel CP -violating phases or mechanisms [36].
- Is there a relationship between flavor physics and physics at the TeV scale? Can flavor physics explain the emergence of electroweak symmetry breaking [37]? Furthermore, if supersymmetry [38–43] exists at a higher energy scale, can flavor physics help us understand how supersymmetry breaks?

If new physics beyond the SM (BSM) exists at or below the TeV scale, it should be observable with the ambit of flavor physics. To examine flavor- and CP -violating interactions, we require highly precise measurements with the maximum possible luminosity. Historical precedents demonstrate this: the top quark was discovered at the Tevatron, but its properties were already anticipated using electroweak measurements and studies of B mesons. Rates for $B \leftrightarrow \bar{B}$ mixing and rare flavor changing neutral current (FCNC) processes, such as $B \rightarrow X_s \gamma$, are sensitive to the top quark mass, which emphasizes this point. Research in kaon and B physics has contributed to our understanding of the top quark's flavor-changing couplings and interactions that lead to CP violation.

Exploring the flavor aspects of BSM physics is not merely an academic exercise. It is a key to answer profound questions about the nature of the universe. In the context of the Belle II experiment [44; 45], precision measurements in the quark sector are crucial for addressing these questions. Belle II's detailed studies of rare decays and B meson properties provide a fertile ground to probe BSM theories, and to perform precision tests of the SM predictions. Through meticulous experiments and data analysis, Belle II significantly contributes to the search for BSM physics, enhancing our understanding of the universe's fundamental principles and phenomena.

One way to move forward in our endeavor to uncover the mysteries of the universe is by calculating observables for interesting physics processes, which can then be measured in colliders. The subsequent measurements can help us better understand the process and verify the theoretical predictions. At times, one can get lucky and obtain an experimental result that deviates significantly from the theory's prediction, implying the existence of BSM physics that the theory does not account for.

One such interesting physics process is the rare radiative transition of B meson to the final state $K^*(892)\gamma$. This was the first FCNC transition to be observed at a collider experiment back in 1993 by CLEO [46], followed by an updated measurement [47]. Just to appreciate the result and how rare it is, only one in 100,000 B mesons decays to the $K^*(892)\gamma$ final state. Thanks to the improvements in accelerators, leading to the new generation of colliders, one can now have enough statistics for $B \rightarrow K^*(892)\gamma$ events to perform precision measurements [48; 49]. A preliminary measurement from the Belle II experiment reported the branching fraction of $B \rightarrow K^*(892)\gamma$ decays using 63 fb^{-1} of data [50]. This thesis aims to measure the observables

of $B \rightarrow K^*(892)\gamma$ decays using Run-1 data collected by Belle II.

The rest of the thesis is organized as follows: Chapter 2 introduces the motivation behind the measurement of radiative B decays within the SM framework. Chapter 3 provides an overview of the experimental setup, offering insights into the SuperKEKB accelerator and the Belle II detector. Chapters 4 and 5 detail the event reconstruction and background suppression techniques, respectively, including a step-by-step analysis and the application of multivariate classifiers. In Chapter 6, the signal extraction procedure is elucidated, employing unbinned maximum-likelihood fit. Chapter 7 delves into the intricate process of estimating systematic uncertainties. Chapter 8 offers the results and a discussion of the measurement. Chapter 9 outlines a method to improve the electron identification performance of the Silicon-strip Vertex Detector of the Belle II experiment. Finally, Chapter 10 closes the thesis with the discussion and future work.

Chapter 2

Theory

“As far as I see, all *a priori* statements in physics have their origin in symmetry.”

– Herman Weyl

The Standard Model (SM) of particle physics [9; 19] is an exceptionally successful theoretical framework that explains the fundamental particles and their interactions, except for gravity. It offers a unified description of the fundamental forces through their respective local gauge symmetries [23; 28]. In this chapter, we will take a look at the SM in a bit more detailed fashion, as well as examine some modern tools that help us calculate observables that can be measured in particle accelerators.

2.1 Building the theory

The description of SM Lagrangian in field theoretic perspective requires two ingredients:

1. The symmetries of the Lagrangian, and
2. The field content of the Lagrangian.

There are two additional requirements imposed on the theory:

1. Renormalizability: Only terms of dimension less than or equal to four in the fields and their derivatives should appear in the Lagrangian.
2. Naturalness: Naturalness entails the inclusion of all terms in the Lagrangian that are not prohibited by symmetry or renormalizability requirements.

2.1.1 Symmetries

We start by characterizing the theory in terms of its internal symmetries. In the case of the SM, the following symmetries play a vital role:

1. Symmetry of spacetime: The Minkowski spacetime can be characterized in terms of a continuous symmetry, which includes spacetime translations and Lorentz transformations. These symmetry operations form the Poincaré group.
2. Gauge symmetries: The SM incorporates a local gauge symmetry of the form: $SU(3)_C \times SU(2)_L \times U(1)_Y$.

The local gauge symmetries introduce force carrier bosons in the theory that mediate the gauge interactions. The $SU(3)_C$ part mediates strong interactions governed by quantum chromodynamics (QCD). The $SU(2)_L \times U(1)_Y$ part is spontaneously broken to obtain $U(1)_{em}$, where $U(1)_{em}$ is responsible for electromagnetic interactions.

3. Discrete symmetries: We have three discrete symmetries, namely, charge conjugation, parity transformation, and time reversal. If we take the individual symmetries, the strong and electromagnetic interactions conserve them all. However, the SM is a chiral theory because the weak interactions are maximally parity-violating. Finally, the combined operation of time reversal, charge conjugation, and parity is an exact symmetry of all the interactions (CPT theorem).
4. Accidental symmetries: The SM exhibits additional symmetries that are not postulated at the onset of its construction; rather, they appear as by-products of gauge invariance and renormalizability constraints. The SM has an accidental symmetry: $U(1)_B \times U(1)_e \times U(1)_\mu \times U(1)_\tau$.

This results in the conservation of the baryon (B) and lepton (e, μ, τ) numbers. Moreover, the SM includes a triangle anomaly in the QCD sector [51; 52], which can give rise to baryon number- and lepton number-violating processes. However, the energy scale required for such processes to be realized is far beyond the reach of current accelerators. Additionally, in these processes, the difference between baryon and lepton numbers ($B - L$) remains conserved.

2.1.2 Particles or fields

The basic building blocks of our theory are point-like objects, which we call particles. Mathematically, these particles are described using quantum field theory, the merger of quantum mechanics and special relativity. Hence, each particle is associated with an all-permeable quantized field, and each of these fields carries a set of quantum numbers.

If we delve into the arguments of symmetry, the understanding of particles stems from the seminal paper by E. Wigner [53]. Particles are identified as irreducible unitary representations of the Poincaré group that have real positive mass. They are indexed by a continuous parameter m , which we call the “mass of the particle”, and a discrete parameter s , which we call the “spin,” or in the case of massless particles, the “helicity.” Thus, we have the following two cases:

$$m^2 > 0, s = 0, \frac{1}{2}, 1, \frac{3}{2} \dots (\text{spin}), \quad (2.1)$$

$$m^2 = 0, s = 0, \pm \frac{1}{2}, \pm 1, \pm \frac{3}{2} \dots (\text{helicity}), \quad (2.2)$$

Apart from the Poincaré symmetry, we also have the internal symmetries, which are associated with the internal quantum numbers of particles. As per the Coleman-Mendula theorem [54], the Poincaré symmetry should combine with internal symmetries in a trivial way, which implies that the full symmetry group is a direct product of the Poincaré and internal symmetry groups. The elementary particles that appear in the SM and their associated quantum numbers are shown in Fig. 2.1.

2.2 The Standard Model Lagrangian

After a brief discourse over the inherent symmetries of the model and the underlying fields, we are ready to discuss the SM Lagrangian. The SM gauge group has five representations of fermions, each of which comes in three generations. Below, we list the five representations:

$$Q_{Li}(3, 2)_{+1/6}, U_{Ri}(3, 1)_{+2/3}, D_{Ri}(3, 1)_{-1/3}, L_{Li}(1, 2)_{-1/2}, \text{ and } E_{Ri}(1, 1)_{-1}, \quad (2.3)$$

Here, ‘ L/R ’ denotes left- or right-handed states, ‘ i ’ represents the three generations, and the numbers indicate their transformation properties under the gauge groups. The first numbers in parentheses tell us the dimension of the irreducible representation of $SU(3)_C$ according to which

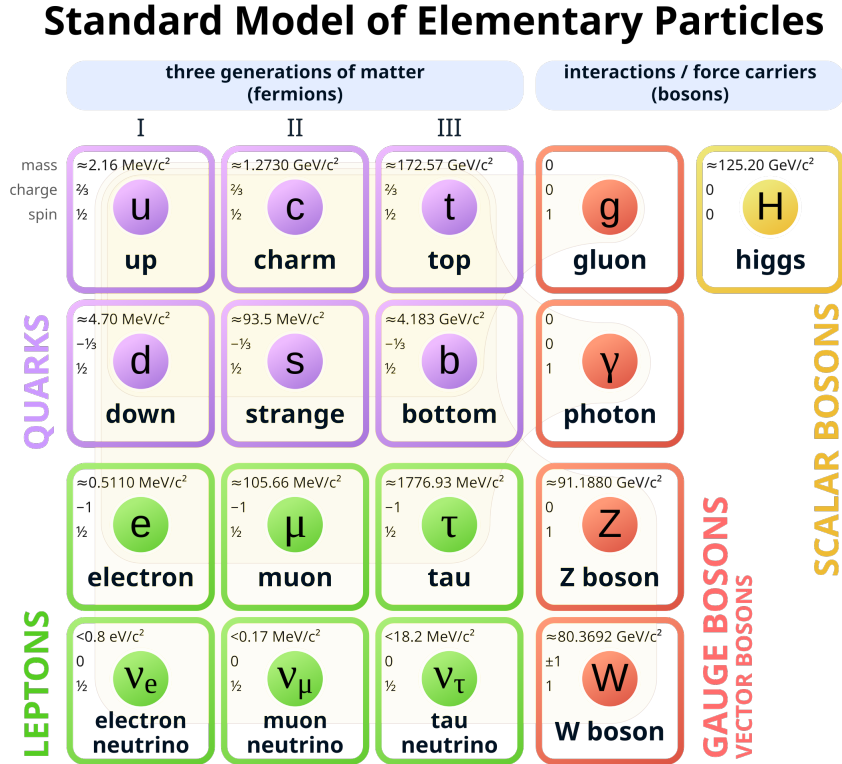


Figure 2.1: Elementary particles. Adapted from Ref. [55]

the field transforms under a gauge transformation, the second numbers similarly describe the transformation under $SU(2)_L$, and the numbers outside parentheses indicates the hypercharge Y , which are the eigenvalues of the $U(1)$ generator. Apart from the fermionic fields, there is also a scalar field known as the Higgs field, which is represented as:

$$\Phi(1, 2)_{+\frac{1}{2}}. \quad (2.4)$$

Once we have described the field content, we can write the most general renormalizable Lagrangian with these fields. The starting point is a simple Lagrangian, with just a kinematic term. The latter gives dynamics to the field and describes the propagation of a free field in spacetime. For a free field, the kinetic term is made up of fields and their derivatives. The gauge symmetry requirement is fulfilled by replacing ordinary derivatives by covariant derivatives and introducing gauge fields. If we consider the SM gauge symmetry, the covariant derivative should take the form:

$$D^\mu = \partial^\mu + ig_s G^{\mu a} L^a + ig W^{\mu b} T^b + ig' B^\mu Y,$$

Here, L^a 's are $SU(3)_C$ generators, T^b 's are $SU(2)_L$ generators, Y 's are the $U(1)_Y$ hypercharges, and g 's are the interaction strengths. As a result of the SM gauge symmetry, we get the following three-gauge boson fields:

$$G_a^\mu(8, 1)_0, W_a^\mu(1, 3)_0, B^\mu(1, 1)_0. \quad (2.5)$$

Here, $G_a^\mu(8, 1)_0$ are the octet gluon fields due to $SU(3)_C$, $W_a^\mu(1, 3)_0$ are the triplet weak boson fields coming from $SU(2)_L$, and $B^\mu(1, 1)_0$ is the singlet hypercharge boson field due to $U(1)_Y$. Adding together the kinetic terms for the fermionic fields (Q, U, D, L, E), the gauge bosons (G, W, B), and the scalar field (Φ), the kinetic term for the SM Lagrangian assumes the following form:

$$\begin{aligned} \mathcal{L}_{\text{Kin}} = & i\bar{Q}_{Li}\gamma_\mu D^\mu Q_{Li} + i\bar{U}_{Ri}\gamma_\mu D^\mu U_{Ri} + i\bar{D}_{Ri}\gamma_\mu D^\mu D_{Ri} + i\bar{L}_{Li}\gamma_\mu D^\mu L_{Li} + i\bar{E}_{Ri}\gamma_\mu D^\mu E_{Ri} \\ & + \frac{1}{4}G_a^{\mu\nu}G_{a\mu\nu} + \frac{1}{4}W_a^{\mu\nu}W_{a\mu\nu} + \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\ & + (D_\mu\Phi)^\dagger(D_\mu\Phi). \end{aligned} \quad (2.6)$$

Till now, we have only discussed the kinetic terms of the fields and introduced gauge interactions through covariant derivatives. If we only examine this part of the Lagrangian, it exhibits exact flavor symmetry and is CP conserving.

Interestingly, the SM Lagrangian is devoid of mass terms for the fermions. Since the fermions are assigned chiral representations of the gauge group, we cannot write Dirac mass terms for them. Furthermore, we cannot write Majorana mass terms for them because they all have nonzero hypercharge. We also note that the gauge bosons a priori are massless, since introducing a mass term will not render the Lagrangian gauge invariant.

The mass terms for the gauge bosons and fermions are introduced by coupling them to the Higgs field. The interaction between the Higgs field and fermions is written as the Yukawa part of the Lagrangian:

$$\mathcal{L}_{\text{Yukawa}} = Y_{ij}^d \bar{Q}_{Li} \Phi D_{Rj} + Y_{ij}^u \bar{Q}_{Li} \tilde{\Phi} U_{Rj} + Y_{ij}^l \bar{L}_{Li} \Phi E_{Rj} + \text{hermitian conjugate terms}. \quad (2.7)$$

Here, $\tilde{\Phi} = i\tau_2\Phi^\dagger$, and Y^a are three 3×3 matrices of dimensionless couplings. This part of

the Lagrangian is, in general, flavor-dependent (that is, $Y^f \neq 1$) and CP-violating. Finally, the remaining piece of the SM Lagrangian is the potential term for the Higgs field:

$$\mathcal{L}_{\text{Higgs}} = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (2.8)$$

2.3 Electroweak unification

The SM Lagrangian, as discussed in the previous section, is devoid of mass terms for fermions and bosons. If we consider the experimental results, the three gauge bosons that mediate the weak interaction are massive. The symmetry requirement forbids the mass term for the gauge bosons; hence, there should be some other mechanism to add mass to the bosons. Similarly, the SM fermions are observed to have nonzero mass; hence, it would be remarkable if we could also introduce mass terms for the SM fermions using the same mechanism. As stated earlier, the SM particles get their masses by interacting with the scalar Higgs boson. In this section, we will discuss the unification of electromagnetic and weak interactions and the mass generation mechanism.

The complex scalar field Φ admits an internal $SU(2)$ global symmetry. If we take $\lambda > 0$, which is necessitated due to the vacuum stability requirement, the scalar potential gets minimized at $|\Phi| = m/\sqrt{\lambda}$. Without loss of generality, one can then choose the vacuum expectation value (VEV) of the Higgs field to be $v = m/\sqrt{\lambda}$ and redefine the field with respect to its VEV as:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix},$$

The lowest energy state for the Higgs doublet must be degenerate, since any $SU(2)$ rotation of the state must also be a minimum energy state. Choosing a particular VEV, i.e., a particular direction in the gauge theory space, breaks the global $SU(2)$ symmetry for the Higgs doublet; this is referred to as spontaneous symmetry breaking (SSB). The SSB also breaks the SM gauge group from $SU(3)_C \times SU(2)_L \times U(1)_Y$ to $SU(3)_C \times U(1)_{\text{EM}}$ and generates mass terms for three of the gauge bosons. Accordingly, we introduce the Higgs VEV into the SM Lagrangian. Considering only the components from the $SU(2)_L \times U(1)_Y$ gauge group, the kinetic term for the Higgs field $|D_\mu \Phi|^2$ can be written as:

$$\begin{aligned}
|D_\mu\Phi|^2 &= \frac{g^2v^2}{8} \left((W_\mu^1)^2 + (W_\mu^2)^2 + \left(\frac{g'}{g}B_\mu - W_\mu^3 \right)^2 \right) \\
&= \frac{1}{2}m_W^2 W_\mu^+ W_\mu^- + \frac{1}{2}m_Z^2 Z^\mu Z_\mu.
\end{aligned}$$

Here, in order to diagonalize the mass terms, the fields have been redefined as:

$$W^\pm = W_\mu^1 \pm iW_\mu^2, \quad Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, \quad \text{and } A_\mu = \frac{g'B_\mu + gW_\mu^3}{\sqrt{g^2 + g'^2}}. \quad (2.9)$$

The SSB mechanism facilitates the W^\pm and Z_μ bosons to acquire mass $m_W = \frac{1}{2}gv$ and $m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}$, respectively, while the A_μ field remains massless. The massless field A_μ is identified with the electromagnetic field. Next, we write the $SU(2)_L$ doublets Q_{Li} and L_{Li} as:

$$Q_{Li} = \begin{pmatrix} U_{Li} \\ D_{Li} \end{pmatrix}, \quad L_{Li} = \begin{pmatrix} \nu_{Li} \\ E_{Li} \end{pmatrix}. \quad (2.10)$$

Here, U_{Li} and D_{Li} are the up- and down-type quarks, and ν_{Li} and E_{Li} are the neutrinos and charged leptons, respectively. Next, we plug in the Higgs VEV into the Yukawa term to get:

$$\mathcal{L}_{\text{Yukawa}} = (M_d)_{ij} \bar{D}_{Li} D_{Rj} + (M_u)_{ij} \bar{U}_{Li} U_{Rj} + (M_e)_{ij} \bar{E}_{Li} E_{Rj},$$

where $M_x = \frac{v}{\sqrt{2}} Y_x$, with $x \in \{U, D, E\}$. Incidentally, we do not get mass terms corresponding to ν_{Li} , by construction, to keep the neutrinos massless. There are ways to introduce mass to the neutrinos; one possibility is adding a ν_{Ri} spinor, but this is beyond the scope of our thesis. Here, the M_x are not necessarily diagonal; without loss of generality, one can do a bi-unitary transformation¹ to diagonalize these matrices:

$$V_{xL} M_x V_{xR}^\dagger = M_x^{\text{diag}}.$$

¹This is true for a general complex matrix; on the contrary, a Hermitian matrix can be diagonalized using the same transformation acting on both sides.

After doing three such bi-unitary transformations, we get:

$$M_U^{\text{diag}} = \text{diag}(m_u^2, m_c^2, m_t^2), M_D^{\text{diag}} = \text{diag}(m_d^2, m_s^2, m_b^2), M_E^{\text{diag}} = \text{diag}(m_e^2, m_\mu^2, m_\tau^2) \quad (2.11)$$

We have now obtained the masses for fundamental fermions. Next, we can transform from the flavor basis of quarks and charged leptons to their mass basis:

$$x_{Li} = (V_{xL})_{ij}x_{Lj}, \quad x_{Ri} = (V_{xR})_{ij}x_{Rj}, \quad \text{with } x \in \{U, D, E\} \quad (2.12)$$

In the mass basis, the interaction between the quarks and W^\pm bosons, widely known as the charged current interactions, takes the following form:

$$\mathcal{L}_{\text{CC}} = \frac{g}{\sqrt{2}} \left(\bar{U}_{Li} \gamma^\mu (V_{\text{CKM}})_{ij} D_{Lj} W_\mu^+ \right) + \frac{g}{\sqrt{2}} \left(\bar{\nu}_{Li} \gamma^\mu E_{Lj} W_\mu^+ \right) + \text{hermitian conjugate terms.} \quad (2.13)$$

Here, $V_{\text{CKM}} = V_{uL} V_{dL}^\dagger$ represents the well-known Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [56; 57]. Notably, this matrix is non-diagonal, causing the W^\pm bosons to couple with different quark generations. Another important observation is that weak interactions maximally violate parity due to the absence of right-handed spinors in terms of Eq. 2.13. Furthermore, charged current interactions couple each charged lepton to the corresponding neutrino, and each neutrino to its charged lepton partner. The couplings of W bosons to each of the lepton doublets $e\nu_e$, $\mu\nu_\mu$, and $\tau\nu_\tau$ are equal, i.e., universal.

The neutral current interactions mediated by the Z boson is given by:

$$-\mathcal{L}_{\text{NC}} = \sqrt{g^2 + g'^2} Z_\mu \left(\sum_{\psi \in U, D, E} \bar{\psi}_{Li} \gamma^\mu T^3 \psi_{Li} - \frac{g'^2}{g^2 + g'^2} J^{\mu\text{EM}} \right) + \text{h.c.} \quad (2.14)$$

Similar to the case of charged current interactions mediated by W bosons, the couplings of Z boson are also chiral and parity-violating. The Z boson couples diagonally, and, as a result of this, there are no Z -mediated flavor-changing neutral current (FCNC) processes at tree level. The couplings of the Z boson to different fermion generations are universal.

The electromagnetic current, mediated by the photon (A_μ), is responsible for electromagnetic interactions. Since the neutrino is electrically neutral, it does not contribute to the corresponding electromagnetic Lagrangian (\mathcal{L}_{em}). In electromagnetic interactions, both right- and

left-handed spinors contribute, leading to the conservation of parity, in contrast to the weak interaction described earlier. The electromagnetic Lagrangian is expressed as:

$$\mathcal{L}_{\text{em}} = \frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu Q_i \sum_{\psi \in \{U, D, E\}} (\bar{\psi}_{Li} \gamma^\mu \psi_{Li} + \bar{\psi}_{Ri} \gamma^\mu \psi_{Ri}) + \text{h.c.} \quad (2.15)$$

$$= e A_\mu J_\mu^{\text{EM}}, \quad (2.16)$$

where $e = Qgg' / \sqrt{g^2 + g'^2}$ represents the elementary electric charge, and J_μ^{EM} is the electromagnetic current.

Before moving forward, for completeness, we add a few lines about the strong interactions. All colored fermions, namely quarks, interact with the gluon fields via the following term:

$$\mathcal{L}_{QCD} = \frac{g_s}{2} \bar{q} \lambda_a G_a q \quad (2.17)$$

The gluon couplings, similar to the case of electromagnetic interactions, are vector-like and parity-conserving. The gluon couples to a fermion-antifermion pair of the same flavor, and the couplings are flavor-universal. The universality of the photon and gluon couplings is a consequence of the $SU(3)_C \times U(1)_{EM}$ gauge invariance, which survives SSB.

2.4 The CKM matrix

One of the key results from the previous section is that the flavor and mass basis of the quarks cannot be diagonalized simultaneously. The matrix needed to go from the flavor to mass eigenstates is the CKM matrix, denoted by V_{CKM} . Considering charged current interactions, the CKM matrix also defines the relative coupling strengths between the different combinations of up- and down-type quarks. Its elements are arranged as follows:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (2.18)$$

The CKM matrix is unitary by construction. The SM has three generations of fermions, which results in a matrix of dimensions 3×3 . A unitary $n \times n$ matrix will have n^2 parameters.

However, in the case of the CKM matrix, we may use our freedom to define the relative phases of the quark fields. In the case of n families, we have n up- and n down-type quarks, leaving us the freedom to choose $2n - 1$ relative phases. Consequently, the number of parameters is:

$$N = n^2 - 2n + 1 = (n - 1)^2 \quad (2.19)$$

Next, one can separate them into the rotation (mixing) angles (N_{angles}) and complex phases (N_{phase}):

$$N_{\text{angles}} = \frac{n(n-1)}{2}, N_{\text{phases}} = \frac{(n-1)(n-2)}{2} \quad (2.20)$$

With $n = 3$, the number of independent parameters for the CKM matrix is four, three mixing angles, and one complex phase. The phase gives rise to CP violation. Thus, the CKM matrix can be parameterized as a product of three real orthogonal rotation matrices (U_{12} , U_{13} , and U_{23}) and one unitary matrix (U_δ). Below, we write all four matrices:

$$U_{12} = \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, U_{13} = \begin{pmatrix} c_{13} & 0 & s_{13} \\ 0 & 1 & 0 \\ -s_{13} & 0 & c_{13} \end{pmatrix}, U_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{13} \end{pmatrix}, U_\delta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\delta_{13}} \end{pmatrix}. \quad (2.21)$$

In Eq. 2.21, the c_{ij} and s_{ij} denote cosines and sines of the three real rotation parameters ij , respectively, and δ_{13} is the complex phase. The standard parameterization of the CKM matrix is given by the product of the three rotation matrices, where the U_{13} matrix is transformed by the matrix U_δ :

$$V_{\text{CKM}} = U_{23}U_\delta^\dagger U_{13}U_\delta U_{12}. \quad (2.22)$$

Now that we have a parameterization of the CKM matrix, we should look at the values of the individual elements. Experimental results suggest that the CKM matrix is almost diagonal, and the off-diagonal elements follow a certain pattern in order of magnitudes:

$$V_{\text{CKM}} = \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}. \quad (2.23)$$

Keeping in mind the pattern followed by the elements of the CKM matrix, we can use a more

convenient parametrization known as the Wolfenstein parametrization. In this parametrization, one starts with four parameters: λ , A , ρ , and η , where λ is the small parameter that can be used for perturbative expansion, and all other parameters are $O(1)$. These parameters are defined relative to the standard parameterization by:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, s_{13} = A\lambda^2 = \lambda \frac{|V_{cb}|}{|V_{us}|}, \text{ and} \quad (2.24)$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = V_{ub}^* = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})}{\sqrt{1 - \lambda^2[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]}}, \quad (2.25)$$

where $\bar{\rho} + i\bar{\eta} = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$ is independent of how we choose to define the phase of the CKM matrix. Here, the ‘bars’ represent the exact values of the parameters, with $\rho + i\eta$ being their expansions about small λ . In the Wolfenstein parametrization, the CKM matrix takes the form:

$$V_{\text{CKM}} \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \quad (2.26)$$

This parametrization demonstrates that the transitions within the same generation, such as $u \rightarrow d$, $c \rightarrow s$, and $t \rightarrow b$, are favored. On the other hand, transitions between the first and second, second and third, and first and third generations are subject to CKM suppressions of order λ , λ^2 , and λ^3 , respectively. These suppressions are an artifact of mixing between the different quark generations. The CKMFitter [58] and UTFit [59] Collaborations provide up-to-date results for the CKM matrix parameters, combining recent experimental results.

2.4.1 Unitary triangles and Jarlskog invariant

The unitarity of the CKM matrix implies that the rows and columns of the matrix are orthonormal. Considering only the off-diagonal elements of $V_{\text{CKM}}V_{\text{CKM}}^\dagger$, we can have relations of the form:

$$\sum_i V_{iq}V_{iq'}^* = 0 \quad (2.27)$$

Here, if q and q' are both up-type quarks, the summation should be over down-type quarks, and vice versa, e.g., $\sum_i^{u,c,t} V_{id}V_{ib}^* = 0$. These relations are represented by triangles in the complex plane, referred to as unitary triangles. The parameters of the unitary triangles (sides and angles)

can be described in terms of CKM matrix elements. These parameters are related to physical observables of weak decays of hadrons. Hence, one of the important research areas in flavor physics is to overconstrain the parameters of the unitary triangle. Owing to the unitarity of the CKM matrix, all 12 unitary triangles have the same area, independent of the phase convention. Mathematically, this is related to the fourth-order rephasing invariants:

$$\Delta_{\alpha\rho}^{(4)} = V_{\beta\sigma} V_{\gamma\tau} V_{\beta\sigma}^* V_{\gamma\tau}^*, \text{ with } \begin{cases} \alpha, \beta, \gamma = u, c, t(\text{cyclic}), \\ \rho, \sigma, \delta = d, s, b, (\text{cyclic}) \end{cases} \quad (2.28)$$

In the CKM matrix, there is only one fourth-order rephasing invariant, Δ . The imaginary part of Δ corresponds to the area of the unitarity triangles, serving as a measure of CP violation. In the Wolfenstein parameterization, the imaginary part of Δ takes the form:

$$\text{Im } \Delta = c_{12} c_{23} c_{13}^2 s_{12} s_{23} s_{13} \sin(\delta) \approx \lambda^6 A^2 \eta \quad (2.29)$$

Finally, CP violation in the SM would also be absent if any of the up- or down-type quarks were degenerate in mass. In this case, one may perform a rotation among the two degenerate quarks, which will remove the CP -violating phase. It is, however, possible to define an invariant measure of CP violation by referring to the mass matrices defined in the previous section. It can be shown that the determinant of the commutator of the two mass matrices:

$$J = \det([\mathcal{M}_u, \mathcal{M}_d]) \quad (2.30)$$

is an invariant measure of CP violation, which is called the Jarlskog invariant [60]. Explicit evaluation reveals that:

$$J = 2\text{Im}\Delta(m_u - m_c)(m_u - m_t)(m_c - m_t)(m_d - m_s)(m_d - m_b)(m_s - m_b). \quad (2.31)$$

The Jarlskog invariant is a parameterization-independent quantity. It depends on all physical mixing angles, the complex phase, and the quark masses. Another observation is that the amount of CP violation in the SM is small, not only because the complex phase δ is small but also due to the mixing angles.

2.4.2 Flavor-changing neutral current

In the SM, the strong and electromagnetic interactions are flavor-conserving. If we look at the weak bosons, the W^\pm bosons mediate flavor-changing charged current interactions, while the neutral current interactions mediated by Z^0 bosons at tree level conserve flavor. We have three mechanisms at work that result in the suppression of FCNC processes in the SM:

1. GIM suppression: The Glashow-Iliopoulos-Maiani [61] mechanism, an artifact of unitarity in the CKM matrix, renders the FCNC processes forbidden at tree level. In order to have FCNC transitions in the SM, one needs to go for higher-order loop diagrams.
2. Loop suppression: The FCNC processes can only proceed via loop diagrams involving W^\pm bosons and quarks. As a general rule in quantum field theory, loop diagrams are typically suppressed compared to their tree-level counterparts due to the presence of additional factors of the coupling constant and the integration over internal momenta.
3. CKM suppression: These loop diagrams also involve off-diagonal elements of the CKM matrix, which are small, and hence the FCNCs are further suppressed.

As an example, let us consider the $b \rightarrow sy$ transition. At the lowest order, this process can occur through a diagram involving an intermediate up-type quark (u, c, t) and a W boson in the loop. The decay amplitude can be written as:

$$\mathcal{A}_{b \rightarrow sy} = \sum_q^{u,c,t} |V_{qb} V_{qs}^*| f(x_q) \quad (2.32)$$

Here, the function $f(x_q)$ with $x_q = m_q^2/m_W^2$ is the result of the loop integration. The amplitude contains off-diagonal elements of the CKM matrix, which are small; hence, the transition is CKM suppressed. If the masses of up-type quarks were equal ($m_u = m_c = m_t = m$), the amplitude would vanish owing to CKM unitarity.

In practice, the amplitude is finite but strongly suppressed, and it is proportional to the mass splitting of up-type quarks. This is the famous GIM mechanism, which results in further suppression of FCNC processes at tree level. Since the mass splitting between down-type quarks is much smaller, the GIM suppression for FCNC processes of up-type quarks is much more effective than for down-type quarks.

2.5 CP violation

In the previous sections, we have discussed various aspects of the SM Lagrangian in detail. In this section, we will briefly describe the theoretical basis of direct CP violation [33–35] in weak decays. Consider an initial state $|i\rangle$ that transforms to some final state $|f\rangle$. Here, $|i\rangle$ can be a B meson, and $|f\rangle$ can be an arbitrary final state. Also, let's assume that the transition can proceed through two different amplitudes, corresponding to two different Feynman diagrams with amplitudes A_1 and A_2 . The total matrix element for such a transition can then be written as:

$$\Gamma_{i \rightarrow f} = A_1 e^{i\alpha_1} e^{i\beta_1} + A_2 e^{i\alpha_2} e^{i\beta_2}. \quad (2.33)$$

Here, α_i and β_i are phases attributed to the weak and strong interaction components of the Hamiltonian. Taking into consideration the complex nature of the CKM matrix, the weak phase is taken to be CP odd. Assuming the absence of any CP violating contribution in the QCD interactions, the strong phase is taken to be CP even. If we consider the CP conjugate decay, the matrix element becomes:

$$\Gamma_{\bar{i} \rightarrow \bar{f}} = A_1 e^{-i\alpha_1} e^{i\beta_1} + A_2 e^{-i\alpha_2} e^{i\beta_2}. \quad (2.34)$$

Next, we define the direct CP asymmetry:

$$A_{CP} = \frac{\Gamma_{i \rightarrow f} - \Gamma_{\bar{i} \rightarrow \bar{f}}}{\Gamma_{i \rightarrow f} + \Gamma_{\bar{i} \rightarrow \bar{f}}} = \frac{-2A_1 A_2 \sin(\Delta\alpha) \sin(\Delta\beta)}{|A_1|^2 + |A_2|^2 + 2A_1 A_2 \cos(\Delta\alpha) \cos(\Delta\beta)}. \quad (2.35)$$

Thus, in order to have a direct CP violation, the following three conditions should be satisfied:

1. There should be two or more interfering amplitudes.
2. Nonzero weak phase difference ($\Delta\alpha \neq 0$)
3. Nonzero strong phase difference ($\Delta\beta \neq 0$)

Furthermore, the two amplitudes should have comparable sizes in order to have a sizeable CP asymmetry.

2.6 Effective field theory

The study of a physical system often focuses on the relevant degrees of freedom. Thus, in the case of particle physics, particles having a Compton wavelength less than that scale are

irrelevant. Such a particle is characterized as a heavy particle, which cannot be generated at this scale. Appelquist and Carazzone’s decoupling theorem [62] shows that such heavy degrees of freedom can be decoupled from the light degrees of freedom at the low energy. One can then write an effective field theory (EFT) [63–73] for such a scenario, where the heavy degrees of freedom have been ‘integrated out’. Thus, the EFT only contains light degrees of freedom as dynamical fields relevant at the given scale. Such a framework considerably simplifies many of our calculations. In this section, we will briefly describe the model-building exercise for an EFT framework, then take a look at the renormalization group flow for the EFT, and towards the end, we will describe the effective Hamiltonian for the $b \rightarrow sy$ transition.

2.6.1 Building a generic EFT

Building an EFT involves several key steps. Let’s assume we have a Hamiltonian (H_{full}), referred to as the full theory, and we want to obtain an EFT (H_{eff}) for the same. First, one must choose a cutoff scale $\Lambda < M$ that separates the fields into light and heavy degrees of freedom, with Λ acting as a “threshold of ignorance” beyond which the theory’s details are not considered. Next, the path integral over heavy degrees of freedom is performed, effectively integrating them out to obtain the “Wilsonian effective action.” This action becomes nonlocal on scales $\Delta x^t \approx 1/\Lambda$ due to the absence of heavy fields. Finally, the effective action is expanded in terms of nonlocal operators composed of light fields through the Wilsonian operator-product expansion (OPE), resulting in an “effective Lagrangian” for the theory.

The effective Lagrangian is composed of an infinite number of terms; each of them can be factored into local operators \mathcal{O}_i and coupling constants C_i ; the latter are known as Wilson coefficients. In principle, the effective Lagrangian contains all operators that are allowed by symmetries of the theory. Thus, a transition from state $|i\rangle$ to $|f\rangle$ can be described using the effective Hamiltonian, which takes the form:

$$H_{\text{eff}} = \sum_i C_i \mathcal{O}_i \quad (2.36)$$

Here, the values of C_i are determined by “matching” the effective theory to the full one. The matching is usually done at the high energy scale Λ , to a given order in the perturbation theory:

$$\langle f | H_{\text{full}} | i \rangle = \sum_j C_j(\Lambda) \langle f | \mathcal{O}_j | i \rangle \quad (2.37)$$

The effective theory contains an infinite number of operators, making truncation essential. If the effective Hamiltonian has a mass dimension of d , each term in the OPE should be of the same dimension d . To simplify power counting, one can factor out the appropriate powers of $1/\Lambda$ to make the Wilson coefficients dimensionless. Thus, the effective Hamiltonian can be rewritten as:

$$\langle f | H_{\text{full}} | i \rangle = \sum_k \frac{1}{\Lambda^k} \sum_j c_{j,k}(\Lambda) \langle f | \mathcal{O}_{j,k} | i \rangle |_{\Lambda} \quad (2.38)$$

Here, k represents the dimension, and we consider that multiple operators (labeled by subscript j) can contribute to a fixed dimension k . Truncating this sum by neglecting operators of mass dimension n corresponds to dropping terms of order $1/\Lambda^{(n-d)}$. The matrix elements calculated from the OPE ($\langle i | \mathcal{O}_{j,k} | f \rangle$) will get contributions only from the long-distance scales of states, and their dimension is given by the energies of states ($E_{i,f}$). This results in a series expansion in powers of $E_{i,f}/\Lambda$. For weak decays of hadrons, this series in powers of m_{hadron}/M_W converges rapidly.

2.6.2 Renormalization group and EFT

The generic method of building an EFT described in the previous section needs some adjustments before we can work with real-life scenarios. As an example, apart from the weak Hamiltonian, the radiative decay of B meson gets additional contributions from QCD effects, which are nonperturbative at low energies ($\mu \approx M_B$), rendering the matching calculations difficult. Thankfully, the QCD coupling decouples at high energies due to asymptotic freedom. Thus, one can resort to perturbative methods at the high energy cut-off scale ($\Lambda = M_W$) to determine the Wilson coefficients. More details regarding the choice of scale, and the effective Hamiltonian for radiative decay of B mesons, is described in the next section. As the next step, we need machinery to calculate the Wilson coefficients at the desired energy ($\mu = M_B$). This is accomplished by renormalization techniques [74–77]. For the OPE derived in the previous section, consider an arbitrary mass scale $\mu \leq \Lambda$ such that the contributions to matrix elements above μ are considered as short-distance pieces. The transition matrix then takes the form:

$$\langle f | H_{\text{full}} | i \rangle = \sum_k \frac{1}{\Lambda^k} \sum_j c_{j,k}(\Lambda/\mu) \langle f | \mathcal{O}_{j,k} | i \rangle \Big|_{\mu} \quad (2.39)$$

Here, the dimensionless coefficients $c_{j,k}(\Lambda/\mu)$ contain the short-distance contributions (i.e., physics above the scale μ) and may depend on the ratio Λ/μ . The matrix elements $\langle f | \mathcal{O}_{j,k} | i \rangle \Big|_{\mu}$

capture long-distance contributions from scales below μ . Changing μ shifts contributions between the coefficient and matrix elements. The renormalizability condition with d dimensional terms in the EFT ensures that no power corrections of order $1/\mu$ appear. Since μ is an arbitrary scale parameter, we require that the matrix elements be μ -independent:

$$0 = \mu \frac{d}{d\mu} \langle f | H_{\text{eff}} | \mu \rangle \quad (2.40)$$

The requirement that the matrix elements be independent of the scale μ leads to the renormalization group equations for the EFT. Changing the scale μ can reshuffle terms between the matrix elements of the local operators into the Wilson coefficients $c_{j,k}$. Next, we perform the differentiation and write the results for each power of:

$$0 = \sum_j \left[\left(\mu \frac{d}{d\mu} c_j(\Lambda/\mu) \right) \langle f | \mathcal{O}_j | i \rangle \Big|_{\mu} + c_j(\Lambda/\mu) \left(\mu \frac{d}{d\mu} \langle f | \mathcal{O}_j | i \rangle \right) \Big|_{\mu} \right] \quad (2.41)$$

To simplify the equations, we have dropped the mass dimension label k , and from here on, the operators \mathcal{O}_j have the same mass dimension. Another thing to note is that the local operators form a closed basis. Thus, for a given mass dimension, the transformed operator $\mathcal{O}'_{j,k}$ at the new scale μ' can be written as a linear combination of the basis operators:

$$\left. \mu \frac{d}{d\mu} \langle f | \mathcal{O}_j | i \rangle \right|_{\mu} = \sum_k \gamma_{jk} \mathcal{O}_k \quad (2.42)$$

where the matrix γ is called the anomalous-dimension matrix, which depends on the scale μ . Inserting this into Eq. 2.41, we obtain:

$$0 = \sum_j \left(\mu \frac{d}{d\mu} c_j(\Lambda/\mu) \right) \langle f | \mathcal{O}_j | i \rangle \Big|_{\mu} + \sum_j \sum_k c_j(\Lambda/\mu) \gamma_{jk} \mathcal{O}_k \quad (2.43)$$

The local operators \mathcal{O}_i form a basis, and they are linearly independent. This leads to further simplification, and we obtain the following result:

$$\sum_i \left[\delta_{ij} \mu \frac{d}{d\mu} + \gamma_{ij}^T \right] c_j(\mu/\Lambda) = 0 \quad (2.44)$$

Next, we consider the case where the renormalization group flow has been induced by the strong interactions. The change of QCD coupling with scale is governed by the equation:

$$\frac{d}{d\mu}\alpha_s(\mu) = \beta(\alpha_s(\mu)) \quad (2.45)$$

The $\beta(\alpha_s(\mu))$ and γ_{ij} depend on μ only through the strong coupling constant α_s . Thus, one can expand them perturbatively in powers of α_s . The dimensionless Wilson coefficients c_i have dependence on μ through both Λ/μ and α_s , due to the coupling being scale-dependent in QCD. Hence, we replace the total derivative with respect to μ by:

$$\mu \frac{d}{d\mu} = \left(\mu \frac{\partial}{\partial \mu} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right) \quad (2.46)$$

to get the renormalization group equation:

$$\sum_i \left[\delta_{ij} \left(\mu \frac{\partial}{\partial \mu} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right) + \gamma_{ij}^T(\alpha_s) \right] c_j(\mu/\Lambda, \alpha_s) = 0 \quad (2.47)$$

In an EFT, effects from scales above Λ are encapsulated into the short-distance coefficients c_i , which are obtained by matching to the full theory. The values of coefficients c_i at $\mu = \Lambda$, denoted by $c_i(\Lambda/\mu = 1, \alpha_s(\Lambda))$, serve as the initial conditions for the renormalization group flow. The EFTs thus provide an effective description of the process at energies far lower than the cutoff scale Λ . The renormalization group flow enables us to gradually shift the Wilson coefficients from the high-energy scale Λ to the low-energy scale of interest μ , ensuring that the physical observables remain scale-independent.

2.7 $B \rightarrow K^*(892)\gamma$ decays

The radiative transition of B meson to the $K^*(892)\gamma$ final state is suppressed at tree level in the SM [78; 79]. The transition proceeds dominantly through a one-loop $b \rightarrow s\gamma$ diagram, as shown in Fig. 2.2.

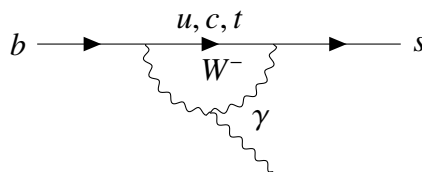


Figure 2.2: $b \rightarrow s\gamma$ loop diagram

The decay rate for $B \rightarrow K^*(892)\gamma$ is of the order of 10^{-5} , thus, it is classified as a rare decay of B meson. In this section, we will outline the theory calculations that are performed to obtain the observables of this decay. We will also shed light on the importance of these observables, the relevant SM contributions, and their importance as probes for BSM effects.

