Search for the decay $B_s^0 \to \eta' X_{s\bar{s}}$ Using a Semi-Inclusive Method at the Belle Experiment

Shawn Dubey

Department of Physics and Astronomy
College of Natural Sciences

Abstract

The decay $B_s^0 \to \eta' X_{s\bar{s}}$ is searched for at Belle, using Belle's 121.4 fb^{-1} integrated luminosity data sample, taken at $\Upsilon(5S)$ resonance. A semi-inclusive reconstruction method whereby the $X_{s\bar{s}}$ is reconstructed as a system of two kaons and up to four pions, with at most one neutral pion, is used. Using the η' sub-decay mode $\eta' \to \eta(\to \gamma\gamma)\pi^+\pi^-$ and examining the $X_{s\bar{s}}$ mass range $M(X_{s\bar{s}}) \leq 2.4 \, \text{GeV}/c^2$, an upper limit at 90% confidence level is set. This mode had been previously unstudied and this analysis partially hopes to motivate future theoretical studies as well as future analyses at the Belle II experiment.

Acknowledgements

There are several people and groups that I owe much thanks and more to, who helped make this possible. First, and most directly relevant to this dissertation, is Tom Browder, who, with astonishing patience and a staggering amount of knowledge, advised this research. It is hard to see how this would have been possible without him.

I also owe a lot to my significant other, Melody Wan, who stood by me through the fire and the flames (and who drew Figs. 1.3, 3.1, and 3.2 with her excellent artistic abilities). There is not enough room here for the things I want to say about her so I will leave it at that.

My grandparents are also owed a lot of thanks for always being supportive. My grandfather, a scientist, is partially responsible for engendering in me an interest in science, and my grandmother is ever-encouraging.

Thanks also goes to my parents and siblings whose support did not wane, and in fact grew, over the time this dissertation was being completed.

I also would like to thank my PhD committee, Belle RC, and Belle and Belle II collaborations and all of the people in them who patiently helped me and answered my questions.

Preface

1. Motivation

This thesis provides the world's first measurement of the process $B_s^0 \to \eta' X_{s\bar{s}}$, using the largest $\Upsilon(5S)$ data sample in the world from e^+e^- collisions. Previously, $B \to \eta' X_s$ and $B \to \eta X_s$ had been measured, and some theoretical results were given, but not for $B_s^0 \to \eta' X_{s\bar{s}}$. It is hoped that the analysis in this thesis will motivate some theoretical research, as well as improved measurements at the Belle II experiment. Further, part of this thesis describes the author's work on the electronics for the imaging Time of Propagation (iTOP) detector, one of the sub-detectors in Belle II partially developed at the University of Hawaii Manoa, Department of Physics and Astronomy, Instrumentation Development Laboratory (IDL). The author was involved in the assembly and laser verification of the iTOP electronic readout modules.

2. Organization

The thesis is organized as follows. Chapter 1 gives an introduction and overview of the KEKB e^+e^- collider and the Belle experiment. Chapter 2 gives an overview of the SuperKEKB e^+e^- collider and Belle II experiment, with more emphasis on the iTOP sub-detector. Chapter 3 gives a brief overview of the relevant theoretical information. Chapters 4 and 5 give the $B_s^0 \to \eta' X_{s\bar{s}}$ analysis method and results, respectively. Finally, the appendices give some supplementary information.

3. Updates

3.1. Erratum

This version has been updated from the version in Ref. [1]. There are minor corrections to the systematic uncertainties in Sec. 4.12 and the text description of the η and η' selections in Sec. 4.5.1 and Sec. 4.8.3. The former does not materially affect the analysis and the final result remains unchanged.

A typo was corrected in equation 3.6: the ∂ in the numerator of the first term was accidentally left out.

A reference to analysis materials is given in Ref. [2]. It is current as of this writing. A statement of this has also been added in Sec. 4.1.

3.2. Publication Reference

The journal publication of this analysis can be found in Ref. [3].

Contents

Chapter 1.

KEKB and Belle

1.1. High Energy Accelerator Research Center and KEKB

1.1.1. High Energy Accelerator Research Center - KEK

The Belle experiment was situated at the interaction point of the KEKB accelerator at the High Energy Accelerator Research Center (Japanese acronym: KEK) in Tsukuba, Ibaraki prefecture, Japan.

1.1.2. KEKB

The KEKB B-factory is an accelerator that produces electron-positron (e^+e^-) collisions with asymmetric energies. The center of mass (CM) energy is tuned to produce states of the Y meson, a bottomonium ($b\bar{b}$) particle, which yields $B\bar{B}$ states with a boost of $\beta\gamma = 0.425$ [4,5].

Electron and positron beams are first accelerated through a linac (LINear ACcelerator) before being injected into the two main rings. Due to space limitations, the linac is not entirely straight but consists of two linacs connected by a $180°$ J-shaped beam transport, the J-arc [6], seen in Fig. 1.1. Figure 1.2 gives a schematic of KEKB.

The electron source is an electron gun that produces electrons from thermionic emission. The positron source is comprised of a target, accelerator sections, solenoids and an e^+e^- separator. The target is a water-cooled tungsten plate that is struck by electrons from the electron source to produce positrons. Electrons and positrons are fed

through bunchers and then are injected into the main KEKB rings, via beam transfer lines [6, 7].
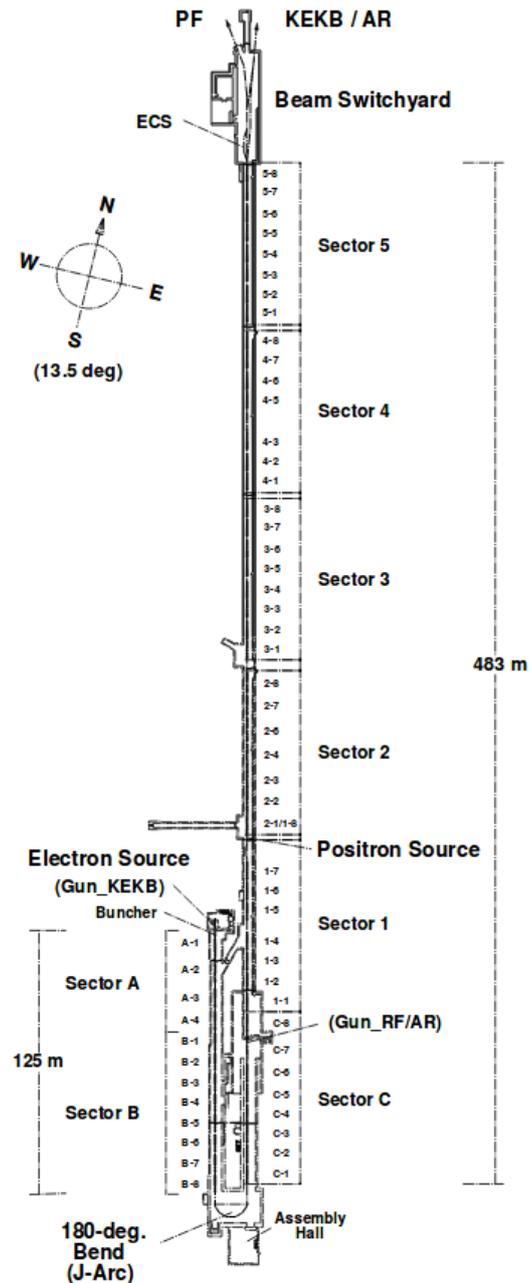
Figure 1.1.: Diagram of the KEKB LINAC