2.7.1 Effective Hamiltonian

The EFT machinery described in the earlier section can be applied to the radiative penguin process $B \rightarrow K^*(892)\gamma$. The effective Hamiltonian [80; 81] is obtained by integrating out the heavy degrees of freedom, in this case the top quark and W boson. The Hamiltonian can be expanded in terms of local operators and undetermined Wilson coefficients. The effective Hamiltonian for the $b \rightarrow s\gamma$ process then takes the form [82]:

$$\mathcal{H}_{\text{eff}}^{\text{SM}}(b \rightarrow s\gamma) = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* [C'_7(\mu)\mathcal{P}_7(\mu) + C'_8(\mu)\mathcal{P}_8(\mu)]. \quad (2.48)$$

The relevant operators for the $b \rightarrow s\gamma$ transition are the following:

$$C_7 = \frac{e}{(4\pi)^2} m_b \left(\bar{s}_{L,\alpha} \sigma_{\mu\nu} P_R b_{R,\alpha} \right) F^{\mu\nu}, \quad (2.49)$$

$$C_8 = \frac{g}{(4\pi)^2} m_b \left(\bar{s}_{L,\alpha} T_{\alpha\beta}^a \sigma_{\mu\nu} P_R b_{R,\beta} \right) G^{a\mu\nu}. \quad (2.50)$$

Here, $\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu]$, e represents the QED coupling constant, and $P_R = (1 + \gamma_5)/2$. The corresponding Wilson coefficients at M_W are obtained by matching to the full theory (SM) and calculating the relevant Feynman diagrams. In this case, the Wilson coefficients turn out to be functions of the ratio $x_q = m_q^2/M_W^2$, where m_q is the mass of the internal quark in the loop. These functions are called Inami-Lim functions [83], and they take the form:

$$C'_7(M_W) = \frac{1}{2} x_t \left[\frac{2x_t^2/3 + 5x_t/12 - 7/12}{(x_t - 1)^3} - \frac{3x_t^2/2 - x_t}{(x_t - 1)^4} \ln x_t \right] \quad (2.51)$$

$$C'_8(M_W) = \frac{1}{2} x_t \left[\frac{x_t^2/4 - 5x_t/4 - 1/2}{(x_t - 1)^3} - \frac{3x_t/2}{(x_t - 1)^4} \ln x_t \right] \quad (2.52)$$

The terms with $q \neq t$ ($x_q \ll 1$) are omitted and the dominant contribution to the $b \rightarrow s\gamma$ process comes from the top quark loop. The renormalization of C_7 and C_8 operators requires adding other operators that are part of the full $b \rightarrow s$ transition basis ($C_1 \dots C_6$). Naturally, we expect mixing between these operators, which depends on the renormalization scheme. One can rather

work with scheme-independent Wilson coefficients:

$$C_7^{\text{eff}}(\mu) = C_7(\mu) + \sum_{i=1}^{i=6} r_i C_i(\mu) \quad (2.53)$$

$$C_8^{\text{eff}}(\mu) = C_8(\mu) + \sum_{i=1}^{i=6} s_i C_i(\mu) \quad (2.54)$$

where r_i and s_i are regularization scheme dependent constants. Solving the renormalization group equations gives us the required Wilson coefficients: $C_7^{\text{eff}}(\mu = 5 \text{ GeV}) = -0.299$ and $C_8^{\text{eff}}(\mu = 5 \text{ GeV}) = -0.143$.

The calculations outlined until now involved the transitions of free quarks. In real life, we would rather work with hadrons or mesons, which are bound states of quarks. For our case, in order to obtain the observables for $B \rightarrow K^*(892)\gamma$ decays, one needs to calculate the hadronic matrix elements of the form $\langle K^*(892)\gamma | C_i | B \rangle$. These matrix elements represent the effects of strong interactions binding the hadrons.

The evaluation of the matrix elements requires another set of EFT methods. One such method is the Heavy Quark Expansion (HQE) [84–98]. The HQE approach exploits the large mass of the heavy quarks (in this case b) to systematically organize corrections to physical observables in powers of $1/m_Q$, where m_Q is the heavy quark mass. Another useful technique is the Light Cone Sum Rule (LCSR) [99–104], which utilizes the concept of light cone dominance in the theory to relate the properties of hadrons (such as form factors, decay constants, and distribution amplitudes) to the underlying quark-gluon dynamics.

Combining HQE with LCSR offers an efficient method to obtain the matrix elements. This method takes advantage of both techniques: the systematic treatment of heavy quark dynamics in HQE and the capacity to handle nonperturbative QCD effects with LCSR. HQE facilitates separating the heavy quark mass scale (M_Q) from the standard QCD scale (Λ_{QCD}). LCSR provides a framework to calculate hadronic form factors, which are essential for understanding the decay amplitudes of processes like $B \rightarrow K^*\gamma$. LCSR uses OPE near the light cone and light-cone distribution amplitudes (LCDAs) to incorporate the nonperturbative QCD effects. In the heavy quark limit, the matrix element can be factorized as [105]:

$$\langle K^*(\epsilon)\gamma | C_i | \bar{B} \rangle = [F^{B \rightarrow K^*}(0) T_i^I + \int_0^1 d\xi d\nu T_i^{II}(\xi, \nu) \phi_B(\xi) \phi_{K^*}(\nu)] \cdot \epsilon, \quad (2.55)$$

where ϵ is the photon polarization 4-vector, $F^{B \rightarrow K^*}(0)$ is the transition form factor, and ϕ_B and ϕ_{K^*} are the light-cone distribution amplitudes of the B and K^* meson, respectively. The form factor $F^{B \rightarrow K^*}$ and the light-cone distribution amplitudes ϕ_B and ϕ_{K^*} are nonperturbative in nature. On the other hand, the hard-scattering kernels T_i^I and T_i^{II} are the perturbative, short-distance interactions, which get factored from the long-distance dynamics of matrix elements.

The transition form factor $F^{B \rightarrow K^*}(0)$ is crucial to determine the decay rate. In the LCSR approach, this form factor is expressed as a sum rule:

$$F^{B \rightarrow K^*}(0) = \int_0^1 du \phi_{K^*}(u) \int_0^1 d\xi \phi_B(\xi) T_H(u, \xi), \quad (2.56)$$

where T_H is the hard-scattering kernel calculable in perturbative QCD. Thus, HQEFT provides a framework to factorize the hard and soft contributions. The soft contributions are encapsulated in the LCDAs ϕ_B and ϕ_{K^*} , and the hard contributions are calculated using perturbative QCD. This factorization is particularly important for handling contributions from the heavy quark and simplifying the calculation of form factors.

2.7.2 Observables

Being a rare FCNC transition that proceeds through a loop diagram, the $B \rightarrow K^*\gamma$ process can get contributions from particles that are postulated in various BSM scenarios [106; 107]. This can potentially alter the observables from their SM predictions. Hence, this radiative decay serves as a promising probe for BSM effects. The transition provides a rich set of observables that can be measured to perform precision tests of SM, as well as searches for BSM effects. In this section, we will describe the various observables related to this transition.

One of the observables that is almost always measured for a rare transition is the branching fraction, defined as the fraction of B mesons that decay to a specific final state. The rare decays tend to have small branching fraction, making it difficult to perform the measurement in the collider environments with enough precision. In the case of $B \rightarrow K^*(892)\gamma$ decay, the SM prediction for branching fraction has large uncertainties of the order of 20-30% due to form factors [108; 109]. From the experimental side, the precision for the branching fraction measurements is around 4%, which has already surpassed the SM precision. The substantial theoretical uncertainty on this observable renders it unsuitable for precision tests of SM, and a probe for BSM effects.

Another commonly discussed observable for B decays is the CP violating asymmetry. In the case of $B \rightarrow K^*(892)\gamma$ decays, it is defined as:

$$\mathcal{A}_{CP} \equiv \frac{\Gamma(\bar{B} \rightarrow \bar{K}^* \gamma) - \Gamma(B \rightarrow K^* \gamma)}{\Gamma(\bar{B} \rightarrow \bar{K}^* \gamma) + \Gamma(B \rightarrow K^* \gamma)} \quad (2.57)$$

In contrast to the branching fraction, \mathcal{A}_{CP} has reduced theoretical uncertainties due to the cancellation of form factor contributions [110; 111]. In the SM, the value of \mathcal{A}_{CP} is of the order of 1% [112; 113], with an uncertainty of sub-percent level. The clean prediction from SM makes \mathcal{A}_{CP} an excellent probe for BSM effects [114]. As per the description provided in Ref. [112], at the order of α_s^0 (i.e, the lowest order), only the top quark-mediated diagram contributes to the amplitude of $B \rightarrow K^* \gamma$ decay. As discussed earlier, one of the requirements of nonzero \mathcal{A}_{CP} is the presence of two interfering amplitudes that have comparable magnitude. In order to obtain a nonzero asymmetry, one needs to consider higher order diagrams in α_s . There are two classes of such diagrams, depending on whether the gluon from b meson probes the spectator quark or the s quark.

Similar to the CP violating asymmetry, we can define another observable, the isospin asymmetry:

$$\Delta_{0+} \equiv \frac{\Gamma(B^0 \rightarrow K^{*0} \gamma) - \Gamma(B^+ \rightarrow K^{*+} \gamma)}{\Gamma(B^0 \rightarrow K^{*0} \gamma) + \Gamma(B^+ \rightarrow K^{*+} \gamma)}, \quad (2.58)$$

where Γ denotes the partial width and K^* denotes a $K^*(892)$ meson. Mirroring \mathcal{A}_{CP} , the SM prediction of isospin asymmetry is also clean, thanks to the cancellation of form factor uncertainties in the ratio. The isospin asymmetry can arise primarily due to long-distance effects such as the emission of photon from the spectator quark that can probe the charge difference between the spectator quarks, and due to the matrix elements involving isotriplet combinations of hadronic operators in the effective Hamiltonian. The dominant SM contribution to isospin asymmetry of $B \rightarrow K^*(892)\gamma$ is a subleading effect in the HQE and depends on the Wilson coefficients $C_1 - C_6$. In the SM, the estimated value of Δ_{0+} ranges from approximately 2.7% [111] to 8.0% [115], with a typical uncertainty of about 2-3%. However, BSM effects have the potential to enhance Δ_{0+} [114].

One can also measure the difference in CP violation asymmetry between the charged and

neutral $B \rightarrow K^*(892)\gamma$ decays, given as:

$$\Delta\mathcal{A}_{CP} = \mathcal{A}_{CP}(B^+ \rightarrow K^{*+}\gamma) - \mathcal{A}_{CP}(B^0 \rightarrow K^{*0}\gamma) \quad (2.59)$$

Theoretically, an estimate of $\Delta\mathcal{A}_{CP}$ has been discussed in the context of the inclusive $B \rightarrow X_s\gamma$ process [116]. It is theorized that the resolved photon contribution can introduce a flavor-dependent term to the CP violation asymmetry, and contribution from BSM effects can result in a $\Delta\mathcal{A}_{CP}$ value as large as 10%. So far, such an observable has not been discussed in the context of the exclusive $B \rightarrow K^*(892)\gamma$ channel. However, following the arguments of Ref. [116], we believe that a measurement of $\Delta\mathcal{A}_{CP}$ with the $B \rightarrow K^*(892)\gamma$ channel can be a relevant alternative, with potential to uncover BSM effects.

In the end, we would like to provide a summary of the measurements performed by various experiments for the $B \rightarrow K^*(892)\gamma$ decay, as listed in Table. 2.1. We report results from BaBar [48] and Belle [49], as well as the world average provided by the Particle Data Group (PDG) [117]. Apart from the branching fractions, CP and isospin asymmetries, Belle [49] also report the measurement of $\Delta\mathcal{A}_{CP} = (2.4 \pm 2.8 \pm 0.5)\%$.

Table 2.1: Summary of measurements performed for the observables of $B \rightarrow K^*\gamma$ decays. The first uncertainty is statistical, and the second is systematic.

Observable	BaBar	Belle	PDG
$\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) (\times 10^5)$	$4.47 \pm 0.10 \pm 0.16$	$3.96 \pm 0.07 \pm 0.14$	4.18 ± 0.25
$\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) (\times 10^5)$	$4.22 \pm 0.14 \pm 0.16$	$3.76 \pm 0.10 \pm 0.12$	3.92 ± 0.22
$\mathcal{A}_{CP}(B^0 \rightarrow K^{*0}\gamma) (\%)$	$-1.6 \pm 2.2 \pm 0.7$	$-1.3 \pm 1.7 \pm 0.4$	-0.6 ± 1.1
$\mathcal{A}_{CP}(B^+ \rightarrow K^{*+}\gamma) (\%)$	$1.8 \pm 2.8 \pm 0.7$	$1.1 \pm 2.3 \pm 0.3$	1.4 ± 1.8
Δ_{0+}	$6.6 \pm 2.1 \pm 2.2$	$6.2 \pm 1.5 \pm 1.3$	6.3 ± 1.7

Chapter 3

Experimental Setup

“A physicist is just an atom’s way of looking at itself.”

– Neils Bohr

We have covered the theoretical basis of particle physics; now we will describe the experimental setup to study B meson decays. Two key components are required: a source to produce B mesons and a detector to detect their decay products. First, we must consider the physical process used to produce B mesons. One effective method is to utilize an e^+e^- collider operating at the $\Upsilon(4S)$ resonance, which is a $b\bar{b}$ bound state that decays almost exclusively into B -meson pairs. The e^+e^- collider offers several advantages for this type of study:

1. **Clean Environment:** The low-background environment of lepton colliders allows for an efficient reconstruction of hadrons with photons in the final state, such as π^0 , ρ^\pm , and η .
2. **Low Track Multiplicities:** Lepton colliders profit from low track and cluster multiplicities and detector occupancy, resulting in high reconstruction efficiency for B , D mesons, and τ leptons with minimal trigger bias.
3. **Precise Luminosity Measurement:** Luminosity can be precisely measured with the Bhabha scattering process, enabling the determination of absolute branching fractions.
4. **Flavor Tagging:** The B mesons originating from $\Upsilon(4S)$ decay are produced in a quantum entangled state. Furthermore, the fraction of events with multiple e^+e^- interactions (pile-up) is negligible. Thus, if one of the B decays has been fully reconstructed (signal-side

B), the remaining tracks and clusters in the event can be attributed to the other B meson in the event (tag-side B). By employing inclusive techniques that make use of different flavor-specific signatures of B decays, it is possible to efficiently identify (tag) the flavor of the tag-side B meson.

5. **Lorentz Boost:** Asymmetric beam energies provide a substantial Lorentz boost to the e^+e^- system, causing the B or D mesons to travel an appreciable distance before decaying. This enables precision measurements of lifetimes, mixing parameters, and time-dependent CP violations.

Thanks to these advantages, the e^+e^- collider operating at the $\Upsilon(4S)$ resonance is an ideal setup for studying B meson decays. Next, we will briefly describe both the accelerator and the detector.

3.1 SuperKEKB accelerator

SuperKEKB [118] is an energy-asymmetric e^+e^- collider located in Tsukuba, Japan. The electron and positron beams have energies of 7 GeV and 4 GeV, respectively, with a center-of-mass energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance. A schematic diagram of the accelerator is presented in Fig. 3.1. The SuperKEKB collider is an upgraded version of the KEKB collider [119]. The target peak luminosity for SuperKEKB is 30 times that of KEKB, which requires significant upgrades to the accelerator complex. The key improvements include doubling of the beam currents and an over 15-fold reduction in the beam size at the collision location; the latter is accomplished via the nano-beam technique. The basic idea behind this scheme is to squeeze the vertical beta function (β_y^*) at the interaction point (IP) by minimizing the longitudinal size of the overlap region of the two beams. This overlap region typically limits the effective minimum value of β_y^* due to the ‘‘Hourglass effect’’ [120]. The size of the overlap region d , considered the effective bunch length for the nano-beam scheme, is much smaller than the actual bunch length (σ_z). The length is $d = \sigma_x^*/\phi$, where σ_x^* is the horizontal beam size and ϕ is the horizontal half crossing angle at the IP. Compared to KEKB, σ_x is reduced by a factor of 15 ($150\ \mu\text{m} \rightarrow 10\ \mu\text{m}$), and the two beams collide at an even larger angle of $\phi = 83$ mrad ($\phi = 22$ mrad in KEKB), as illustrated in Fig. 3.3. Unlike KEKB, which collided 8 GeV electrons with 3.5 GeV positrons, SuperKEKB uses a lower beam energy asymmetry to reduce beam losses due to Touschek scattering [121; 122]. While this decreases the spatial separation between the two B -mesons for time-dependent CP violation studies, it slightly improves the

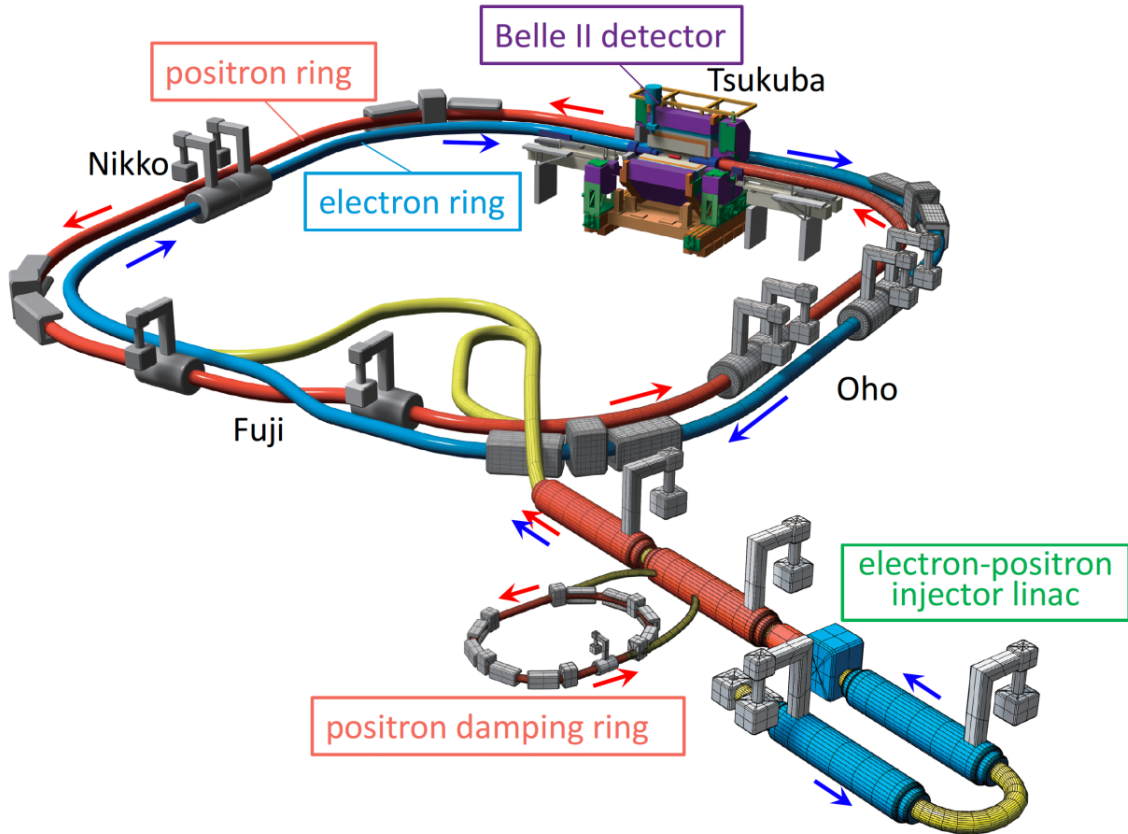


Figure 3.1: A schematic diagram of the SuperKEKB accelerator. Adapted from Ref. [118].

solid angle acceptance for missing energy decays.

The modifications to the accelerator complex include a new electron injection gun, a new target for positron production, and a new additional damping ring for the positron beam. The accelerator upgrade also includes a redesign of the lattices of the low- and high-energy rings, replacing short dipoles with longer ones (in the low-energy ring), installing TiN-coated beam pipes with ante-chambers, modifications to the RF system, and a completely redesigned interaction region.

3.2 Belle II Detector

Belle II [44; 45] is a hermetic detector located at the collision point of the SuperKEKB accelerator. It is a state-of-the-art particle physics experiment designed to study the e^+e^- collision events with high precision. The symmetry axis of the detector, defined as the z -axis, bisects the electron and positron beam axes. The x -axis is horizontal and points towards the center of the SuperKEKB storage ring, and the y -axis is vertical, pointing upwards. The polar and azimuthal angles are defined with respect to the z - and the x -axis, respectively. The detector consists of

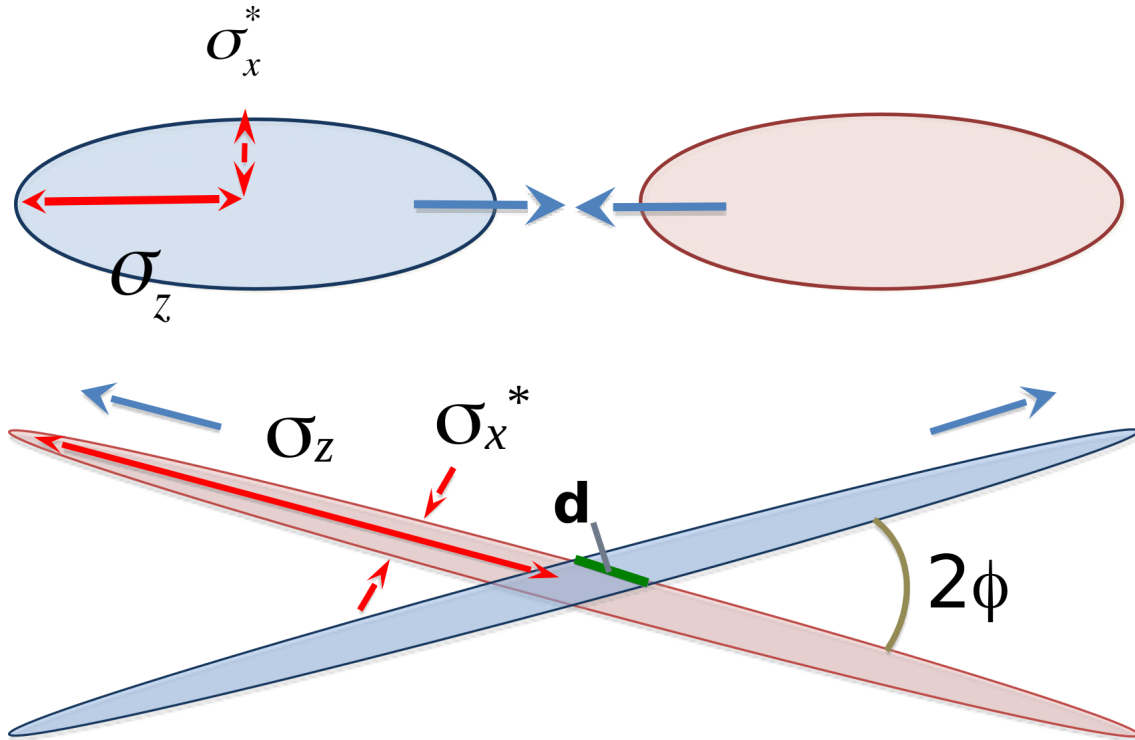


Figure 3.2: Pictorial presentation of the difference between head-on collision scheme (top) and nano-beam scheme (bottom).

several specialized subdetectors that work together to identify and measure particles produced in the collisions:

- Vertex Detector (PXD and SVD): These subdetectors provide precise position measurements near the IP, which is crucial for identifying the origin of particles.
- Central Drift Chamber (CDC): The CDC measures the momenta of charged particles as they curve in the magnetic field, allowing for accurate momentum reconstruction.
- Time-of-Propagation (TOP) subdetector: The TOP subdetector uses Cherenkov radiation in quartz to identify charged particles, particularly to distinguish between pions, kaons, and protons.
- Aerogel Ring Imaging Cherenkov (ARICH) subdetector: The ARICH subdetector identifies particles by their Cherenkov radiation in aerogel, which is crucial to distinguish between pions and kaons over a wide momentum range.
- Electromagnetic Calorimeter (ECL): The ECL measures the energy and angle of photons and electrons, which is essential for identifying and reconstructing particle energies.

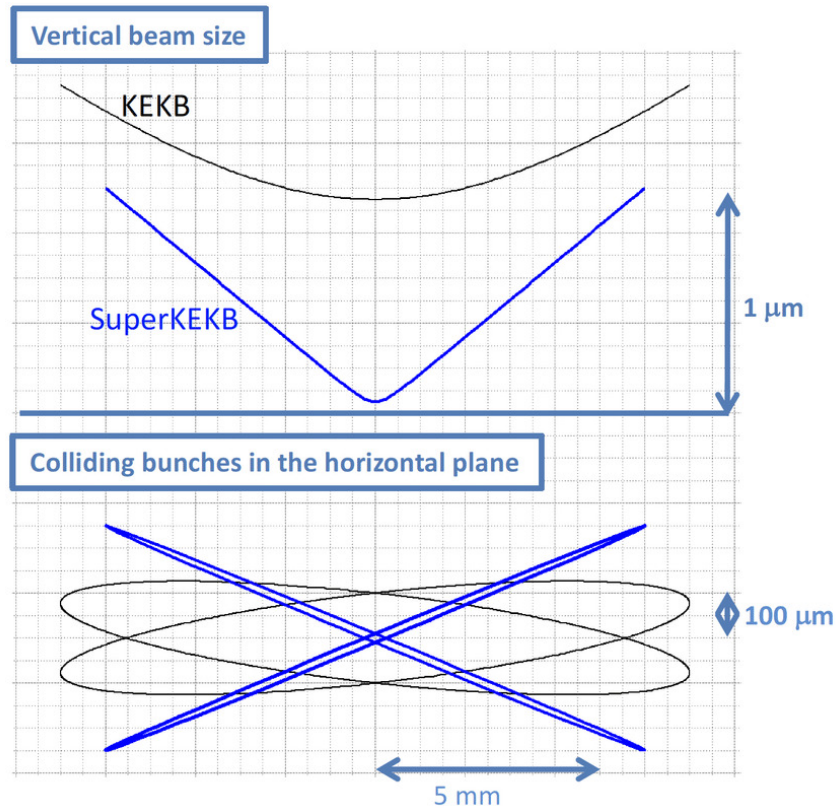


Figure 3.3: Beam profile of nano-beam scheme at SuperKEKB (blue) and KEKB (black) is portrayed. Adapted from Ref.[118].

- KLong and Muon (KLM) subdetector: The KLM subdetector detects muons and neutral kaons by their penetration through layers of iron and detection in resistive plate chambers (RPCs) and scintillator strips.
- Trigger and Data Acquisition (TDAQ) System: The TDAQ system is a critical component responsible for selecting and recording the relevant collision events.

Each subdetector is specialized in its function, working synergistically to provide comprehensive particle identification, momentum measurement, and energy reconstruction capabilities. The Belle II detector's design and performance are optimized for high precision and efficiency, enabling it to perform precision tests of SM as well as to look for new physics effects. A schematic diagram of Belle II is shown in Fig. 3.4. Next, we will describe the various subdetectors in more detail.

3.2.1 Pixel Detector

The pixel detector is part of the Belle II vertex detector (VXD), situated closest to the IP where e^+e^- collisions occur. The VXD consists of two sub-detectors: the Pixel Detector

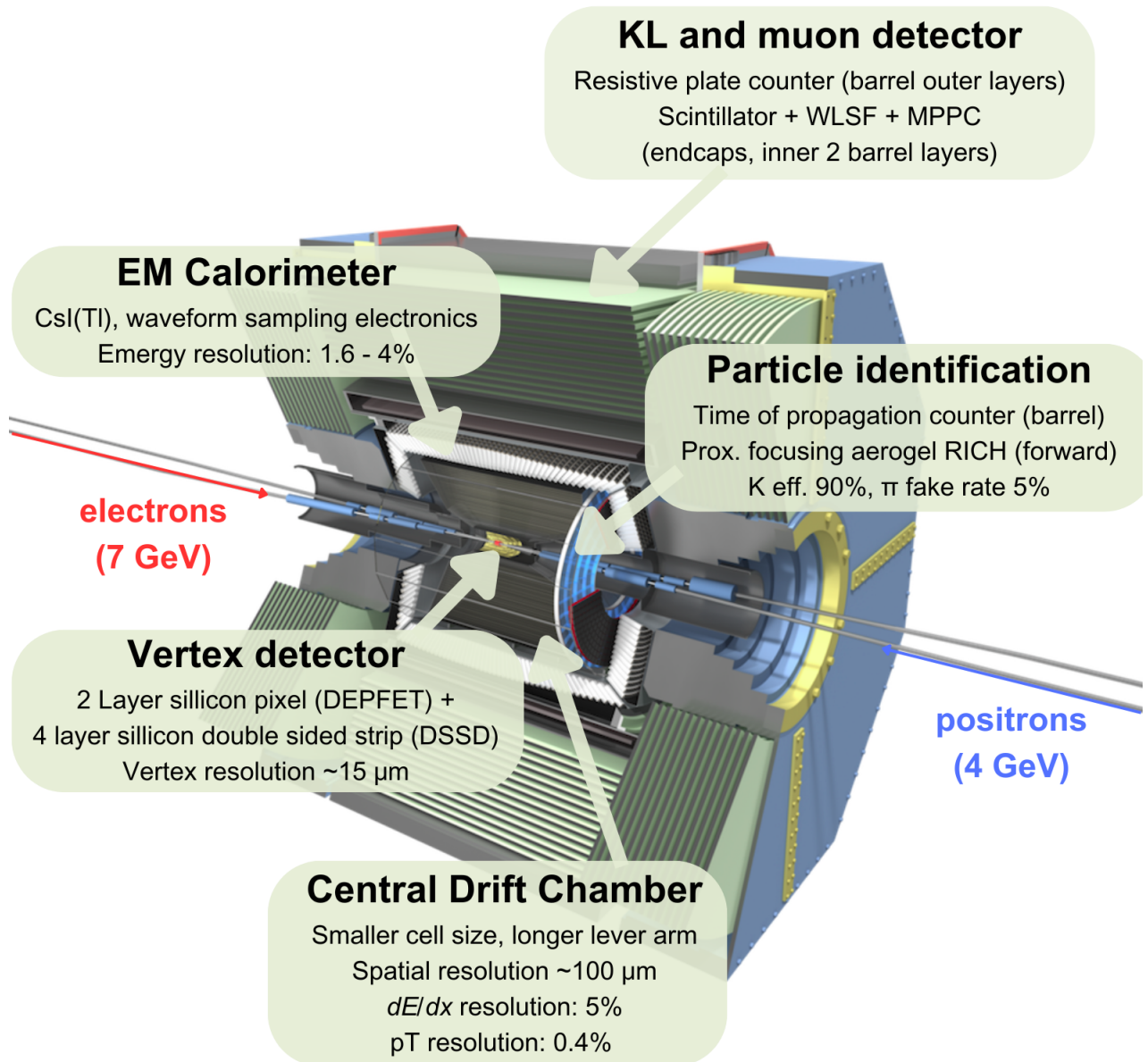


Figure 3.4: Belle II detector

(PXD) and the Silicon-strip Vertex Detector (SVD). The PXD is the innermost layer and provides high-precision measurements necessary for accurate vertex reconstruction. It employs DEPFET (DEpleted P-channel Field Effect Transistor) technology [123], chosen for its excellent signal-to-noise ratio and precise spatial resolution. Each DEPFET pixel consists of a MOSFET (Metal-Oxide Semiconductor Field-Effect Transistor) embedded in a fully depleted silicon substrate.

The PXD utilizes pixel sensors with a typical size of $50 \times 75 \mu\text{m}^2$, carefully chosen to strike

a balance between spatial resolution and readout speed. The sensors are organized in matrices, each containing thousands of pixels, facilitating precise tracking of particles. The DEPFET structure integrated within the sensors enables in-pixel amplification, which reduces noise and enhances the accuracy of charge extraction from individual pixels. The PXD consists of two layers of these advanced pixel sensors: Layer 1, situated at a radius of approximately 14 mm from the IP, and Layer 2, positioned at around 22 mm from the IP.

The sensors are mounted on lightweight, low-mass mechanical supports to minimize material budget and reduce multiple scattering effects, crucial for maintaining high tracking accuracy. The sensors are organized into rectangular modules, referred to as "ladders." The inner layer consists of 8 ladders, while the outer layer contains 12 ladders. Only one-sixth of the second pixel layer was installed for the data analyzed in this thesis.

The PXD boasts impressive performance specifications that are critical for its role in precise particle tracking. It offers a spatial resolution of approximately $14(15)\mu\text{m}$ in the transverse (longitudinal) plane [124], which is essential for accurate vertex determination. Additionally, the timing resolution of the PXD is engineered to be on the order of tens of nanoseconds, enabling efficient separation of events that occur in close succession.

3.2.2 Silicon Vertex Detector

The SVD is a crucial component of the Belle II detector, working in conjunction with the PXD to provide high-precision tracking and vertex reconstruction capabilities essential for studying final-state particles. The SVD comprises several cylindrically arranged DSSD (double-sided silicon strip detector) sensors placed around the IP. It consists of four layers arranged concentrically around the IP to enhance tracking precision. The innermost layer, Layer 3, is positioned just outside the PXD at a radius of approximately 38 mm from the IP. Layer 4 is situated at a radius of about 80 mm, while Layer 5 is placed further out at around 115 mm. The outermost layer, Layer 6, is located at a radius of about 140 mm. These layers collectively ensure comprehensive tracking coverage and contribute to the precise reconstruction of particle trajectories. Each layer consists of several ladder structures that support the DSSD sensors, providing mechanical stability while minimizing material budget to reduce multiple scattering and improve tracking accuracy.

The SVD utilizes DSSDs to precisely measure the positions of passing charged particles in two dimensions, which is crucial for accurate particle trajectory reconstruction. Each DSSD

sensor is designed with either a rectangular or trapezoidal shape and has a thickness ranging from 300 to 320 μm . The sensors feature an n-type bulk with distinct doping on each side: one side is doped with acceptors (p- or U-side) and the other with donors (n- or V-side). The longer strips on the p-side measure the r - ϕ coordinate, while the shorter strips on the n-side measure the z coordinate. The readout pitch varies across layers: in Layer 3, it is 50 μm in the r - ϕ plane and 160 μm along the z axis, whereas in Layers 4, 5, and 6, it is 75 μm in the r - ϕ plane and 240 μm along the z axis. This configuration ensures high precision in tracking and particle identification.

In an SVD sensor, when a charged particle traverses the reverse-biased silicon sensor material, it imparts energy, generating electron-hole pairs. Due to the applied electric field, the electrons move towards the p-side of the sensor, while the positively charged holes move towards the n-side. This differential motion of charges creates a current signal, which is detected by the p- and n-side strips on the sensor.

By analyzing the current signal from these orthogonal p- and n-side strips, the SVD can precisely determine the two-dimensional coordinates of the particle's trajectory on a specific layer of the detector. This information, when combined with data from multiple layers, allows the reconstruction of the particle's three-dimensional path. These reconstructed trajectories are essential for particle tracking and the identification of decay vertices.

The SVD is designed to offer exceptional performance specifications crucial for precise particle tracking. It provides a spatial resolution of approximately 10–20 μm in both the radial and azimuthal directions, enabling highly accurate vertex reconstruction, which is essential for understanding particle interactions and decays. Additionally, the silicon sensors and readout electronics are engineered to be highly radiation-hard, allowing them to endure the intense radiation environment of Belle II. This robust design ensures that the SVD maintains its performance and reliability over the long term, contributing significantly to the overall success of the experiment.

3.2.3 Central Drift Chamber

The CDC forms an integral part of the Belle II tracking system, working in conjunction with the PXD and SVD subdetectors. Its primary function is to detect and measure the trajectories of charged particles by tracking their ionization trails in a gas-filled volume. Beyond tracking, the CDC also aids in particle identification for particles with momenta less than 1 GeV/c, leverag-

ing their ionization loss as a distinguishing feature. The CDC is designed as a large cylindrical subdetector situated between the VXD and the outer subdetectors, such as the ECL and the KLM. The CDC extends approximately 2.3 m along the beam axis and has an outer radius of about 1.1 m. It contains 56 layers of wire detectors arranged in a hexagonal pattern. These layers are strategically positioned radially around the IP, spanning from 0.168 to 1.111 m.

Within the CDC, two types of wires play pivotal roles. Axial wires run parallel to the beam axis, providing information in the $r - \phi$ plane, perpendicular to the beam line. On the other hand, stereo wires are slightly tilted relative to the beam axis to gather information along the z axis. This dual-wire setup enables the CDC to capture precise three-dimensional positions as charged particles traverse the chamber. The layers of axial and stereo wires are grouped into superlayers. The first superlayer, primarily axial, encompasses eight layers of wires, while the subsequent superlayers consist of six layers of wires each.

Thin gold-plated tungsten sense wires, approximately $30\ \mu\text{m}$ in diameter, are employed to detect ionization signals generated by passing particles. These wires are connected to the readout electronics, enabling precise signal capture. Complementing the sense wires, thicker aluminum field wires, around $120\ \mu\text{m}$ in diameter, are used to shape the electric field within the chamber. This field configuration is essential for guiding the ionized electrons toward the sense wires.

The CDC is filled with a specific gas mixture optimized for ionization and drift properties. In particular, the detection volume is filled with a $\text{C}_2\text{H}_6 + \text{He}$ gas mixture, with small amounts of isobutane added to improve quenching properties. This mixture provides a good balance between drift velocity, spatial resolution, and minimal multiple scattering.

As charged particles traverse the gas medium, they ionize it, generating electron-ion pairs. Guided by the electric field within the CDC volume, the drifting electrons are directed towards the wire detectors. Here, their timing and position are meticulously recorded. This wealth of information allows for the reconstruction of particle trajectories and the precise determination of their momenta in the context of high-energy collision events.

The CDC excels at providing high-performance tracking capabilities essential for precise particle measurements. It offers very good spatial resolution, typically around $100\text{--}150\ \mu\text{m}$ in the transverse plane, which is critical for accurate tracking and momentum measurement.

The momentum resolution achieved by the CDC is $\Delta p/p \approx 0.1\%$ for high-momentum tracks, thanks to the precise measurement of the curvature of particle trajectories in the magnetic field. Additionally, the CDC plays a key role in particle identification by measuring the ionization energy loss (dE/dx) along the particle's path. This capability allows for a distinction between various types of charged particles, such as pions, kaons, and protons, enhancing the overall analysis of particle interactions.

3.2.4 Time of Propagation Subdetector

The TOP subdetector, located in the barrel section, and the ARICH subdetector in the end-cap region play a crucial role in particle identification, particularly for distinguishing between kaons and pions. Particle identification aids in determining the particle velocity, which, when combined with the track momentum measured by the tracking system, allows for a precise particle identification.

The principle underlying particle identification in these subdetectors is Cherenkov radiation. When charged particles travel through a clear medium like aerogel or quartz at a velocity greater than the speed of light in that medium, they emit Cherenkov radiation. The angle of emission (θ_C) is dependent on the particle's velocity (β) and the refractive index (n) of the medium, as described by the equation:

$$\beta = \frac{1}{n \cos \theta_C}. \quad (3.1)$$

The TOP subdetector is primarily constructed using an array of fused silica (quartz) bars, typically measuring about 270 cm in length, 4.5 cm in width, and 2.0 cm in thickness. The Cherenkov light emitted within these bars undergoes total internal reflection and is retained within the material. On the forward end of the subdetector, spherical mirrors are used to focus and direct the emitted Cherenkov light. On the opposite end, Micro-Channel Plate Photomultipliers (MCP-PMTs) are mounted to measure the timing of photons. These photodetectors were selected for their exceptional time resolution and their capability to function effectively within the magnetic field present in the Belle II experiment.

As charged particles pass through the quartz bars, they emit Cherenkov photons that propagate toward the array of MCP-PMTs, as shown in Fig. 3.5. By measuring the arrival times and positions of these photons at the MCP-PMTs, the subdetector accurately determines the

emission angle, providing critical information for particle identification at Belle II.

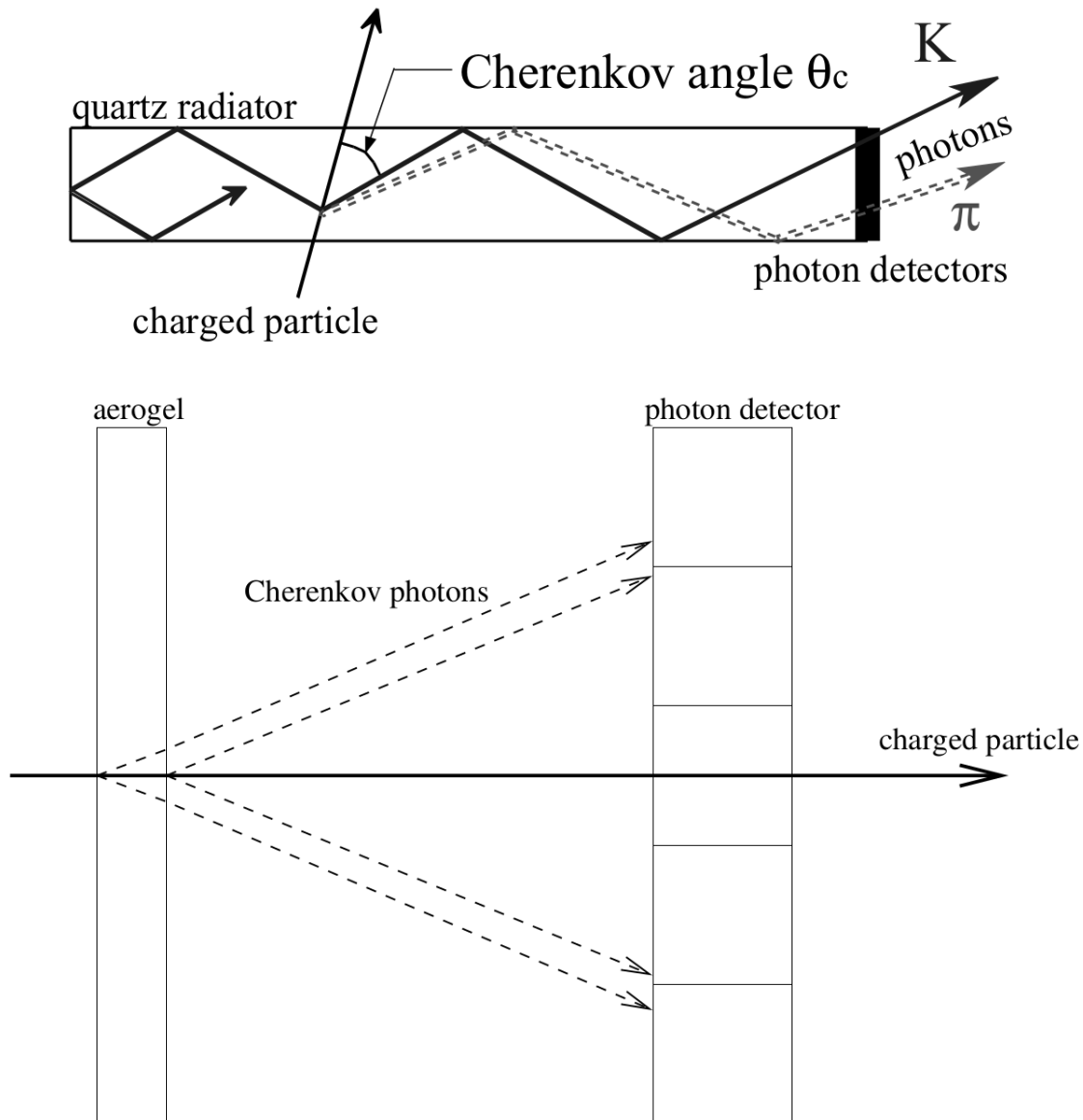


Figure 3.5: Working principles of TOP and ARICH subdetectors are depicted in the top and bottom plots, respectively. Adapted from Ref. [44].

The TOP subdetector excels in particle identification, particularly for particles with momenta up to a few GeV/c . It distinguishes between different types of charged particles with high efficiency by analyzing the Cherenkov angle, which is reconstructed from the precise measurement of Cherenkov photon's path length and time of flight. The detector's time resolution, which includes contributions from both the MCP-PMTs and the readout electronics, is approximately 100 ps. This high level of precision supports accurate and reliable particle identification.

3.2.5 Aerogel Ring Imaging Cherenkov Subdetector

The ARICH subdetector, located in the forward endcap region outside the CDC, serves the essential purpose of identifying kaons, pions and protons within a momentum range of 0.5 to 4.0 GeV/ c , as well as muons and electrons below 1.0 GeV/ c . This discrimination is achieved through a meticulously designed setup consisting of an aerogel radiator, where Cherenkov photons are generated by the passage of charged particles, as shown in Fig. 3.5. These emitted photons subsequently travel through an array of imaging devices, precisely measuring both their positions and arrival times. By analyzing these data, the ARICH subdetector effectively calculates the Cherenkov angle and, in turn, the velocity of particles using Eq. 3.1.

Aerogel, a low-density material, serves as the radiator medium in the ARICH. When charged particles traverse the aerogel, they emit Cherenkov photons due to the material's optical properties. The refractive index of the aerogel, which typically ranges from 1.03 to 1.05, is specifically chosen to optimize the Cherenkov angle for the momentum range of interest in the detector. The aerogel is arranged in tiles, each a few cm thick, which are laid out in a mosaic pattern to cover the active area of the subdetector effectively.

The detection of Cherenkov photons in the ARICH subdetector is facilitated by an array of photodetectors positioned behind the aerogel tiles. Hybrid Avalanche PhotoDiodes (HAPDs) are used for this purpose due to their high quantum efficiency and excellent spatial resolution. Each HAPD is composed of a photocathode, a microchannel plate (MCP) for amplification, and a silicon sensor for readout. The HAPDs provide a pixelated readout, allowing precise measurement of the position and number of Cherenkov photons.

The ARICH subdetector identifies particles by analyzing the Cherenkov radiation they emit. The angle of Cherenkov photon emission is dependent on the particle's velocity, which is influenced by its momentum and mass. By measuring the Cherenkov angle, the ARICH subdetector can differentiate between various particle types. The emitted Cherenkov photons create a ring pattern on the photodetector array. The radius of this ring correlates with the Cherenkov angle, and pattern recognition algorithms are employed to reconstruct these rings and identify the particles.

In terms of performance specifications, the Cherenkov angle measurement has a typical resolution of around 14 mrad, which is adequate to distinguish between pions and kaons. Additionally, the HAPDs boast a high photon detection efficiency of about 30%, ensuring that

a sufficient number of Cherenkov photons are captured for accurate ring reconstruction.

3.2.6 Electromagnetic Calorimeter

The ECL serves a central role in our experiment, efficiently detecting photons, precisely measuring their energy and angular coordinates, identifying electrons, and contributing to event triggering and luminosity measurements, both in real-time and offline data analysis.

The ECL is constructed with CsI(Tl) crystals. These crystals are chosen for their excellent scintillation properties, with the thallium dopant enhancing the scintillation efficiency by modifying the energy bands and improving the emission of photons from excited electrons. Each CsI(Tl) crystal is designed to convert the energy of incoming particles into visible light, which is then detected by photodetectors.

The crystals used in the ECL are typically shaped as truncated pyramids, with a cross-sectional area of approximately $6 \times 6 \text{ cm}^2$ and a length of 30 cm. CsI(Tl) crystals exhibit a high light yield of about 54 photons per keV of deposited energy and a relatively fast decay time of around 1 μs . These properties make them well-suited for high-precision energy measurements, ensuring that the detector can accurately measure the energy of incoming particles.

The ECL provides extensive coverage, encompassing both the barrel and endcap sections of the detector. It covers a polar angle range from 12.4° to 155.1° , offering about 90% coverage over a wide solid angle. In the barrel region, 6624 CsI(Tl) crystals are arranged to cover the polar angle range from approximately 32° to 128° . The forward and backward endcaps, which together add 2112 CsI(Tl) crystals, extend the coverage to the forward and backward directions, resulting in a total of 8736 crystals.

When high-energy photons or electrons enter the CsI(Tl) crystals, they generate electromagnetic showers—cascades of secondary particles, namely electrons, positrons, and photons. These showers result from the interaction of high-energy photons or electrons with the material, creating new e^+e^- pairs and emitting photons. The crystals produce scintillation light in response to these interactions, which is then detected by the photodetectors. The length of the crystals, which is 16.1 times the radiation length, ensures that all the energy of high-energy photons is absorbed.

The visible light produced by the scintillating CsI(Tl) crystals is detected by photodetectors attached to each crystal. Each crystal is connected to two photodiodes, which are responsible

for collecting the signals. The photodiodes detect the scintillation light and convert it into an electrical signal, enabling accurate measurement of the particle energy.

The energy resolution of the ECL is high, typically around 1.6% at 1 GeV. This precision is critical for accurate measurements of photon and electron energies. The spatial resolution of the ECL is a few mm, allowing for precise determination of the position of electromagnetic showers. Additionally, the ECL has a timing resolution of a few ns, which aids in the efficient separation of closely occurring events and enhances event reconstruction capabilities. The overall performance specifications ensure that the ECL plays a vital role in the precise measurement and identification of high-energy particles.

3.2.7 Solenoid Magnet

Encircling the ECL, there is a superconducting solenoid that generates a uniform magnetic field of 1.5 T within a cylindrical volume measuring 3.4 m in diameter and 4.4 m in length. The solenoid coil is made of a niobium-titanium (NbTi) alloy, which exhibits superconductivity at cryogenic temperatures. This choice of material is crucial to achieve the high magnetic field strength required for particle tracking. Charged particles passing through the magnetic field follow curved trajectories. By measuring the curvature of these trajectories through data collected from the tracking system, the momentum of the particles can be accurately determined using the formula:

$$p_T[\text{GeV}/c] = 0.3B[\text{T}]r[\text{m}], \quad (3.2)$$

where B represents the magnetic field strength and r denotes the radius of curvature.

To maintain the superconducting state of the solenoid, it is housed in a vacuum-insulated cryostat and cooled to cryogenic temperatures. The cryostat provides the necessary insulation to maintain low temperatures required for superconductivity. The cooling system relies on liquid helium, which circulates through the coil windings to keep the NbTi alloy at temperatures below its critical temperature of approximately 9.2° K. This helium-based cooling system is essential for ensuring that the solenoid remains in its superconducting state and operates efficiently.

3.2.8 KLong and Muon Subdetector

The KLM subdetector constitutes the outermost component of Belle II. Its primary role is to identify and track muons and neutral long-lived kaons (K_L^0), particularly those that exhibit

partial showering in the ECL. The KLM is essential for particle identification and background suppression, significantly contributing to the physics goals of Belle II. It consists of alternating layers of RPCs and scintillator strips, embedded within the iron flux return yoke of the Belle II solenoid magnet. The detector covers the polar angle range of $20^\circ < \theta < 155^\circ$.

The KLM subdetector is designed with a dual-section structure: a barrel section and endcaps, which together provide comprehensive coverage of the specified polar angle range. This design incorporates multiple layers of active detection elements, including RPCs and scintillators, interleaved with iron plates that act as absorbers. The iron plates filter out low-energy particles, ensuring that only penetrating particles such as muons and K_L^0 reach the active layers of the subdetector. The barrel section consists of 15 detector layers interspersed with 14 iron plates, while the endcap section features 14 detector layers and 14 iron plates. The use of plastic scintillators in the endcaps, as opposed to RPCs used in KEKB, addresses the higher beam background conditions expected at SuperKEKB.

RPCs are a key component of the KLM subdetector, consisting of two parallel resistive plates, typically made from materials such as glass or bakelite, separated by a narrow gas gap. The gas mixture used within the RPCs generally includes argon, isobutane, and Freon. When a charged particle traverses the RPC, it ionizes the gas, leading to an avalanche of charge that generates a signal on external readout electrodes. This process enables the RPCs to achieve a time resolution on the order of a few ns, which is crucial for accurate particle detection.

The scintillator strips used in the KLM subdetector are made from plastic scintillator, which emits light when charged particles pass through. This emitted light is collected by wavelength-shifting fibers embedded within the strips and then transported to photodetectors, such as silicon photomultipliers (SiPMs) or PMTs. Scintillators are valued for their excellent timing and spatial resolution, which significantly enhances the performance of the KLM subdetector by improving both the accuracy and precision of particle detection.

The KLM subdetector identifies muons and K_L^0 mesons by detecting their penetration through the iron absorber layers. Muons, due to their high penetration power, pass through the iron layers and generate signals in the RPCs or scintillator strips. The pattern and timing of these signals are used to identify muons accurately. For K_L^0 mesons, which are neutral, interactions with the iron plates produce secondary charged particles. These secondary particles are detected by the RPCs or scintillators, and their signatures help in identifying K_L^0 events.

The performance of the KLM subdetector is characterized by several key specifications. The spatial resolution, determined by the granularity of the RPCs or scintillator strips, is typically on the order of a few cm. Timing resolution is critical to distinguish between signals from different particles to minimize background noise; RPCs provide timing resolution on the order of a few ns, while scintillator strips offer sub-ns timing precision. The KLM also boasts high detection efficiency for both muons and K_L^0 mesons, ensuring reliable particle identification and effective background suppression.

3.2.9 TDAQ System

The Belle II TDAQ system is designed to efficiently select interesting events from the vast amount of data generated by e^+e^- collisions at the SuperKEKB accelerator and store them on the tape. This system plays a crucial role in preserving significant physics events while filtering out noise and background, thereby ensuring that valuable information is retained for further analysis. It operates in multiple layers, each playing a specific role in event selection and data management.

The Level-1 (L1) trigger represents the initial stage of event selection. It operates with a fixed latency of $5\mu\text{s}$ and utilizes a field-programmable gate array (FPGA)-based system for rapid data processing. The L1 trigger processes information from the CDC to count tracks, as well as information from the ECL to evaluate energy deposition and clusters. This information is used to classify events into two main categories: physics triggers, which focus on retaining important events like $\Upsilon(4S)$ decays and $e^+e^- \rightarrow q\bar{q}$ processes by identifying multiple tracks or clusters, and calibration triggers, which record specific events such as Bhabha scattering and $e^+e^- \rightarrow \gamma\gamma$ processes for luminosity measurements.

The High-Level Trigger (HLT) is the second stage in the trigger system, designed to handle and analyze the extensive data generated by the L1 trigger. Operating on a μs timescale, the HLT performs detailed event reconstruction to identify and isolate important events. By examining detailed particle tracks, energy deposits, and other key features, the HLT makes real-time decision about which events to retain for further analysis. This stage significantly reduces the data rate, ensuring that only the most relevant information is preserved and processed, thereby enhancing the overall efficiency of the data analysis workflow.

The DAQ architecture is modular, enabling both scalability and flexibility in data processing. It employs high-speed optical links to ensure quick data transfer from the front-end elec-

tronics to the storage systems. Key components of the data flow include the unified data link known as Belle2Link, the common readout platform called COPPER, the event builder system, and the HLT. The system follows a hierarchical structure, where the L1 trigger rapidly selects pertinent events through swift data analysis, while the HLT conducts more detailed event reconstruction and selection in real time. Ensuring data integrity and synchronization is crucial, prompting the DAQ to utilize advanced timing systems that maintain precise timing across all components. After significant events are identified, the DAQ temporarily stores the data in a buffer before transferring them to permanent storage systems.

Chapter 4

Event reconstruction

The earlier chapters discussed the standard model of particle physics, the theoretical motivation for measuring observables of $B \rightarrow K^*(892)\gamma$ decays, and provided a detailed overview of the Belle II detector. This section describes the event kinematics and the reconstruction of B candidates in the $K^*(892)\gamma$ final state. We will delve into the specific methodologies and techniques employed to reconstruct and filter these events from the background.

4.1 Analysis strategy

As mentioned earlier, Belle II functions as a B factory, producing B mesons through e^+e^- collisions. The center-of-mass energy is tuned to produce an $\Upsilon(4S)$ resonance, which decays into a pair of B mesons. We reconstruct $B \rightarrow K^*(892)\gamma$ decays via the channels $K^{*0} \rightarrow K^+\pi^-$, $K^{*0} \rightarrow K_S^0\pi^0$, $K^{*+} \rightarrow K^+\pi^0$, and $K^{*+} \rightarrow K_S^0\pi^+$. The B meson reconstructed as $K^*(892)\gamma$ is referred to as the signal-side B candidate, while the recoiling B meson is denoted as the tag-side B candidate. A schematic diagram of the event is shown in Fig. 4.1. Henceforth, K^* will refer to a $K^*(892)$ meson, and the inclusion of charge-conjugate processes is implied unless stated otherwise.

4.1.1 Blind analysis methodology

Our study employs a blind analysis strategy, where all selection criteria and the fitting procedure are finalized with simulated samples prior to examining the data in the interesting kinematic or signal region. This approach is a critical aspect in high-energy physics research, ensuring the integrity of experimental findings. Blind analysis safeguards against unconscious

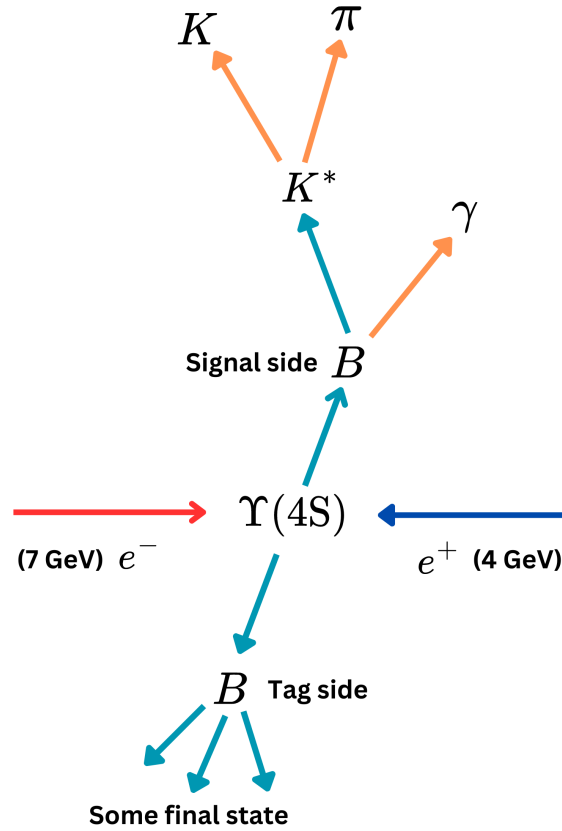


Figure 4.1: Schematic of a $B \rightarrow K^* \gamma$ event

biases or prior assumptions that might influence the interpretation of results. It is particularly crucial in precision measurements and searches for new phenomena, where even minor biases can lead to inaccurate conclusions. By employing blind analysis, the credibility and robustness of the findings are enhanced, guaranteeing that any discoveries or measurements accurately reflect the underlying physics rather than the artifacts of the analysis process.

4.2 Data and simulated samples

The data used in this analysis were collected by Belle II from 2019 to 2022 at a center-of-mass (c.m.) energy corresponding to the $\Upsilon(4S)$ resonance and 60 MeV below the resonance. The integrated luminosities for these on- and off-resonance data are $365.3 \pm 1.7 \text{ fb}^{-1}$ and $42.7 \pm 0.2 \text{ fb}^{-1}$, respectively [125].

To optimize selection criteria, calculate reconstruction efficiencies, and study backgrounds, we use Monte Carlo simulated events equivalent to 1.4 ab^{-1} of integrated luminosity. For studying signal properties, we utilize simulated samples of $\Upsilon(4S) \rightarrow B\bar{B}$, where one of the B mesons decays to the $K^* \gamma$ final state. To study possible background contamination from

other B decays, we use inclusive simulated samples of $\Upsilon(4S) \rightarrow B\bar{B}$. These simulations are generated with the EVTGEN [126] package. We also account for background contributions from $e^+e^- \rightarrow q\bar{q}$ processes, where q represents $u, d, s,$ and c quarks. These events are generated with KKMC[127] interfaced to PYTHIA[128]. The production cross-sections for the $\Upsilon(4S) \rightarrow B\bar{B}$ and $e^+e^- \rightarrow q\bar{q}$ events are listed in Table 4.1. As the analysis focuses on B decays, low-multiplicity events are typically excluded from our consideration by means of trigger requirements. The response of the Belle II detector is simulated with GEANT4[129]. Both simulated and real data samples are processed with the Belle II software framework[130; 131].

Process	Cross-section (nb)
$\Upsilon(4S) \rightarrow B\bar{B}$	1.05 [132]
$e^+e^- \rightarrow u\bar{u}$	1.61 [127]
$e^+e^- \rightarrow d\bar{d}$	0.40 [127]
$e^+e^- \rightarrow s\bar{s}$	0.38 [127]
$e^+e^- \rightarrow u\bar{u}$	1.30 [127]

Table 4.1: Cross-sections of various physics processes at the $\Upsilon(4S)$ resonance.

4.3 B meson reconstruction

Particles with sufficiently long lifetimes to be detected in one or some of the subdetectors of the Belle II experiment are regarded as primary particles. Primary particles (γ, K^+, π^+) are reconstructed using information from the ECL, particle identification (PID), and tracking systems. Particles with shorter lifetimes that decay into primary particles within the detector are referred to as secondary particles, such as $\pi^0, K_S^0,$ and J/ψ .

Reconstructing the B meson in the $K^*(892)\gamma$ final state requires a hierarchical approach. First, the primary particles $K^+, \pi^-,$ and photon are reconstructed. Next, these primary particles are combined to reconstruct secondary particles. For example, two photons are combined to form a π^0 , and two pion tracks are combined to form a K_S^0 . A kaon (K^+ or K_S^0) and a pion (π^+ or π^0) are then used to reconstruct a K^* candidate, retaining those with an invariant mass within $75 \text{ MeV}/c^2$ of its known mass [117]. Finally, a K^* is combined with a photon to reconstruct a B meson.

4.3.1 M_{bc} and ΔE

At this stage, selections are applied to the two kinematic variables M_{bc} and ΔE : $5.20 < M_{bc} < 5.29 \text{ GeV}/c^2$ and $-0.5 < \Delta E < 0.3 \text{ GeV}$ to suppress the contribution of combinatorial

background. The M_{bc} and ΔE variables are defined as follows:

$$M_{bc} = \sqrt{s/4 - \vec{p}_B^{*2}}, \quad (4.1)$$

$$\Delta E = E_B^* - \sqrt{s}/2 \quad (4.2)$$

Here, E_B^* and \vec{p}_B^* are the energy and momentum of the reconstructed B candidate calculated in the c.m. frame, with \sqrt{s} representing the total c.m. energy. Henceforth, * indicates that the quantity is calculated in the c.m. frame. The M_{bc} and ΔE variables are key to distinguish signal from background events coming from continuum processes and other B meson decays. The M_{bc} represents the invariant mass of the B meson candidate, constrained by the initial e^+e^- collision momentum. On the other hand, the ΔE measures the energy difference between the reconstructed B meson and the beam energy in the c.m. frame.

The beam energy is tuned to produce a pair of B mesons in an event. Thus, for correctly reconstructed B candidates, the M_{bc} distribution peaks at the known B meson mass [117], and the ΔE distribution peaks around zero. Continuum background events follow either a flat polynomial-like distribution or a falling exponential-like shape for ΔE . For M_{bc} , the continuum background shows a threshold-like distribution, generally a broad polynomial shape that drops to zero at the beam-energy threshold, $M_{bc} \approx 5.29 \text{ GeV}/c^2$.

The M_{bc} variable depends on the c.m. momentum of the B candidate. Therefore, even misreconstructed B decays overlap somewhat with the distribution of correctly reconstructed signal B decays but have a broader shape. On the other hand, ΔE is a better discriminator as it depends on the B meson's energy. If the tracks and clusters used to reconstruct the B meson have incorrect mass hypotheses, the ΔE values will deviate from zero. This efficiently separates misreconstructed B candidates from the signal. For example, misreconstructed decays involving contributions from multibody final states or higher kaonic decays of $B \rightarrow X_s \gamma$ transitions will shift towards negative ΔE values.

To extract the signal yield and determine various physical observables, an unbinned maximum-likelihood fit will be performed with the M_{bc} and ΔE distributions. This method combines the probability distributions of signal and background events, considering detector resolution effects and efficiencies.

As an illustration, Fig 4.2 shows the typical shapes of various components for M_{bc} and ΔE

in the case of $B \rightarrow K^* \gamma$ decay.

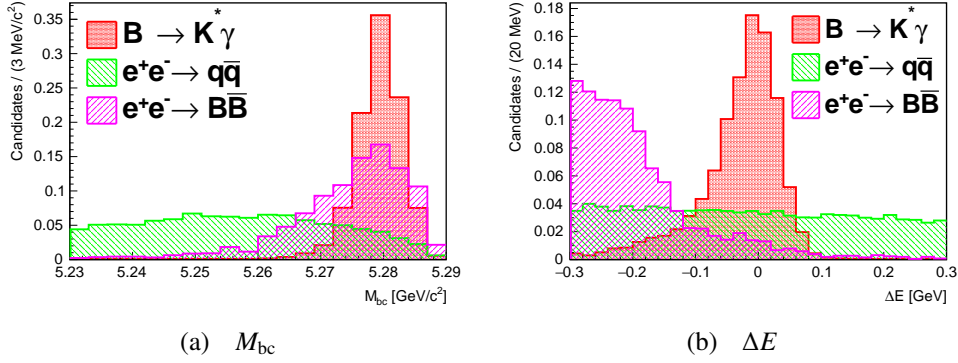


Figure 4.2: A schematic representation of (a) M_{bc} and (b) ΔE for $B \rightarrow K^* \gamma$ decays. The plots are prepared using MC sample.

4.3.2 Modification for M_{bc}

The conventional definition of M_{bc} uses the momentum of the B meson, calculated as the sum of the momenta of its decay products. However, the momentum of high-energy photon candidates often exhibits a large asymmetric tail due to leakage in the ECL. As a result, the M_{bc} calculated using the conventional definition is correlated with the ΔE variable for the signal events. To compensate for the ECL leakage effect and improve the resolution of M_{bc} variable, one can scale the magnitude of the photon energy using kinematic constraints and then calculate the modified momentum of the B candidate. In this constraint, the magnitude of the photon energy is set equal to the difference between the beam energy and the energy of the $K-\pi$ system, while keeping the direction unchanged. The expression for the modified momentum of the B meson is as follows:

$$\vec{p}_B^* = \vec{p}_{K\pi}^* + \frac{\vec{p}_\gamma^*}{|\vec{p}_\gamma^*|} \times (\sqrt{s}/2 - E_{K\pi}) \quad (4.3)$$

The M_{bc} variable, calculated using the modified momentum, exhibits a reduced correlation with the ΔE variable. Hence, we use the modified M_{bc} variable in the analysis. The comparison of the two M_{bc} variables for $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ mode is shown in Fig. 4.3. Table 4.2 compares the correlation of $M_{bc} - \Delta E$ variables of correctly reconstructed signal components for various modes. The M_{bc} resolution improves by around 16% (6%) for channels without (with) a π^0 in the final state.

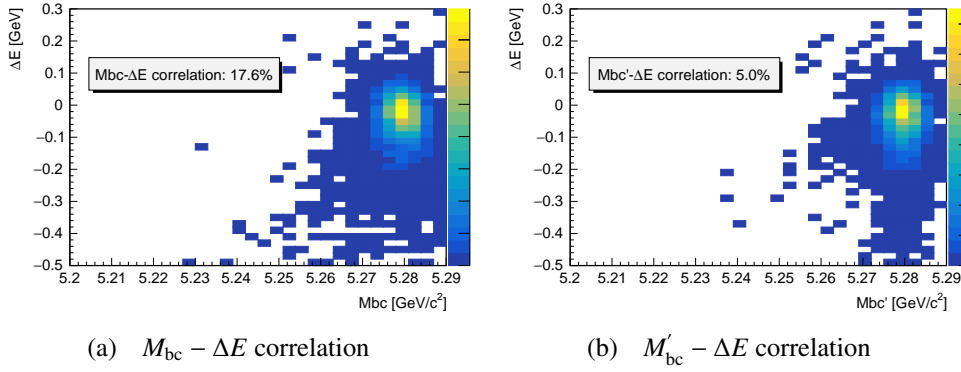


Figure 4.3: (a) $M_{bc} - \Delta E$ and (b) modified $M_{bc} - \Delta E$ correlations for correctly reconstructed $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ signal events. The plots are prepared using MC sample.

Mode	$M_{bc} - \Delta E$ (%)	Modified $M_{bc} - \Delta E$ (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	27.3	1.8
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	17.6	5.0
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	22.7	10.1
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	29.7	2.6

Table 4.2: Correlation for $M_{bc} - \Delta E$ calculated for correctly reconstructed signal events.

4.3.3 Vertex fit

A vertex fit [133] is applied to the entire B decay chain. Since the final state includes a high-energy photon, we impose additional requirements to constrain the vertex. To achieve this, we apply an IP constraint to the B meson vertex. This IP constraint is defined as an ellipsoid with a long axis along the boost direction and dimensions comparable to the beam spot in the orthogonal directions. The IP constraint improves the overall spatial resolution of the vertex reconstruction, even in cases like $B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$, where only a single track is available.

For well-reconstructed tracks and clusters that accurately converge to a common vertex points, the p-value or the χ^2 probability of vertex fit (`chiProb`) should follow a uniform distribution between 0 and 1. This uniformity arises because, in the absence of any systematic bias or significant discrepancy, the observed p-values should statistically vary within this range due to random fluctuations. In the case of misreconstruction, one encounters large χ^2 values, which result in small p-values for such events. Thus, we retain candidates with `chiProb`>0.1% for further analysis. This selection essentially filters out misreconstructed B meson candidates.

4.3.4 Rest-of-event

After reconstructing the signal-side B meson, we store information about the remaining tracks and clusters in the event that are not associated with the signal-side B candidate. This composite object is referred to as the rest-of-event, or ROE for short. Throughout various stages of the analysis, a set of stochastic gradient-boosted decision trees (BDTs) [134] are trained for background suppression, using information from both the signal-side B and the ROE.

Tracks included in the ROE are required to register at least one hit in the CDC (`nCDCHits`) and have momenta $p^* < 3.2 \text{ GeV}/c$. Similarly, clusters included in the ROE must have energies (E) between 0.05 to 3.20 GeV. Once a B candidate is reconstructed, a set of selection criteria are applied to suppress misreconstructed background events.

The vertex position of the tag-side B meson is determined by using tracks from the ROE. In particular, a vertex fit [135] is performed to precisely determine the location of the tag-side B vertex. Information about the vertex position of the tag-side B meson is used in a latter section to suppress background events.

4.3.5 Selection optimization strategy

The selection criteria are optimized with simulated datasets. These datasets are classified into correctly reconstructed signal and misreconstructed background events through a procedure called truth matching. The truth matching algorithm examines the entire decay chain of the B candidate and uses generator-level information to assign flags for reconstruction. These flags indicate whether an event is a signal or background event, allowing for effective selection from the simulated dataset.

The goal of our study is to perform a precision measurement of $B \rightarrow K^*\gamma$ observables. To achieve this, we need an appropriate Figure of Merit (FOM) [117] to optimize the selections. In a counting experiment involving both signal and background processes, we can define the signal strength μ as $\mu = s_{\text{obs}}/s$, where s_{obs} denotes the number of observed signal events and s is the expected number of signal events based on a theoretical model. Thus, the total number of observed events n is given by $n = \mu s + b$, where b represents the number of background events observed in the experiment. Using the Poisson probability, the likelihood for the quantity μ is expressed as:

$$\mathcal{L}(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)} \quad (4.4)$$

The maximum-likelihood estimator for μ is $\hat{\mu} = (n - b)/s$, and the variance associated with this estimate is:

$$V[\hat{\mu}] = V\left[\frac{n - b}{s}\right] = \frac{1}{s^2} V[n] = \frac{n}{s^2} = \frac{s + b}{s^2} \quad (4.5)$$

In the last step, we assume $\mu = 1$. The standard deviation of $\hat{\mu}$ is $\sigma_{\hat{\mu}} = \sqrt{s + b}/s$. For a precision measurement, we aim to minimize the variance, which is equivalent to achieving the best measurement accuracy. Thus, we use the FOM defined as $s/\sqrt{s + b}$ and maximize it to obtain the optimal selection criteria for the discriminating variables.

In the case of $B \rightarrow K^*\gamma$, most signal events are within the kinematic region defined by $M_{bc} > 5.27 \text{ GeV}/c^2$ and $-0.15 < \Delta E < 0.07 \text{ GeV}$, corresponding to approximately a $\pm 3\sigma$ interval. The optimization procedure is performed using events within this so-called signal region.

A summary of the selection criteria for the B meson is provided in Table 4.3. Details regarding the selections applied to primary and secondary particles are described in the subsequent sections. The combined set of selection criteria from this section and the following one is collectively referred to as the preselection.

Table 4.3: Summary of preselection criteria applied on B meson.

Variable	Selection Criteria
Energy difference	$-0.5 < \Delta E < 0.3 \text{ GeV}$
Beam-constrained mass	$5.20 < M_{bc} < 5.29 \text{ GeV}/c^2$
Vertex fit p-value	$\chi^2 > 0.1\%$
ROE tracks	$n\text{CDCHits} > 1$ and $p^* < 3.2 \text{ GeV}/c^2$
ROE clusters	$0.05 < E < 3.20 \text{ GeV}$

4.4 K^+ and π^+ selection

Charged final-state particles are reconstructed using track-level information from the tracking system. We select tracks with a transverse momentum $p_T > 0.1 \text{ GeV}/c$, energy $E > 0.1 \text{ GeV}$, and momentum $p < 5.5 \text{ GeV}/c$. Additionally, we require the distance of closest approach in the transverse plane (x - y plane) to satisfy $|d_0| < 0.5 \text{ cm}$ and along the z axis to satisfy

$|z_0| < 2.0$ cm. These criteria ensure that the charged particles originate from a region near the e^+e^- collision point. Identification of the charged particle as a kaon or a pion is performed using likelihood-based PID variables:

$$\mathcal{P}(K) = \frac{\mathcal{L}_K}{\sum_i \mathcal{L}_i}, \mathcal{P}(\pi) = \frac{\mathcal{L}_\pi}{\sum_i \mathcal{L}_i}, \quad (4.6)$$

Here, \mathcal{L}_π (\mathcal{L}_K) denotes the likelihood for a track to be identified as a π^+ (K^+). The summation in the denominator includes all particle hypotheses, calculated based on inputs from PID subdetectors. We apply a selection criterion of $\mathcal{P}(\pi) > 0.6$ ($\mathcal{P}(K) > 0.6$) to select π^\pm (K^\pm) candidates. The efficiency of these selections to correctly identify a K^+ (π^+) track is approximately 79% (69%), and the misidentification rate for $K \rightarrow \pi$ ($\pi \rightarrow K$) is around 4% (5%) in the data. The efficiency and misidentification rates are derived using results from $D^{*+} \rightarrow D^0[K^-\pi^+]\pi^+$ sample, more details are provided in Section 7. The selection criteria for charged final-state particles are summarized in Table 4.4.

Table 4.4: Summary of preselection criteria for charged tracks.

Variable	Selection
Impact parameters	$dr < 0.5$ cm, $ dz < 2$ cm
Energy and momentum	$p_T > 0.1$ GeV/c, $E > 0.1$ GeV, $p < 5.5$ GeV/c
PID K^+ (π^+)	$\mathcal{P}(K/\pi) > 0.6$ ($\mathcal{P}(\pi/K) > 0.6$)

4.5 γ selection

Photon candidates are reconstructed using cluster-level information from the ECL. The $B \rightarrow K^*\gamma$ process is effectively a two-body decay, where the high-energy photon from the B decay should have an energy approximately equal to half of the B meson mass (≈ 2.5 GeV) in the c.m. frame. Therefore, we select high-energy photons within the range of $1.4 < E_\gamma^* < 3.4$ GeV.

4.5.1 Photon timing selection

To ensure the quality of photon reconstruction, we impose several additional criteria. First, we require that the difference between the photon time, determined by the fitted time of the highest energy crystal’s recorded waveform in the cluster, and the event time (`clusterTiming`), be less than 200 ns. Additionally, we apply a criterion on the ratio of `clusterTiming` and its uncertainty (`clusterErrorTiming`) to be less than 2. These criteria effectively suppress photons originating from out-of-time beam background processes.

4.5.2 Selection on shower shape

Each particle exhibits a characteristic shower shape in the ECL depending on its interaction with the detector material. Photons typically produce symmetrical showers, whereas showers from neutral hadrons tend to be more scattered due to hadronic split-offs.

To distinguish photon showers from those originating from hadrons, we employ a criterion on the ratio $E9/E21$, where $E9$ and $E21$ represent the energies deposited in the inner 3×3 and 5×5 (excluding the corner) crystals of the ECL centered around the highest energy crystal, respectively. In particular, we require $E9/E21 > 0.9$. This ratio is effective because photon showers predominantly deposit their energy in the central crystals, resulting in $E9/E21$ values close to 1. This characteristic helps in suppressing showers from hadronic sources.

Furthermore, to distinguish high-energy photon candidates from K_L^0 clusters, we utilize a dedicated stochastic gradient-boosted decision tree (BDT)[134]. This BDT is trained with a set of 11 Zernike moment variables[136] extracted from the ECL showers. The output of this BDT, referred to as `clusterZernikeMVA`, tends to approach 1 for photon showers and 0 for K_L^0 showers. We impose an optimized criterion of `clusterZernikeMVA > 0.76` to suppress contamination from K_L^0 clusters, where the optimization is based on the procedure described earlier in Section 4.3. The distribution of `clusterZernikeMVA` and the corresponding FOM plot can be found in Fig. 4.4.

4.5.3 Veto photons from π^0 and η decays

Photons originating from asymmetric decays of π^0 or η mesons into $\gamma\gamma$ final states pose a significant background for the analysis. A photon emitted along the flight direction of a π^0 (η) can mimic a signal-side hard photon candidate, leading to misreconstruction. To mitigate this, we employ a strategy where low-energy photons reconstructed in the ECL are paired with high-energy photons to form a π^0 (η)-like particle. A high-energy photon candidate is rejected if any of these pairs are found to be consistent with a π^0 or η decay, using two dedicated BDT classifiers known as the π^0 and η veto.

The classifiers use as inputs the diphoton invariant mass and various quantities for the low-energy photon: its energy, polar angle, ECL shower shape, the closest distance between the ECL cluster and tracks extrapolated from the CDC, the ratio $E9/E21$, and `ZernikeBDT`. The π^0 veto also uses the cosine of the angle in the π^0 rest frame between the signal photon and the

boost direction of the lab frame.

The outputs of the π^0 and η veto classifiers are denoted as `pi0Prob` and `etaProb`, respectively. For photons originating from the decay of π^0 (η) mesons, the value of `pi0Prob` (`etaProb`) approaches 1, while for signal photons, it tends towards 0. We apply selection criteria of `pi0Prob` < 0.8 and `etaProb` < 0.72 to reject background photons from π^0 and η decays. These criteria are determined by optimizing the FOM, as described in Section 4.3. It has been found that these criteria reject approximately 64% of background photons from π^0 or η decays while retaining 89% of the signal. The distributions of `pi0Prob` and `etaProb`, along with their respective FOM plots, are shown in Fig. 4.5.

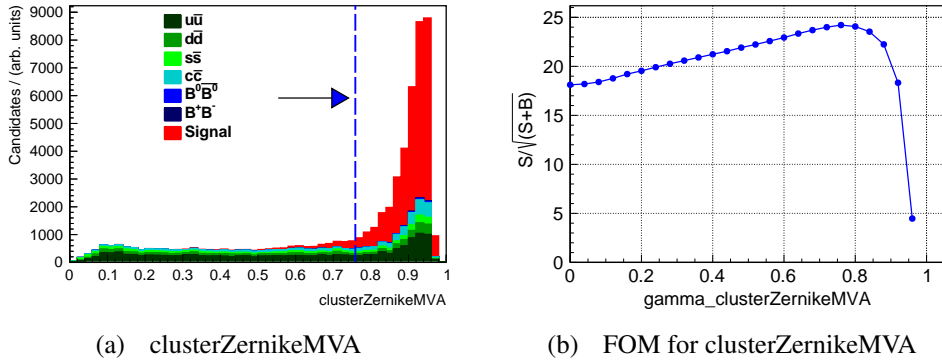


Figure 4.4: Distributions of (a) `clusterZernikeMVA` and (b) FOM for `clusterZernikeMVA` to select photons. The dashed blue line shows the cut applied on the variable, and the arrow shows the region of interest selected by the cut.

4.5.4 Impact of ECL region on ΔE Resolution

The impact of photons reconstructed from different regions of the ECL on the kinematics of B candidates is assessed by studying the resolution of ΔE using MC samples. Three distinct histograms depicting the ΔE distribution of correctly reconstructed B candidates are plotted on the same canvas (Fig. 4.6). Each of them corresponds to the case where the photon of the B candidate is reconstructed from the forward endcap, backward endcap, or barrel region of the ECL. All histograms are normalized to unit area to facilitate direct shape comparison. The resolution of ΔE is found to be superior for photons reconstructed from the barrel region compared to those from the endcap regions. Within the endcap regions, photons from the forward endcap exhibit relatively better resolution than those from the backward endcap. As a result, we select photon candidates solely from the forward endcap and barrel regions of the ECL. Excluding photons from the backward endcap leads to a marginal loss of approximately

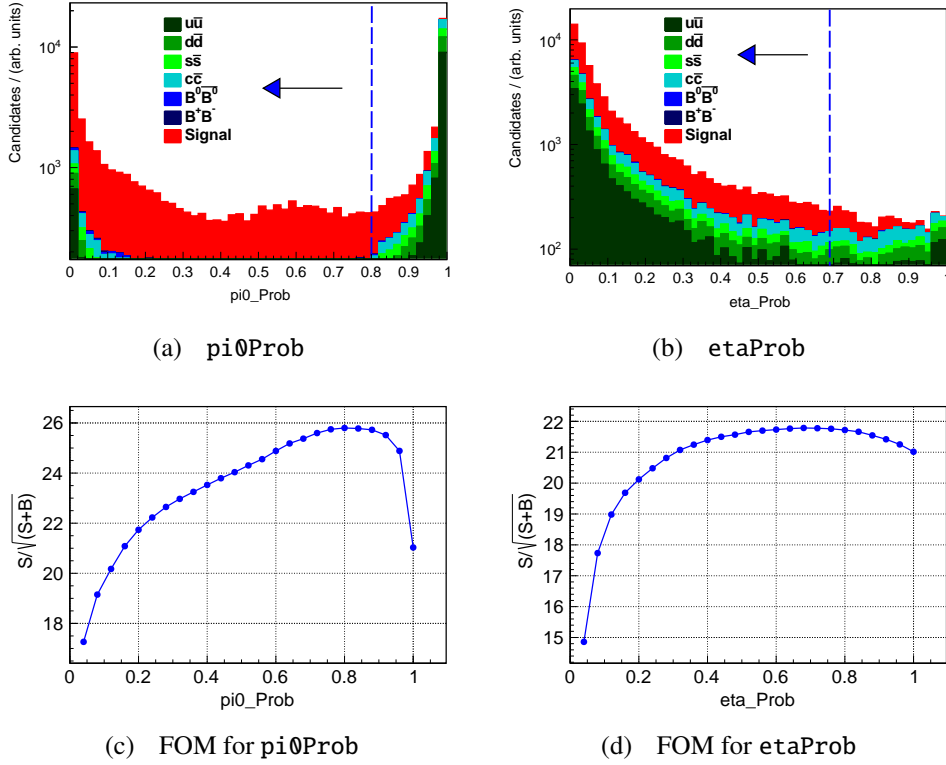


Figure 4.5: Distributions of (a) π^0 and (b) η probability for photons, and the FOM for (c) π^0 and (d) η probability. Here, the dashed blue line shows the veto criterion, and the arrow shows the region of interest retained by the veto.

5% (4%) of correctly reconstructed signal events for neutral (charged) modes.

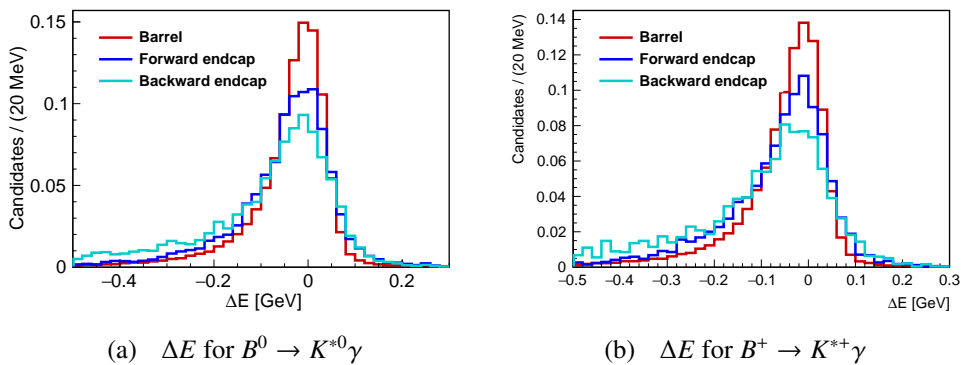


Figure 4.6: Distributions of ΔE for (a) $B^0 \rightarrow K^{*0}\gamma$ and (b) $B^+ \rightarrow K^{*+}\gamma$ modes.

The selection criteria applied to the high-energy primary photons are summarized in Table 4.5. In the following section, we describe the reconstruction of secondary particles and their selection criteria.

Table 4.5: Summary of selection criteria for primary photons coming from B meson.

Variable	Selection Criteria
Energy	$1.4 < E^* < 3.4$ GeV
Shower shape	$E9/E21 > 0.9$
Calorimeter region	Barrel + forward endcap
Timing	$\text{clusterTiming} < 200$ ns, $\text{clusterTiming}/\text{clusterErrorTiming} < 2$
Zernike shower shape	$\text{clusterZernikeMVA} > 0.76$
π^0/η veto	$\text{pi0Prob} < 0.8$ and $\text{etaProb} < 0.72$

4.6 K_S^0 selection

K_S^0 mesons are reconstructed from pairs of oppositely charged tracks assumed to be pions and constrained to come from a common vertex. Candidates that fail the vertex fit are rejected. Selections on dr and dz , as well as PID criteria described earlier are not applied to these tracks. The invariant mass of the K_S^0 candidate is required to be within $10 \text{ MeV}/c^2$ of the known K_S^0 mass [117].

Additional selections are applied to K_S^0 candidates, incorporating momentum-dependent requirements on several variables: dR , $d\phi$, z_dist , and fl . The variable dR is defined as the minimum of the impact parameters (dr_i) of the two pion tracks. The azimuthal angle between the K_S^0 momentum and the vector pointing to the K_S^0 vertex from the e^+e^- collision point is denoted as $d\phi$. The distance between the K_S^0 vertex and the e^+e^- collision point along the z -axis is referred to as z_dist . Finally, the flight distance of the K_S^0 in the transverse plane is represented by fl . These selection criteria are detailed in Table 4.6. After applying these criteria, the K_S^0 selection efficiency is approximately 94%.