The beams are brought into collision at an angle of $\pm$ 11 mrad at the interaction point (IP), inside the beryllium beam pipe, underground. When this happens, sev-
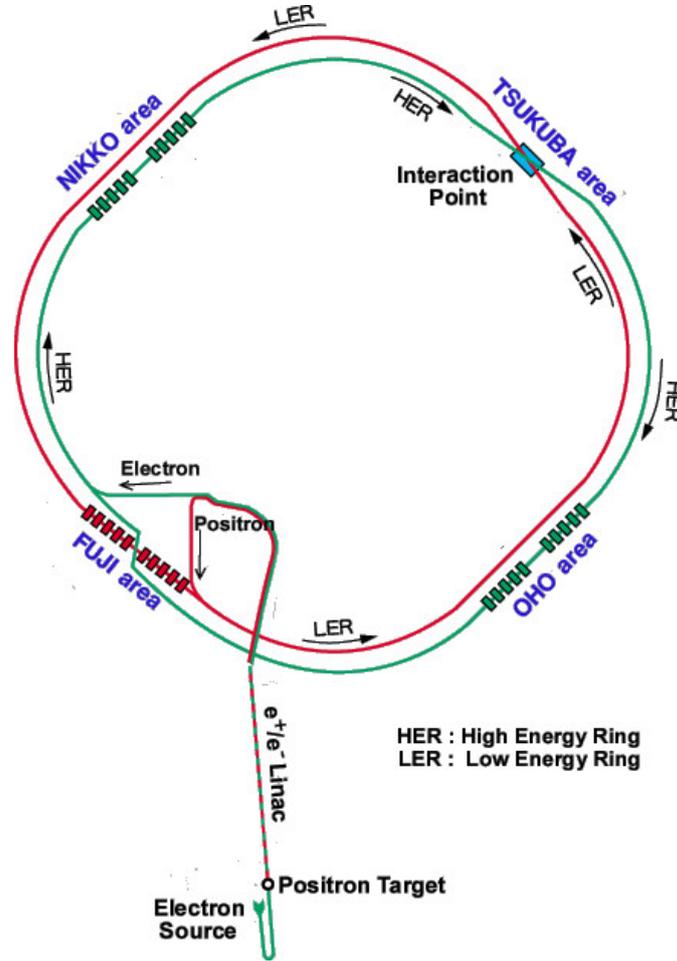
Figure 1.2.: Schematic of the KEKB accelerator. Belle is located at the Tsukuba IP.

eral processes can occur, including the production of Υ mesons. When the $\Upsilon(5S)$ is produced, it decays to $B_s^0 - \bar{B}_s^0$ pairs as shown in Fig. 1.3 [5] [1].

At the $\Upsilon(5S)$ resonance the electrons, fed through the high energy ring (HER), have an energy of 8.2 GeV and the positrons, fed through the low energy ring (LER) have an energy of 3.6 GeV. This corresponds to a CM energy of approximately 10.866 GeV.

The integrated luminosity collected at the $\Upsilon(5S)$ resonance is 121.4 fb^{-1}. The number of $B_s^{0(*)}$ mesons produced is given by $N_{B_s^0} = 2f_s N_{b\bar{b}}$, where the factor of 2 accounts for the fact that they are produced in pairs; f_s is the fraction of B_s^0 events compared to all $b\bar{b}$ events and is 0.172 ± 0.030 [8] in the e^+e^- collisions, and $N_{b\bar{b}}$ is given by

---

[1] For most of the running time relevant to the Belle experiment, to be discussed subsequently, KEKB was used to produce the $\Upsilon(4S)$ and non-strange-B-mesons. However, for this analysis the $\Upsilon(5S)$ meson is the most relevant particle.
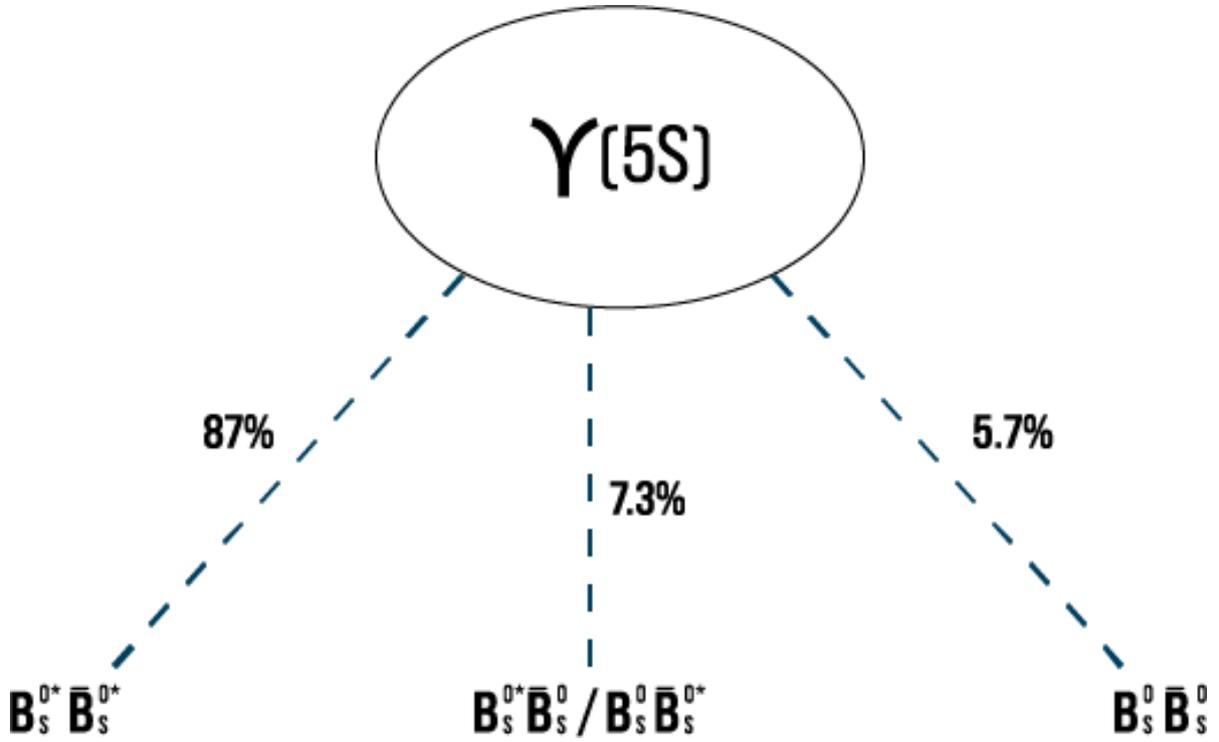
Figure 1.3.: This shows the $\Upsilon(5S)$ decay chain for the $\Upsilon(5S)$ decaying to B_s^0 mesons.

$$N_{b\bar{b}} = \sigma_{b\bar{b}}^{\Upsilon(5S)} L_{\text{int}} \tag{1.1}$$

where $\sigma_{b\bar{b}}^{\Upsilon(5S)} = \sigma(e^+e^- \to b\bar{b}) = (0.340 \pm 0.016)$ nb, and L_{int} is the integrated luminosity.