Table 4.6: Additional selection criteria for the K_S^0 candidates

Momentum (GeV/ c)	dR (cm)	$d\phi$ (rad)	z_dist (cm)	fl (cm)
< 0.5	> 0.05	< 0.3	< 0.8	–
$0.5 - 1.5$	> 0.03	< 0.1	< 1.8	> 0.08
> 1.5	> 0.2	< 0.03	< 2.4	> 0.22

4.7 π^0 selection

The π^0 candidates are reconstructed from a pair of photons. We require π^0 candidates to have a diphoton invariant mass ($m_{\gamma\gamma}$) in the range $120 < m_{\gamma\gamma} < 145 \text{ MeV}/c^2$. The photons used to reconstruct π^0 candidates must also satisfy the same timing criteria applied to high-energy photons. Additionally, depending on whether the photon is detected in the forward, barrel, or

backward region of the ECL, they should have energies greater than 80 MeV, 30 MeV, or 60 MeV, respectively. Further requirements are imposed on variables related to the π^0 photons, including the ECL shower shape, polar and azimuthal angles, and the opening angle between the photons.

The selection criteria for the π^0 are presented in Table 4.7. Here, `clusterNHits` denotes the sum of weights of crystals in the ECL cluster of the photon, `clusterTheta` denotes the polar angle of the photon, γ_1^ϕ (γ_2^ϕ) is the azimuthal angle for the first (second) decay product of π^0 , and $\alpha_{\gamma_1\gamma_2}$ is the opening angle between the two decay products. In addition, a mass constrained fit is performed for the π^0 candidates to improve their momentum resolution. The overall π^0 selection efficiency is approximately 75%.

Table 4.7: Selection criteria of π^0 .

Candidate	Selection
π_γ^0	<code>clusterNHits</code> > 1.5 $0.2967 < \text{clusterTheta} < 2.6180$ $E_\gamma > 80 \text{ MeV}$ (forward endcap) $E_\gamma > 30 \text{ MeV}$ (barrel) $E_\gamma > 60 \text{ MeV}$ (backward endcap)
π^0	$ \gamma_1^\phi - \gamma_2^\phi < 1.5 \text{ rad}$, $\alpha_{\gamma_1\gamma_2} < 1.4 \text{ rad}$, and $120 < M_{\gamma\gamma} < 145 \text{ MeV}/c^2$

Up to this point, we have described the reconstruction procedure for the B meson and provided a detailed overview of the preselection criteria. In the next section, we will discuss the methodology to suppress background from light-quark pair events and the classification of residual background from B decays.

Chapter 5

Background study

The cross-section for hadronic events produced in e^+e^- collisions can be expressed as:

$$\sigma(e^+e^- \rightarrow \text{hadrons}) = 3 \sum_f z_f^2 \frac{4\pi\alpha^2}{3s} (\hbar c)^2, \quad (5.1)$$

where α is the fine-structure constant, and z_f represents the fractional charge of the quark pairs produced in e^+e^- collisions. The summation over f includes the different quark flavors that can be produced given the available \sqrt{s} . Consequently, a significant number of events produced in e^+e^- collisions come from light quarks. Figure 5.1 shows the relative contributions of various e^+e^- final states at $\sqrt{s} = 10.58$ GeV.

The term ‘continuum’ refers to background events that exhibit a continuous distribution of energy or momentum spectra over a wide range. At $\sqrt{s} = 10.58$ GeV, $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events are part of the $\Upsilon(4S)$ resonance, while events originating from light quarks ($q\bar{q}$, $q \in (u, d, s, c)$) are classified as continuum.

After applying the preselection criteria, the dominant background contribution comes from the continuum background. Light quarks have a significantly smaller mass than the B meson, causing the final-state particles to be highly boosted in the c.m. frame, resulting in a jet-like topology. In contrast, B meson pairs are produced almost at rest in the c.m. frame, with their decay products exhibiting a spherical topology. Figure 5.2 illustrates the topological differences between $e^+e^- \rightarrow q\bar{q}$ and $e^+e^- \rightarrow B\bar{B}$ events. These differences are exploited to suppress the continuum background.

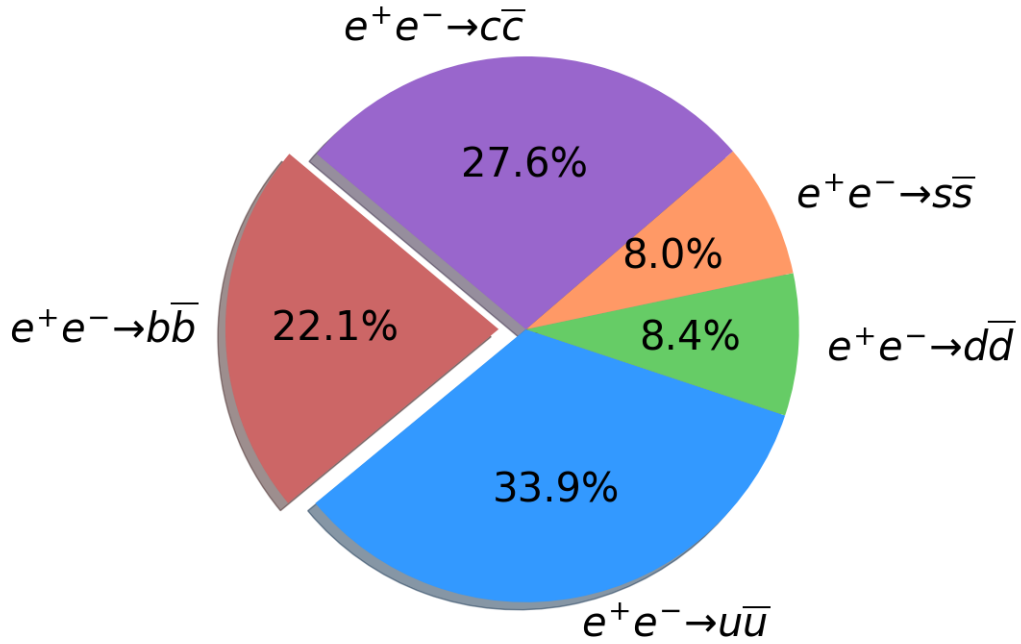


Figure 5.1: Pie chart illustrating the approximate relative contribution of various events originating from e^+e^- collisions at $\sqrt{s} = 10.58$ GeV.

5.1 Continuum suppression

To suppress the continuum background, a Boosted Decision Tree (BDT) [134], referred to as CSBDT, is independently trained for each $B \rightarrow K^*\gamma$ channel using simulated datasets. The simulated events are used to create two identical but statistically independent datasets. Each of these datasets includes an equal number of signal and continuum background events. One sample is used as the training dataset for the CSBDT, while the other serves as the testing dataset to evaluate the CSBDT performance. A set of 24 variables are taken as input features for the CSBDT. These variables, chosen for their discrimination power between signal and continuum background, also show a reasonable agreement between data and simulation. In the next section, we will briefly describe each of these input variables.

5.1.1 Input variables

The following input variables were employed to train the CSBDT:

- Kakuno Super Fox-Wolfram moments (KSF \bar{W}): The KSF \bar{W} moments [137] are derived by modifying the original Fox-Wolfram moment variables [138]. These moments are designed to reflect the topology and momentum flow within an event.

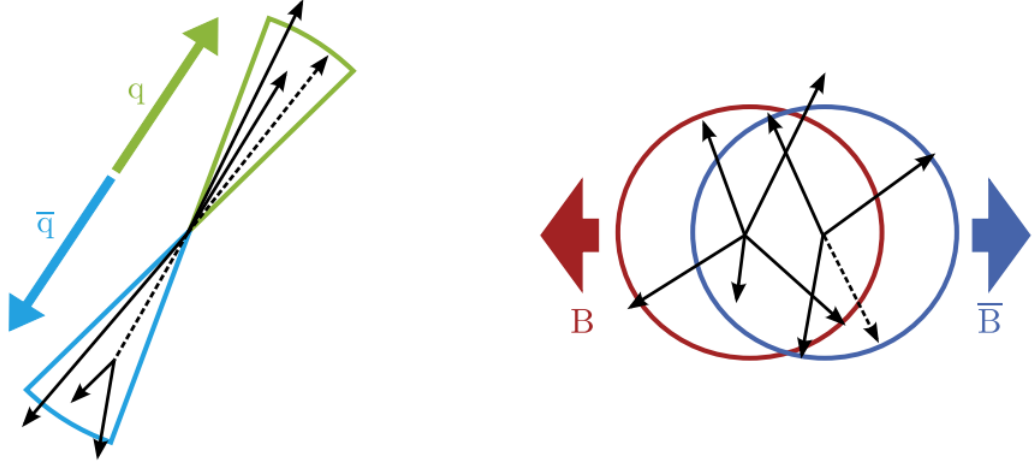


Figure 5.2: Schematic of $e^+e^- \rightarrow q\bar{q}$ process with jet-like topology (left), and $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events with spherical topology (right). Credit: Markus Röhrken.

The KSFW moments are calculated by dividing the set of reconstructed particles into two groups: the decay products of the reconstructed B candidate, denoted as ‘ s ’, and the tracks and clusters from the rest-of-event (ROE), denoted as ‘ o ’. The H_{xl}^{so} moments are further categorized based on whether the reconstructed particle is charged ($x = c$), neutral ($x = n$), or missing ($x = m$), with the missing momentum treated as a separate particle. Here, ‘ l ’ represents the order of the KSFW moment. The H_{xl}^{so} moments are defined using the relations:

$$H_{xl}^{so} = \sum_i \sum_{j,x} |\vec{p}_{jx}| P_l(\cos \theta_{i,jx}), \quad (l \text{ even})$$

$$H_{cl}^{so} = \sum_i \sum_{j,x} Q_i Q_{jx} |\vec{p}_{jx}| P_l(\cos \theta_{i,jx}), \quad (l \text{ odd})$$

Here, i runs over the decay products of the B candidate, and j, x run over the ROE of category x ; \vec{p}_{jx} is the momentum of particle jx ; $P_l(\cos \theta_{i,jx})$ is the l^{th} order Legendre polynomial of the cosine of the angle between particles i and jx ; Q_i and Q_{jx} denote the charges of particles i and jx , respectively.

For odd values of l , $H_{nl}^{so} = H_{ml}^{so} = 0$, resulting in a total of 11 H_{xl}^{so} moments: five from H_{cl}^{so} ($l \in [0, 4]$) and three each from H_{ml}^{so} and H_{nl}^{so} ($l \in [0, 2, 4]$).

Additionally, five H_l^{oo} moments are used, defined as:

$$H_l^{oo} = \sum_j \sum_k |\vec{p}_j| |\vec{p}_k| P_l(\cos \theta_{j,k}), \quad (l \text{ even})$$

$$H_l^{oo} = \sum_j \sum_k Q_j Q_k |\vec{p}_j| |\vec{p}_k| P_l(\cos \theta_{j,k}), \quad (l \text{ odd})$$

Here, j, k run over particles in the ROE.

To ensure independence from ΔE , the H_{xl}^{so} and H_l^{oo} moments are normalized to H_0^{\max} and $(H_0^{\max})^2$, respectively, where $H_0^{\max} = 2(E_{\text{beam}}^* - \Delta E)$. A total of 13 KSFW moments are used as input features to train the CSBDT.

- **CLEO cones**: These cone variables [139] were introduced by the CLEO collaboration. They are based on the sum of the magnitudes of momenta of particles within angular regions around the thrust axis, divided into intervals of 10° . This results in nine concentric cones centered on the thrust axis. All final-state particles, including both charged tracks and neutral clusters, are used to calculate the CLEO cone variables. These variables provide a comprehensive measure of the event shape and distribution of particles relative to the thrust axis.
- **B meson flavor tagger (qrOutput (FBDT))**: The flavor tagger [140] is a BDT trained to determine the flavor of the tag-side B meson. Many decay modes of B mesons exhibit flavor-specific signatures that correlate with the charge of the b quark involved. The flavor tagger uses variables derived from kinematic properties, track-hit patterns, and PID information of particles originating from the tag side. This information helps identify whether the tag-side B meson decayed into a particle containing a b quark or its anti-quark, aiding in the determination of the flavor-specific characteristics of the signal-side B meson decay. In $q\bar{q}$ continuum events (where a quark and an antiquark are produced), the production mechanism is different from $B\bar{B}$ events. Typically, continuum events do not exhibit the flavor correlation seen in $B\bar{B}$ events. Therefore, by utilizing the flavor tagger to identify the flavor of the tag-side B meson, one can distinguish $B\bar{B}$ from continuum events.
- **Vertex fit χ^2 p-value (chiProb)**: The χ^2 probability of the B vertex fit is an effective discriminating variable to separate continuum events. As discussed earlier, the χ^2 value for correctly reconstructed B candidates results in a uniform p-value distribution between 0 and 1. In contrast, for misreconstructed continuum background events, the tracks from the primary and secondary vertices are less likely to converge perfectly due to the jet-like nature of the final-state particles. This discrepancy leads to higher χ^2 values and lower

p-values (closer to 0) from the vertex fit.

- **Vertex separation (DeltaZ):** The `DeltaZ` variable denotes the distance along the z axis between the tag-side B meson vertex and the signal-side B meson vertex, serving as a crucial discriminator between continuum events and signal decays. In e^+e^- collisions, B mesons typically travel a finite distance before decaying, leading to a broader distribution of `DeltaZ` for $B\bar{B}$ events. Conversely, continuum events (prompt decay) exhibit a smaller flight distance, resulting in a narrower distribution of `DeltaZ`.

We derive observables for $B \rightarrow K^*\gamma$ modes through a 2D fit to $M_{bc} - \Delta E$ variables. This involves assessing the correlation of input variables with the distributions of correctly reconstructed signal events in $M_{bc} - \Delta E$, to avoid sculpting. To optimize the CSBDT, we exclude the first CLEO cone variable, `KSFW(hoo0)`, and `KSFW(et)` due to their correlation exceeding 5% with $M_{bc} - \Delta E$. By training the CSBDT with these variables removed, we ensure that the correlation with $M_{bc} - \Delta E$ remains below 5%.

In detail, the CSBDT incorporates 13 KSFW moments, eight CLEO cones, two topological variables (`chiProb` and `DeltaZ`), and the flavor tagger output. A concise description of these discriminating variables is available in Ref. [45]. The distributions of CSBDT input variables for signal and background events are provided in Appendix B. It also contains plots illustrating the correlation among CSBDT input variables for both signal and background events, and presents the importance of these variables across all four modes.

5.1.2 Hyperparameter Optimization

The values of hyperparameters for the BDT training are listed in Table 5.1. Here, `nTrees` fixes the number of trees, `nCuts` fixes the granularity of the variables for determining split, and `nLevels` is the number of stages of split in a tree. The learning rate of the classifier is labeled as `nShrinkage`; the value of `nShrinkage` is between 0.05 – 0.1, depending on mode. Since we don't observe any overtraining for the sample (discussed in the next section), the hyperparameters of BDT are kept fixed to their default values.

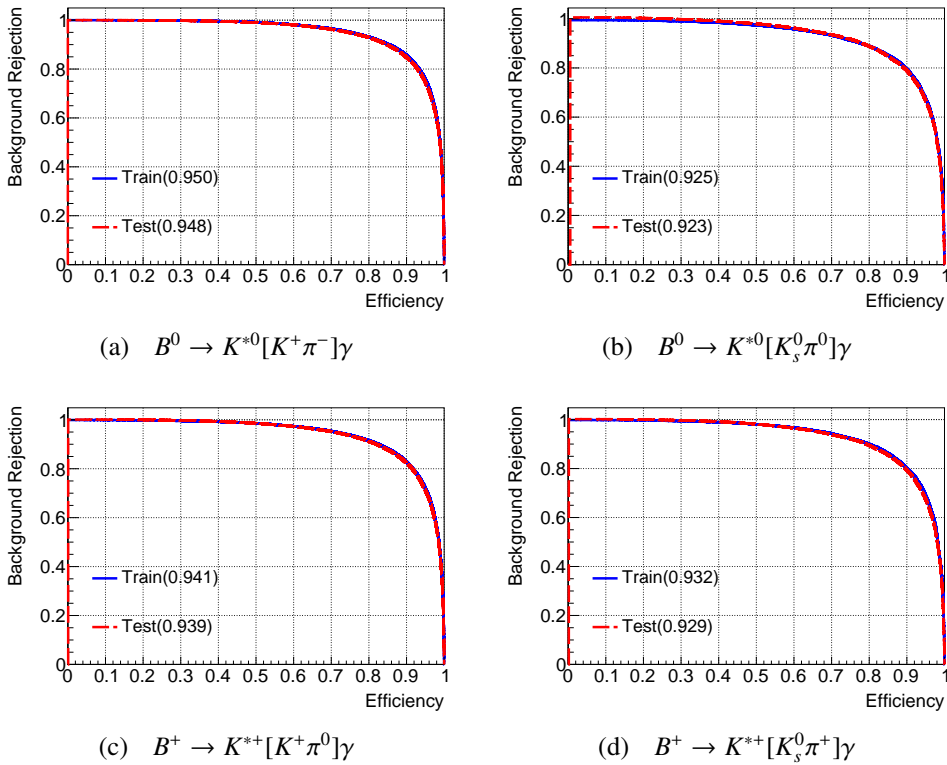
5.1.3 Overtraining check

To ensure that the CSBDT has not learned the statistical fluctuations of the training dataset (overtraining), we perform two kinds of checks. The first check involves plotting efficiency vs. background rejection (ROC curves) for the train and test samples together. If the classifier is

Table 5.1: Hyperparameters of the CSBDT training.

Parameter	Value
nTrees	200
nCuts	8
nLevels	3
nShrinkage	0.05–0.1

not overtrained, the two ROC curves will be identical, implying similar performance of CSBDT on the train and test datasets. On the contrary, if the classifier is overtrained, the two ROC curves will not be identical. The ROC curves of all modes are shown in Fig. 5.3. We observe that the ROC curves of the train and test datasets are identical, implying that the classifier is not overtrained.

Figure 5.3: ROC curves for $B \rightarrow K^*\gamma$ modes.

The second check is a Kolmogorov-Smirnov (KS) test. We prepare two sets of histograms of the classifier output (CSBDT) distribution, one each for the train and test datasets. We calculate the KS test probability (p-value) between the CSBDT distributions of the signal events. A similar exercise is performed for the background events. The KS test is helpful in determining the probability of whether the two given histograms come from the same parent distribution, i.e., they are identical or not. A brief summary of the KS test algorithm is as follows:

- Take two histograms as input and obtain their cumulative density function (CDF).
- The CDF for a localized finite histogram should have a sigmoid shape. Find the largest vertical separation between the two CDFs.
- Convert this distance to a p-value. When the separation is close to zero, the p-value is close to one, implying that the two histograms come from the same parent distribution.

If the classifier is overtrained, the CSBDT distributions of the train and test datasets will not be identical, and the p-value will be close to zero. A p-value threshold of 0.05 has been chosen to identify overtraining of the classifier. Scenarios where the p-value is less than 0.05 are rejected at the 95% confidence level. The results of the KS test for all modes are shown in Fig 5.4. The p-values are larger than 0.05 for all modes; hence, we conclude that none of the classifiers are overtrained.

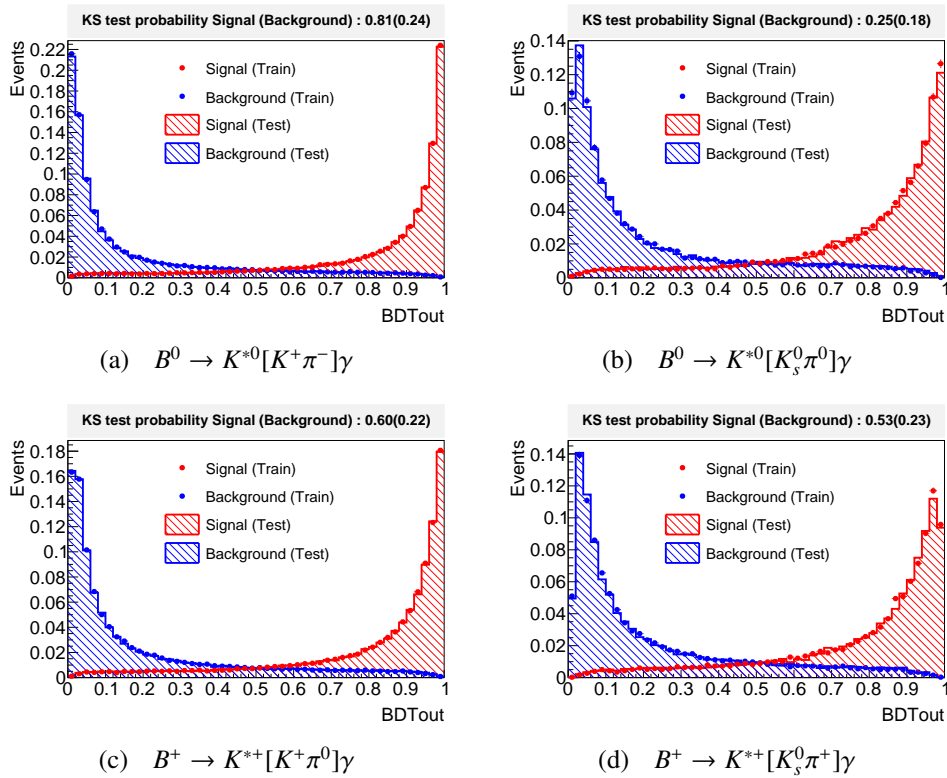


Figure 5.4: KS test results for $B \rightarrow K^*\gamma$ modes.

5.1.4 CSBDT performance

The FOM optimization for the CSBDT is performed for each mode, where we identify a selection to be applied to the variables corresponding to the FOM maximum. The FOM plots are shown in Fig. 5.5. The application of CSBDT selection helps in rejecting around 69–83% of

the continuum background, with a signal loss of 5–16% depending on mode, calculated on the simulated dataset. The selection cut of CSBDT, signal loss, and background rejection are listed in Table. 5.2.

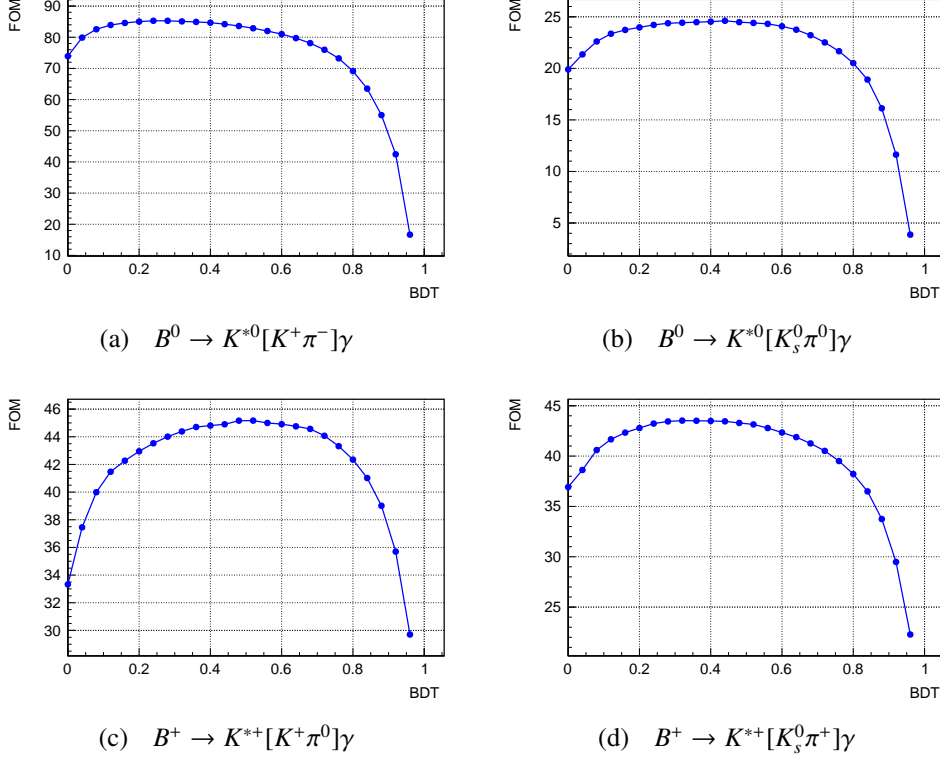


Figure 5.5: Plots of FOM for CSBDT of $B \rightarrow K^*\gamma$ modes.

Table 5.2: CSBDT cut, signal loss, and background rejection for $B \rightarrow K^*\gamma$ modes.

Mode	CSBDT cut	Signal loss (%)	Continuum rejection (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.28	5.0	69.1
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	0.56	15.8	80.8
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.52	14.3	82.6
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.36	8.5	70.9

5.2 Best Candidate Selection

Applying a selection to the CSBDT effectively reduces the majority of continuum events. To further suppress background contributions, we impose additional criteria requiring $M_{bc} > 5.23 \text{ GeV}/c^2$ and $\Delta E > -0.3 \text{ GeV}$. These criteria result in approximately 50% fewer background events, albeit with a loss of around 2–6% of signal events depending on the decay mode.

Despite these stringent selections, multiple reconstructed B candidates can still exist in a given event. In such cases, we implement a best candidate selection (BCS) strategy, retaining

the candidate with the highest CSBDT output. Our BCS choice is guided by an assumption that only one correctly reconstructed signal candidate should predominate per event, thereby minimizing ambiguity stemming from reconstruction complexities or background processes. The application of BCS not only resolves ambiguity but also enhances background suppression. Background events typically lack a distinct best candidate due to their disparate characteristics or imperfect reconstruction compared to genuine signal events.

Figure 5.6 illustrates the average candidate multiplicity across various decay modes in data. The efficiency of BCS, quantified as the fraction of correctly reconstructed signal B candidates selected in events where multiple candidates are present, is listed in Table 5.3 evaluated using MC. Fig 5.7 shows the distributions of M_{bc} and ΔE variables of all modes after BCS.

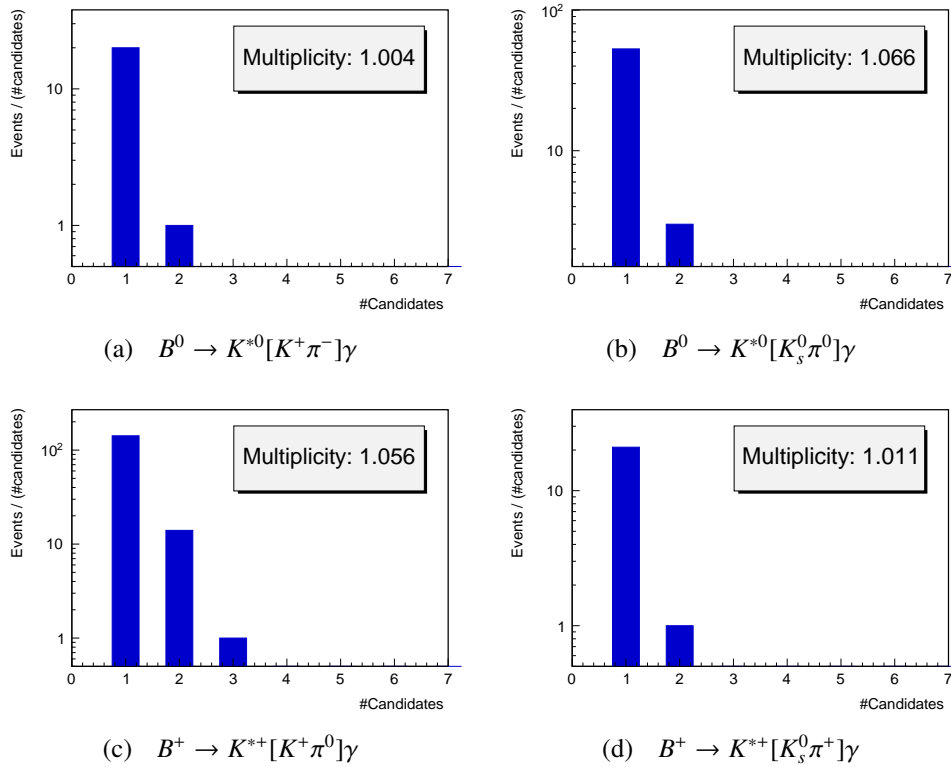


Figure 5.6: Plots for candidate multiplicity of $B \rightarrow K^*\gamma$ modes obtained from simulation.

Table 5.3: Results for best candidate selection.

Mode	Candidate multiplicity	BCS efficiency
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	1.004	75.0%
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	1.066	65.5%
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	1.056	62.4%
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	1.011	71.9%

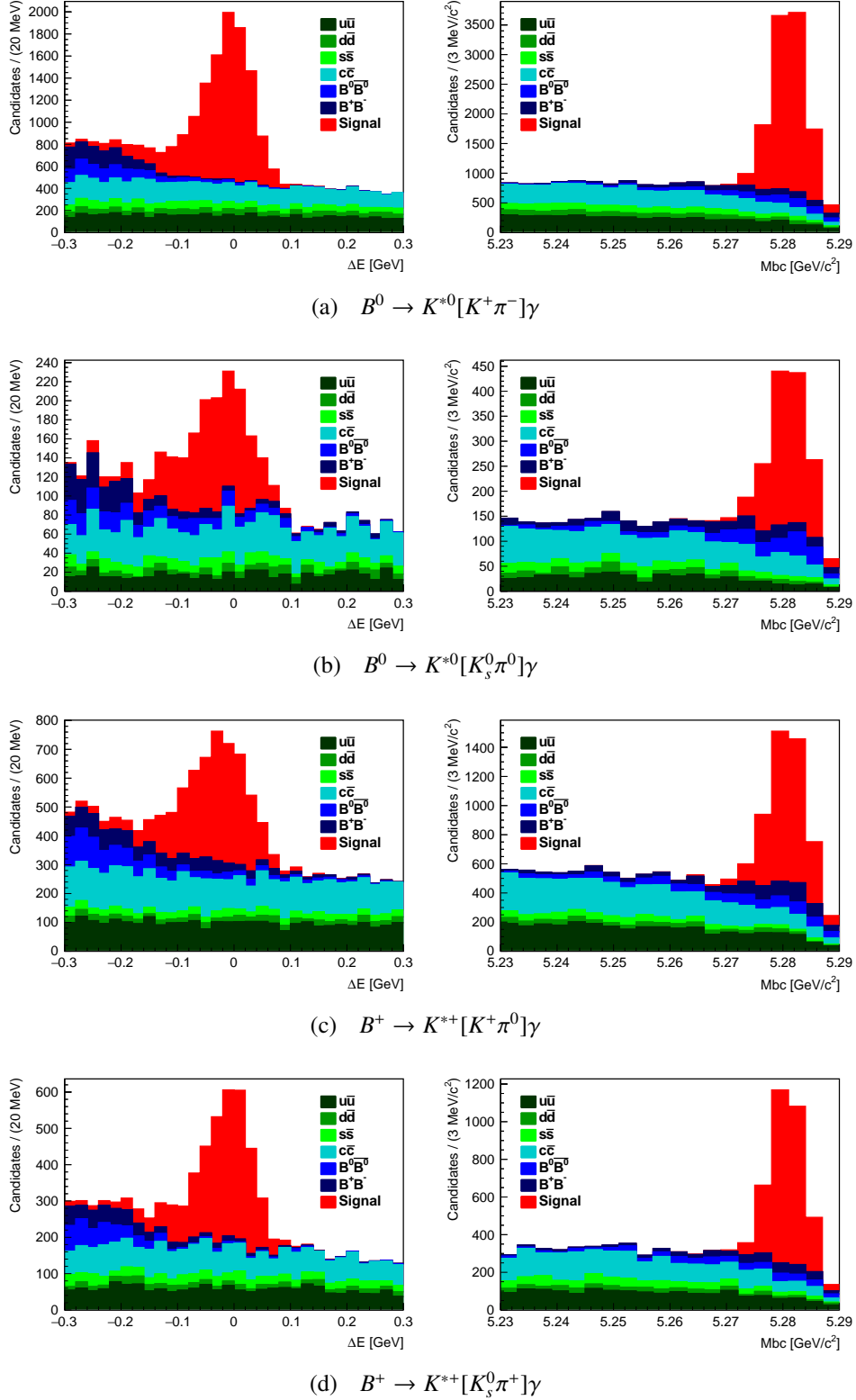


Figure 5.7: Distributions of ΔE and M_{bc} for $B \rightarrow K^*\gamma$ modes after applying all selection cuts.

5.3 $\bar{B}\bar{B}$ Background

A dedicated study is conducted to find out residual backgrounds originating from B decays after the application of all selection criteria. This study aims to identify the sources of various

background contributions, as well as to check if any background component shows a peaking structure in the ΔE variable. The residual B backgrounds can be classified into the following four categories:

1. $B \rightarrow X\gamma$, where X is the quark fragmentation system X_{sd}/X_{su} .
2. $B \rightarrow K^n\gamma$, where K^n represents higher kaonic resonances.
3. $B \rightarrow K^*\pi^0$ and $B \rightarrow K^*\eta$, due to photons evading the π^0/η veto.
4. $B \rightarrow K^*\gamma$ signal events that are misreconstructed; referred to as self-crossfeed.

The ΔE variable is sensitive to the mass hypothesis of reconstructed charged tracks and the energy of ECL clusters. Hence, except for the cross-feed events, all other B backgrounds are shifted towards the negative region of ΔE . The plots of ΔE for the $B\bar{B}$ backgrounds other than the $B \rightarrow K^*\gamma$ cross-feed are shown in Fig. 5.8. We do not observe any peaking contribution in the signal region of ΔE from these $B\bar{B}$ events. The self-crossfeed events can show some peaking contribution in the ΔE distribution. However, the fractional contribution of such self-crossfeed events with respect to the signal events is small, of the order of 2-7%. The self-crossfeed component is included as part of the signal while extracting observables through fit. We assign systematics in the fit model to incorporate possible data-MC differences for the $B\bar{B}$ background and self-crossfeed.

Among all the residual $B\bar{B}$ background events, a possible peaking contribution to the ΔE can be observed in the $B \rightarrow K_2^*(1430)\gamma$ mode, where the $K_2^*(1430)$ resonance is known to decay to the $K\pi$ final state approximately 50% of the time [117]. The investigation of this background mode is conducted by reconstructing 1 million events, confirming that $B \rightarrow K_2^*(1430)\gamma$ is harmless, with no significant peaking contribution observed in the signal region of the ΔE distribution.

5.4 Impact of Helicity variable

The helicity variable $\cos(\theta_{\text{hel}})$ is explored for its potential to suppress backgrounds from $B \rightarrow X\eta$ and $B \rightarrow X\pi^0$ transitions. This variable is the cosine of the angle between the kaon momentum and the direction opposite to the B -meson momentum, calculated in the K^* rest frame. Events from the decay process $B \rightarrow K^*\gamma$ typically exhibit a quadratic shape in $\cos(\theta_{\text{hel}})$, centered around zero. Such shape reflects the isotropic nature of the photon emitted in the K^*

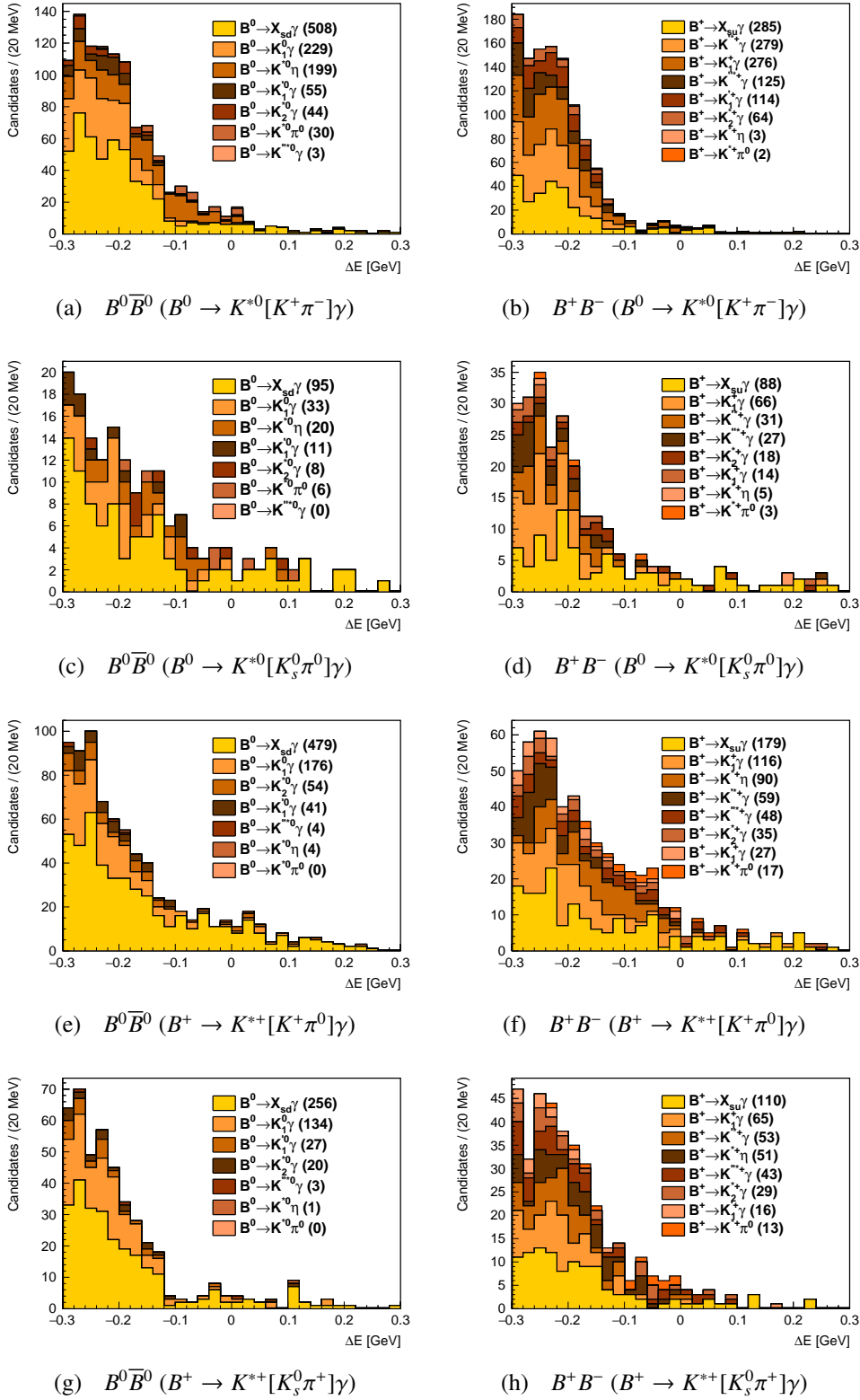


Figure 5.8: Distribution of ΔE for $B\bar{B}$ background events, the number in parentheses is the number of MC events selected for the specific decay.

rest frame. In contrast, backgrounds such as $B \rightarrow K^*\pi^0$ or $B \rightarrow K^*\eta$, which involve additional neutral mesons (π^0 or η), tend to produce decays where the photons are preferentially emitted

along the K^* momentum direction. Consequently, these backgrounds exhibit peaks near the extremities of the $\cos(\theta_{\text{hel}})$ distribution.

After applying all selection cuts, we conducted an optimization study using $\cos(\theta_{\text{hel}})$ based on the FOM. However, we did not find any improvement in the FOM when applying selections on $\cos(\theta_{\text{hel}})$. Consequently, no selection based on this variable is applied; instead, it is later used to validate our analysis by comparing the $s\mathcal{P}lot$ [141] background-subtracted distribution in data with the correctly reconstructed signal events from simulations. A brief description of the $s\mathcal{P}lot$ technique can be found in Appendix A.

Chapter 6

Measurement of observables

6.1 Observables

In this study, we measure the following seven quantities:

- $\mathcal{B}(B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma), \mathcal{A}_{CP}(B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma)$
- $\mathcal{B}(B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma)$
- $\mathcal{B}(B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma), \mathcal{A}_{CP}(B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma)$
- $\mathcal{B}(B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma), \mathcal{A}_{CP}(B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma)$

To measure these quantities, we use a two-dimensional unbinned maximum-likelihood fit to the M_{bc} and ΔE variables. For all transitions but for $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$, the final state includes a charged particle (K^+ or π^+), which helps tag the parent B meson flavor. We use the charge of these tracks to create two data samples: one for B and the other for \bar{B} .

The branching fraction and CP asymmetry are calculated as follows:

$$\mathcal{B} = \frac{N_{\bar{B}}/\epsilon_{\bar{B}} + N_B/\epsilon_B}{2 \times N_{B\bar{B}}} \quad (6.1)$$

$$\mathcal{A}_{CP} = \frac{N_{\bar{B}}/\epsilon_{\bar{B}} - N_B/\epsilon_B}{N_{\bar{B}}/\epsilon_{\bar{B}} + N_B/\epsilon_B} \quad (6.2)$$

where ϵ_B ($\epsilon_{\bar{B}}$) is the signal selection efficiency for the B (\bar{B}) sample, N_B ($N_{\bar{B}}$) is the signal yield from the B (\bar{B}) sample, and $N_{B\bar{B}}$ is the number of $B\bar{B}$ events.

In general, we treat the yield of a specific event type (signal or background) as an observable in our fit model and minimize the likelihood to determine its value. For this case, the yields N_B and $N_{\bar{B}}$ are expressed as functions of the branching fraction and \mathcal{A}_{CP} as shown below:

$$N_B = 2 \times N_{B\bar{B}} \times \epsilon_B \times \mathcal{B} \times (1 + \mathcal{A}_{CP}) \quad (6.3)$$

$$N_{\bar{B}} = 2 \times N_{B\bar{B}} \times \epsilon_{\bar{B}} \times \mathcal{B} \times (1 - \mathcal{A}_{CP}) \quad (6.4)$$

For the $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ mode, since the final state does not include any charged particles that can be used to tag the parent B . We cannot measure \mathcal{A}_{CP} using the above method. Instead, a dedicated approach that involves tagging the flavor of the tag-side B meson, is required to measure \mathcal{A}_{CP} ; however, such a study is beyond the scope of this analysis. Therefore, we only measure the branching fraction for this transition, which is given by:

$$\mathcal{B} = \frac{N_{B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma}}{2 \times N_{B^0\bar{B}^0} \times \epsilon_{B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma}} \quad (6.5)$$

Using the quantities described in Eqs. 6.1, 6.2, and 6.5, we can calculate six more observables, which are: $\mathcal{B}(B^0 \rightarrow K^{*0}\gamma)$, $\mathcal{B}(B^+ \rightarrow K^{*+}\gamma)$, $\mathcal{A}_{CP}(B^0 \rightarrow K^{*0}\gamma)$, $\mathcal{A}_{CP}(B^+ \rightarrow K^{*+}\gamma)$, Δ_{0+} , and $\Delta\mathcal{A}_{CP}$. The combined \mathcal{A}_{CP} s and \mathcal{B} s are calculated by simply taking the averages of \mathcal{A}_{CP} s and \mathcal{B} s of individual modes obtained from the fit and using the full correlation matrix method. Δ_{0+} and $\Delta\mathcal{A}_{CP}$ are calculated from the relations:

$$\Delta_{0+} = \frac{(\tau_+/\tau_0)\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) - \mathcal{B}(B^+ \rightarrow K^{*+}\gamma)}{(\tau_+/\tau_0)\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) + \mathcal{B}(B^+ \rightarrow K^{*+}\gamma)}, \quad (6.6)$$

$$\Delta\mathcal{A}_{CP} = \mathcal{A}_{CP}(B^+ \rightarrow K^{*+}\gamma) - \mathcal{A}_{CP}(B^0 \rightarrow K^{*0}\gamma) \quad (6.7)$$

where τ_+ and τ_0 are the known lifetimes of B^+ and B^0 , respectively.

6.2 Fit strategy

We determine the signal yield from a two-dimensional extended maximum-likelihood fit to the unbinned M_{bc} and ΔE distributions. The fit incorporates three event components: correctly reconstructed signal, continuum, and $B\bar{B}$ background. The probability density function (PDF)

for each component is determined by fitting simulated events. The likelihood function is written as

$$\mathcal{L}(\vec{\alpha}) = \frac{e^{-\sum_j n_j}}{N!} \prod_{i=1}^N \sum_j n_j \times \mathcal{P}_j(\Delta E^i, M_{bc}^i : \vec{\alpha}), \quad (6.8)$$

where n_j and \mathcal{P}_j are the number of events and PDF for component j , and N is the total number of events. The arguments ΔE^i and M_{bc}^i denote the values of ΔE and M_{bc} for the i^{th} candidate. Similarly, $\vec{\alpha}$ are the PDF shape parameters. The likelihood function is maximized with respect to n_j and various PDF shape parameters, as described below.

6.2.1 $B \rightarrow K^* \gamma$ signal

The shape of signal component is empirically determined by fitting correctly reconstructed events from the signal MC sample. The M_{bc} distribution of signal events is modeled with a Crystal Ball (CB) PDF [142] defined as:

$$f(x = M_{bc}, \mu_{M_{bc}}, \sigma_{M_{bc}}, \alpha_{M_{bc}}, n_{M_{bc}}) = \begin{cases} \exp\left(-\frac{1}{2} \times \left[\frac{x - \mu_{M_{bc}}}{\sigma_{M_{bc}}}\right]^2\right), & \text{for } \frac{x - \mu_{M_{bc}}}{\sigma_{M_{bc}}} \geq -\alpha \\ A \times \left(B - \frac{x - \mu_{M_{bc}}}{\sigma_{M_{bc}}}\right)^{-n}, & \text{for } \frac{x - \mu_{M_{bc}}}{\sigma_{M_{bc}}} \leq -\alpha \end{cases} \quad (6.9)$$

where,

$$A = \left(\frac{n}{|\alpha|}\right)^n \times \exp\left(-\frac{|\alpha|^2}{2}\right) \text{ and } B = \frac{n}{|\alpha|} - |\alpha| \quad (6.10)$$

The signal PDF for ΔE variable is modeled with an extension of the CB function, a double-sided CB PDF defined as:

$$f(x = \Delta E, \mu_{\Delta E}, \sigma_L, \sigma_R, \alpha_L, n_L, \alpha_R, n_R) = \begin{cases} A_L \times \left(B_L - \frac{x - \mu_{\Delta E}}{\sigma_L}\right)^{-n_L}, & \text{for } \frac{x - \mu_{\Delta E}}{\sigma_L} < -\alpha_L \\ \exp\left(-\frac{1}{2} \times \left[\frac{x - \mu_{\Delta E}}{\sigma_L}\right]^2\right), & \text{for } \frac{x - \mu_{\Delta E}}{\sigma_L} \leq 0 \\ \exp\left(-\frac{1}{2} \times \left[\frac{x - \mu_{\Delta E}}{\sigma_R}\right]^2\right), & \text{for } \frac{x - \mu_{\Delta E}}{\sigma_R} \leq \alpha_R \\ A_R \times \left(B_R - \frac{x - \mu_{\Delta E}}{\sigma_R}\right)^{-n_R}, & \text{otherwise} \end{cases} \quad (6.11)$$

where,

$$A_i = \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \times \exp\left(-\frac{|\alpha_i|^2}{2}\right) \text{ and } B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i| \text{ with } i = L, R \quad (6.12)$$

The fit projections for ΔE and M_{bc} variables of signal component are presented in Fig 6.1. Now that we have modeled the signal component, let's discuss how we construct the fit model to fit the data sample. The main idea is to fix certain shape parameters of the signal component to the values determined using signal MC samples, while allowing a restricted set of shape parameters to vary.

In our fit model, we allow the mean ($\mu_{M_{bc}}$) and sigma ($\sigma_{M_{bc}}$) parameters to float. The other two parameters (α and n) are kept fixed at the values determined from signal MC sample. We expect that allowing $\mu_{M_{bc}}$ and $\sigma_{M_{bc}}$ to float, we can account for data-MC differences in the mean position and resolution of signal events.

For the ΔE variable, the mean parameter of the double-sided CB function ($\mu_{\Delta E}$) is also allowed to float. We modify the sigma parameters of this function using a common multiplicative factor, such that $\sigma_L \rightarrow f_{\Delta E} \times \sigma_L$ and $\sigma_R \rightarrow f_{\Delta E} \times \sigma_R$. Here, σ_L and σ_R are fixed to signal MC values, while the $f_{\Delta E}$ parameter is allowed to float. All other parameters of the double-sided CB function are fixed to signal MC values. By allowing the mean parameter to float and introducing a common factor for the sigma parameters, we aim to account for data-MC differences in the mean position and resolution of the ΔE distributions for signal events.

The signal component includes the self-crossfeed (SCF) component. The latter is modeled with a two-dimensional Kernel Density Estimation (KDE) PDF [143]. The KDE models the distribution of an input dataset as a superposition of Gaussian kernels, with each data point contributing $1/N$ to the total integral of the PDF. We use the adaptive mode for the kernels, where the width of each Gaussian is calculated based on the local density of events. The PDF is configured to mirror data points at the boundaries to improve the modeling. The total signal PDF can then be written as:

$$\mathcal{P}_{\text{sig}}^{\text{tot}} = (1 - f_{\text{SCF}}) \times \mathcal{P}_{\text{cor}} + f_{\text{SCF}} \times \mathcal{P}_{\text{SCF}},$$

where, \mathcal{P}_{cor} is the PDF of correctly reconstructed signal events, \mathcal{P}_{SCF} is the PDF of the SCF component, and f_{SCF} is the ratio of the number of SCF to total events obtained from the signal MC sample. The fit projections of the SCF component are shown in Fig. 6.2.

6.2.2 Continuum background

For the continuum background, the M_{bc} and ΔE distributions are modeled using an ARGUS function and a first-order polynomial, respectively. Similar to signal, we use the product of the one-dimensional PDF of the continuum background in the fit. The ARGUS function is defined as follows:

$$f(x = M_{bc}, \chi, \xi) = \frac{\chi^3}{\sqrt{2\pi}\Psi(\chi)} \times \frac{M_{bc}}{\xi^2} \times \sqrt{1 - \frac{M_{bc}^2}{\xi^2}} \times \exp\left(-\frac{1}{2}\chi^2 \sqrt{1 - \frac{M_{bc}^2}{\xi^2}}\right) \quad (6.13)$$

The normalization factor $\Psi(\chi)$ is expressed using the CDF and PDF of the standard normal distribution, $\Phi(\chi)$ and $\phi(\chi)$, as:

$$\Psi(\chi) = \Phi(\chi) - \chi\phi(\chi) - \frac{1}{2} \quad (6.14)$$

The fit projections for the ΔE and M_{bc} variables of the continuum background are shown in Fig 6.3. In the fitting process, the shape parameter of the first-order polynomial is allowed to float. The threshold of the ARGUS function (ξ) is fixed to the beam energy threshold of 5.29 GeV/ c^2 , while the other parameter (χ) is allowed to float.

6.2.3 $B\bar{B}$ background

All events reconstructed from inclusive $B\bar{B}$ MC samples that are not $B \rightarrow K^*\gamma$ signal events are considered as $B\bar{B}$ background. Similar to the self-crossfeed, this background is modeled using a two-dimensional KDE PDF. The fit projections for ΔE and M_{bc} of the $B\bar{B}$ background are shown in Fig 6.4.

An overall summary of the fit model is provided in Table 6.1. Tables 6.2-6.3 summarize the status of various parameters of the fit model for M_{bc} and ΔE , respectively.

Table 6.1: Summary of the fit model.

Component	Shape M_{bc}	Shape ΔE	Yield
Continuum	ARGUS	Chebychev	Floated
$\overline{B\overline{B}}$		2D KDE	Floated
Signal	Crystal Ball	Double-sided CB	Floated
Self-crossfeed		2D KDE	Floated

Table 6.2: Status of parameters of the fit model for M_{bc} .

Component	Shape M_{bc}	Parameter	Status
Continuum	ARGUS	ξ	Fixed
		χ	Floated
$\overline{B\overline{B}}$	2D KDE		Fixed
Signal	Crystal Ball	$\mu_{M_{bc}}$	Floated
		$\sigma_{M_{bc}}$	Floated
		$\alpha_{M_{bc}}$	Fixed
		$n_{M_{bc}}$	Fixed
Self-crossfeed	2D KDE		Fixed

Table 6.3: Status of parameters of the fit model for ΔE variable.

Component	Shape ΔE	Parameter	Status
Continuum	Chebychev	$\alpha_{\text{Chebychev}}$	Floated
$\overline{B\overline{B}}$	2D KDE		Fixed
Signal	Double-sided CB	$\mu_{\Delta E}$	Floated
		$f_{\Delta E}$	Floated
		σ_L	Fixed
		σ_R	Fixed
		α_L	Fixed
		α_R	Fixed
		n_L	Fixed
		n_R	Fixed
Self-crossfeed	2D KDE		Fixed

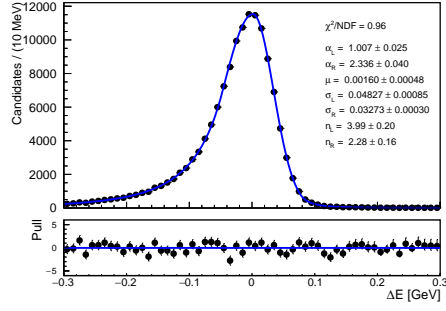
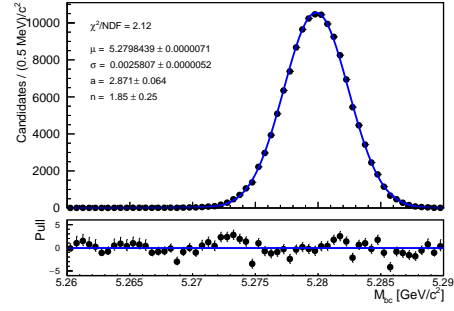
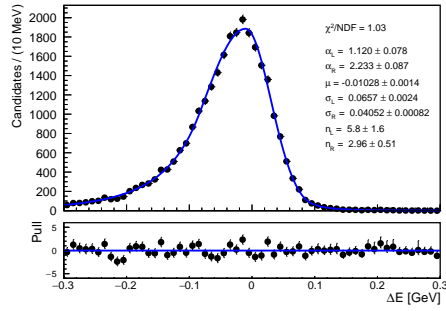
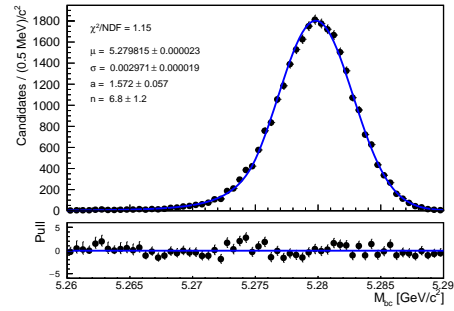
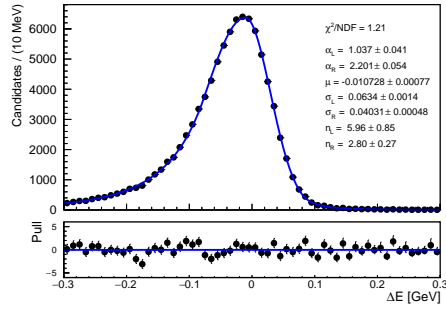
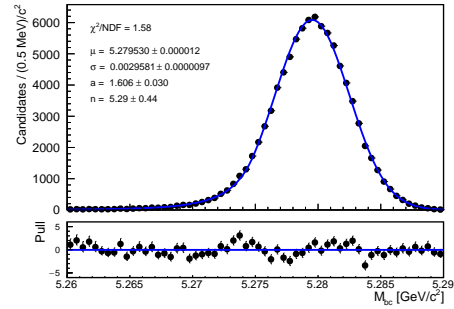
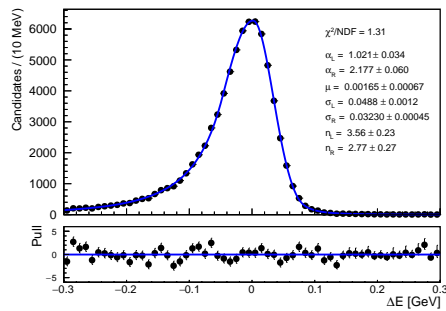
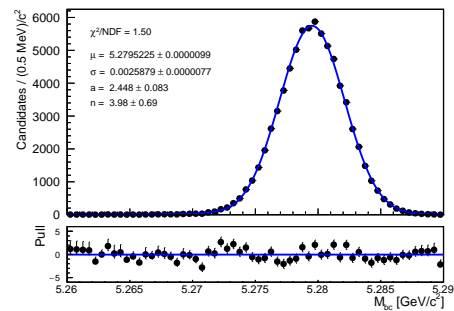
(a) $\Delta E B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (b) $M_{bc} B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (c) $\Delta E B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (d) $M_{bc} B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (e) $\Delta E B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (f) $M_{bc} B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (g) $\Delta E B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (h) $M_{bc} B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$

Figure 6.1: Fit to ΔE and M_{bc} distributions for correctly reconstructed $B \rightarrow K^*\gamma$ signal events from the MC sample after applying all selection criteria.

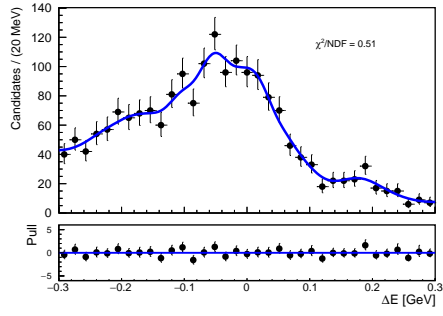
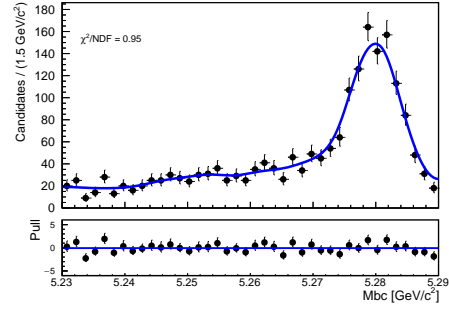
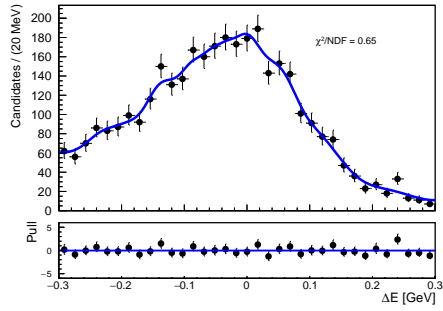
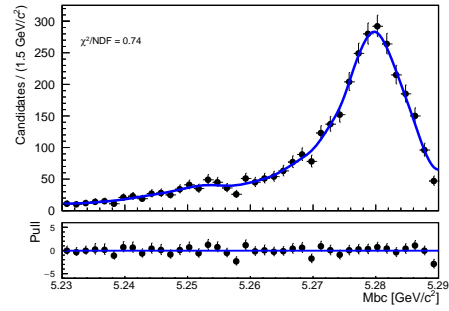
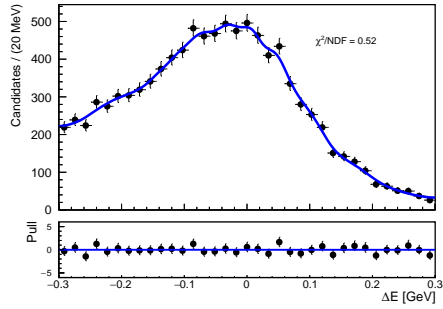
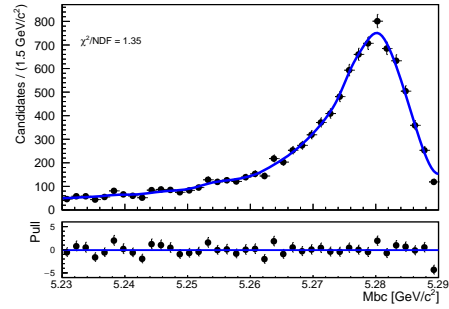
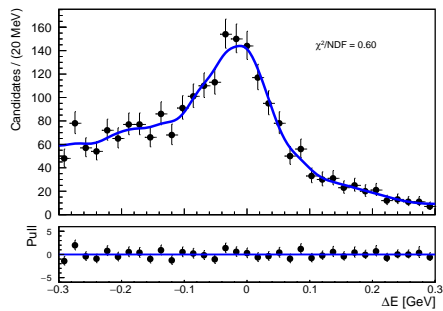
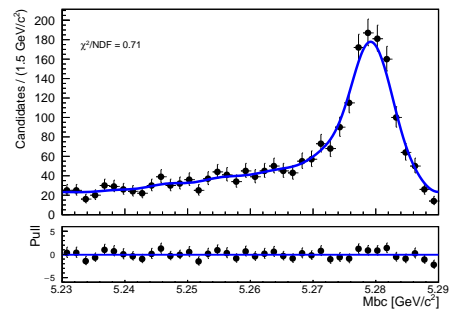
(a) $\Delta E B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (b) $M_{bc} B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (c) $\Delta E B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (d) $M_{bc} B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (e) $\Delta E B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (f) $M_{bc} B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (g) $\Delta E B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (h) $M_{bc} B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$

Figure 6.2: Fit to ΔE and M_{bc} distributions for self-crossfeed events from the MC sample after applying all selection criteria.

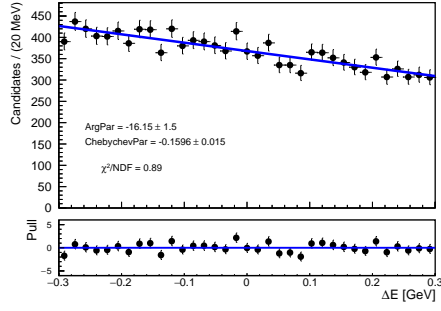
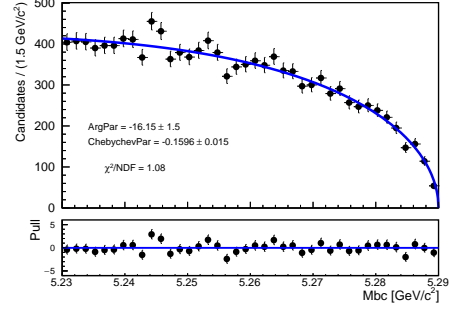
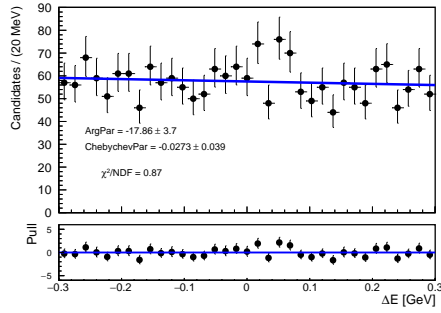
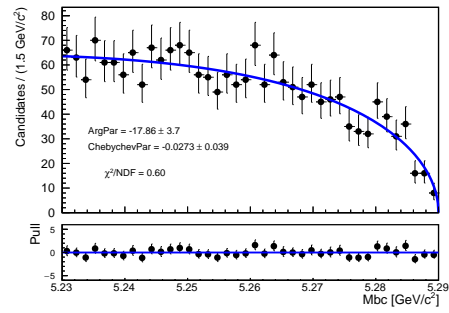
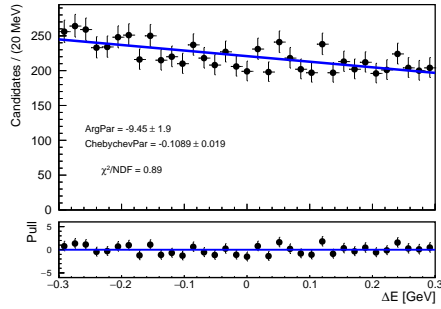
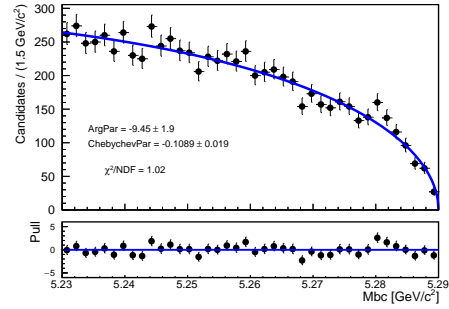
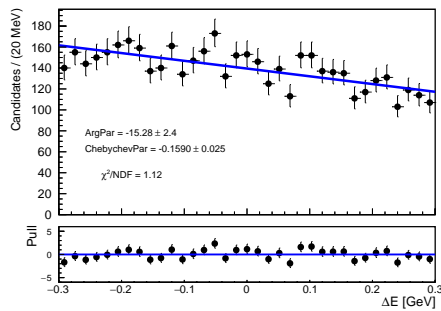
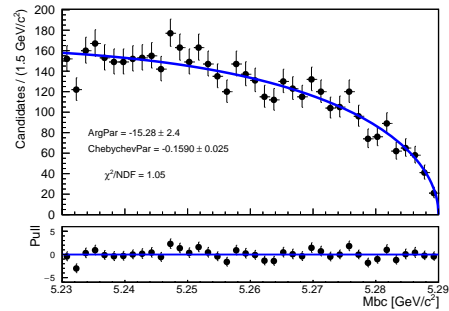
(a) $\Delta E B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (b) $M_{bc} B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (c) $\Delta E B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (d) $M_{bc} B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (e) $\Delta E B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (f) $M_{bc} B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (g) $\Delta E B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (h) $M_{bc} B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$

Figure 6.3: Fit to ΔE and M_{bc} distributions for continuum background events from the MC sample after applying all selection criteria.

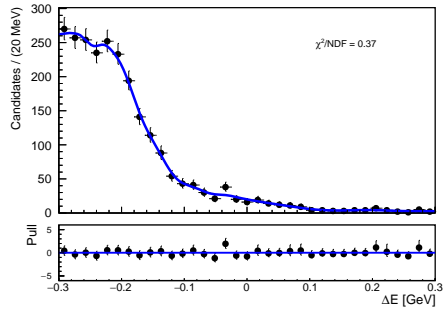
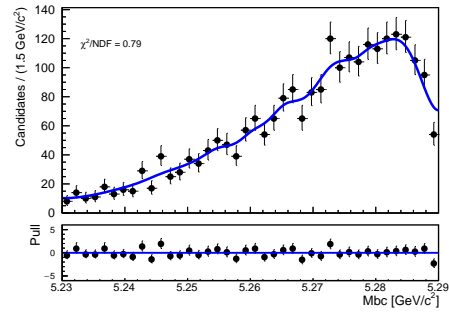
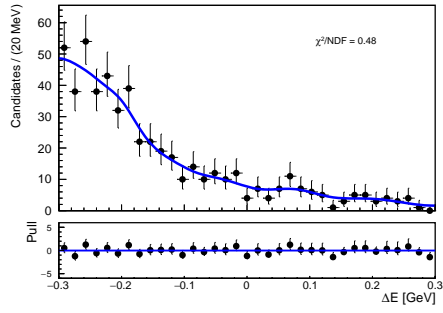
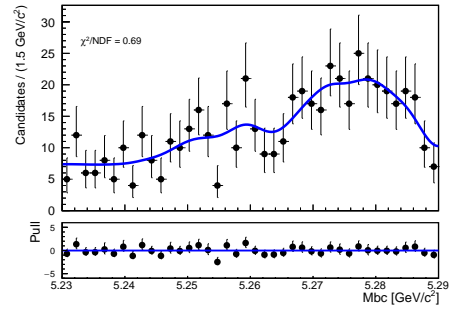
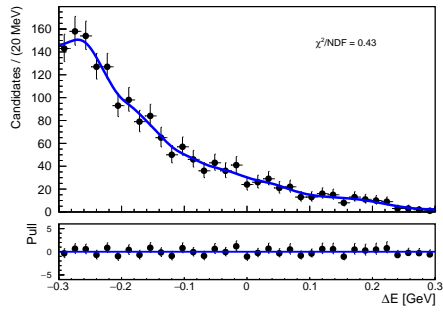
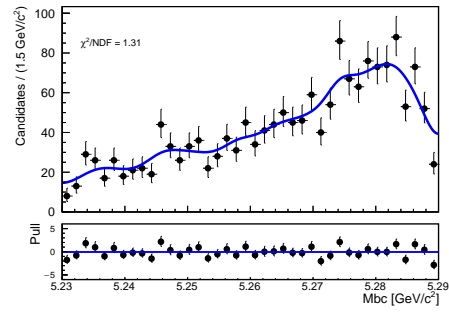
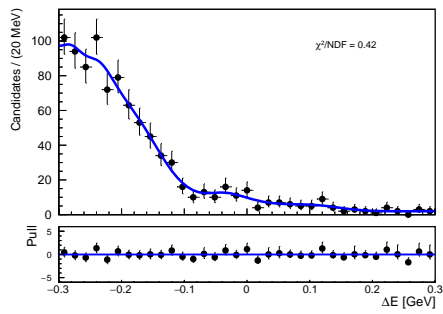
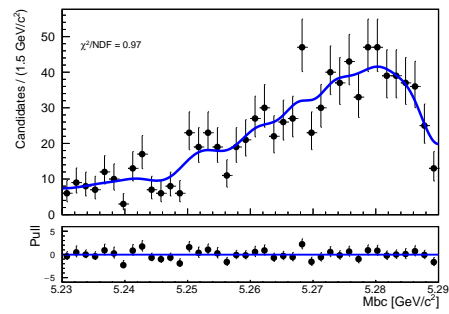
(a) $\Delta E B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (b) $M_{bc} B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ (c) $\Delta E B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (d) $M_{bc} B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ (e) $\Delta E B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (f) $M_{bc} B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (g) $\Delta E B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (h) $M_{bc} B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$

Figure 6.4: Fit to ΔE and M_{bc} distributions for $B\bar{B}$ background events from the MC sample after applying all selection criteria.

6.3 Fit to MC events

A dress rehearsal of the fit model is conducted with an MC dataset that is roughly equivalent to four times the data luminosity. The fit projections of the M_{bc} and ΔE distributions for the $B^0 \rightarrow K^0\gamma$ and $B^+ \rightarrow K^+\gamma$ modes are shown in Figs. 6.5 and 6.6, respectively.

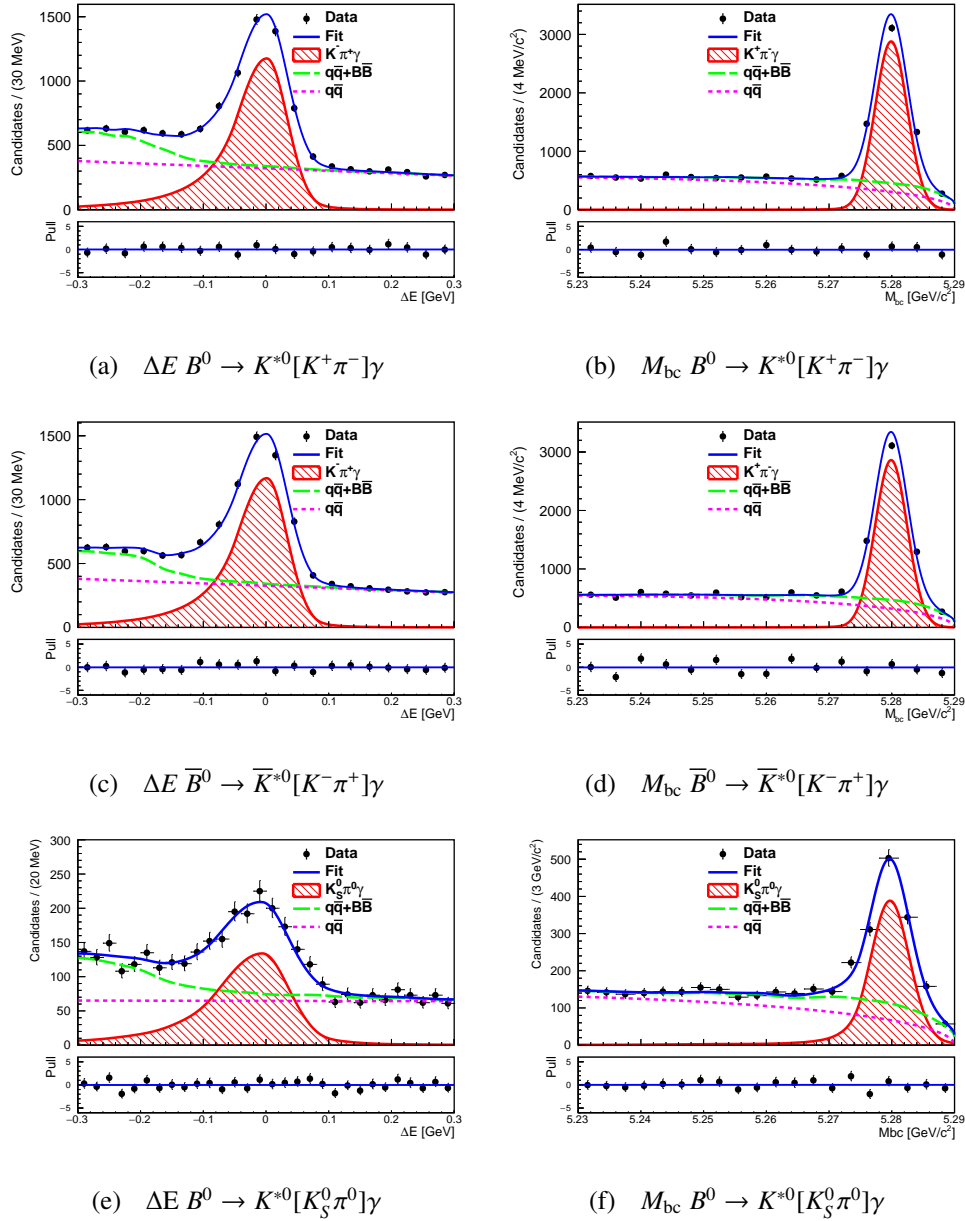


Figure 6.5: Fit projections for $M_{bc} - \Delta E$ distributions of $B^0 \rightarrow K^{*0}\gamma$ modes. The fit is performed on an MC dataset of around 4 times the data luminosity. The fit projections in (e-f) include both B^0 and \bar{B}^0 events.

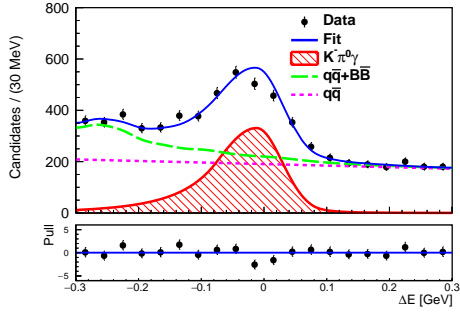
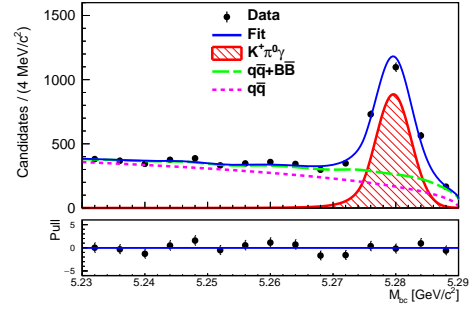
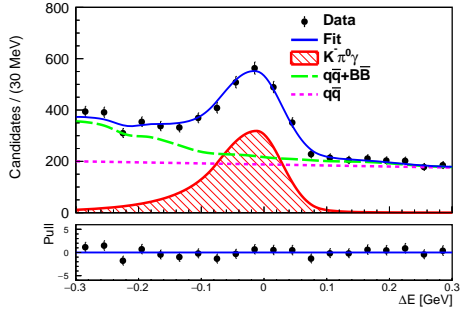
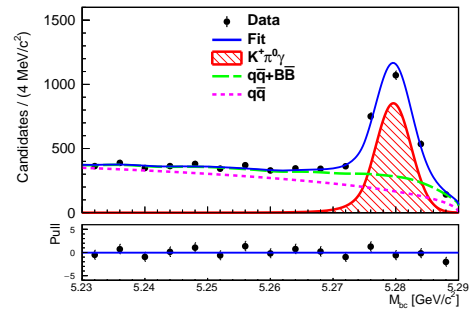
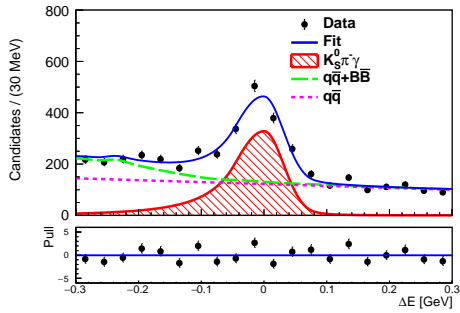
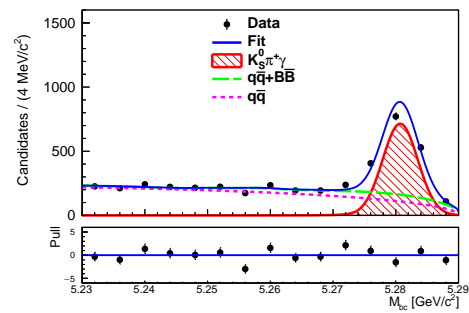
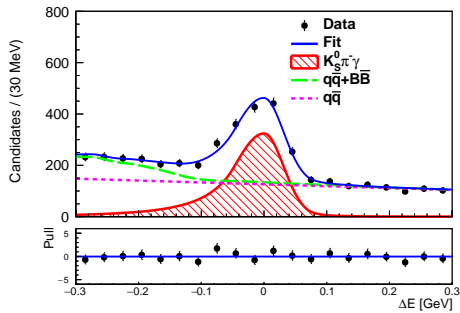
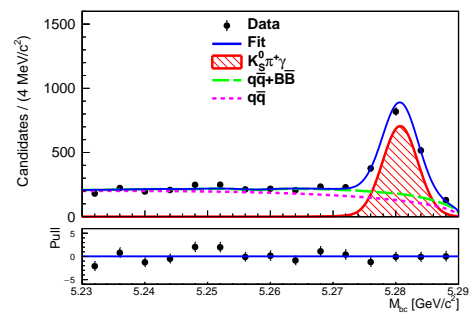
(a) $\Delta E B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (b) $M_{bc} B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ (c) $\Delta E B^- \rightarrow K^{*-}[K^-\pi^0]\gamma$ (d) $M_{bc} B^- \rightarrow K^{*-}[K^-\pi^0]\gamma$ (e) $\Delta E B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (f) $M_{bc} B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ (g) $\Delta E B^- \rightarrow K^{*-}[K_S^0\pi^-]\gamma$ (h) $M_{bc} B^- \rightarrow K^{*-}[K_S^0\pi^-]\gamma$

Figure 6.6: Fit projections for $M_{bc}-\Delta E$ distributions of $B^+ \rightarrow K^{*+}\gamma$ modes. The fit is performed on an MC sample of around 4 times the data luminosity.

The results of the observables obtained from the fit to the generic MC sample are provided in Table 6.4. We find that the values of observables obtained from the fit are consistent with their expected values from the MC decay file within the uncertainties. The values of these observables in the decay file are also consistent to the world average values [117].

Table 6.4: Fit results for an MC equivalent to $\mathcal{L} = 1.4 \text{ ab}^{-1}$.

Mode	\mathcal{B} (fit)	\mathcal{B} (MC expectation)	A_{CP} (%) (fit)	A_{CP} (%) (MC expectation)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	4.18 ± 0.05	4.18	0.5 ± 1.1	0.0
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	4.38 ± 0.18	4.18	–	–
$B^0 \rightarrow K^{*0}\gamma$	4.19 ± 0.05	4.18	0.5 ± 1.1	0.0
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	3.96 ± 0.09	3.92	1.3 ± 2.1	0.0
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	3.86 ± 0.09	3.92	-0.3 ± 2.2	0.0
$B^+ \rightarrow K^{*+}\gamma$	3.91 ± 0.06	3.92	0.5 ± 1.5	0.0
Mode	Δ_{0+} (fit)	Δ_{0+} (.dec)	$\Delta\mathcal{A}_{CP}$ (%) (fit)	$\Delta\mathcal{A}_{CP}$ (%) (.dec)
$B \rightarrow K^*\gamma$	7.1 ± 1.0	6.9	0.0 ± 1.9	0.0

As a cross-check, we plot in Fig. 6.7 the *sPlot* [141] background-subtracted distribution of the $\cos\theta_{\text{hel}}$ variable for the $B \rightarrow K^*\gamma$ modes and compare it with the distribution of correctly reconstructed signal events from the same MC sample. We observe a good agreement between the two distributions. This confirms that the fit model effectively selects the $B \rightarrow K^*\gamma$ candidates.

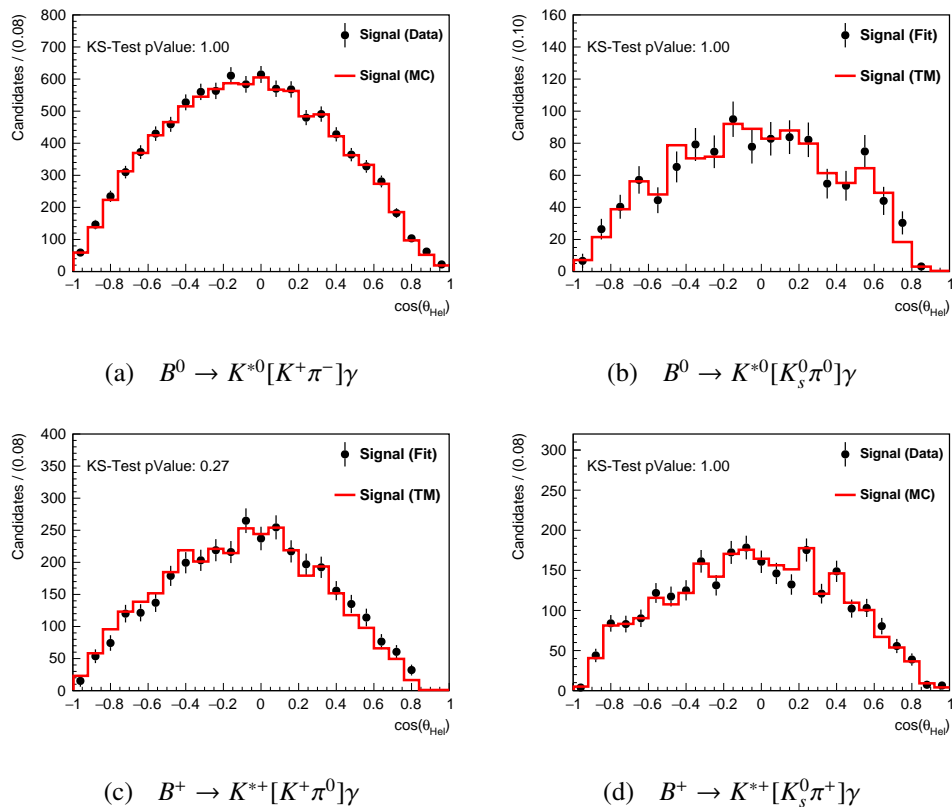


Figure 6.7: Comparison between $sPlot$ background-subtracted distribution of $\cos\theta_{\text{hel}}$ in the MC events (black points with error bars) with that obtained from correctly reconstructed signal events (red histograms). The plots are obtained from a fit performed on an MC sample of around 4 times the data luminosity.

Chapter 7

Systematic Uncertainty

“ Most important part of doing physics is the knowledge of approximation.”

– Lev Landau

In high-energy physics, the analysis relies on inputs from various sources that are crucial for measuring observables. To accurately determine the observables, it is essential to incorporate uncertainties arising from these external inputs, referred to as systematic uncertainties.

External inputs include parameters calculated from secondary measurements, such as the fraction of $\Upsilon(4S)$ candidates decaying into charged or neutral B meson pairs. Systematic uncertainties also arise from differences in selection efficiency between data and simulation, namely the PID requirements on charged hadrons, reconstruction efficiencies for secondary particles like π^0 and K_S^0 candidates, and the CSBDT threshold applied to B candidates. Furthermore, uncertainties related to the maximum-likelihood fit must be accounted for, as the fitting process can introduce biases affecting the final results.

Estimating these systematic uncertainties involves a comprehensive approach. Dedicated studies are conducted to quantify each source of uncertainty, often using control channels that mimic the signal's kinematics. These channels help isolate specific effects and provide a clear understanding of how different factors can influence the measurements. For instance, control samples of well-known processes can be used to calibrate and validate the performance of selection criteria, ensuring that simulations accurately reflect real data.

We cross-check systematic uncertainties using different methods and datasets to confirm

their robustness. Techniques such as bootstrapping and toy MC studies can help assess the stability and reliability of fit results under various conditions. In this chapter, we will describe systematic uncertainties in detail, outlining the methodologies used to quantify and incorporate them into our analysis. By carefully addressing these uncertainties, we ensure that our measurements are as accurate and reliable as possible.

7.1 Double-ratio method

To evaluate the systematic uncertainties for a given selection, we use well-known control samples with high statistics and low background contamination. These samples should have a topology similar to signal events and be reconstructed with almost the same selection criteria. First, we select a specific criterion for which the systematic uncertainty needs to be estimated. By performing a maximum-likelihood fit or counting the event yield, we determine the selection efficiency (ϵ) for the control sample events. We then calculate the ratio of efficiency between data and simulation, taken as the correction factor $\epsilon_{\text{data/MC}} = \epsilon(\text{data})/\epsilon(\text{MC})$. This double ratio is used to adjust the efficiency obtained from simulation to match that of the data, while the uncertainty on the ratio is taken as the systematic uncertainty.

In some cases, the correction factor is determined in bins of kinematic variables such as the momentum and cosine of the polar angle to account for differences in the detector's performance. The data-MC correction factor for each bin is weighted by the relative fraction of signal events in that bin, as calculated in simulation. The overall correction is the sum of these weighted correction factors. The corresponding uncertainty arises due to the uncertainty on the normalized bin entries and bin-wise uncertainties on the correction factors.

7.2 External parameters

Some experimental parameters are obtained from external sources, such as the B meson lifetime ratio $\tau_{B^0}/\tau_{B^+} = 1.076 \pm 0.004$, the $\Upsilon(4S)$ fragmentation fraction $f^{00} = 0.487 \pm 0.013$ [144], and $f^{\pm}/f^{00} = 1.065 \pm 0.052$ [145].

Multiplying the fragmentation fraction f^{00} by the ratio f^{\pm}/f^{00} yields $f^{\pm} = 0.519 \pm 0.029$. This approach introduces an asymmetry in the uncertainties of f^{00} and f^{\pm} . Furthermore, directly using the measured result of f^{\pm}/f^{00} would result in inflated systematic uncertainties for isospin asymmetry measurement. To address this, we perform a χ^2 fit to f^{00} and f^{\pm} , subject

to the constraint $f^{00} + f^\pm = 1$, and use the results from the fit. The values obtained from the fit are: $f^{00} = 0.486 \pm 0.008$, $f^\pm = 0.514 \pm 0.008$ and $f^{00}/f^\pm = 1.058 \pm 0.024$.