1.2. Belle

1.2.1. Belle Overview

The Belle detector is a general purpose magnetic spectrometer, situated at the IP in Tsukuba Hall, at KEK. It was initially conceived of to study CP violation in B-meson systems [9, 10]. Asymmetric energy collisions take place with an energy of 3.5 GeV for the positrons and 8.0 GeV for the electrons, giving a boost of $\beta\gamma = 0.425$. These energies produce the $\Upsilon(4S)$ meson. Data was also taken at $\Upsilon(5S)$ energy and at the nearby continuum. Belle collected data from 1999-2010 with a total integrated luminosity of 1 ab^{-1} collected. Here integrated luminosity is given by $L_{\text{int}} = \int \mathcal{L}dt$, where $\mathcal{L}$ is the instantaneous luminosity. KEKB reached a peak luminosity of 2.11 $\times$ 10^{34} cm^{-2}s^{-1} [11].

Belle consists of several sub-detectors: the Silicon Vertex Detector (SVD), Aerogel Cherenkov Coutners (ACC), Time of Flight (TOF) counter, Central Drift Chamber (CDC), Electromagnetic Calorimeter (ECL), Extreme Forward Calorimeter (EFC), and the K_L^0 and muon detector (KLM) [12]. These sub-detectors gather information from the B-meson decays (and other processes) to be able to perform physics analyses.

1.2.2. Silicon Vertex Detector

The silicon vertex detector (SVD) measured the decay vertices of B-meson decays. It was located around the beam pipe, at the IP. The design resolution of the SVD required to resolve the difference in z-vertex positions of the B-mesons was 100 μm. It also had to be tolerant to the large beam backgrounds that came from the high luminosity of KEKB [12, 13].

The SVD came in two main stages: SVD1, the original SVD configuration with a beampipe radius of $r = 2.0$ cm, and SVD2 with a beampipe radius of $r = 1.5$ cm.

SVD1

SVD1 consisted of three cylindrical layers around the beam pipe and covered a solid angle, θ, the angle from the beam axis, of $23° < \theta < 139°$. The radius of the outermost
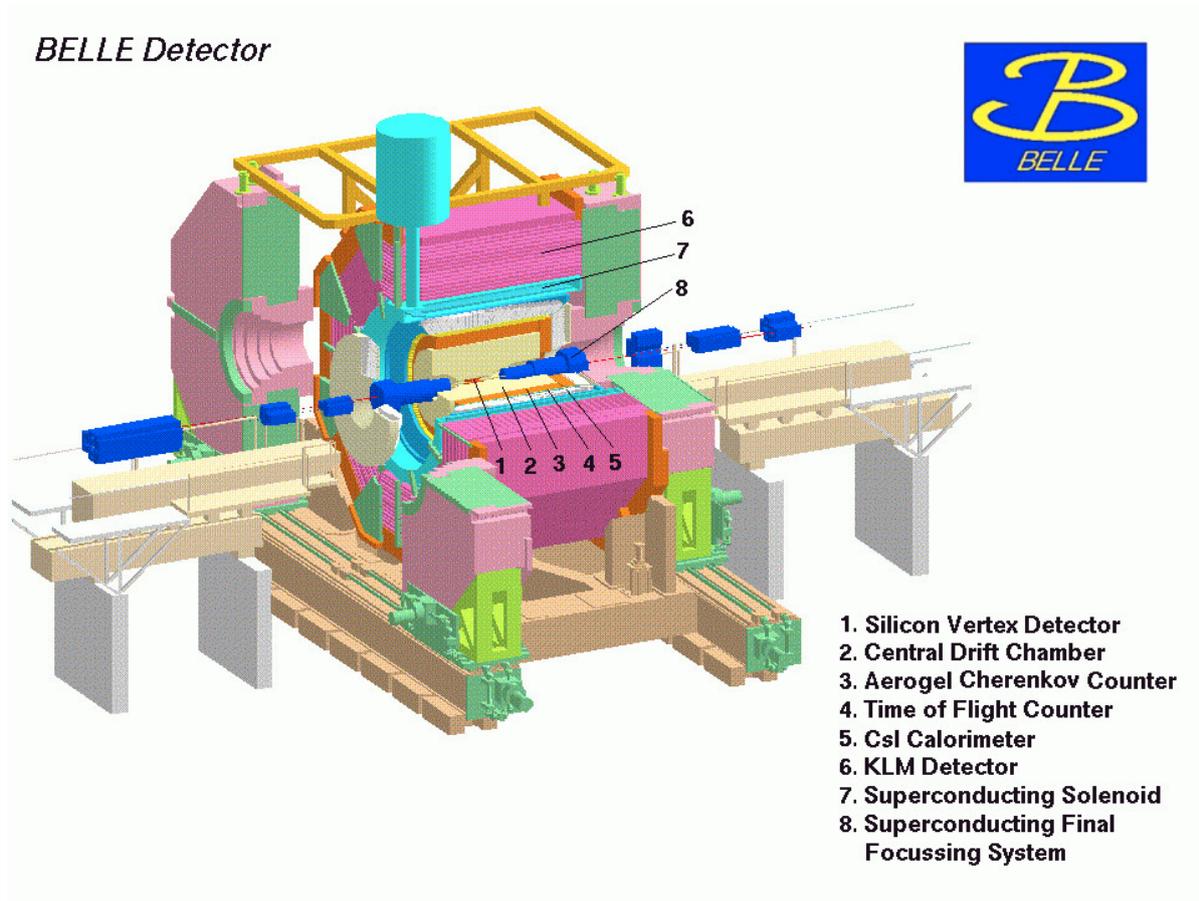
Figure 1.4.: Rendering of the Belle detector with all sub-detectors labeled

layer was 60.5 mm, the radius of the middle layer is 45.5 mm, and the radius of the innermost layer was 30.0 mm [13].

The layers were comprised of "ladders". The ladders were made up of double-sided silicon strip detectors (DSSDs) with boron-nitride ribs for support. A charged particle passed through a DSSD, creating electron-hole pairs that induce a current, from which event information could be obtained.