7.3 B counting

The number of $B\bar{B}$ events produced in e^+e^- collisions is determined with a data-driven approach. First, events with hadronic final states are selected by requiring the presence of at least three good quality tracks with $p_T > 100 \text{ MeV}/c$, $|d_r| < 1 \text{ cm}$, and $|d_z| < 2 \text{ cm}$, as well as at least three good quality clusters with $E > 100 \text{ MeV}$ and polar angle θ within the CDC acceptance. Additionally, the sum of track and photon energies must exceed 4.0 GeV , the energy measured in the ECL must satisfy $2.0 < E_{\text{ECL}} < 7.0 \text{ GeV}$, and the sum of the z component of momentum for tracks and clusters must be $0 < p_z < 5.0 \text{ GeV}/c$. The number of $B\bar{B}$ events is then calculated by subtracting the contribution of $q\bar{q}$ background, which is estimated by scaling the off-resonance data.

The Belle II performance group estimates the number of $B\bar{B}$ pairs to be $N_{B\bar{B}} = (387 \pm 6) \times 10^6$, contributing 1.55% to the systematic uncertainties. The number of charged and neutral B pairs is defined as $N_{B^+B^-} = f^\pm N_{B\bar{B}}$ and $N_{B^0\bar{B}^0} = f^{00} N_{B\bar{B}}$, respectively. Each of these quantities, f^\pm and f^{00} , has a relative uncertainty of around 1.6%.

7.4 Tracking

The systematic uncertainty due to track reconstruction is derived from studies of $e^+e^- \rightarrow \tau^+\tau^-$ events. A tag and probe method has been used to estimate this systematics. The study uses events where one of the τ leptons decays leptonically ($\tau \rightarrow \ell\nu_\tau\bar{\nu}_\ell$, $\ell = e, \mu$) and the other decays to three pions ($\tau \rightarrow 3\pi\nu_\tau$). Three tracks are used to tag this event, while the existence of the fourth track ('probe') is inferred from charge conservation. The tracking efficiency is defined as the fraction of events where this fourth track was reconstructed. The tracking efficiency is found to be consistent between data and simulation, with an uncertainty of 0.24%. Accordingly, we do not correct for efficiency, instead assign a systematic uncertainty of 0.24% per track, based on the prescription of Section 7.1. This uncertainty is considered to be fully correlated among the tracks. Consequently, we assign a systematic uncertainty of 0.48% for the neutral $B \rightarrow K^*\gamma$ modes, each of which has two tracks. For the charged modes, a systematic uncertainty of 0.24% is assigned to $B^+ \rightarrow K^+[K^+\pi^0]\gamma$ and 0.72% to $B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$.

7.5 Charged hadron identification

The difference in efficiency between data and simulation due to PID selection for charged hadrons is studied with $D^{*+} \rightarrow D^0[K^-\pi^+]\pi^+$ events. Correction factors are calculated in bins of the momentum and cosine of the polar angle of the track. We use the method described in Section. 7.1 to calculate these factors. The obtained factors are used to correct for the signal efficiency calculated with simulation, and the uncertainty on the efficiency is taken as systematic. We assign a systematic uncertainty of 0.2% for the PID selection on pions and 0.4% for kaons. The correction factors and corresponding systematic uncertainties are summarized in Table 7.1. We also check for the charge dependency of the correction factors and find no significant asymmetry, allowing us to use the charge-combined systematics from Table 7.1.

Table 7.1: Data-MC correction factors and assorted systematics due to charged hadron identification. (Here, NA denotes not applicable.)

Mode	kaon	pion
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.946 ± 0.004	0.951 ± 0.002
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	NA	NA
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.946 ± 0.004	NA
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	NA	0.952 ± 0.002

7.6 Photon selection

The systematics due to data-MC difference in photon selection is evaluated by matching the missing momentum and its direction to photon clusters in the ECL arising in $e^+e^- \rightarrow \mu^+\mu^-\gamma$ events. The correction factors are evaluated in bins of polar and azimuthal angles, and the momentum of the recoil system. The correction factor comes out to be $\epsilon_{\text{data/MC}} = 1.005 \pm 0.009$. Its central value is used to correct for the signal efficiency and the uncertainty is considered as systematic. Therefore, the uncertainty due to the selection of high-energy photon candidates is assigned to be 0.9%.

7.7 K_S^0 reconstruction

The systematic uncertainty due to K_S^0 reconstruction is estimated with the $D^+ \rightarrow K_S^0\pi^+$ control sample. The K_S^0 and charged pions are reconstructed with the same selection criteria as

the signal channel. To reconstruct D^+ candidates, K_S^0 and π^+ are combined within the invariant-mass window $|M_{K_S^0\pi^+} - M_{D^+}| < 20 \text{ MeV}/c^2$, where M_{D^+} is the known D^+ mass [117]. The momentum of the K_S^0 candidate is restricted to the range $0.5 < p < 3.0 \text{ GeV}/c$ to match the signal channel. An unbinned maximum-likelihood fit is performed on the D^+ invariant-mass distribution to obtain the D^+ yields in bins of the K_S^0 flight length. The signal is modeled with a double-sided CB function, and the background is modeled with a first-order polynomial. The fit provides the number of K_S^0 candidates in the respective bins reconstructed with D^+ .

In this case, the efficiency of control-sample events for data and simulation is calculated as follows: The yield of D^+ events in bins of K_S^0 flight length is normalized to the first bin, and the normalized quantity is referred to as ϵ_{MC}^i (ϵ_{data}^i) for the i^{th} bin. Next, we calculate the correction factor for each bin using these normalized efficiencies and propagate the relevant uncertainties.

By comparing the normalized efficiencies of K_S^0 as a function of the flight distance in data and simulation, we estimate the systematic uncertainty due to K_S^0 selection to be 1.4%. The corresponding correction factor is $\epsilon_{\text{data/MC}} = 0.985 \pm 0.004$ (stat) ± 0.013 (norm) ± 0.003 (sys). Here, the first uncertainty is the uncorrelated statistical uncertainty of each bin, the second is the fully correlated uncertainty due to normalization with respect to the first bin, and the third is systematic due to fixed PDF shape parameters. The uncertainty on the correction factor (1.4%) is taken as a systematic, while the central value is used to correct for the signal efficiency in data.

7.8 π^0 selection

A systematic uncertainty of 3.9% is assigned due to the difference between data and simulation for the π^0 reconstruction efficiency. This is obtained from studies of $D^0 \rightarrow K^-\pi^+\pi^0$ and $D^0 \rightarrow K^-\pi^+$ decays. The data-MC correction factors are calculated in bins of the momentum and cosine of the polar angle of π^0 . The correction factor is $\epsilon_{\text{data/MC}} = 1.014 \pm 0.040$. The uncertainty on the correction factor is taken as a systematic, and the central value is used to correct for the signal efficiency.

7.9 χ^2 selection

The difference in efficiency between data and simulation for vertex quality criterion is studied with $B^- \rightarrow D^0[\rightarrow K^+\pi^-]\pi^-$ and $\bar{B}^0 \rightarrow D^+[\rightarrow K_S^0\pi^+]\pi^-$ control samples. The selection criteria for $B \rightarrow D\pi$ decays mirror those for the $B \rightarrow K^*\gamma$ channels. We select D^+ and D^0 can-

didates within an invariant-mass window of $|m_{K\pi} - m_D| < 10 \text{ MeV}/c^2$, where m_D is the known D meson mass [117]. A D meson is then paired with a prompt pion to form a B candidate. As in the signal channel, the range of M_{bc} is restricted to $5.23 < M_{bc} < 5.29 \text{ GeV}/c^2$. However, since there is no photon in the final state, ΔE is restricted to a narrower range ($-0.1 < \Delta E < 0.2 \text{ GeV}$) to suppress partially reconstructed $B \rightarrow D^*\pi$ decays that appear at lower ΔE values. To mimic the topology of $B \rightarrow K^*\gamma$ events, we do not include the prompt pion from the B meson in the vertex fit. The systematic uncertainty is obtained by comparing the relative efficiency of the vertex quality selection in data and simulation. The results from $\bar{B}^0 \rightarrow D^+[K_S^0\pi^+]\pi^-$ and $B^- \rightarrow D^0[K^-\pi^+]\pi^-$ decays are used for channels with and without K_S^0 in the final state, respectively. The corresponding systematics are 1.0% for channels containing a K_S^0 and 0.2% for non- K_S^0 channels.

7.10 π^0/η Veto

To evaluate systematic uncertainties due to the π^0/η veto, we apply the same strategy described in the previous section. To avoid inflated systematic uncertainties due to limited statistics, we focus on the $B^- \rightarrow D^0[K^-\pi^+]\pi^-$ control sample for these studies. The high-momentum charged pion coming directly from the B^- meson is treated as a high-energy photon, similar to the signal mode, and is then paired with a low-energy photon from the same event to create a fictitious π^0 or an η candidate. We apply the trained π^0/η veto to the control sample and include the difference in efficiency between data and simulation as a systematic uncertainty due to the veto.

The data-MC correction factor due to application of the π^0 veto is $\epsilon_{\text{data/MC}} = 1.043 \pm 0.007$, and for the η veto, it is $\epsilon_{\text{data/MC}} = 0.996 \pm 0.002$. The central value of these correction factors is used to correct for the signal selection efficiency, and the uncertainty is used to calculate the systematic uncertainty. A systematic uncertainty of 0.7% (0.2%) is assigned due to application of the $\pi^0(\eta)$ veto.

7.11 Continuum suppression BDT

The systematic uncertainty for each of the four CSBDTs trained on different $B \rightarrow K^*\gamma$ channels is determined by applying the CSBDT criteria to the $B^- \rightarrow D^0[K^-\pi^+]\pi^-$ control sample and evaluating the data-simulation difference due to the CSBDT efficiency. The data-MC correction factors calculated from the control modes are listed in Table 7.2. We assign a

systematic uncertainty of 0.3–0.4% for the application of CSBDT, depending on the channel.

Table 7.2: Data-MC correction factors and systematics for the MVA selection.

Mode	$R_{\text{data/MC}}$	Systematics (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	1.002 ± 0.002	0.3
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	1.002 ± 0.004	0.4
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	1.000 ± 0.004	0.4
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.999 ± 0.003	0.3

7.12 Fit bias

The maximum likelihood fitter, used to determine the \mathcal{A}_{CP} and branching fraction for each $B \rightarrow K^*\gamma$ mode, may exhibit some inherent bias. We estimate this bias using 1000 toy MC datasets, generated through bootstrapping [146] MC events. Each of these datasets corresponds to a target data luminosity of 362 fb^{-1} . Poisson smearing is applied to the yields of individual fit components (signal, continuum, and $B\bar{B}$) in the toy datasets. These datasets are then fitted using the respective fit model of the $B \rightarrow K^*\gamma$ modes. In order to assess the stability of the fitter, we define the pull for a floated parameter α in the fit as:

$$\alpha_{\text{pull}} = \frac{\alpha_{\text{fit}} - \alpha_{\text{true}}}{\alpha_{\text{error}}} \quad (7.1)$$

where α_{fit} is the parameter value obtained from the fit, α_{true} is the parameter value used to generate the toy dataset, and α_{error} is the fit error on the parameter. Invoking the central limit theorem [147], for an unbiased fitter, the pull distribution should resemble a normal distribution of zero mean and unit width. We also examine scatter plots of the observables, comparing their MC expectations (true values) with the fitted values obtained from the toy datasets. These scatter plots are fitted with a first-order polynomial $y(x) = mx + c$. The intercept c of an unbiased fitter should be close to zero, and the slope m should be close to one.

The pull distributions and scatter plots of observables, namely the branching fraction and \mathcal{A}_{CP} for the flavor eigenstates, and the branching fraction for $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$, are presented in Appendix D. The fitted values for the mean and sigma of pull distributions indicate consistency with normal distributions within the fit uncertainties, demonstrating that the fitter is unbiased. The values of slope and intercept for the scatter plots also confirm the observation.

Since the mean values obtained from the toy study are consistent with zero within uncertainties, we assign systematic due to fit bias based on the uncertainty in the mean value of the pull distributions of observables. The systematic uncertainty for the observables is defined as:

$$\text{Systematics for } \mathcal{B}(\%) = \frac{\sigma_{\mathcal{B}} \times 2 \times \Delta\mu}{\mathcal{B}} \times 100 \quad (7.2)$$

$$\text{Systematics for } \mathcal{A}_{CP}(\%) = \sigma_{\mathcal{A}_{CP}} \times 2 \times \Delta\mu \times 100 \quad (7.3)$$

In these equations, \mathcal{B} represents the branching fraction, where the denominator in Eq. 7.2 uses the world average values from the PDG for $\mathcal{B}(B^0 \rightarrow K^{*0}\gamma)$, specifically 4.18 ± 0.25 and 3.92 ± 0.22 . The terms $\sigma_{\mathcal{B}}$ and $\sigma_{\mathcal{A}_{CP}}$ denote the uncertainties on these observables, as reported in Table 6.3. The factor 2 accounts for the fact that the luminosity of the MC sample is four times that of the data, necessitating a scaling of the uncertainties. The quantity $\Delta\mu$ represents the uncertainty on the mean of the pull distributions. The systematic uncertainties due to fit bias are summarized in Table 7.3.

Table 7.3: Systematics due to fit bias.

Mode	$\mathcal{B}(\%)$	$\mathcal{A}_{CP}(\%)$
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.1	0.1
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	0.3	-
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.2	0.1
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.2	0.1

7.13 Limited MC statistics

The uncertainty due to limited MC statistics is considered a source of systematic error. The signal selection process can be likened to a binomial experiment, where a certain fraction of events is selected from the total number of events. If the signal selection efficiency is ϵ (in %), this corresponds to a fraction $p = \epsilon/100$ of events being selected, with p representing the binomial probability. The standard deviation for p is defined as:

$$\sigma_p = \sqrt{\frac{p(1-p)}{N}}, \quad (7.4)$$

where N is the total number of signal MC events generated. Using this information, the relative systematic uncertainty can be calculated as $\sigma_p \times 100/p$ (%). The systematic uncertainties due to limited MC statistics are presented in Table 7.4.

Table 7.4: Systematic due to limited signal MC statistics.

Mode	p	σ_p	$\sigma_p \times 100/p$ (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.150	0.00037	0.2
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	0.015	0.00013	0.8
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.052	0.00023	0.4
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.042	0.00021	0.5

7.14 Fit model

In our analysis, we consider two sources of systematic uncertainty due to the fit model:

- Fixed parameters of signal PDF shapes
- Modeling of $B\bar{B}$ background using KDE

The treatment of each source is described below:

7.14.1 Signal PDF shape

The parameters of signal PDF shapes, which are fixed to MC values during the fit, are smeared with Gaussian functions. The mean values of these Gaussians are set to the fixed parameters, and the widths are set to their corresponding uncertainties from the fit to MC samples. The data sample is then refitted with these new parameters. The results of the refit are used to create histograms for each observable, which resemble a Gaussian. The width of these histograms is used to assign a systematic uncertainty, defined as:

$$\text{Systematics for } \mathcal{B} \text{ (\%)} = \frac{\sigma_{\mathcal{B}}}{\mathcal{B}} \times 100 \quad (7.5)$$

$$\text{Systematics for } \mathcal{A}_{CP} \text{ (\%)} = \sigma_{\mathcal{A}_{CP}} \times 100 \quad (7.6)$$

Here, the branching fraction \mathcal{B} in the denominator of Eq. 7.5 is taken as the PDG [117] world average $\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = 4.18 \pm 0.25$ and $\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = 3.92 \pm 0.22$ to prevent accidental unblinding of the results. The systematics due to fixed parameters of PDF shapes are listed in Table 7.5. A conservative systematic uncertainty of 0.1% is assigned if the calculated systematic is smaller than 0.1%.

Table 7.5: Systematic due to fixed PDF shape parameters.

Mode	\mathcal{B} (%)	\mathcal{A}_{CP} (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.2	0.1
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	0.4	–
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.4	0.1
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.3	0.1

Self-crossfeed

As described in Section 6.2, the SCF component of the signal is modeled with a KDE. Using MC samples, we find the SCF contribution in the signal region to be around 6-7% for the two modes with π^0 in the final state and negligible for the other two modes. To account for the potential mismodeling of this component in data, we fluctuate the SCF fraction (f_{SCF}) by $\pm 50\%$ in the total signal PDF and assign the relative deviation of the observables as an additional systematic. Such a variation only affects the branching fractions of the modes and the respective systematics is around 0.2–2.2%, depending on the mode.

7.14.2 KDE Model for $B\bar{B}$

To evaluate the systematic uncertainties due to the parametrization of the $B\bar{B}$ background, we use the bootstrap technique to create 1000 toy MC $B\bar{B}$ datasets, with Poisson smearing. Newly created KDE PDFs are then used to refit the data sample. The results of the refit are used to create histograms that resemble Gaussians. The systematic uncertainty is assigned based on the resolution of these histograms, similar to the way uncertainties were assigned for the fixed PDF shape in the previous section. Table 7.6 presents the systematic uncertainties due to parametrization of the $B\bar{B}$ background.

Table 7.6: Systematic due to parametrization of $B\bar{B}$ background using KDE PDF.

Mode	\mathcal{B} (%)	\mathcal{A}_{CP} (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.2	0.1
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	0.4	–
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.4	0.1
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.3	0.1

7.15 Best Candidate Selection

In cases where multiple reconstructed B candidates are present in an event, we perform BCS using the output of the continuum suppression BDT. This selection process could potentially introduce bias into the analysis. To evaluate the systematic uncertainty associated with the BCS procedure, the data are refitted after randomly selecting the best candidate. Table 7.7 presents the systematic due to BCS,

Table 7.7: Systematic due to the best candidate selection.

Mode	\mathcal{B} (%)	\mathcal{A}_{CP} (%)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	0.1	0.1
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	1.0	–
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	0.6	0.5
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	0.2	0.2

7.16 Detector asymmetry

The interaction of charged hadrons with detector material can lead to asymmetries in track reconstruction efficiency. The systematics due to asymmetry in the reconstruction of K^+ and π^+ tracks are derived from studies on $D^0 \rightarrow K^-\pi^+$ and $D^+ \rightarrow K_S^0\pi^+$ control samples. We measure the yield (N) of these channels for a charge-specific final state and its CP conjugate (\bar{N}) to calculate the asymmetry $\mathcal{A} = (N - \bar{N})/(N + \bar{N})$.

The asymmetry \mathcal{A} can have three contributions: the \mathcal{A}_{CP} of the control channel, a forward-backward asymmetry in the production of D and \bar{D} mesons due to interference in $e^+e^- \rightarrow c\bar{c}$ processes [148–150], and the instrumental asymmetry. The forward-backward asymmetry is an antisymmetric function of the cosine of the D meson polar angle in the center-of-mass frame ($\cos\theta_D^*$). We remove the effect of forward-backward asymmetry by averaging \mathcal{A} in opposite bins of $\cos\theta_D^*$. The instrumental asymmetry is then obtained by subtracting the \mathcal{A}_{CP} values of the control channel [117] from \mathcal{A} . The uncertainty of the instrumental asymmetry is taken as a systematic uncertainty, as listed in Table 7.8.

Table 7.8: Systematics due to detector asymmetry.

Source	Systematics (%)
π^+	0.56
K^+	0.60
$K^+\pi^-$	0.32

The calculation of systematics for Δ_{0^+} is based on the principle of error propagation. A brief description of this method and a breakdown of the covariance matrix into individual sources is provided in Appendix A. The various systematic uncertainties for the branching fraction and Δ_{0^+} measurements are summarized in Table 7.9, and for the \mathcal{A}_{CP} measurements in Table 7.10.

Comparing the various sources of systematic uncertainties listed in Tables 7.9-7.10 with the previous Belle result [49], we find that most of them like B counting, π^0/η veto, tracking, and CSBDT selection and K_S^0 selection efficiency have similar contribution for both Belle and Belle II. The systematics gets halved in the case of Belle II for hadron identification and photon selection. The π^0 selection efficiency has a significantly larger contribution to the systematics in the case of Belle II, compared to Belle. The fit bias and PDF systematics are found to be comparable to the Belle result.

Table 7.9: Systematic uncertainties (%) on the branching fraction and Δ_{0+} .

Source	$K^{*0}[K^+\pi^-]\gamma$	$K^{*0}[K_S^0\pi^0]\gamma$	$K^{*+}[K^+\pi^0]\gamma$	$K^{*+}[K_S^0\pi^+]\gamma$	$K^{*0}\gamma$	$K^{*+}\gamma$	Δ_{0+}
B counting	1.6	1.6	1.6	1.6	1.6	1.6	–
f^\pm/f^{00}	1.6	1.6	1.6	1.6	1.6	1.6	1.2
τ_{B^0}/τ_{B^+}	–	–	–	–	–	–	0.2
γ selection	0.9	0.9	0.9	0.9	0.9	0.9	–
π^0 veto	0.7	0.7	0.7	0.7	0.7	0.7	–
η veto	0.2	0.2	0.2	0.2	0.2	0.2	–
Tracking efficiency	0.5	0.5	0.2	0.7	0.5	0.5	–
pion selection	0.2	–	–	0.2	0.2	0.1	–
kaon selection	0.4	–	0.4	–	0.4	0.2	0.1
K_S^0 reconstruction	–	1.4	–	1.4	0.1	0.8	0.4
π^0 reconstruction	–	3.9	3.9	–	0.2	1.5	0.6
chiProb selection	0.2	1.0	0.2	1.0	0.2	0.7	0.3
BDT selection	0.3	0.4	0.4	0.3	0.1	0.2	0.2
BCS	0.1	1.0	0.5	0.2	0.1	0.2	0.1
Fit bias	0.1	0.3	0.2	0.2	0.1	0.1	0.1
Signal PDF model	0.2	0.4	0.3	0.2	0.2	0.2	0.1
KDE PDF model	0.2	0.4	0.4	0.3	0.1	0.2	0.2
Limited MC stat.	0.2	0.8	0.4	0.5	0.2	0.3	0.2
Self-crossfeed	0.2	2.2	1.8	0.5	0.1	0.8	0.4
Total	2.7	5.7	5.1	3.3	2.7	3.3	1.5

Table 7.10: Systematic uncertainties (%) on the \mathcal{A}_{CP} measurement.

Source	$\mathcal{A}_{CP}(K^{*0}[K^+\pi^-]\gamma)$	$\mathcal{A}_{CP}(K^{*+}[K^+\pi^0]\gamma)$	$\mathcal{A}_{CP}(K^{*+}[K_S^0\pi^+]\gamma)$	$\mathcal{A}_{CP}(K^{*+}\gamma)$	$\Delta\mathcal{A}_{CP}$
Fit bias	0.1	0.1	0.1	0.1	0.1
Signal PDF model	0.1	0.1	0.1	0.1	0.1
KDE modeling	0.1	0.4	0.2	0.2	0.2
BCS	0.1	0.5	0.2	0.3	0.3
K^+ asymmetry	–	0.6	–	0.3	0.3
π^+ asymmetry	–	–	0.6	0.3	0.3
$K^+\pi^-$ asymmetry	0.3	–	–	–	0.3
Total	0.4	0.9	0.7	0.6	0.7

Chapter 8

Results

Up to this point, we have discussed the theoretical motivation behind studying $B \rightarrow K^* \gamma$ decays, provided a brief description of the detector, and outlined the strategy to suppress background contributions. We have also described the fitting procedure used to extract signal yields, calculate the observables, and account for systematic uncertainties. In this section, we will discuss the strategy for unblinding the dataset and present the results obtained.

8.1 Background Modeling

To evaluate how well the simulation models the background, we compare the distributions of the M_{bc} and ΔE variables between data and MC events using both on-resonance sidebands and off-resonance samples. Initially, we compare the on-resonance continuum MC events with off-resonance data to ensure the continuum is well modeled. Once established, we compare the on-resonance sideband of data with MC events to assess the agreement of the $B\bar{B}$ background between the two samples.

8.1.1 Off-resonance data

Since the off-resonance dataset was collected 60 MeV below the $\Upsilon(4S)$ resonance energy, it predominantly contains background from continuum events. To check the modeling of the continuum component, we compare the distributions of the M_{bc} and ΔE between off-resonance data and on-resonance continuum MC events. The resulting plots are shown in Appendix C as Fig C.2, where the MC histogram is scaled to match the number of events in the data histogram. We observe good data-MC agreement for continuum events in M_{bc} and ΔE .

8.1.2 On-resonance Sideband

After confirming that the continuum background is well modeled, we proceed to evaluate the $B\bar{B}$ background. We compare the distributions of M_{bc} and ΔE after applying all selections for the on-resonance sideband. We have already ascertained that the continuum background is well modelled from checks performed with the off-resonance sample. The on-resonance sideband includes the contribution of both continuum and $B\bar{B}$ events, thus any discrepancy observed can be attributed to the modeling of $B\bar{B}$ background. For this study, the signal region is blinded with a ΔE selection of $\Delta E < -0.15$ or $0.07 < \Delta E$ GeV. The MC histograms are scaled to the number of events in the data histogram. The resulting plots are presented in Appendix C as Fig C.3. Additionally, we prepare another set of data-MC comparison plots by blinding the signal region using an $M_{bc} < 5.27$ GeV/ c^2 selection, which are shown in Appendix C as Fig C.4. We observe good agreement between data and simulation in the sideband for M_{bc} and ΔE .

8.2 Unblinding Strategy

After the sideband validation, we proceed to unblind the data. To ensure robustness of the results, before fully unblinding the signal region, we conduct further cross-checks. In particular, we compare the statistical uncertainties of observables obtained from the fit between the data and simulation while keeping the central values of the data results blinded. The results of this comparison are listed in Table 8.1. We find that the precision expected from the MC events is consistent with that obtained from the data. Next, we unblind a partial dataset, which is roughly one-fourth of the total luminosity. The measured values of observables with the partial dataset are listed in Table. 8.2. We observe that the results are consistent with the world average. Finally, we proceed to the full box opening.

Table 8.1: Comparison of statistical uncertainties (%) between data and simulation for various observables.

Mode	\mathcal{B} (MC)	\mathcal{B} (Data)	\mathcal{A}_{CP} (MC)	\mathcal{A}_{CP} (Data)
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	2.4	2.4	2.2	2.3
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	8.2	8.1	–	–
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	4.5	4.3	4.2	4.0
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	4.7	4.4	4.4	4.3

Table 8.2: Fit results for the partial dataset of 89 fb^{-1} luminosity. The quoted uncertainties are statistical only.

Mode	\mathcal{A}_{CP} (%)	$\mathcal{B} \times 10^5$
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	2.8 ± 4.7	4.09 ± 0.20
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	–	4.32 ± 0.63
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	14.3 ± 8.4	3.81 ± 0.35
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	-5.2 ± 8.3	4.19 ± 0.36

8.3 Results

After verifying the results with the partial dataset, we proceed to analyze the full data. The fit projections for the complete dataset are shown in Figs. 8.1 and 8.2. The measured values of the observables are listed in Table 8.3. We combine the branching fractions and CP asymmetries for the charged and neutral channels by calculating the weighted average, while taking into account the correlations among systematic uncertainties [151; 152]. The combination procedure is briefly described in Appendix A.

As a cross-check, we compare the $s\mathcal{P}$ lot background-subtracted distributions of the $\cos(\theta_{\text{hel}})$ variable from the data alongside those from signal MC sample. The result of this comparison is presented in Appendix C as Fig C.1. We observe good agreement between data and simulation.

Table 8.3: Fit results for full dataset equivalent to 365 fb^{-1} . The first uncertainty is statistical, the second is systematic, and the third uncertainty on Δ_{0+} is due to f^\pm/f^{00} .

Mode	\mathcal{A}_{CP} (%)	$\mathcal{B} \times 10^5$
$B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$	$-3.3 \pm 2.3 \pm 0.4$	$4.14 \pm 0.10 \pm 0.10$
$B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$	–	$4.07 \pm 0.33 \pm 0.23$
$B^0 \rightarrow K^{*0}\gamma$	$-3.3 \pm 2.3 \pm 0.4$	$4.14 \pm 0.10 \pm 0.10$
$B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$	$1.7 \pm 4.0 \pm 0.8$	$3.97 \pm 0.17 \pm 0.20$
$B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$	$-3.5 \pm 4.3 \pm 0.7$	$4.06 \pm 0.18 \pm 0.13$
$B^+ \rightarrow K^{*+}\gamma$	$-0.7 \pm 2.9 \pm 0.6$	$4.02 \pm 0.13 \pm 0.13$
	$\Delta\mathcal{A}_{CP}$ (%)	Δ_{0+} (%)
$B \rightarrow K^*\gamma$	$2.6 \pm 3.8 \pm 0.7$	$5.0 \pm 2.0 \pm 1.0 \pm 1.2$

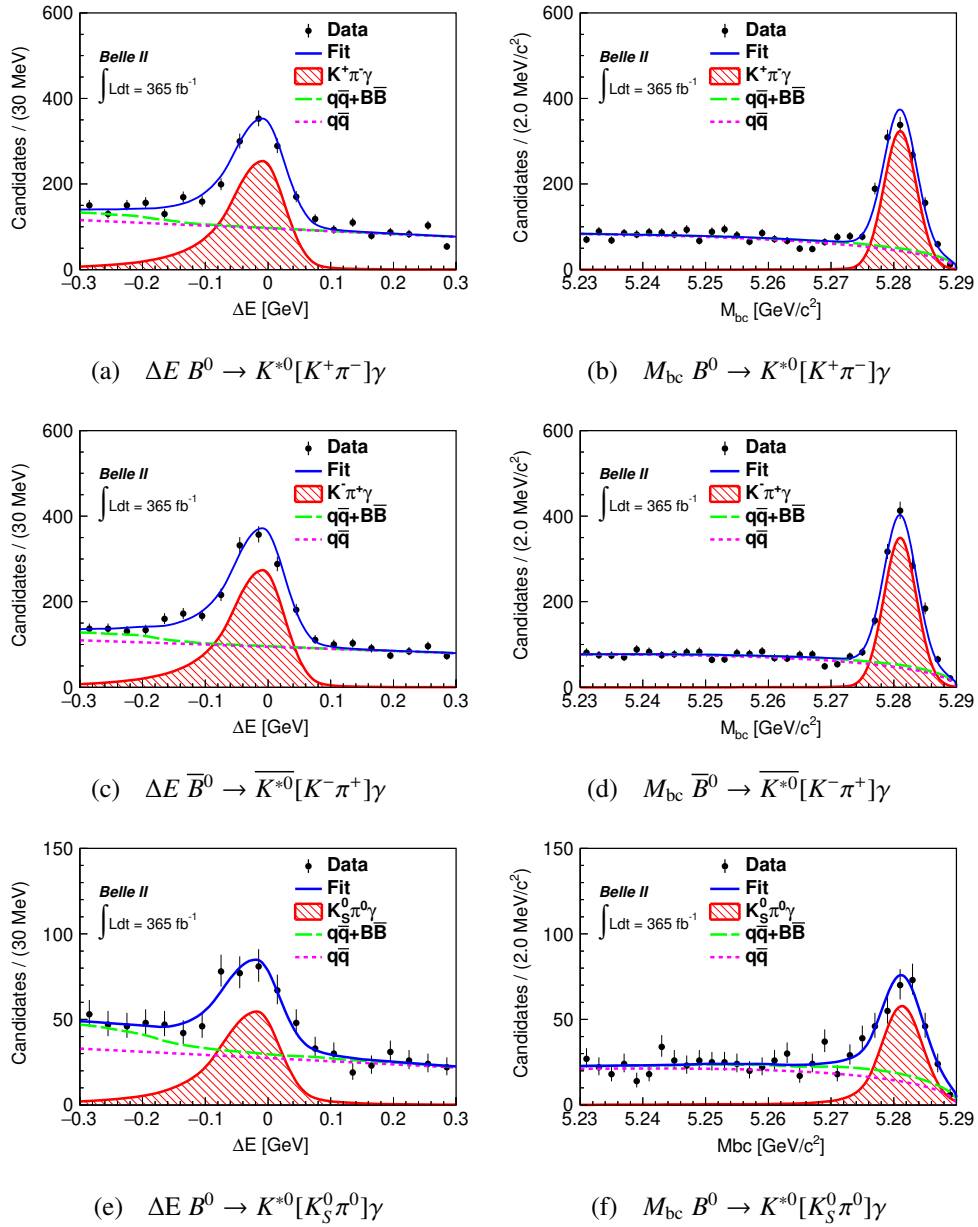


Figure 8.1: Fit projections for $M_{bc} - \Delta E$ distributions of $B^0 \rightarrow K^{*0}\gamma$ modes. The fit is performed on the full dataset, equivalent to 365 fb^{-1} . The fit results for (e-f) include both B^0 and \bar{B}^0 events

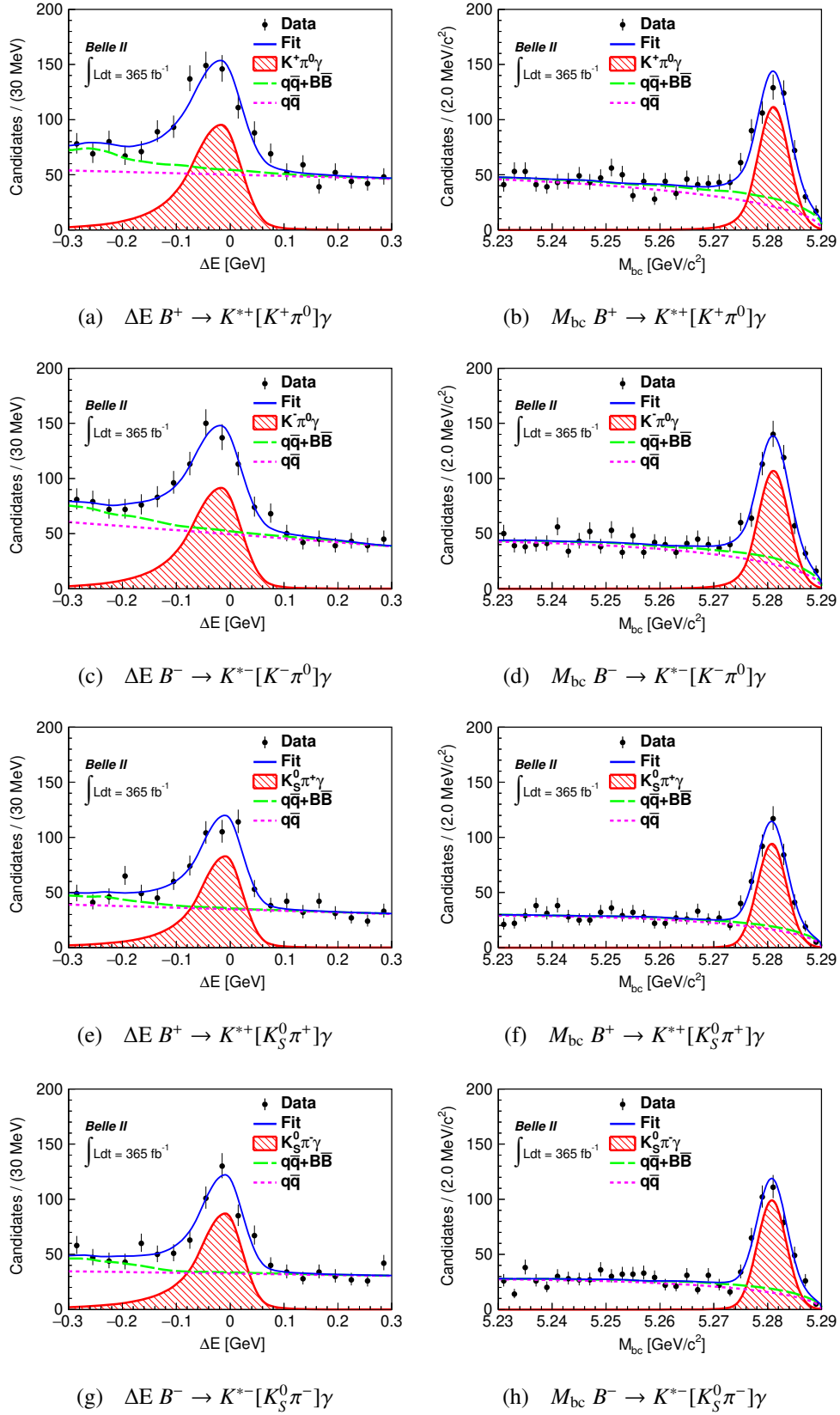


Figure 8.2: Fit projections for $M_{bc}-\Delta E$ distributions of $B^+ \rightarrow K^{*+}\gamma$ modes. The fit is performed on the full dataset, equivalent to 365 fb^{-1} .

A comparison of results with the world average values [117] is provided in Table. 8.4. We observe that measured values of branching fractions and CP asymmetries are in agreement with the world average.

Table 8.4: Comparison of results with PDG world averages. The first uncertainty is statistical, the second is systematic, and the third uncertainty on Δ_{0+} is due to f^\pm/f^{00} .

Observable	Belle II	PDG [117]
$\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) \times 10^5$	$4.14 \pm 0.10 \pm 0.10$	4.18 ± 0.25
$\mathcal{A}_{CP}(B^0 \rightarrow K^{*0}\gamma) (\%)$	$-3.3 \pm 2.3 \pm 0.4$	-0.6 ± 1.1
$\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) \times 10^5$	$4.02 \pm 0.13 \pm 0.13$	3.92 ± 0.22
$\mathcal{A}_{CP}(B^+ \rightarrow K^{*+}\gamma) (\%)$	$-0.7 \pm 2.9 \pm 0.6$	1.4 ± 1.8
$\mathcal{B}(B \rightarrow K^*\gamma) \times 10^5$	$4.10 \pm 0.08 \pm 0.10$	4.2 ± 0.6
$\mathcal{A}_{CP}(B \rightarrow K^*\gamma) (\%)$	$-2.4 \pm 1.9 \pm 0.3$	-0.3 ± 1.1
$\Delta_{0+} (\%)$	$5.0 \pm 2.0 \pm 1.0 \pm 1.2$	6.3 ± 1.7
$\Delta\mathcal{A}_{CP} (\%)$	$2.6 \pm 3.8 \pm 0.7$	2.4 ± 2.8

Chapter 9

Lepton identification with Belle II SVD

9.1 Introduction

“A method is more important than a discovery, since the right method will lead to new and even more important discoveries.”

– Lev Landau

Identification of charged particles produced in physics processes is crucial for any flavor physics experiment, such as Belle II and LHCb. The information from a number of subdetectors is combined to identify the final-state particles. At Belle II, we use a likelihood-based PID algorithm to combine information from various subdetectors towards identifying a charged track, e.g., lepton or hadron [45]. This algorithm heavily relies on the information provided by the ECL to identify electrons and the KLM to identify muons. A transverse momentum (p_T) threshold of 0.3 and 0.7 GeV/ c is required to reach the ECL and KLM, respectively. This requirement limits the lepton ID performance below the momentum threshold. This is where the specific ionization information from the CDC and SVD can play an important role, as it is capable of providing PID information in such a low-momentum region.

In this chapter, we describe how, by using the specific ionization information from the SVD, one can improve the identification performance for low-momentum electrons. Our study will aid the analyses involving such electrons in rare semileptonic decays of B mesons: $B \rightarrow K^{(*)}e^+e^-$ and $B \rightarrow K^{(*)}\tau^+\tau^-$, as well as in lepton-flavor violating decays of tau leptons: $\tau \rightarrow e\ell\ell$.

Figure 9.1 shows the generator-level p_T distributions of electrons from two such decays, where a good number of electron candidates are found to have a p_T less than $0.3 \text{ GeV}/c$.

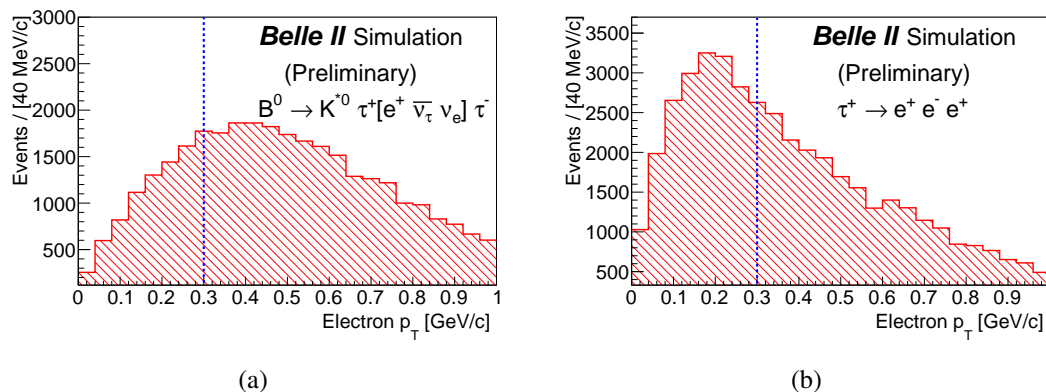


Figure 9.1: Generator-level p_T distributions for low-momentum electrons coming from (a) $B \rightarrow K^* \tau \tau$ and (b) $\tau \rightarrow e e e$ decays. The dashed vertical lines indicate the minimum p_T value required to reach the ECL.

9.2 Reconstruction and extraction of likelihood

Electrons originating from photon conversions ($\gamma \rightarrow e^+ e^-$) that occur within detector material of the two inner tracking systems, namely the PXD and SVD, are used to get the PID likelihood information of SVD. A converted photon candidate is reconstructed by combining two oppositely charged tracks assumed to be electrons. The sample purity is enhanced by applying a dielectron invariant mass criterion of $3 < M_{ee} < 15 \text{ MeV}/c^2$ and performing a vertex fit, where only photon candidates with a vertex fit $\chi^2 > 0.1\%$ are retained. The magnitude of the z coordinate of the photon production vertex is also required to be less than 8.0 cm to suppress secondary electrons. To purify the sample further, we optimize the FOM on the dr coordinate of the $e^+ e^-$ production vertex to select events arising from the vicinity of PXD and SVD material. The $s\mathcal{P}$ lot [141] technique is then used to subtract the residual background by performing a fit to the M_{ee} distribution. The $s\mathcal{P}$ lot extracted two-dimensional distribution of dE/dx vs. momentum is used as the SVD input to the total PID likelihood for electron tracks.

9.3 Performance

The total PID likelihood is constructed by combining information from all subdetectors for different particle hypotheses. This can be written as: $\mathcal{L} = \prod_{\text{det}} \mathcal{L}_{\text{det}}$, where the product is taken over the individual likelihoods of each subdetector. The SVD's contribution to the total likelihood is obtained as described in Ref. [153], with the likelihood for a particle hypothesis j

defined as:

$$\mathcal{L}_m(dE/dx, p) = \prod_i \mathcal{P}_m[(dE/dx)_i, p] \quad (9.1)$$

where m represents different particle types (e.g., e, π, K) and i runs over all dE/dx values assigned to a track.

We compare the performance of the binary PID variables $\mathcal{L}_{e/\pi} = \mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_\pi)$ and $\mathcal{L}_{e/K} = \mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_K)$, evaluated with and without the SVD dE/dx information for low-momentum tracks ($p < 1 \text{ GeV}/c$). To calculate the electron identification efficiency, we use electrons from the $\gamma \rightarrow e^+e^-$ sample. To determine the probability of a pion (kaon) being misidentified as an electron, we use pion (kaon) tracks from the $D^{*+} \rightarrow D^0[\rightarrow K^-\pi^+]\pi^+$ sample.

The efficiency for the PID variable to select a track under the particle hypothesis i is defined as $\epsilon_i = N_i/K_i$. Similarly, the probability of the PID variable to misidentify a track belonging to the particle hypothesis j as that to the hypothesis i is defined as $f_{j \rightarrow i} = N_i/K_j$. In both expressions, N_i denotes the number of tracks selected by the PID variable under the hypothesis i and K_i denotes the number of tracks kinematically selected under the hypothesis i .