For SVD1 there were 8 ladders in the first layer, 10 ladders in the middle layer, and 14 ladders in the outer layer. Between the ladders in a layer there was some overlap in the r-ϕ plane that was purposefully introduced to help with the alignment of the sensors [12, 13].
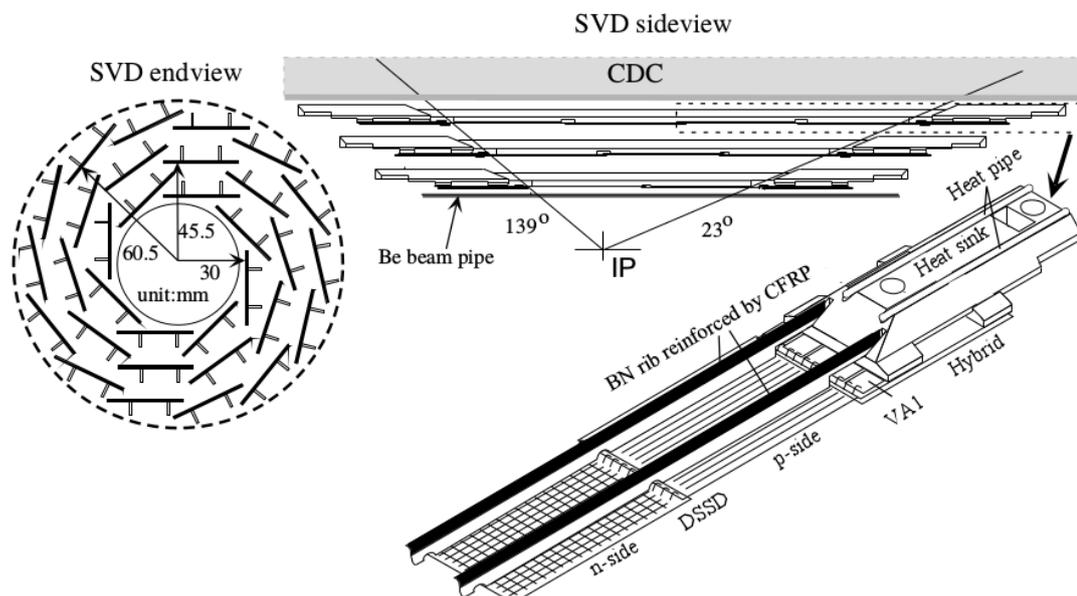
Figure 1.5.: Schematic of the SVD; cross-sectional image (left); module view (bottom right); detector situation (top right)

SVD2

SVD2, installed in 2003, was an upgrade from SVD1 that included four layers. The inner most layer had a radius of 20.0 mm, the second layer had a radius of 43.5 mm, the third layer had a radius of 70.0 mm, and the fourth and outermost layer had a radius of 88.0 mm. The addition of a fourth layer allowed for the reconstruction of charged tracks with just the SVD [13]. See Figure 1.6.

The beam pipe radius was reduced from 20 mm to 15 mm so that the SVD could be placed closer to the IP. This allowed for better vertex resolution. The SVD2 upgrade also improved the angular acceptance to $17° < \theta < 150°$.

SVD Performance Metrics

Critical to SVD performance were the SVD-CDC track matching efficiency and impact parameter resolution, the resolution of the point of closest approach to the IP. The SVD-CDC track matching efficiency was defined as "the probability that a CDC track within the SVD acceptance has associated SVD hits in at least two layers, and in at least
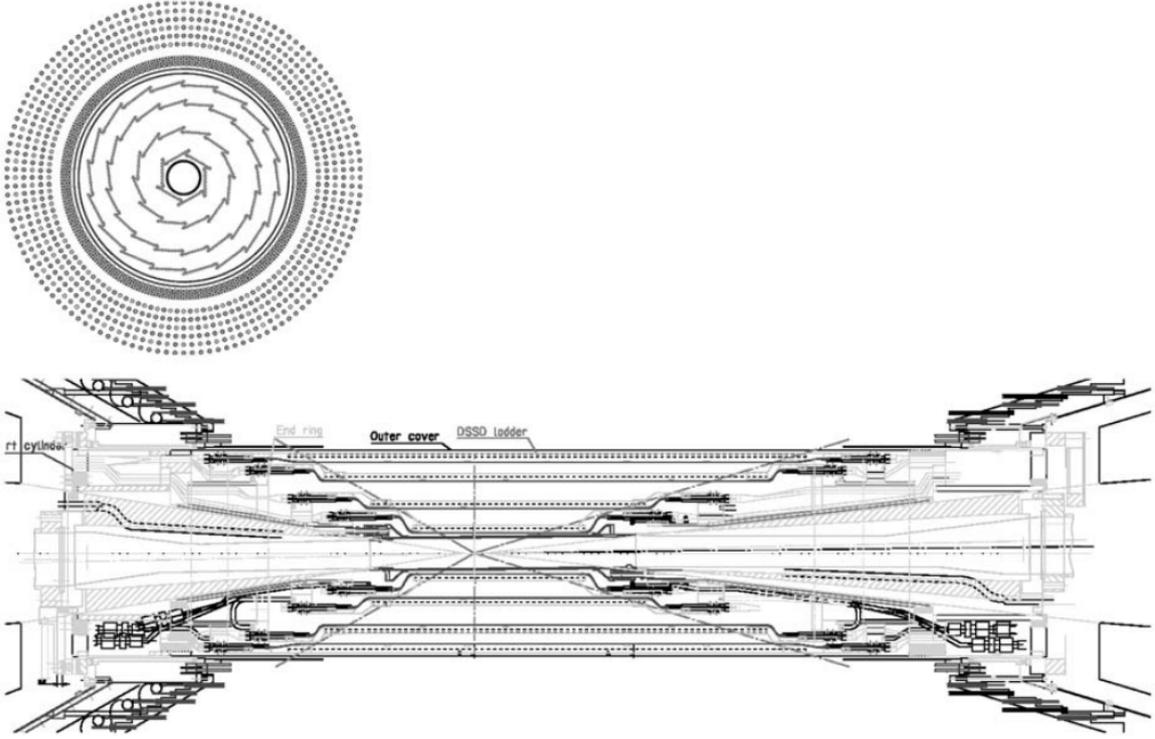
Figure 1.6.: Schematic of the SVD2; cross-sectional image (top); module side view (bottom)

one layer with both the *r-φ* and *r-z* information" [12]. This efficiency was determined to be greater than 98% [12, 14].

The impact parameter resolutions for SVD1 are given [12] by

$$\sigma_{xy}^{\text{SVD1}} = 19 \oplus 50/p\beta \sin^{3/2}\theta \mu m \tag{1.2}$$

$$\sigma_{z}^{\text{SVD1}} = 36 \oplus 42/p\beta \sin^{5/2}\theta \tag{1.3}$$

For SVD2, the impact parameter resolutions are given [14] by

$$\sigma_{xy}^{\text{SVD2}} = 17 \oplus 34/p\beta \sin^{3/2}\theta \mu m \tag{1.4}$$

$$\sigma_z^{\text{SVD2}} = 26 \oplus 33/p\beta \sin^{5/2}\theta \tag{1.5}$$

where p is the track momentum and θ is the angle of incidence.

1.2.3. Aerogel Chereknov Counter

For $K^\pm$ and $\pi^\pm$ separation, silica (SiO_2) aerogel Cherenkov counters (ACC) were deployed as part of the Belle particle identification (PID) system. They were based on the well known Cherenkov effect, whereby a particle traveling faster than the speed of light in a medium emits electromagnetic radiation. As properties of Cherenkov light depend on particle velocities, charged kaons and charged pions with different momenta were discriminated when passing through the ACC. The ACC covered a momentum range of up to 3.5 GeV/c for $K^\pm$ and $\pi^\pm$ [12, 15]. Figure 1.8 shows the ACC performance in separating kaons from Bhabha backgrounds.
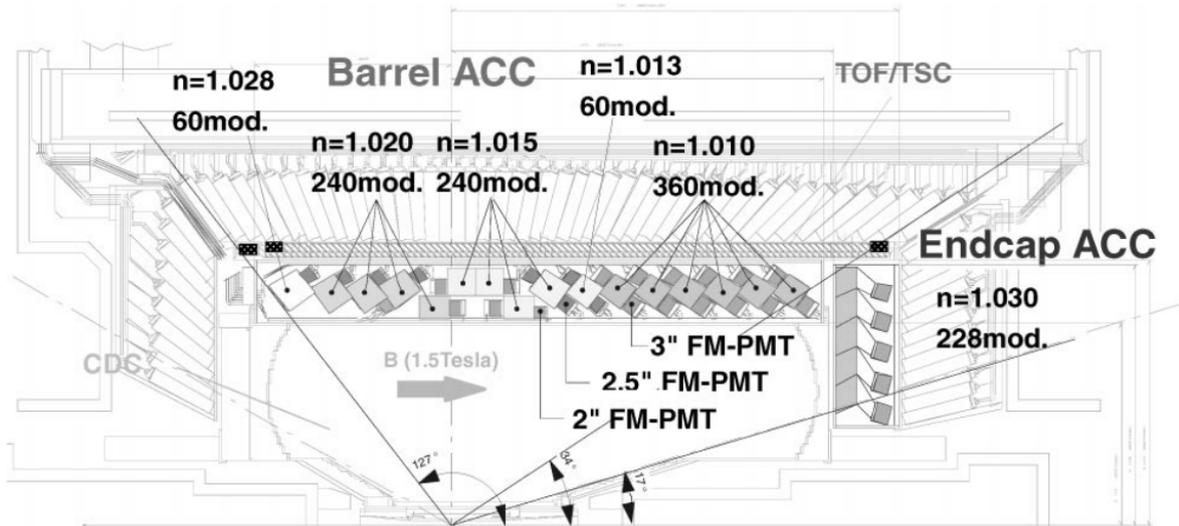
Figure 1.7.: Diagram of the ACC

There were 1080 ACC modules in total: 960 in the barrel region, arranged in the ϕ direction, and 228 in the forward end-cap arranged in concentric rings. All modules pointed toward the IP. To cover the momentum range [1.2, 3.5] GeV/c^2, the aerogels used refractive indices that ranged from 1.01 - 1.03 [12, 15].
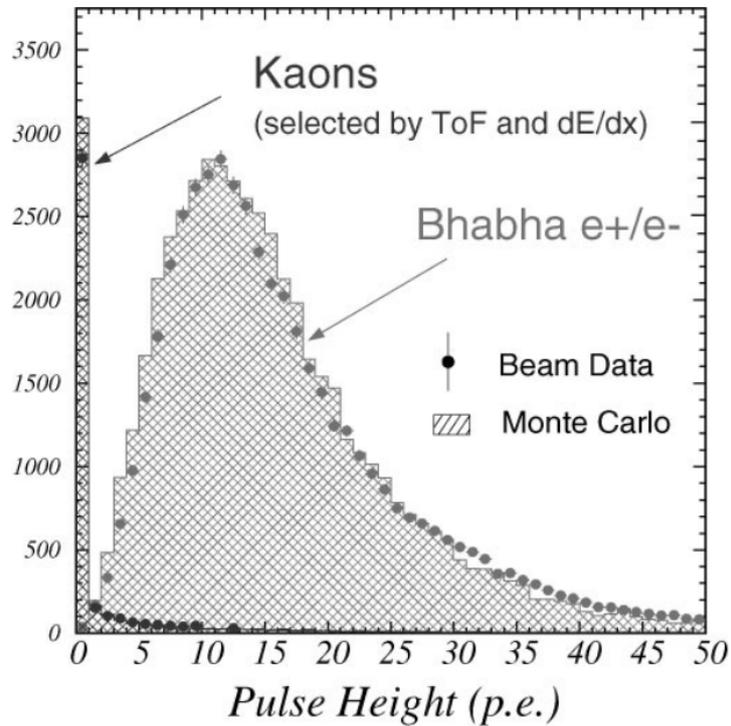
Figure 1.8.: Hadron separation performance with the ACC
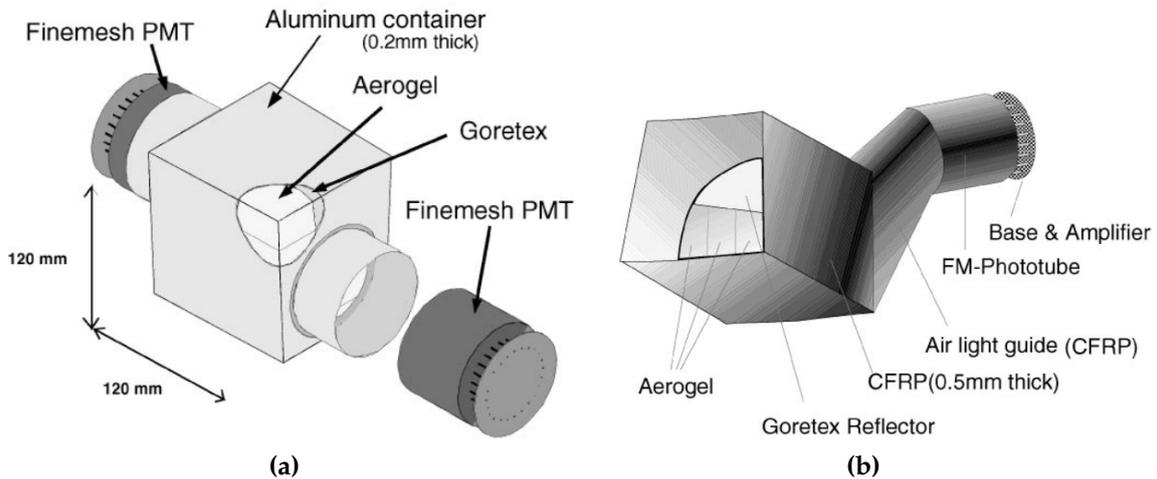
Figure 1.9.: PMTs for the ACC; Barrel ACC (left), Endcap ACC (right)

The ACC boxes were made of aluminum and inside were stacked aerogel tiles that were read out by fine-mesh photomultiplier tubes (FM-PMTs). The average quantum efficiency was approximately 25% at a wavelength of 400 nm. The FM-PMTs operated in the 1.5 T magnetic field [15]. Figure 1.9 shows the PMT configurations.