Figure 9.2 shows the improvement in electron ID performance with the introduction of SVD information. We find the electron efficiency to increase from around 80 to 90% for a fixed 20% $\pi \rightarrow e$ misidentification probability. Similarly, for a fixed 40% $K \rightarrow e$ misidentification probability, the electron efficiency increases from around 94 to 98%.

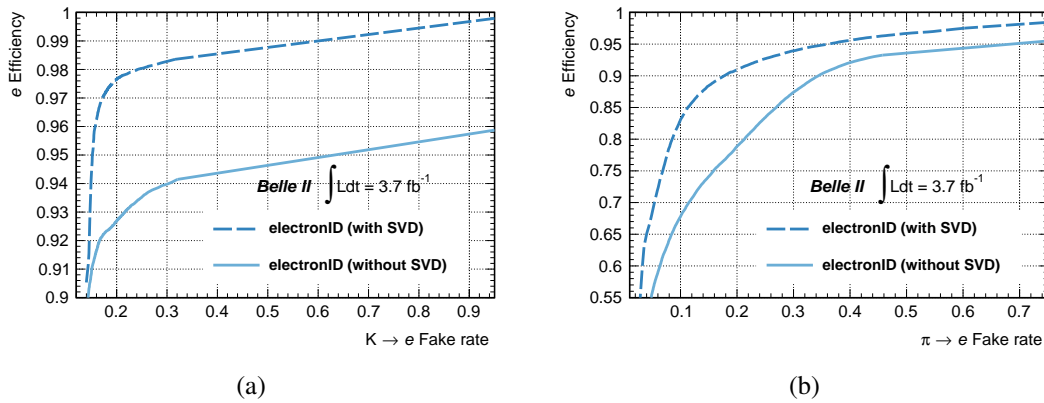


Figure 9.2: Electron efficiency vs. (a) $K \rightarrow e$ and (b) $\pi \rightarrow e$ misidentification probability for different criteria on the total PID likelihood.

9.4 Tests for muons

The performance of the SVD dE/dx information to distinguish muons and pions is studied with simulated samples where muons are reconstructed from the two-photon process $e^+e^- \rightarrow e^+e^-\mu^+\mu^-$ and pions from the $D^{*+} \rightarrow D^0[\rightarrow K^-\pi^+]\pi^+$ decay. In the four-lepton final state, the tracks corresponding to the electron and positron are mostly lost in the beam pipe, while the remaining two tracks are recorded in the detector. The dE/dx vs. p distributions for various charged particles obtained from simulated samples are shown in Fig. 9.3(a). Here, the band for the electron hypothesis is obtained from a simulated photon conversion sample.

For simulated charged particles of $p < 1 \text{ GeV}/c$, the dE/dx distributions are obtained in the momentum bins of $50 \text{ MeV}/c$. The distributions are fitted with a Gaussian function to obtain the mean (m) and width (w) values, which are used to define a lepton (i) vs. hadron (j) separation metric as $s_{i/j} = (m_i - m_j) / \sqrt{w_i^2 + w_j^2}$. The metric plotted as a function of the particle momentum is shown in Fig. 9.3(b). It shows that the muon and pion tracks cannot be well separated in the SVD, owing to their small mass difference (only $30 \text{ MeV}/c^2$). Therefore, the same dE/dx vs. momentum distribution can be used as the SVD information for the total PID likelihood for both muons and pions. The dE/dx vs. momentum distribution for pions is obtained using $D^{*+} \rightarrow D^0[\rightarrow K^-\pi^+]\pi^+$, which is an experimentally clean and high statistics sample.

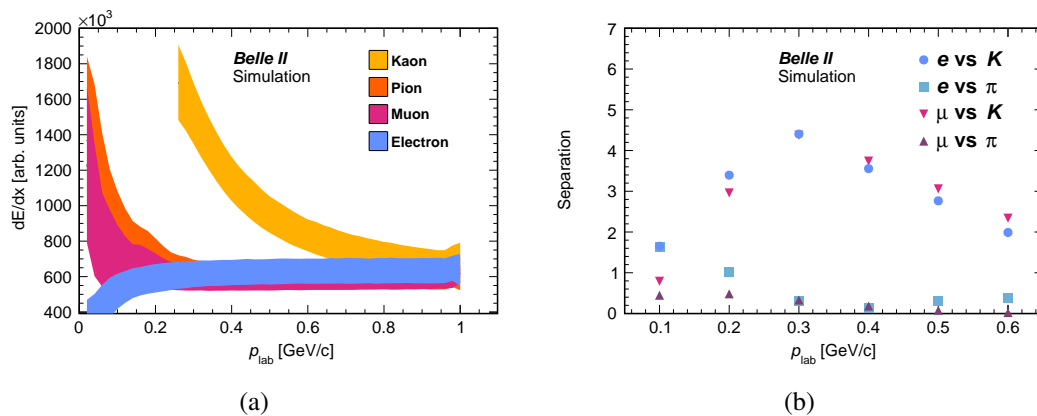


Figure 9.3: SVD-based (a) simulated dE/dx vs. p distributions for leptons and hadrons and (b) separation metric $s_{i/j}$ as a function of p . The dE/dx distributions for charged particles are obtained in bins of momentum of $100 \text{ MeV}/c$ width.

Chapter 10

Summary

10.1 Discussion

The core of this thesis is a measurement of $B \rightarrow K^*\gamma$ decays with the Run 1 dataset from the Belle II experiment. We begin by discussing the motivation behind studying these decays, followed by a brief description of the Belle II detector. The procedure for reconstructing the B decay is detailed along with strategies to suppress background contributions. To measure the observables, a maximum likelihood fit is performed on the kinematic variables M_{bc} and ΔE . Systematic effects are studied to estimate the uncertainties on the obtained observables.

Earlier, Belle II measured the branching fraction of $B \rightarrow K^*\gamma$ decays by fitting the unbinned ΔE distribution with 63 fb^{-1} of data [50]. Based on that preliminary result and the published Belle measurement [49], an estimate is made about how the relative precision of these observables will change for different luminosity scenarios at Belle II [154]. We find that the precision expected for CP asymmetries will continue to improve until it surpasses the precision of the SM prediction. On the other hand, the isospin asymmetry will remain strongly sensitive to BSM contributions up to the final dataset size.

The study presented in this thesis supersedes the preliminary results on branching fractions [50], with significant improvements coming from the background photon rejection and BDT-based continuum suppression. A two-dimensional fit to M_{bc} and ΔE yields both CP asymmetries and branching fractions with excellent control on backgrounds. Variables sensitive to BSM effects, such as isospin asymmetry and $\Delta\mathcal{A}_{CP}$, are also measured. The use of dedicated control samples helps to calculate systematic uncertainties more accurately. The results are

consistent with world average values [117] as well as SM expectations. Importantly, they exhibit comparable sensitivity to the Belle result, with only half of the dataset.

In the last chapter, we discuss a method to improve the performance of the identification algorithm for low-momentum electrons using dE/dx information from the SVD. This study will help with analyses involving low-momentum electrons, such as $B \rightarrow K^* \tau \tau$ decays.

10.2 Future Work

The thesis presents an improved measurement of observables for $B \rightarrow K^* \gamma$ decays using Belle II data. There are still opportunities for improvement. One area is the fitting variable ΔE , which exhibits a significant tail due to energy leakage in the ECL. The data-MC agreement could have been better for ΔE if correction factors from $e^+e^- \rightarrow \mu^+\mu^-\gamma$ decays were used to calibrate the energy of the prompt photons. This would in turn have made the analysis more solid. Secondary particles namely π^0 and K_S^0 are currently reconstructed with various kinematic selections. Another way to make selections work better is to use MVA methods based on variables such as the invariant mass of secondary particles and the kinematic properties of their decay products. For example, the Belle analysis used a likelihood-based π^0/η veto, whereas the Belle II implementation with MVA methods showed superior performance. The present study employs a two-dimensional fit model using ΔE and M_{bc} ; future improvements could include adding the cosine of the helicity angle as the third dimension in the fit model. This would help isolate nonresonant contributions better, improving the separation of background from B decays.

As for the observables, the statistical and systematic uncertainties are comparable for the branching fraction measurements. To make these results more accurate, future studies should focus on reducing systematic contributions, especially from sources like π^0 and K_S^0 reconstruction and B counting. For CP asymmetries, further improvements are anticipated as more data are collected, since statistical uncertainties are the dominant source of uncertainty in the current analysis. Additionally, the systematic uncertainty in isospin asymmetry could be reduced through a more precise determination of f^\pm/f^{00} .

Not only are $B \rightarrow K^* \gamma$ decays a prime candidate for precision tests of the SM, but they also serve as a key calibration channel for other radiative penguin decays of B mesons. This highlights the importance of revisiting analyses of these decays with larger datasets. System-

atic uncertainties arising from sources such as the π^0/η veto, continuum suppression MVA, or vertex fitting could be better constrained by using $B \rightarrow K^*\gamma$ decays as a control sample. Specifically, for $B \rightarrow \rho\gamma$ decays—where the transition is further suppressed by CKM factors—the former constitutes a dominant peaking background. A simultaneous analysis of these two decay modes, where the peaking contribution of $B \rightarrow K^*\gamma$ is directly estimated based on the $K \rightarrow \pi$ misidentification rate in the data, could lead to more precise results. This contrasts the current analysis strategy of $B \rightarrow \rho\gamma$ where the yield of $B \rightarrow K^*\gamma$ is fixed to the value calculated using the world average value of branching fractions [117].

Appendix A

Statistical tools

“It is often said that the language of science is mathematics. It could well be said that the language of experimental science is statistics.”

– Kyle Cranmer [155]

A typical analysis in high energy physics requires the use of a variety of statistical tools [156–159]. These tools are employed for a broad range of cases, starting from the discrimination based on likelihoods and multivariate methods, to calculating the variance and combining measurements. In this section, we provide a brief description of the statistical tools used in our study.

A.1 Likelihood discriminator

The likelihood discriminator [156–159] is widely used in experimental high-energy physics. It is defined as a function that calculates the probability of observing a given set of data under different hypothetical parameter values. Mathematically, if X represents the observed data and θ represents the parameters of a statistical model, the likelihood function $\mathcal{L}(\theta|X)$ is given by the probability $\mathcal{P}(X|\theta)$. This function plays a crucial role in parameter estimation.

Often, the likelihood function is used as a test statistic to distinguish between different hypotheses. For instance, when searching for a new particle, physicists compare the likelihoods between the null hypothesis (no new particle) and the alternative hypothesis (presence of a new particle). The ratio of the two likelihoods, known as the likelihood ratio, provides a powerful

test statistic for hypothesis testing. By comparing the observed likelihood ratio to its expected distribution under the null hypothesis, physicists can quantify the significance of the observations and determine whether they have discovered new particle.

Wilks' theorem [160] plays a significant role in this context, providing a theoretical foundation to use the likelihood ratio as a test statistic. According to the theorem, under the null hypothesis, the distribution of the likelihood-ratio test statistic asymptotically follows a χ^2 distribution, with the degrees of freedom equal to the difference in the number of parameters between the two hypotheses. This result is important because it allows physicists to determine the p-value or significance level of their test without relying on extensive MC simulations. By leveraging Wilks' theorem, experimentalists can more efficiently and accurately assess the evidence for new phenomena in their data.

A.1.1 Particle identification

The likelihood estimators are also extensively used in particle identification [45], where information from various subdetectors is combined to distinguish between different types of particles. Each subdetector provides its own likelihood for a particle hypothesis based on its specific measurements e.g., energy deposits, track curvature, timing information. These individual likelihoods are then combined to form an overall likelihood for each particle hypothesis. For example, in an experiment like Belle II, the likelihoods from the SVD, CDC, and ECL can be combined to create a comprehensive likelihood to accurately identify electrons.

The likelihood ratio can be transformed into a variable bounded between zero and one, which is used as a PID variable. We take the logarithm of the likelihood ratio for two particle hypotheses, "a" and "b", as

$$t(a/b) = \log(\mathcal{L}_a/\mathcal{L}_b)$$

The logarithm of the likelihood ratio is an equally effective test statistic, as Wilks' theorem can be applied in such cases. Since $t(a/b)$ ranges from negative infinity to positive infinity, we use the sigmoid function to convert $t(a/b)$ into a variable bounded between zero and one:

$$\mathcal{P}(a/b) = \frac{1}{1 + e^{-t(a/b)}} = \frac{1}{1 + e^{-\log(\mathcal{L}_a/\mathcal{L}_b)}} = \frac{\mathcal{L}_a}{\mathcal{L}_a + \mathcal{L}_b}$$

The value of $\mathcal{P}(a/b)$ is close to zero if the likelihood for hypothesis "a" is small and that for

hypothesis "b" is large. Similarly, its value is close to one if the converse is true. This variable acts as the PID variable to identify particles according to a given hypothesis.

A.1.2 Maximum-likelihood fit

The maximum-likelihood estimation is another standard statistical method that is used to estimate the parameters of a model by maximizing the likelihood function. The likelihood represents the probability of observing the given data as a function of the model parameters. In practice, we often work with the negative log-likelihood because it transforms the product of probabilities into a sum, making the optimization process more straightforward computation-wise. To find the parameter estimates, we minimize the negative log-likelihood using numerical optimization techniques. The uncertainties on the fitted parameters are derived from the second derivative (the Hessian matrix) of the negative log-likelihood function at the optimum. Specifically, the inverse of the Hessian matrix provides an estimate of the covariance matrix of the parameter estimates, from which the standard errors (uncertainties) can be extracted. This method is widely used in high-energy physics due to its statistical efficiency and ability to handle complex models and large datasets.

A.2 Background subtraction

The *sPlot* technique [141] is a background subtraction method employed in high-energy physics to obtain the distribution of kinematic variables for the signal component. The core idea behind *sPlot* is to use a discriminating variable that can effectively separate signal from background through a fitting procedure. Ideally, this variable should be uncorrelated with the target variable whose shape we want to discern, ensuring that the fit does not introduce biases in the distribution.

Once the fit is performed on the discriminating variable, *sWeights*, or signal weights, are calculated. By applying *sWeights* to the distribution of target variables, one can effectively separate signal from background. The *sWeight* (w_i) for signal, denoted as the i^{th} component in the fit, is defined as:

$$w_i(x) = \frac{\sum_j V_{ij} f_j(x)}{\sum_k N_k f_k(x)}$$

where V_{ij} denotes the covariance matrix obtained from the fit, $f_j(x)$ is the PDF for the j^{th} component, and N_k is the yield of the k^{th} component. In this manner, the *sPlot* technique allows for the extraction of the distribution of target variables for signal.

A.3 Multivariate methods

Multivariate methods are crucial in high-energy physics to discriminate between signal and background [157], especially given the complex and high-dimensional nature of the data. One widely used technique is the decision tree, a supervised learning method that divides data into subsets based on feature values to make predictions. A decision tree splits the feature space into regions based on decision rules, creating a tree-like structure where each internal node represents a feature test, each branch represents a test outcome, and each leaf node represents a class label or prediction. This method allows for a clear and interpretable model of how different features influence the classification outcome.

However, a single decision tree, while straightforward, often struggles with overfitting and limited predictive power. It is considered a “weak learner” because its performance can be sub-optimal due to its simplicity and tendency to capture noise in the data rather than generalizable patterns. To address these limitations, more advanced techniques, like gradient boosting [161], are employed.

Stochastic Gradient Boosted Decision Trees (GBDTs) [162] enhance predictive power by building an ensemble of decision trees in a sequential manner. In this approach, each tree corrects the errors made by the previous ones. GBDT involves training decision trees on residuals—the difference between the actual outcomes and the predictions of the ensemble—rather than directly on the original data. Additionally, stochastic methods introduce randomness into the training process by subsampling the data, which helps mitigate overfitting and improve generalization. The combined effect of these trees, often trained with gradient descent to minimize a loss function, results in a robust and highly accurate model for distinguishing between signal and background events. This technique is particularly effective in high-energy physics experiments, where the ability to separate a rare signal from copious background events is key for precision measurements or discovery. Most Belle II analyses use a speed optimized and cache-friendly implementation of GBDT [134], which enhances the traditional GBDT by focusing on computational efficiency without sacrificing classification performance.

A.4 Variance for observables

When dealing with a relation of the form of $y = f(x_i)$, where x_i represents measurements with associated uncertainties and y is a function of these measurements, the uncertainty (vari-

ance) on y can be obtained using an approximate variance formula relying on the first-order second moment method [156–159]. The challenge of finding the variance for a nonlinear function y is overcome by linearizing the function about its mean and expanding it in terms of a Taylor series. The uncertainty (σ_y) is calculated from the covariance matrix of measurements (V), and the weight vector comprising partial derivatives of the function $W = (\frac{\partial f}{\partial x_1} \dots \frac{\partial f}{\partial x_n})$ using the expression:

$$\sigma_y^2 = W \times V \times W^T \quad (\text{A.1})$$

For a general case, where the measurements exhibit some level of correlation, the covariance matrix is given by:

$$V = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad (\text{A.2})$$

with ρ representing the correlation coefficient between the measurements.

When dealing with measurements affected by various independent sources of uncertainty, it is possible to construct separate covariance matrices for each of them. The total covariance matrix is then obtained by simply adding these individual covariance matrices together. Two primary sources of uncertainty are statistical and systematic. The covariance matrix for statistical uncertainty, denoted as V_{stat} , is a diagonal matrix:

$$V_{\text{stat}} = \begin{pmatrix} \sigma_1^2(\text{stat}) & 0 \\ 0 & \sigma_1^2(\text{stat}) \end{pmatrix} \quad (\text{A.3})$$

The covariance matrix of systematic uncertainty is obtained by summing over the covariance matrix of individual systematic components,

$$V_{\text{sys}} = \sum_i V_{\text{sys}}^i \quad (\text{A.4})$$

Some systematic uncertainties are taken to be fully correlated across the different signal modes; in those cases we fix $\rho = 1$. Examples of such correlated systematic sources include π^0/η veto, photon reconstruction, and hadron identification, where each mode uses correction factors derived from the same control sample study. In contrast, sources such as the application of BDT, fit bias, and limited simulation sample are treated as uncorrelated sources of uncertainty and we take $\rho = 0$.

The total covariance matrix V is then simply the sum of V_{stat} and V_{sys} . Once we have the full covariance matrix, we can calculate the variance on the observable σ_y as well as decompose the contribution of individual sources to the total variance using the formula:

$$(\sigma_y^i)^2 = W \times V_i \times W^T \quad (\text{A.5})$$

To estimate the uncertainty on isospin asymmetry, we go back to its definition:

$$\Delta_{0+} = \frac{(\tau \times f \times N^0) - N^+}{(\tau \times f \times N^0) + N^+} \quad (\text{A.6})$$

where, we have taken $\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = N^0/f^{00}$ and $\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = N^+/f^{\pm}$; as well as substituted $\tau = \frac{\tau_{B^0}}{\tau_{B^+}}$ and $f = \frac{f^{\pm}}{f^{00}}$.

The partial derivatives are:

$$\frac{\partial \Delta_{0+}}{\partial \tau} = \frac{(1 - \Delta_{0+}) \times f \times N^0}{(\tau \times f \times N^0) + N^+} \quad (\text{A.7})$$

$$\frac{\partial \Delta_{0+}}{\partial f} = \frac{(1 - \Delta_{0+}) \times \tau \times N^0}{(\tau \times f \times N^0) + N^+} \quad (\text{A.8})$$

$$\frac{\partial \Delta_{0+}}{\partial N^0} = \frac{\tau \times f \times (1 - \Delta_{0+})}{(\tau \times f \times N^0) + N^+} \quad (\text{A.9})$$

$$\frac{\partial \Delta_{0+}}{\partial N^+} = \frac{-(1 + \Delta_{0+})}{(\tau \times f \times N^0) + N^+} \quad (\text{A.10})$$

A.5 Combining measurements

When we have a set of independent measurements, denoted as x_i , for a particular physical quantity, we can create a linear combination \hat{X} by summing these measurements with associated coefficients β_i , expressed as $\hat{X} = \sum_i^N \beta_i x_i$. Among the various linear combinations, we need to identify the best linear unbiased estimate (BLUE) [151; 152] for the given physical quantity.

To ensure that the estimate (\hat{X}) remains unbiased, it is important that the coefficients β_i

satisfy the condition $\sum_i^N \beta_i = 1$, assuming that the individual measurements themselves are unbiased. The variance of this estimate, denoted as σ_X^2 , can be calculated using the error propagation formula. The variance of the estimate takes the form:

$$\sigma_X^2 = \beta^T \times V \times \beta, \quad (\text{A.11})$$

where β represents the weight vector containing the entries β_i and V denotes the covariance matrix associated with the measurements.

To achieve the best estimate, it is necessary to determine a set of weights, β_i , that minimizes the variance (σ_X) while adhering to the constraint of $\sum_i^N \beta_i = 1$. The optimization process is equivalent to construct a weighted sum of squares, given by:

$$\chi^2 = \sum_i^N \sum_j^N (x_i - \hat{X}) \times (V^{-1})_{ij} \times (\hat{X} - x_j), \quad (\text{A.12})$$

and to minimize it with respect to the estimate \hat{X} . The resulting weight vector β that satisfies these criteria:

$$\beta = V^{-1} \times U \times (U^T \times V \times U)^{-1}, \quad (\text{A.13})$$

where U is an $N \times 1$ column vector with all entries set to 1. Once we have determined β , we can calculate \hat{X} and then apply the error propagation formula to determine the uncertainty on \hat{X} as well as break it down into individual sources.

Appendix B

Variables for continuum suppression

The relative importance of CSBDT input variables for all four modes are shown in Fig B.1. The density plots of the CSBDT training variables are shown in Figs B.2-B.5. The histograms of density plots are normalized to unit area for the comparison.

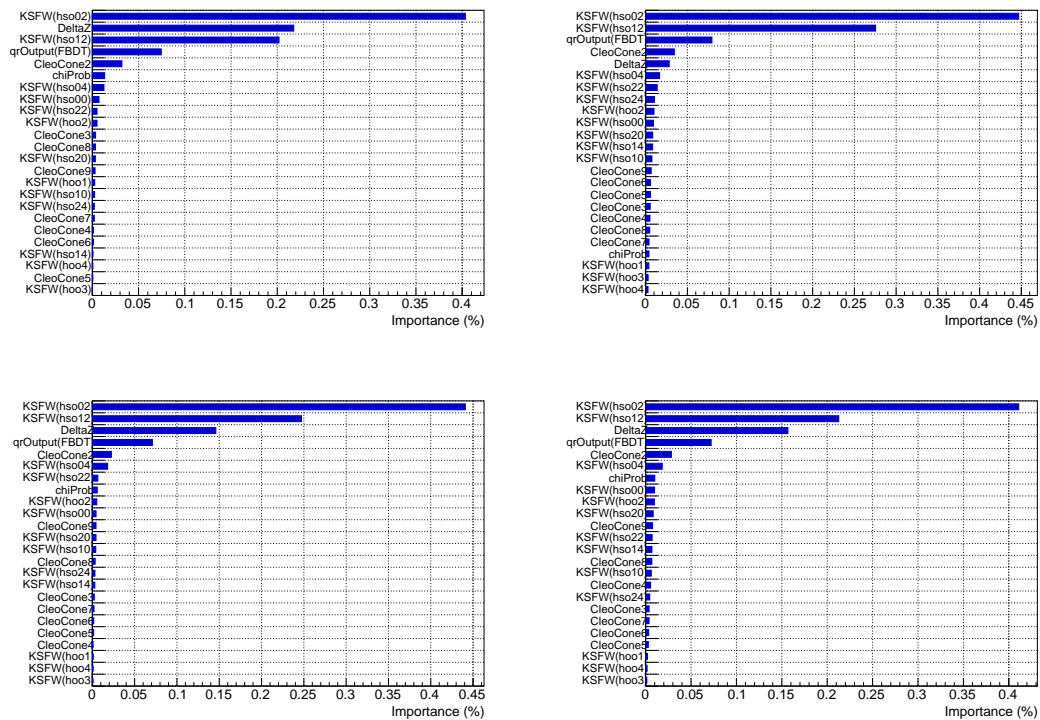


Figure B.1: Variable importance

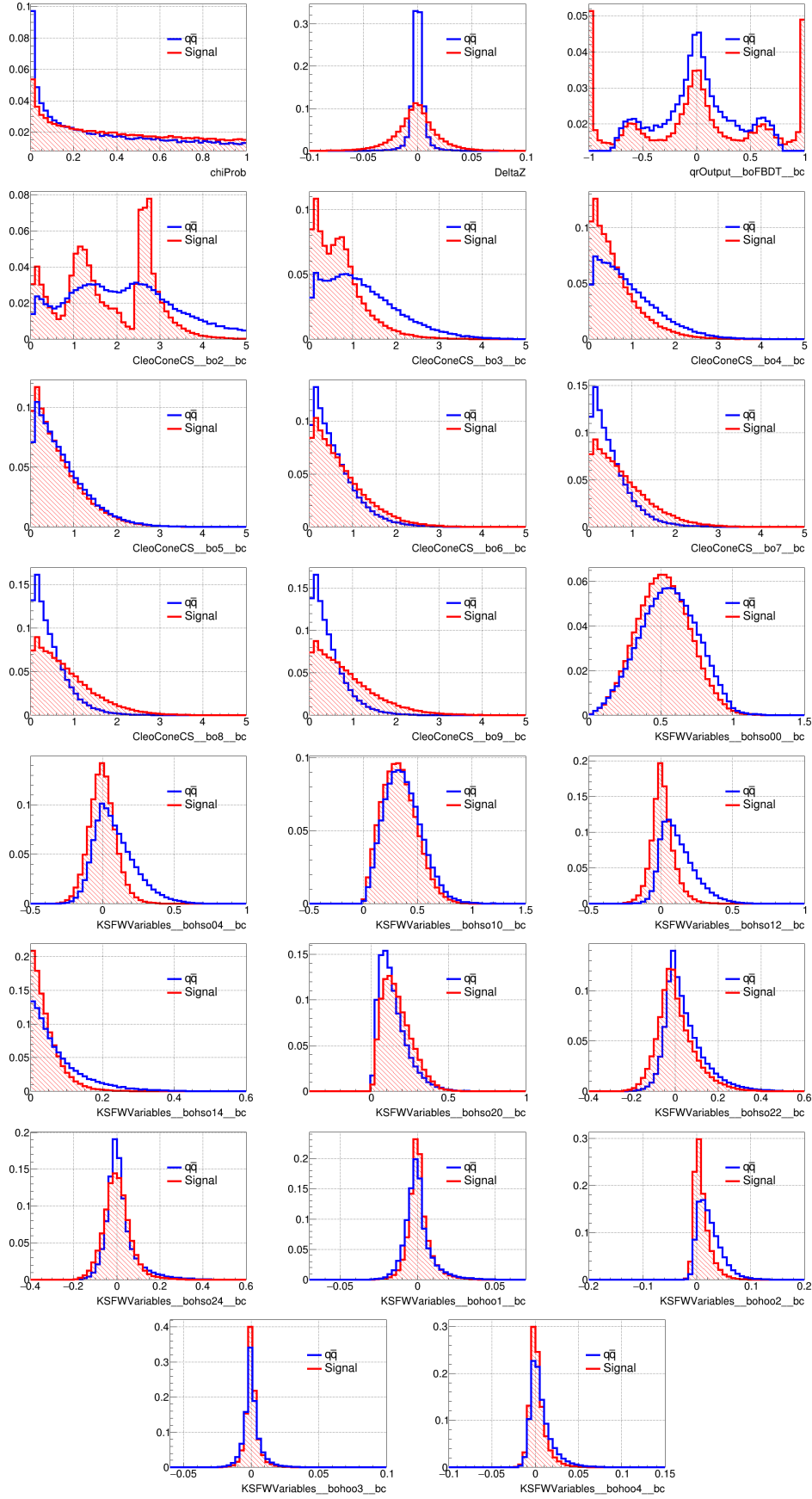


Figure B.2: Distribution of CSBDT input variables for $B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$ modes .

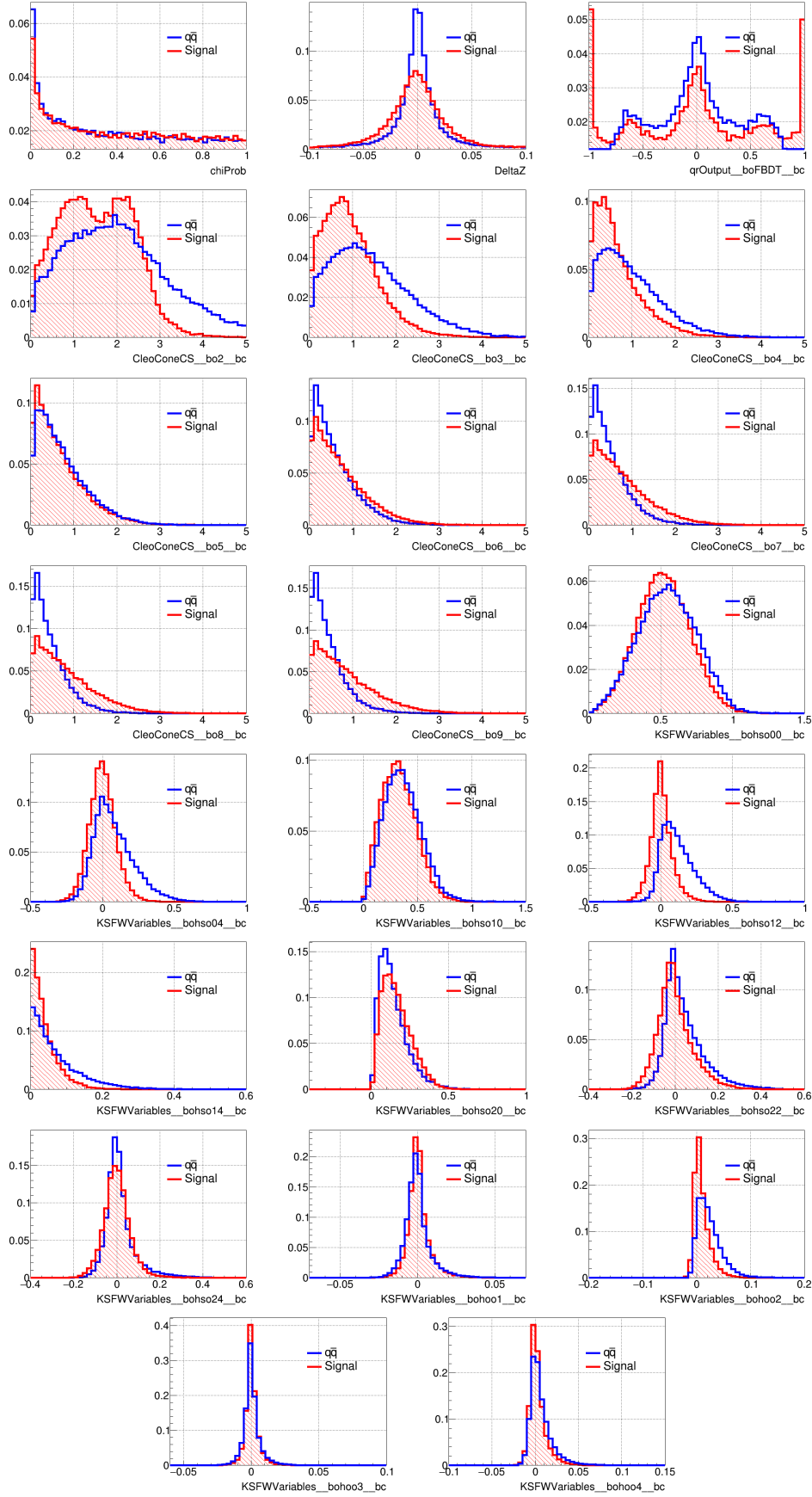


Figure B.3: Distribution of CSBDT input variables for $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$ modes .

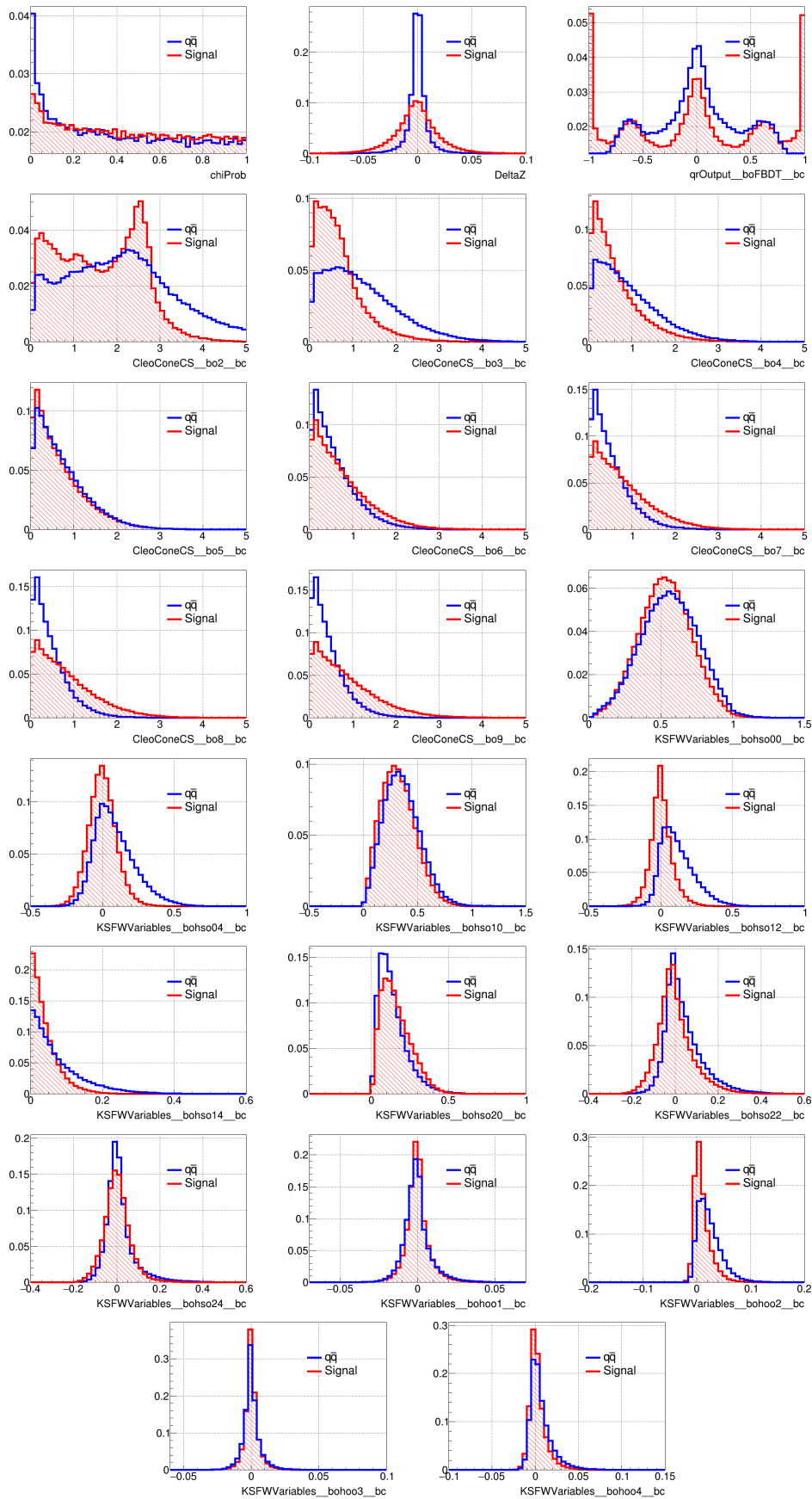


Figure B.4: Distribution of CSBDT input variables for $B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$ modes .

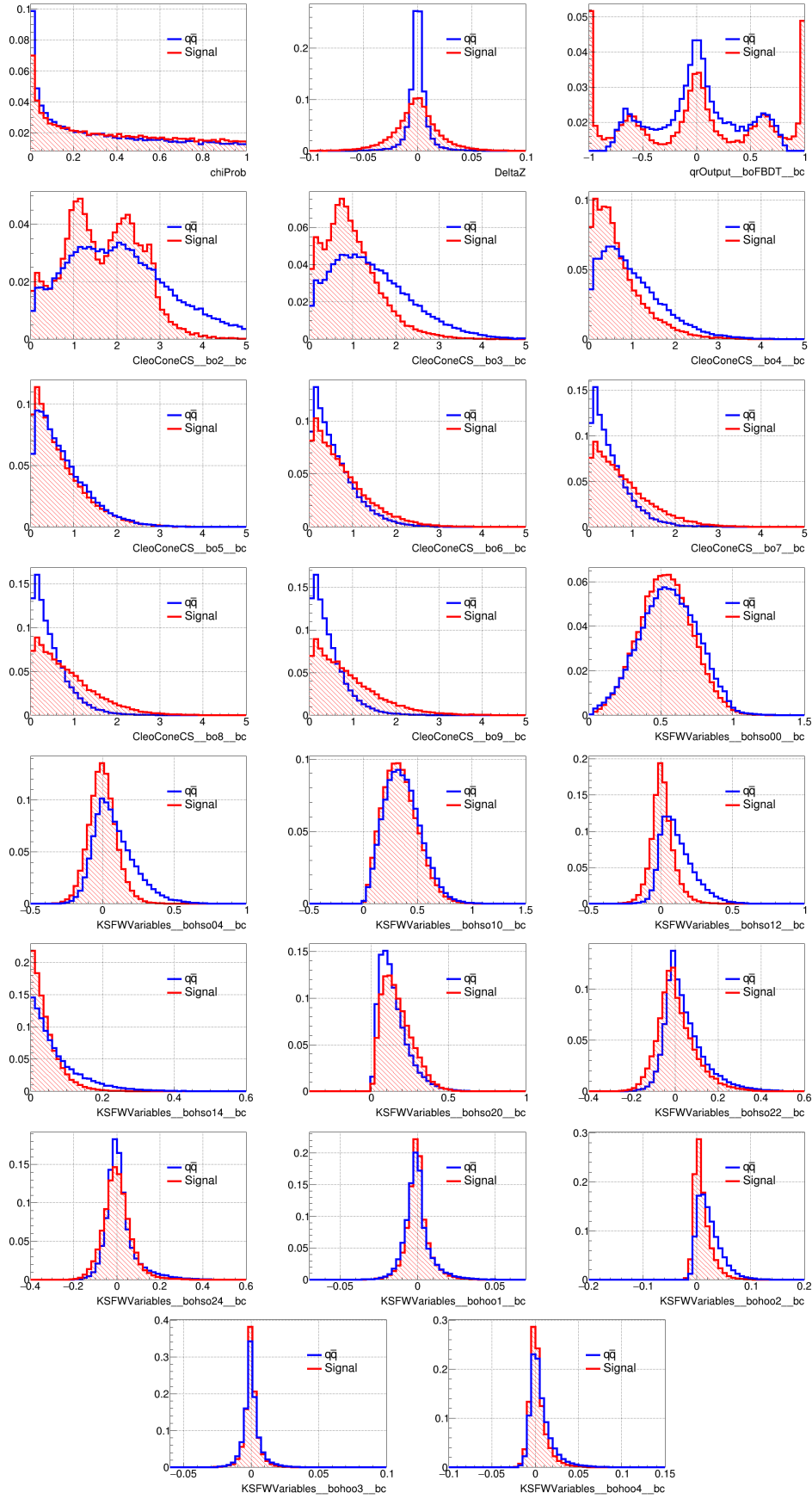


Figure B.5: Distribution of CSBDT input variables for $B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$ modes .

Appendix C

Plots for various cross-checks

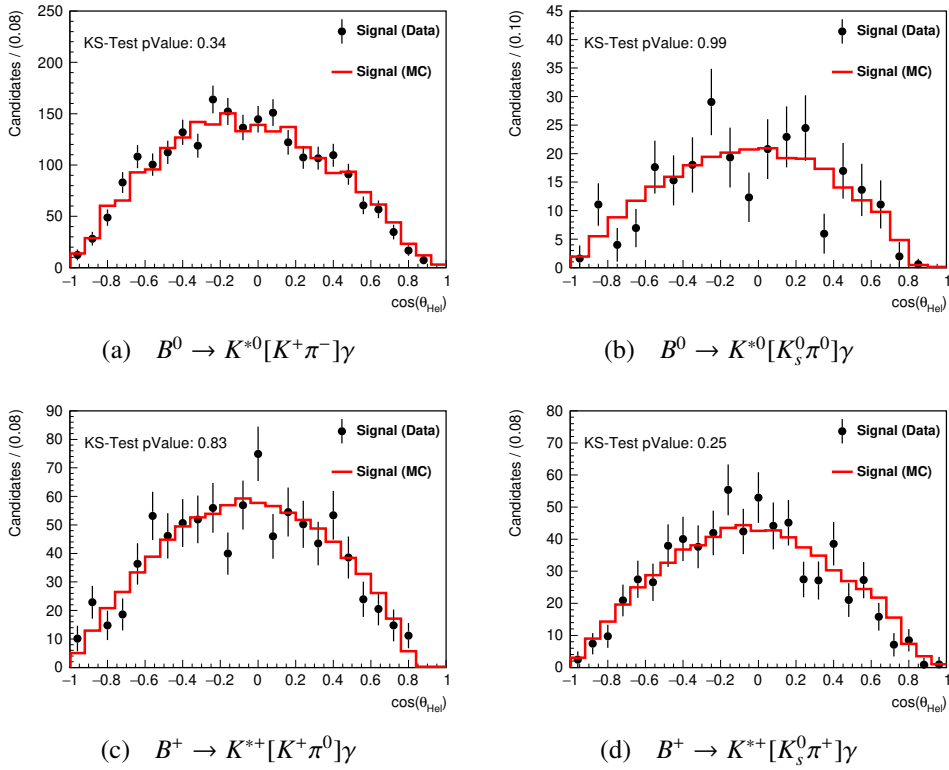


Figure C.1: Comparison between \mathcal{P} Plot background subtracted distribution of $\cos\theta$ helicity variable (blue data points with error) for data ($\int \mathcal{L} = 365 \text{ fb}^{-1}$) with that obtained from correctly reconstructed signal events (red solid line) of MC sample.