1.2.4. Time of Flight

The time-of-flight (TOF) sub-detector was part of the PID system that was used for $K^{\pm}/\pi^{\pm}$ separation for the momentum range below $1.2\,\mathrm{GeV}/c$. Around 90% of events at the $\Upsilon(4S)$ resonance contain kaons or pions in this momentum range. To achieve the PID goals of the experiment, a timing resolution of 100 ps was desired. Large area FM-PMTs and a fast scintillator material were used to achieve the 100 ps timing resolution goal [12, 16].
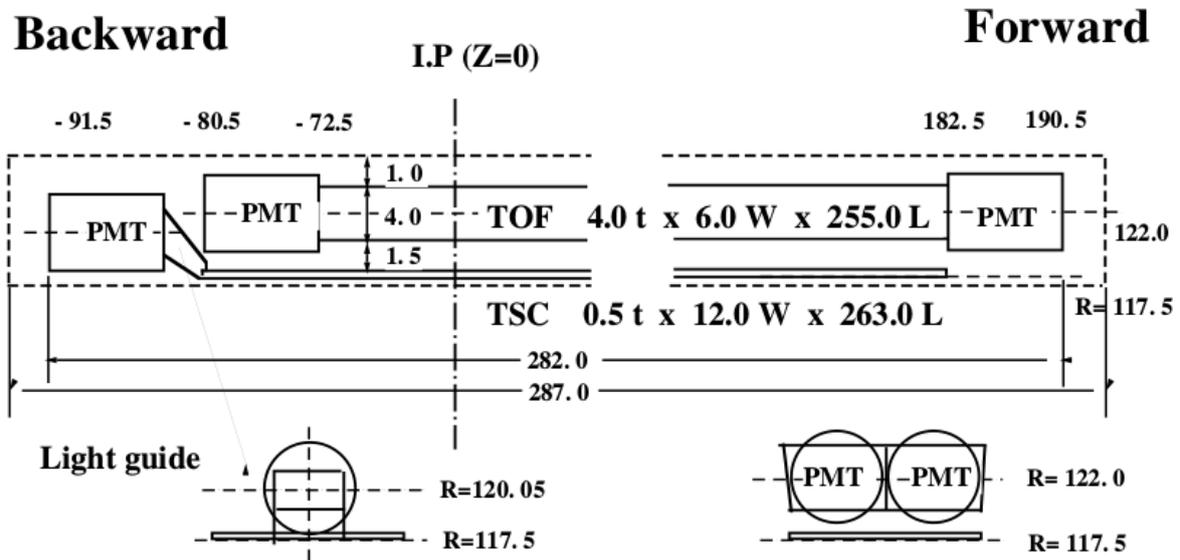
Figure 1.10.: Diagram of the TOF system

One TOF module was composed of two TOF counters and one trigger scintillator counter (TSC). The entire TOF system had 128 TOF counters and 64 TSCs, making a total of 64 TOF+TSC modules. The modules were situated radially around the IP at a distance of 1.2 m from it. There was a 1.5 cm radial gap between the TOF counters and the TSC of one TOF+TSC module. The polar angle acceptance was $33° < \theta < 121°$. The 1.5 cm gap helped to reduce photon conversion backgrounds by vetoing events that were coincident in the TSC and TOF counter components [16].

The FM-PMTs were mounted at the ends of the TOF and TSC component. On TOF counter ends the FM-PMTs were mounted with a 0.1 mm air gap, and on the TSCs the FM-PMTs were glued to the light guides that were situated on the backward ends. The TOF system was located in the 1.5 T magnetic field. The 0.1 mm gap helped to reduce saturation in the FM-PMTs from large pulses. Low energy conversions curl up in the gap. TOF signal were read out by the TOF front end electronics (TOFFEE) [16].

Overall the TOF system gave good separation between charged particles in hadronic events, and results were in good agreement with Monte Carlo (MC) studies, giving a timing resolution of approximately 120 ps. See Figs. 1.11 and 1.12 for particle separation and timing resolution performance, respectively [16].
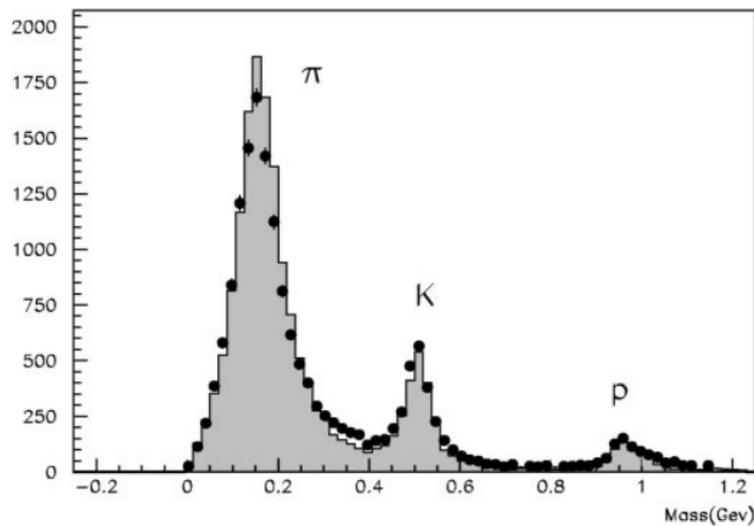
Figure 1.11.: Separation of charged particles given by TOF measurements. Circles are physics data and the solid shaded region is MC.
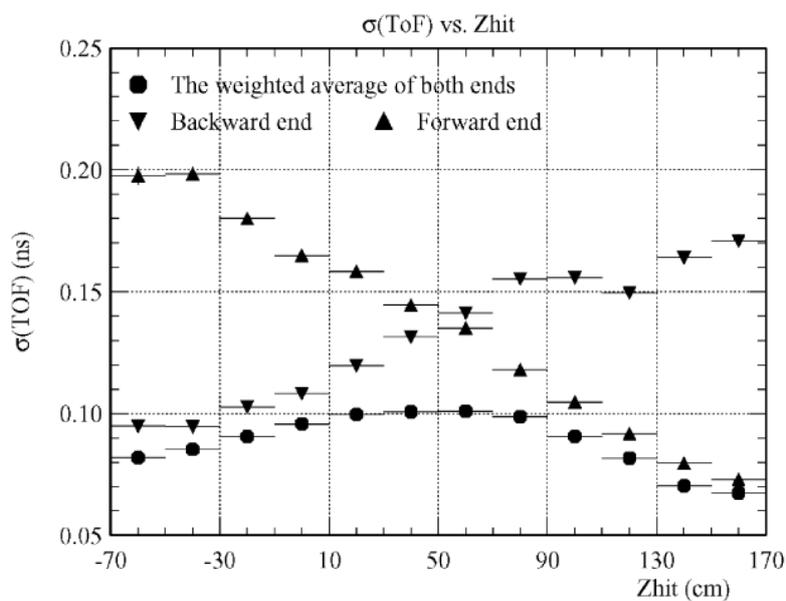
Figure 1.12.: TOF timing resolution for muon pair events