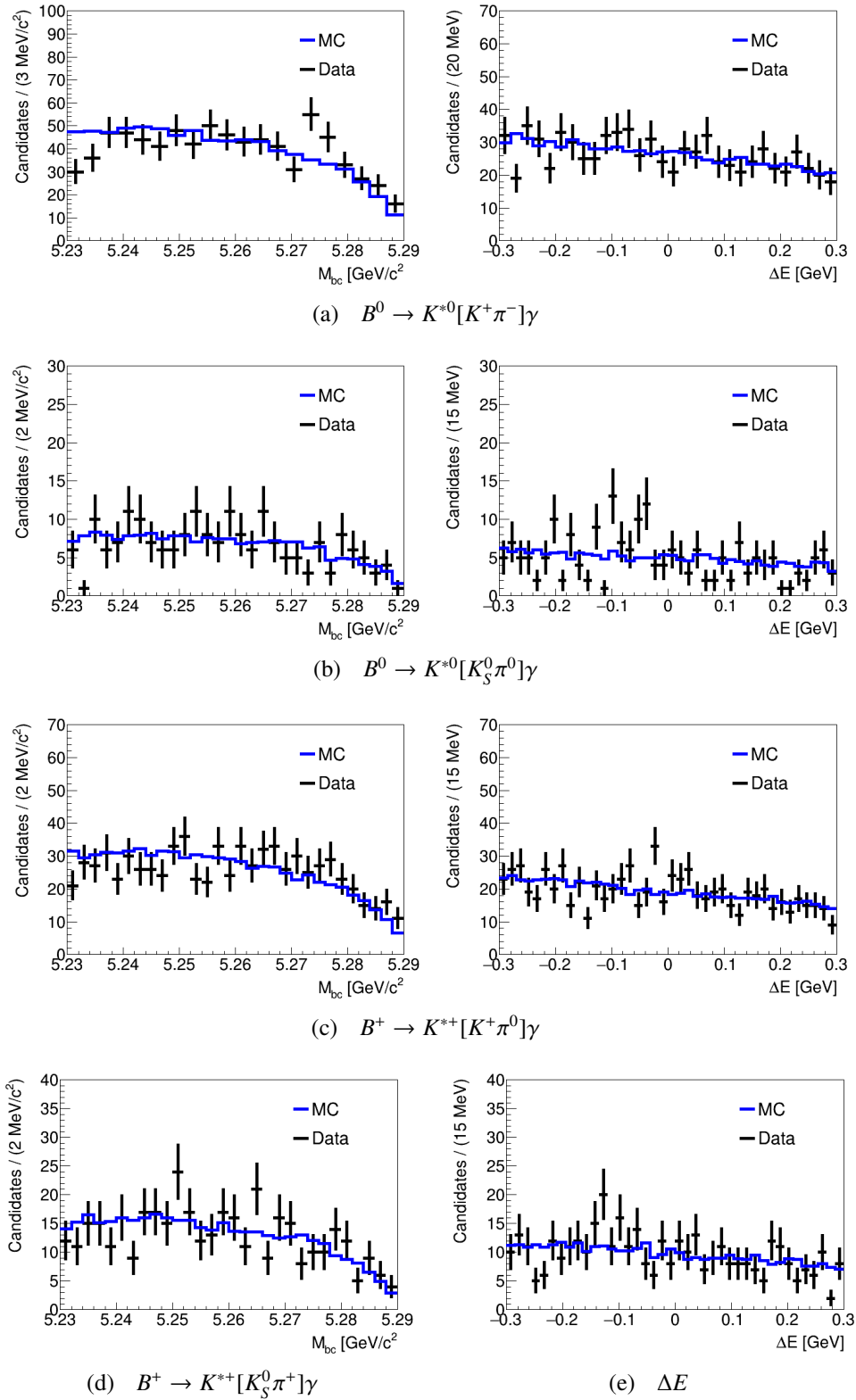


Figure C.2: Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes in the off-resonance sideband. The MC histogram is scaled down to the number of events in the data histogram.

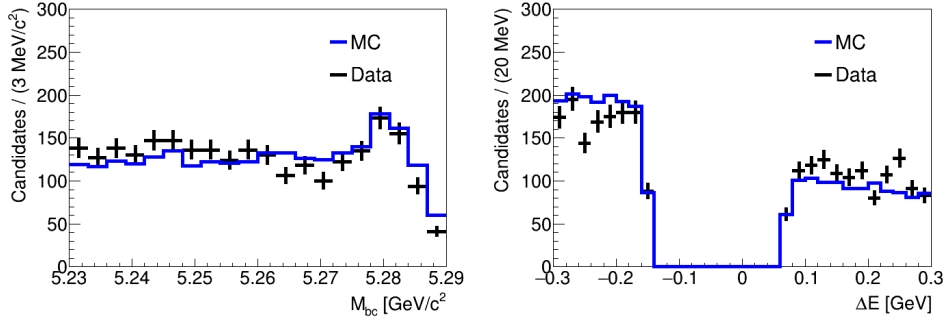
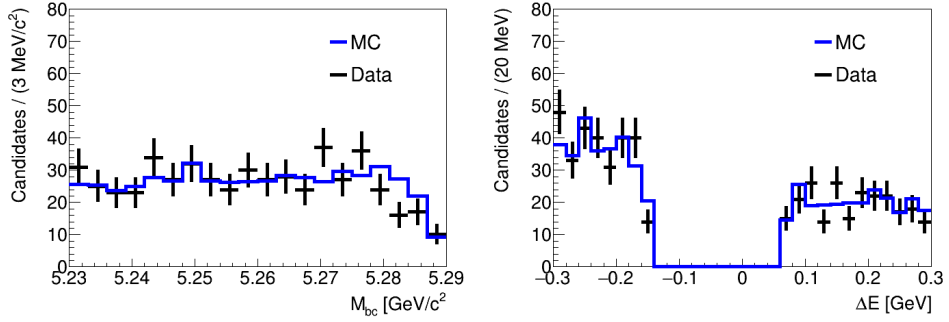
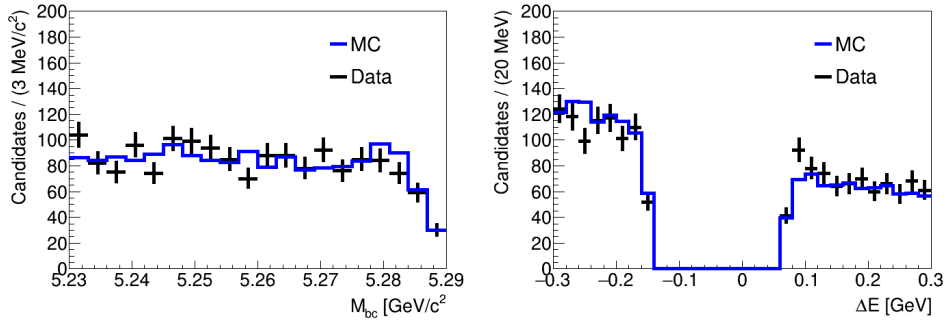
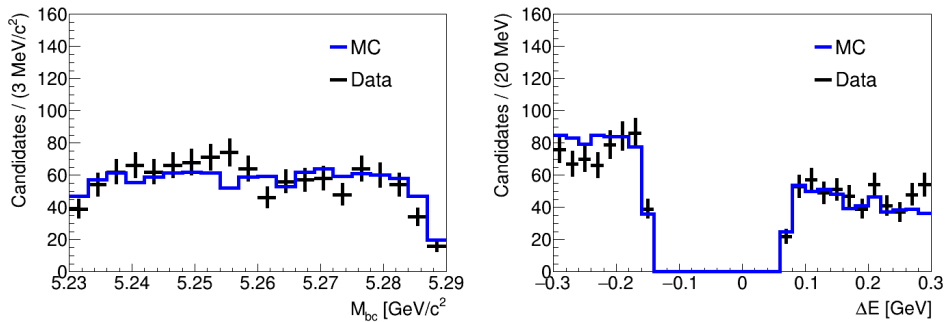

 (a) $B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma$

 (b) $B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma$

 (c) $B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma$

 (d) $B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma$

Figure C.3: Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes in the on-resonance sideband. Here we blind the signal region with a ΔE selection of $\Delta E < -0.15 \text{ || } 0.07 < \Delta E$. The MC histogram is scaled down to the number of events in the data histogram.

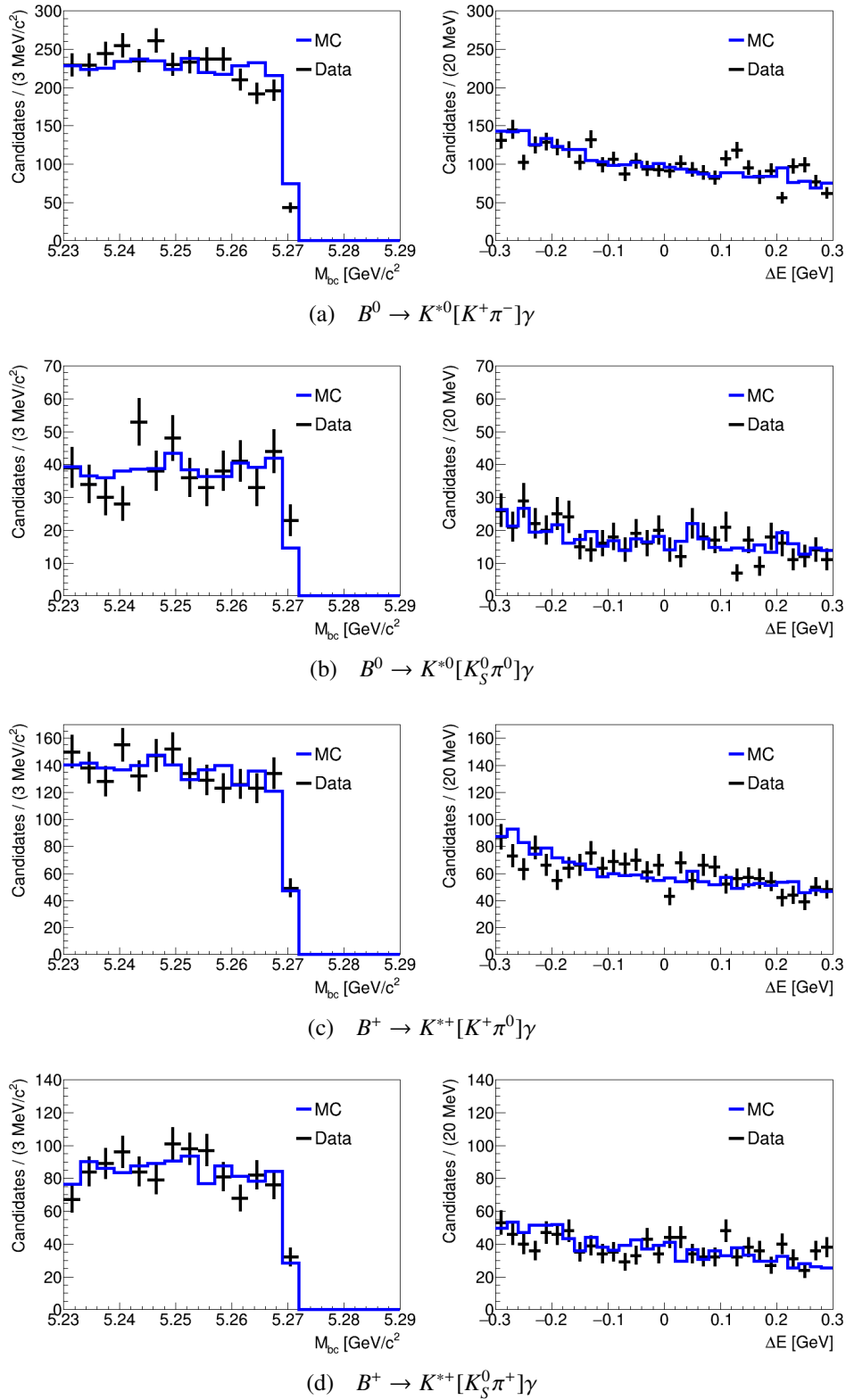
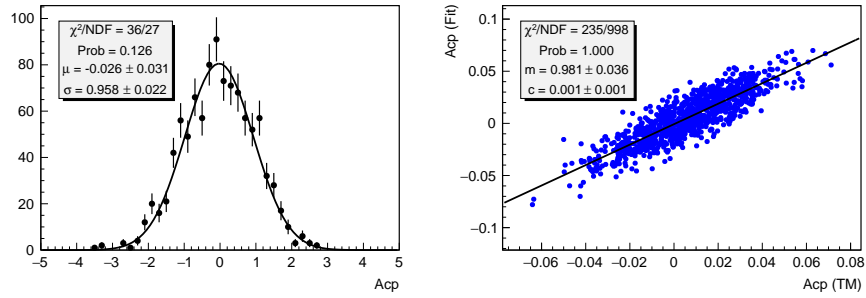


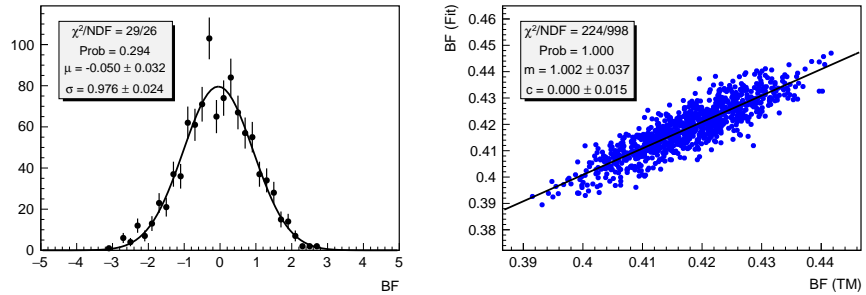
Figure C.4: Data vs. MC comparison of M_{bc} and ΔE distributions for $B \rightarrow K^*\gamma$ modes. Here we blind the signal region with $M_{bc} < 5.27 \text{ GeV}/c^2$ selection. The MC histogram is scaled down to the number of events in the data histogram.

Appendix D

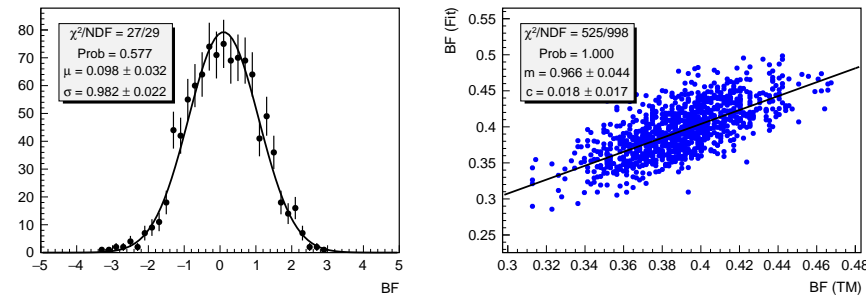
Plots for fit bias study



(a) $\mathcal{A}_{CP}(B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma)$

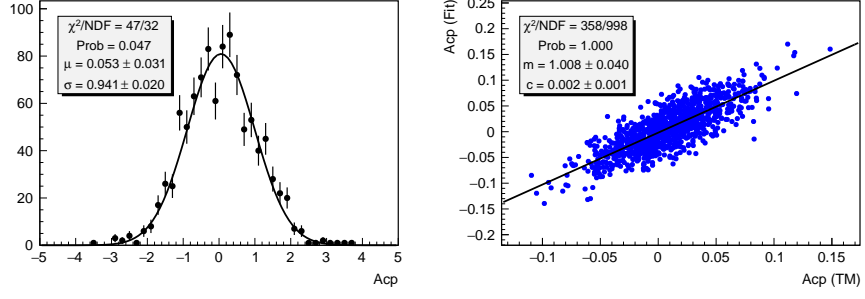


(b) $\mathcal{B}(B^0 \rightarrow K^{*0}[K^+\pi^-]\gamma)$

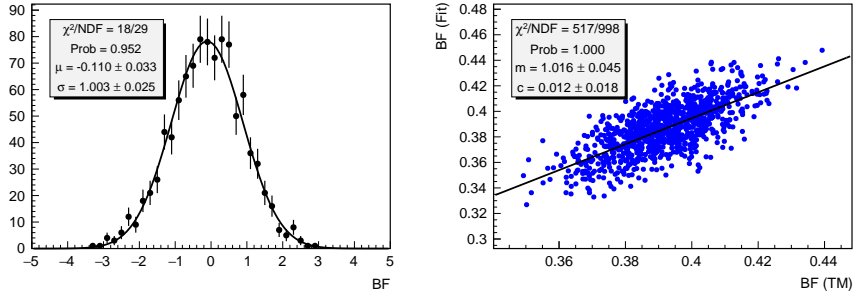


(c) $\mathcal{B}(B^0 \rightarrow K^{*0}[K_S^0\pi^0]\gamma)$

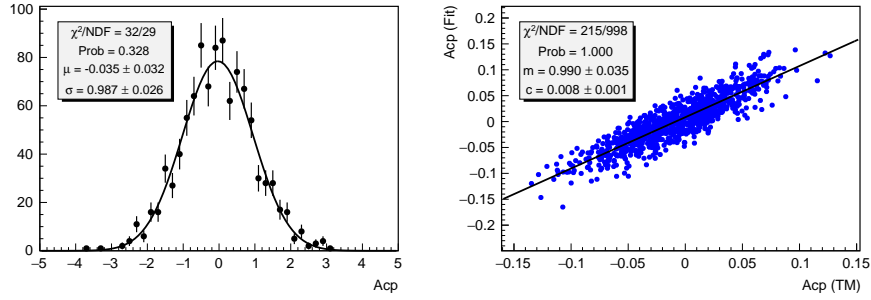
Figure D.1: Pull distributions and scatter plots for \mathcal{A}_{CP} and \mathcal{B} for the fit model of $B^0 \rightarrow K^{*0}\gamma$ modes.



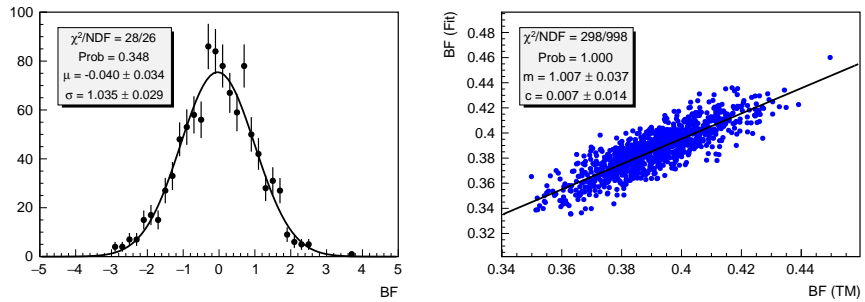
(a) $\mathcal{A}_{CP}(B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma)$



(b) $\mathcal{B}(B^+ \rightarrow K^{*+}[K^+\pi^0]\gamma)$



(c) $\mathcal{A}_{CP}(B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma)$



(d) $\mathcal{B}(B^+ \rightarrow K^{*+}[K_S^0\pi^+]\gamma)$

Figure D.2: Pull distributions and scatter plots for fit parameters \mathcal{A}_{CP} and \mathcal{B} for the fit model of $B^+ \rightarrow K^{*+}\gamma$ modes.

Bibliography

- [1] E. Rutherford, *The scattering of α and β particles by matter and the structure of the atom*, *Philosophical Magazine* **21** (1911) 669.
- [2] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons (1984), [10.1088/0031-9112/36/2/033](https://doi.org/10.1088/0031-9112/36/2/033).
- [3] O. Nachtmann, *Elementary Particle Physics*, Springer Texts and Monographs in Physics, Springer, Berlin, Heidelberg (1989),
https://doi.org/10.1007/978-3-642-61281-7_7.
- [4] D. Griffiths, *Introduction to Elementary Particles*, Wiley-VCH, Germany, 2 enl and rev ed ed. (2008).
- [5] S. Coleman, *Notes from Sidney Coleman's Physics 253a: Quantum Field Theory*,
[arXiv:1110.5013](https://arxiv.org/abs/1110.5013).
- [6] M. Srednicki, *Quantum field theory*, Cambridge University Press (1, 2007),
[10.1017/CBO9780511813917](https://doi.org/10.1017/CBO9780511813917).
- [7] L.H. Ryder, *Quantum Field Theory*, Cambridge University Press, 2 ed. (1996),
<https://doi.org/10.1017/CBO9780511813900>.
- [8] M.D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press (2013), <https://doi.org/10.1017/9781139540940>.
- [9] A. Salam and J.C. Ward, *Weak and electromagnetic interactions*, *Nuovo Cim.* **11** (1959) 568.
- [10] S.L. Glashow, *The renormalizability of vector meson interactions*, *Nucl. Phys.* **10** (1959) 107.

- [11] S.L. Glashow, *Partial-symmetries of weak interactions*, *Nuclear Physics* **22** (1961) 579.
- [12] A. Salam and J.C. Ward, *Electromagnetic and weak interactions*, *Phys. Lett.* **13** (1964) 168.
- [13] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [14] P.W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [15] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [16] S. Weinberg, *A model of leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [17] P.W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, *Phys. Rev.* **145** (1966) 1156.
- [18] T.W.B. Kibble, *Symmetry breaking in non-abelian gauge theories*, *Phys. Rev.* **155** (1967) 1554.
- [19] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) 367.
- [20] O. Greenberg, *Resource letter q-1: Quarks*, *Am. J. Phys.* **50** (1982) 1074.
- [21] Y. Nambu, *The Confinement of Quarks*, *Sci. Am.* **235N5** (1976) 48.
- [22] D.B. Lichtenberg and S.P. Rosen, *Developments in the quark theory of hadrons*, 1980, <https://api.semanticscholar.org/CorpusID:118133251>.
- [23] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, no. 56 in *Frontiers in Physics*, Addison-Wesley, Redwood City (1983), <https://doi.org/10.1201/9780429034978>.
- [24] K. Huang, *Quarks, Leptons and Gauge Fields*, World Scientific, Singapore (1992), <https://doi.org/10.1142/1409>.
- [25] P. Renton, *Electroweak Interactions*, Cambridge University Press, Cambridge (1990), <https://doi.org/10.1017/CBO9780511608148>.
- [26] J.F. Donoghue, E. Golowich and B.R. Holstein, *Dynamics of the Standard Model*,

- Cambridge Monographs on Particle Physics, Cambridge University Press, Cambridge (1992), <https://doi.org/10.1017/cbo9780511803512>.
- [27] M. Böhm, A. Denner and H. Joos, *Gauge Theories of the Strong and Electroweak Interaction*, Teubner, Stuttgart (2001), https://doi.org/10.1007/978-3-322-80160-9_4.
- [28] L.D. Faddeev and A.A. Slavnov, *Gauge Fields. Introduction To Quantum Theory*, no. 83 in *Frontiers in Physics*, Addison-Wesley, Redwood City (1990), <https://doi.org/10.1201/9780429493829>.
- [29] A.J. Buras, *Flavour physics and CP violation*, *arXiv* (2005) [[hep-ph/0505175](https://arxiv.org/abs/hep-ph/0505175)].
- [30] G. Isidori, *Flavor physics and CP violation*, *arXiv* (2013) [[1302.0661](https://arxiv.org/abs/1302.0661)].
- [31] Y. Grossman and P. Tanedo, *Just a taste: Lectures on flavor physics*, in *Anticipating the Next Discoveries in Particle Physics*, WORLD SCIENTIFIC, May, 2018, DOI.
- [32] A.D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32.
- [33] G.C. Branco, L. Lavoura and J.P. Silva, *CP Violation*, vol. 103, Oxford University Press (1999), [10.1093/oso/9780198503996.001.0001](https://doi.org/10.1093/oso/9780198503996.001.0001).
- [34] I.I. Bigi and A.I. Sanda, *CP violation*, vol. 9, Cambridge University Press (9, 2009), [10.1017/CBO9780511581014](https://doi.org/10.1017/CBO9780511581014).
- [35] Y. Nir, *Flavour physics and CP violation*, *CERN* (2015) .
- [36] A. Riotto, *Theories of baryogenesis*, in *ICTP Summer School in High-Energy Physics and Cosmology*, pp. 326–436, 7, 1998 [[hep-ph/9807454](https://arxiv.org/abs/hep-ph/9807454)].
- [37] R.S. Chivukula, E.H. Simmons and J. Terning, *Phenomenology of electroweak symmetry breaking*, *Phys. Rept.* **267** (1996) 339 [[hep-ph/9506427](https://arxiv.org/abs/hep-ph/9506427)].
- [38] J. Wess and J. Bagger, *Supersymmetry and Supergravity*, Princeton University Press, 2nd ed. (1992).
- [39] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*, Cambridge University Press (2007).

- [40] M. Dine, *Supersymmetry and String Theory: Beyond the Standard Model*, Cambridge University Press (2007).
- [41] S.P. Martin, *A supersymmetry primer*, *Adv. Ser. Direct. High Energy Phys.* **21** (2010) 1 [[hep-ph/9709356](#)].
- [42] G.L. Kane, ed., *Perspectives on Supersymmetry*, World Scientific Publishing (1998).
- [43] P. Ramond, *Supersymmetry and unification of fundamental interactions*, *Phys. Rep.* **227** (1993) 121.
- [44] BELLE II collaboration, *Belle II Technical Design Report*, [arXiv:1011.0352](#).
- [45] BELLE II collaboration, *The Belle II Physics Book*, *PTEP* **2019** (2019) 123C01 [[1808.10567](#)].
- [46] CLEO collaboration, *Evidence for penguins: First observation of $B \rightarrow K^*(892)\gamma$* , *Phys. Rev. Lett.* **71** (1993) 674.
- [47] CLEO collaboration, *Study of exclusive radiative B meson decays*, *Phys. Rev. Lett.* **84** (2000) 5283 [[hep-ex/9912057](#)].
- [48] BABAR collaboration, *Measurement of Branching Fractions and CP and Isospin Asymmetries in $B \rightarrow K^*(892)\gamma$ Decays*, *Phys. Rev. Lett.* **103** (2009) 211802 [[0906.2177](#)].
- [49] BELLE collaboration, *Evidence for Isospin Violation and Measurement of CP Asymmetries in $B \rightarrow K^*(892)\gamma$* , *Phys. Rev. Lett.* **119** (2017) 191802 [[1707.00394](#)].
- [50] BELLE II collaboration, *Measurements of the branching fractions for $B \rightarrow K^*\gamma$ decays at Belle II*, [arXiv:2110.08219](#).
- [51] S.L. Adler, *Axial vector vertex in spinor electrodynamics*, *Phys. Rev.* **177** (1969) 2426.
- [52] J.S. Bell and R. Jackiw, *A PCAC puzzle: $\pi^0 \rightarrow \gamma\gamma$ in the σ model*, *Nuovo Cim. A* **60** (1969) 47.
- [53] E.P. Wigner, *On Unitary Representations of the Inhomogeneous Lorentz Group*, *Annals Math.* **40** (1939) 149.

- [54] S. Coleman and J. Mandula, *All possible symmetries of the s matrix*, *Phys. Rev.* **159** (1967) 1251.
- [55] Wikipedia contributors, “Standard Model of Elementary Particles.”
https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, Accessed: 15-Jul-2024.
- [56] N. Cabibbo, *Unitary symmetry and leptonic decays*, *Phys. Rev. Lett.* **10** (1963) 531.
- [57] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, *Progress of Theoretical Physics* **49** (1973) 652
[<https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf>].
- [58] J. Charles, A. Hocker, H. Lacker, F.R. Le Diberder, M. Pivk and S. T’Jampens, *Cp violation and the ckm matrix: Assessing the impact of the asymmetric b factories*, *Eur. Phys. J. C* **41** (2005) 1 [[hep-ph/0406184](https://arxiv.org/abs/hep-ph/0406184)].
- [59] M. Bona, M. Ciuchini, E. Franco, V. Lubicz, G. Martinelli, F. Parodi et al., *The unitarity triangle fit in the standard model and constraints on new physics*, *JHEP* **03** (2006) 080 [[hep-ph/0501199](https://arxiv.org/abs/hep-ph/0501199)].
- [60] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation*, *Phys. Rev. Lett.* **55** (1985) 1039.
- [61] S.L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285.
- [62] T. Appelquist and J. Carazzone, *Infrared singularities and massive fields*, *Phys. Rev. D* **11** (1975) 2856.
- [63] S. Weinberg, *Phenomenological Lagrangians*, *Physica A* **96** (1979) 327.
- [64] J. Polchinski, *Renormalization and Effective Lagrangians*, *Nucl. Phys. B* **231** (1984) 269.
- [65] H. Georgi, *Effective field theory*, *Ann. Rev. Nucl. Part. Sci.* **43** (1993) 209.
- [66] H.M. Georgi, *Effective quantum field theories*, in *The New Physics*, P. Davies, ed., (Cambridge), pp. 446–457, Cambridge University Press (1989).
- [67] G.W. Leibbrandt, *Effective Field Theories*, World Scientific Publishing (1994).

- [68] A.V. Manohar, *Effective field theories*, in *Perturbative and Nonperturbative Aspects of Quantum Field Theory*, H. Latal and W. Schweiger, eds., (Berlin, Heidelberg), pp. 311–362, Springer Berlin Heidelberg, 1997.
- [69] R.D. Peccei and C. Schweiger, eds., *Modern Effective Field Theory: Central Themes and Philosophical Issues*, World Scientific Publishing (1999).
- [70] T. Mannel, *Effective Field Theories in Flavor Physics*, *Springer Tracts Mod. Phys.* **203** (2004) 1.
- [71] M. NEUBERT, *Effective field theory and heavy quark physics*, in *Physics In $D \geq 4$ Tasi 2004*, WORLD SCIENTIFIC, July, 2006, DOI.
- [72] C.P. Burgess, *Introduction to effective field theory*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 329.
- [73] U.G. Meissner, ed., *Effective Field Theories of the Standard Model*, World Scientific Publishing (2007).
- [74] R.D. Miller and B.H. McKellar, *Effective field theory and weak non-leptonic interactions*, *Physics Reports* **106** (1984) 169.
- [75] K.G. Wilson and J. Kogut, *The renormalization group and the ϵ expansion*, *Physics reports* **12** (1974) 75.
- [76] J.C. Collins, *Renormalization: An Introduction to Renormalization, the Renormalization Group and the Operator-Product Expansion*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (1984).
- [77] S. Weinberg, *New approach to the renormalization group*, *Phys. Rev. D* **8** (1973) 3497.
- [78] S.W. Bosch and G. Buchalla, *The radiative decays $B \rightarrow V\gamma$ at next-to-leading order in QCD*, *Nucl. Phys. B* **621** (2002) 459 [[hep-ph/0106081](#)].
- [79] B. Grinstein and D. Pirjol, *Long-distance effects in $B \rightarrow V\gamma$ radiative weak decays*, *Phys. Rev. D* **62** (2000) 093002 [[hep-ph/0002216](#)].
- [80] A.J. Buras, *Weak Hamiltonian, CP violation and rare decays*, in *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*, pp. 281–539, 6, 1998 [[hep-ph/9806471](#)].

- [81] A.J. Buras, *Gauge Theory of Weak Decays: The Standard Model and the Expedition to New Physics Summits*, Cambridge University Press (2020).
- [82] G. Buchalla, A.J. Buras and M.E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125.
- [83] T. Inami and C.S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \rightarrow \mu\bar{\mu}$, $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$* , *Progress of Theoretical Physics* **65** (1981) 297
[<https://academic.oup.com/ptp/article-pdf/65/1/297/5252099/65-1-297.pdf>].
- [84] N. Isgur and M.B. Wise, *Weak Decays of Heavy Mesons in the Static Quark Approximation*, *Phys. Lett. B* **232** (1989) 113.
- [85] N. Isgur and M. Wise, *Spectra and decay constants of heavy mesons in the heavy quark effective theory*, *Phys. Rev. Lett.* **66** (1989) 1130.
- [86] N. Isgur and M.B. Wise, *Weak transition form factors between heavy mesons*, *Physics Letters B* **237** (1990) 527.
- [87] M. Voloshin and M. Shifman, *On production of d^* and d mesons in b -meson decays*, *Sov. J. Nucl. Phys.(Engl. Transl.);(United States)* **47** (1988) .
- [88] M.B. Voloshin and M.A. Shifman, *Heavy quarks and heavy quark symmetry*, in *Proceedings of the 1988 International Symposium on Lepton and Photon Interactions at High Energies*, pp. 218–220, World Scientific Publishing, 1988.
- [89] B. Grinstein, *The Static Quark Effective Theory*, *Nucl. Phys. B* **339** (1990) 253.
- [90] H. Georgi, *An Effective Field Theory for Heavy Quarks at Low-energies*, *Phys. Lett. B* **240** (1990) 447.
- [91] H. Georgi, *Applications of heavy quark symmetry*, *Phys. Lett. B* **264** (1991) 447.
- [92] A.F. Falk, B. Grinstein and M.E. Luke, *Leading mass corrections to the heavy quark effective theory*, *Nucl. Phys. B* **357** (1991) 185.
- [93] M. Neubert, *Heavy quark symmetry*, *Phys. Rept.* **245** (1994) 259 [hep-ph/9306320].
- [94] M. Neubert, *Heavy quark effective theory*, *Phys. Rept.* **245** (1994) 259.

- [95] M. Shifman, *The heavy quark expansion in the non-perturbative qcd regime*, *Int. J. Mod. Phys. A* **10** (1995) 2379.
- [96] T. Mannel, *Heavy-quark effective field theory*, *Reports on Progress in Physics* **60** (1997) 1113.
- [97] I.I.Y. Bigi, M.A. Shifman and N. Uraltsev, *Aspects of heavy quark theory*, *Annual Review of Nuclear and Particle Science* **47** (1997) 591.
- [98] A. Manohar and M. Wise, *Heavy Quark Physics*, vol. 10 of *Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology*, Cambridge University Press (2000).
- [99] A.K. V. M. Braun and M. Maul, *Qcd sum rules on the light cone and heavy quark effective theory*, *Phys. Rev. D* **49** (1994) 2114.
- [100] A. Khodjamirian and A.V. Radyushkin, *Light-cone sum rules*, *Phys. Lett. B* **457** (1998) 245.
- [101] P. Ball and V.M. Braun, *Exclusive semileptonic and rare b meson decays in qcd*, *Phys. Rev. D* **58** (1998) 094016.
- [102] M.A. Shifman, *QCD Sum Rules*, vol. 10 of *Current Physics - Sources and Comments*, World Scientific Publishing (1998).
- [103] A. Khodjamirian, *Light-cone sum rules in qcd*, in *Proceedings of the Workshop on Continuous Advances in QCD*, pp. 58–71, World Scientific Publishing, 2000.
- [104] T.M. A. Khodjamirian and M. Melcher, *Light-cone sum rules for heavy flavor physics*, *Phys. Rev. D* **74** (2006) 074005.
- [105] S.W. Bosch and G. Buchalla, *The radiative decays $b \rightarrow v\gamma$ at next-to-leading order in qcd*, *Nuclear Physics B* **621** (2002) 459.
- [106] M. Benzke, S.J. Lee, M. Neubert and G. Paz, *Long-Distance Dominance of the CP Asymmetry in $\bar{B} \rightarrow X_{s,d}\gamma$ Decays*, *Phys. Rev. Lett.* **106** (2011) 141801 [[arXiv:1012.3167](https://arxiv.org/abs/1012.3167)].
- [107] J. Lyon and R. Zwicky, *Isospin asymmetries in $B \rightarrow (K^*, \rho)\gamma/l^+l^-$ and $B \rightarrow Kl^+l^-$ in and beyond the standard model*, *Phys. Rev. D* **88** (2013) 094004 [[arXiv:1305.4797](https://arxiv.org/abs/1305.4797)].

- [108] A. Ali, B.D. Pecjak and C. Greub, *Towards $B \rightarrow V\gamma$ decays at NNLO in SCET*, *Eur. Phys. J. C* **55** (2008) 577 [[arXiv:0709.4422](#)].
- [109] A. Paul and D.M. Straub, *Constraints on new physics from radiative B decays*, *JHEP* **2017** (2017) 27 [[arXiv:1608.02556](#)].
- [110] W. Altmannshofer and D.M. Straub, *New Physics in $b \rightarrow s$ Transitions after LHC Run 1*, *Eur. Phys. J. C* **75** (2015) 382 [[arXiv:1411.3161](#)].
- [111] M. Matsumori, A.I. Sanda and Y.-Y. Keum, *CP asymmetry, branching ratios, and isospin breaking effects of $B \rightarrow K^*\gamma$ with the perturbative QCD approach*, *Phys. Rev. D* **72** (2005) 014013 [[arXiv:hep-ph/0406055](#)].
- [112] C. Dariescu and M.-A. Dariescu, *$B \rightarrow K^*\gamma$ decay within MSSM*, in *1st International Conference on Hadron Physics*, 10, 2007 [[arXiv:0710.3819](#)].
- [113] C. Greub, H. Simma and D. Wyler, *Branching ratio and direct CP violating rate asymmetry of the rare decays $B \rightarrow K^*\gamma$ and $B \rightarrow \rho\gamma$* , *Nucl. Phys. B* **434** (1995) 39 [[arXiv:hep-ph/9406421](#)].
- [114] M. Jung, X.-Q. Li and A. Pich, *Exclusive radiative B -meson decays within the aligned two-Higgs-doublet model*, *JHEP* **10** (2012) 063 [[arXiv:1208.1251](#)].
- [115] A.L. Kagan and M. Neubert, *Isospin breaking in $B \rightarrow K^*\gamma$ decays*, *Phys. Lett. B* **539** (2002) 227 [[arXiv:hep-ph/0110078](#)].
- [116] M. Benzke, S.J. Lee, M. Neubert and G. Paz, *Long-distance dominance of the cp asymmetry in $\bar{B} \rightarrow X_{s,d}\gamma$ decays*, *Phys. Rev. Lett.* **106** (2011) 141801.
- [117] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [118] SUPERKEKB collaboration, *SuperKEKB Collider*, *Nucl. Instrum. Meth. A* **907** (2018) 188 [[arXiv:1809.01958](#)].
- [119] K. Collaboration, *KEKB B-Factory Design Report*, Tech. Rep. **KEK Report 95-7**, KEK (1995).
- [120] M.A. Furman, *Hourglass effects for asymmetric colliders*, *Conf. Proc. C* **910506** (1991) 422.

- [121] C. Bernardini, G.F. Corazza, G. Di Giugno, G. Ghigo, J. Haissinski, P. Marin et al., *Lifetime and beam size in a storage ring*, *Phys. Rev. Lett.* **10** (1963) 407.
- [122] H. Bruck, *Circular Particle Accelerators*, LA-TR, Los Alamos Scientific Laboratory (1972).
- [123] J. Kemmer and G. Lutz, *New detector concepts*, *Nucl. Instrum. Meth. A* **253** (1987) 365.
- [124] B. Wang, F. Abudinen, K. Ackermann, P. Ahlburg, M. Albalawi, O. Alonso et al., *Operational experience of the belle ii pixel detector*, *Nucl. Instrum. Meth. A* **1032** (2022) 166631.
- [125] BELLE II collaboration, *Measurement of the integrated luminosity of data samples collected during 2019-2022 by the Belle II experiment*, [arXiv:2407.00965](https://arxiv.org/abs/2407.00965).
- [126] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.
- [127] S. Jadach, B.F.L. Ward and Z. Was, *The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions*, *Comput. Phys. Commun.* **130** (2000) 260 [[arXiv:hep-ph/9912214](https://arxiv.org/abs/hep-ph/9912214)].
- [128] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[arXiv:1410.3012](https://arxiv.org/abs/1410.3012)].
- [129] GEANT4 collaboration, *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [130] Belle II Framework Software Group, *The Belle II Core Software*, *Comput. Softw. Big Sci.* **3** (2019) 1 [[arXiv:1809.04299](https://arxiv.org/abs/1809.04299)].
- [131] Belle II collaboration, “Belle II Analysis Software Framework (basf2).” <https://doi.org/10.5281/zenodo.5574115>.
- [132] BABAR collaboration, D. Boutigny et al., *The BABAR physics book: Physics at an asymmetric B factory*, Stanford Linear Acceleration Center (10, 1998), [10.2172/979931](https://arxiv.org/abs/10.2172/979931).
- [133] BELLE II collaboration, *Global decay chain vertex fitting at Belle II*, *Nucl. Instrum. Meth. A* **976** (2020) 164269 [[arXiv:1901.11198](https://arxiv.org/abs/1901.11198)].

- [134] T. Keck, *Fastbdt: A speed-optimized multivariate classification algorithm for the belle ii experiment*, *Comput. Softw. Big Sci.* **1** (2017) 2 [[arXiv:1609.06119](#)].
- [135] W. Waltenberger, *Adaptive vertex reconstruction*, Tech. Rep. , CERN (7, 2008).
- [136] A. Khotanzad and Y. Hong, *Invariant image recognition by zernike moments*, *IEEE Trans. Pattern Anal. Mach. Intell.* **12** (1990) 489.
- [137] BELLE collaboration, *Evidence for $B^0 \rightarrow \pi^0\pi^0$* , *Phys. Rev. Lett.* **91** (2003) 261801 [[arXiv:hep-ex/0308040](#)].
- [138] G.C. Fox and S. Wolfram, *Observables for the analysis of event shapes in e^+e^- annihilation and other processes*, *Phys. Rev. Lett.* **41** (1978) 1581.
- [139] CLEO collaboration, *Search for exclusive charmless hadronic b decays*, *Phys. Rev. D* **53** (1996) 1039 [[arXiv:hep-ex/9508004](#)].
- [140] BELLE II collaboration, *B-flavor tagging at Belle II*, *Eur. Phys. J. C* **82** (2022) 283 [[arXiv:2110.00790](#)].
- [141] M. Pivk and F. Le Diberder, *Plots: A statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555** (2005) 356.
- [142] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, Ph.D. thesis, Cracow, INP, 1986.
- [143] K. Cranmer, *Kernel estimation in high-energy physics*, *Comput. Phys. Commun.* **136** (2001) 198 [[arXiv:hep-ex/0011057](#)].
- [144] BABAR collaboration, *Measurement of the branching fraction of $\Upsilon(4S) \rightarrow B^0\bar{B}^0$* , *Phys. Rev. Lett.* **95** (2005) 042001 [[arXiv:hep-ex/0504001](#)].
- [145] BELLE collaboration, *Measurement of the B^+ / B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \rightarrow J/\psi(\ell\ell)K$ decays at Belle*, *Phys. Rev. D* **107** (2023) L031102 [[arXiv:2207.01194](#)].
- [146] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, *The Annals of Statistics* **7** (1979) 1 .
- [147] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, John Wiley & Sons (1957).

- [148] F. Berends, K. Gaemers and R. Gastmans, α^3 -contribution to the angular asymmetry in $e^+e^- \rightarrow \mu^+\mu^-$, *Nucl. Phys. B* **63** (1973) 381.
- [149] R. Brown, V. Cung, K. Mikaelian and E. Paschos, *Electromagnetic background in the search for neutral weak currents via $e^+e^- \rightarrow \mu^+\mu^-$* , *Phys. Lett. B* **43** (1973) 403.
- [150] R.J. Cashmore, C.M. Hawkes, B.W. Lynn and R.G. Stuart, *The Forward - Backward Asymmetry in $e^+e^- \rightarrow \mu^+\mu^-$ Comparisons Between the Theoretical Calculations at the One Loop Level in the Standard Model and with the Experimental Measurements*, *Z. Phys. C* **30** (1986) 125.
- [151] L. Lyons, D. Gibaut and P. Clifford, *How to combine correlated estimates of a single physical quantity*, *Nucl. Instrum. Meth. A* **270** (1988) 110.
- [152] A. Valassi, *Combining correlated measurements of several different physical quantities*, *Nucl. Instrum. Meth. A* **500** (2003) 391.
- [153] C. Pulvermacher, *dE/dx particle identification and pixel detector data reduction for the Belle II experiment*, 2012.
- [154] BELLE II collaboration, *Snowmass White Paper: Belle II physics reach and plans for the next decade and beyond*, [arXiv:2207.06307](https://arxiv.org/abs/2207.06307).
- [155] K. Cranmer, *Practical statistics for the lhc*, [arXiv:1503.07622](https://arxiv.org/abs/1503.07622).
- [156] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, Wiley (1993).
- [157] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*, Springer, second edition ed. (2017, Volume 941).
- [158] O. Behnke, K. Kröninger, T. Schörner-Sadenius and G. Schott, eds., *Data analysis in high energy physics: A practical guide to statistical methods*, Wiley-VCH, Weinheim, Germany (2013).
- [159] F. James, *Statistical Methods in Experimental Physics*, World Scientific Publishing Co. Pte. Ltd., 2nd ed. (2006).
- [160] S.S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, *The Annals of Mathematical Statistics* **9** (1938) 60.

-
- [161] J.H. Friedman, *Greedy function approximation: a gradient boosting machine*, *Annals of Statistics* **29** (2001) 1189.
- [162] J.H. Friedman, *Stochastic gradient boosting*, *Computational Statistics & Data Analysis* **38** (2002) 367.