1.2.5. Central Drift Chamber

The central drift chamber (CDC) was another sub-detector in Belle that was also part of the PID system. The CDC design was meant to provide information for PID through dE/dx measurements of charged tracks in the CDC helium-ethane gas, precise measurement of charged particle three-momentum vectors, and information used to discriminate against beam-induced backgrounds, at the trigger level. It was located inside the 1.5 T magnetic field [12, 17].
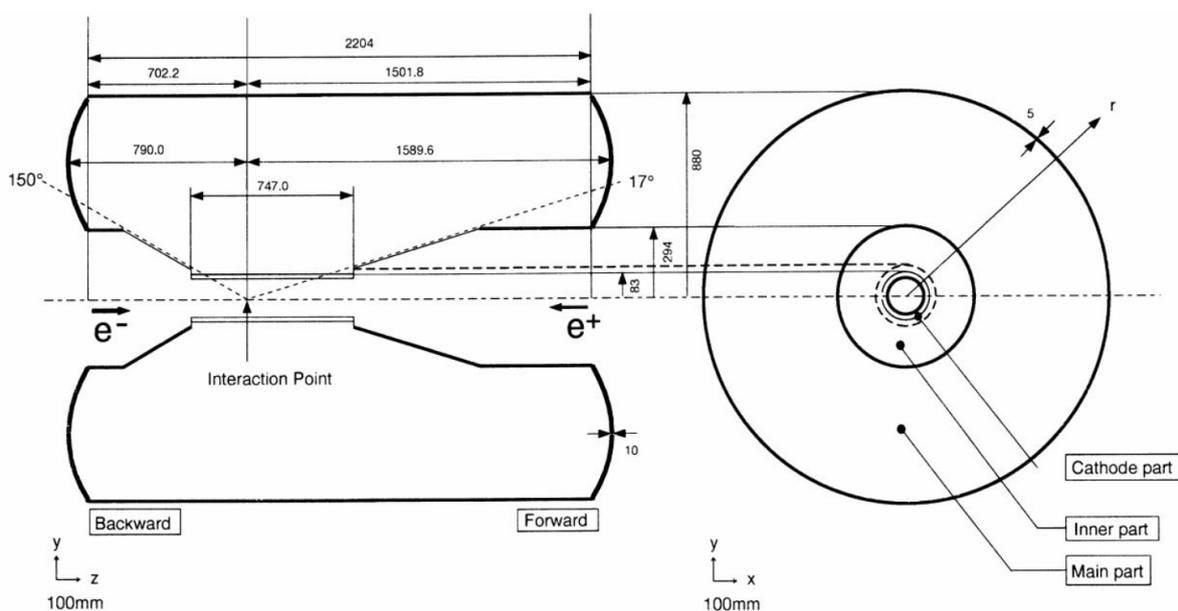
Figure 1.13.: Diagram of the CDC

The CDC was a cylindrical drift chamber with an inner and outer radius. Its length was 2400 mm with an inner radius of 83 mm and an outer radius of 874 mm. The CDC was asymmetric in the z-direction to account for the $\beta\gamma = 0.425$ boost from the accelerator. Its acceptance was $17° < \theta < 150°$ [12].

The chamber contained 8400 drift cells. The cells were arranged into six axial and five stereo layers so as to provide reliable z-position measurements. Each cell was approximately rectangular and had one sense wire surrounded by eight field wires. There were 33344 field wires [17].

The He-C_2H_6 (helium-ethane) gas was chosen because it is a low-Z gas. Low-Z gases are desirable as they can provide better momentum resolution due to reduced multiple Coulomb scattering. Compared to other gases, low-Z gases also provide less
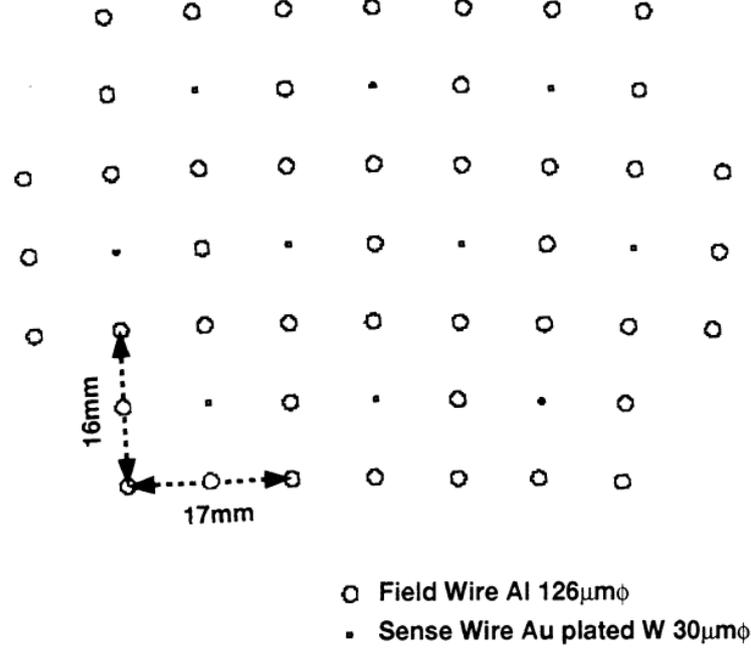
Figure 1.14.: Diagram of the wire configuration inside a drift cell of the CDC

background from synchrotron radiation. The CDC gas was a 50-50 mixture of helium and ethane. The ethane constituent of the the gas gave good dE/dx resolution [17].

Particle identification was performed through charged particle energy loss from gas ionization, dE/dx. Charged particles going through the CDC had a helical trajectory as they were inside the 1.5 T field. The charge of the particle was given by the parity of the helix. When a charged particle passed through the gas it ionized the gas. This ionization produced further ionizations that were detected on the sense wires. Timing and drift-velocity information was combined to reconstruct the charged particle helix parameters [17]. The p_t measurement was determined by the helix curvature and p_z by the helix slope. The resolution for transverse momentum, p_t, is given [9] as

$$\sigma_{p_t}/p_t = 0.0019 p_t[\text{GeV}/c] \oplus 0.0030/\beta \tag{1.6}$$

1.2.6. Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) was an important sub-detector for identifying electrons and measuring neutral particles that decay to photons ($\to \gamma\gamma$). It was required to cover an energy range from approximately 20 MeV to 8 GeV. This range covered events that included background to photons emitted from B-meson decays. An energy resolution of approximately 5% for a 100 MeV photon was required and made rejecting hadronic background more efficient. The ECL was situated inside the 1.5 T magnetic field and was constructed from segmented CsI(Tl) crystals for scintillation, as they have a relatively good photon yield. They were read out by silicon photodiodes [12, 18].

The ECL was composed of the barrel and endcaps. The barrel was 3 m in length and had an inner radius of 1.25 m. The endcaps were at z = −1.0 m and z = +2.0 m from the IP. The barrel had 6624 crystals, the forward endcap had 1152 crystals, and the backward endcap had 960 crystals. The total polar angle coverage for the ECL was $17° < \theta < 150°$ [12]. The support structure was made of aluminum and stainless steel [18].

Figure 1.16 from [18] shows the reconstruction of the π^0 and η mesons from two-photon combinations in hadronic events, in the ECL.

Extreme Forward Calorimeter

The extreme forward calorimeter (EFC) was used to extend the acceptance of the ECL. The acceptance was extended to the regions $6.4° < \theta < 11.5°$ in the forward region and $163.3° < \theta < 171.2°$ in the backward region [12].

The EFC was also used to reduce beam backgrounds coming into the CDC and as a beam monitor with information fed to the KEKB control group and Belle [12].

The EFC was designed to be radiation hard as it was situated near the IP, close to the beam pipe. It was constructed of $Bi_4Ge_3O_{12}$ (bismuth germanate, BGO) [12].
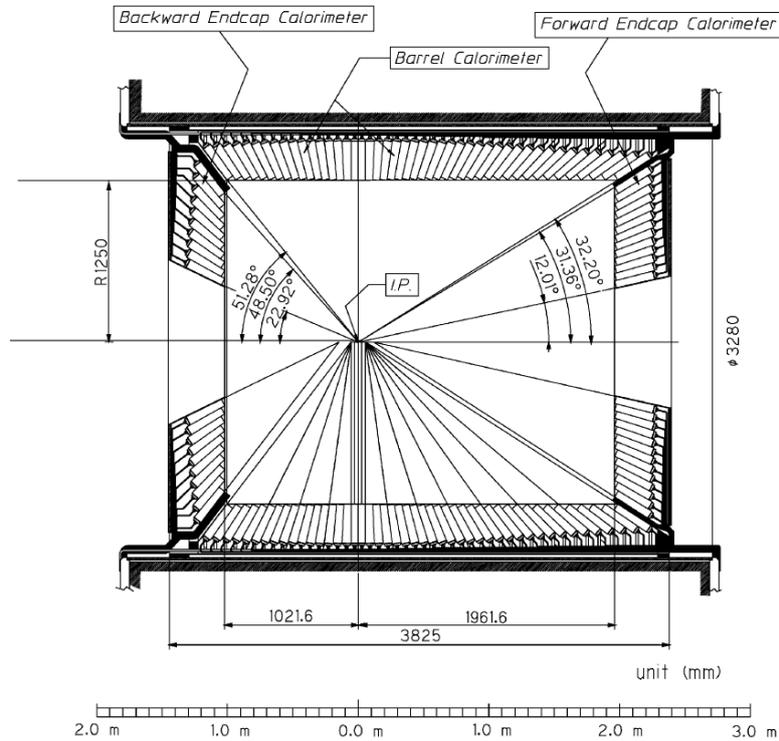
Figure 1.15.: Diagram of the ECL as shown from a side-view
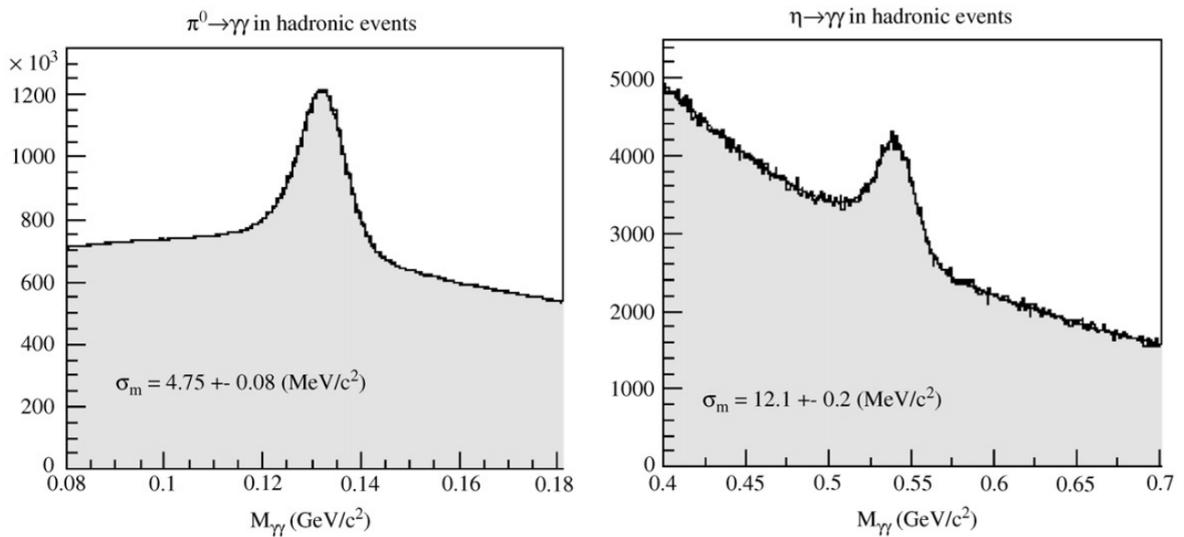
Figure 1.16.: Reconstruction of π^0 and η invariant masses from $\gamma\gamma$ pairs
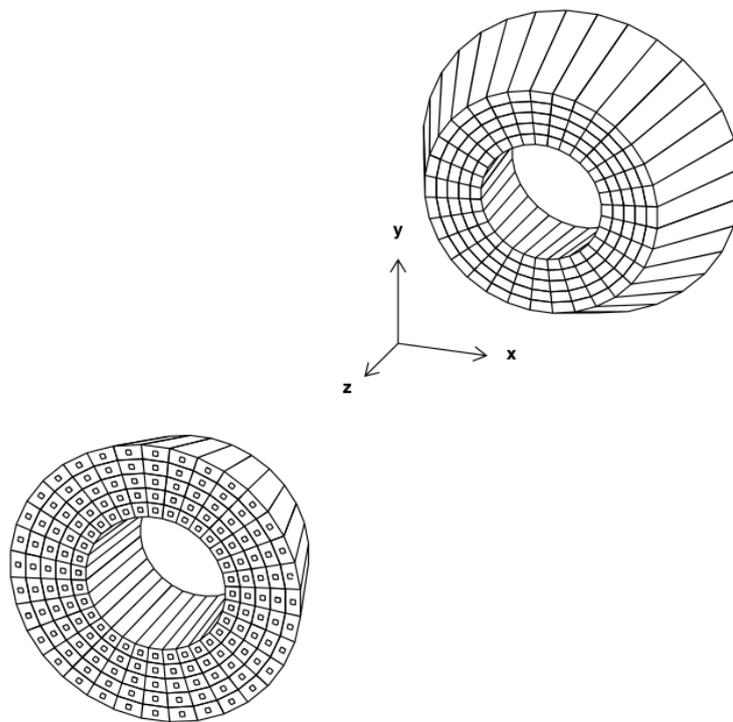
Figure 1.17.: Diagram of the EFC crystals in the forward and backward regions

1.2.7. K_L^0 and Muon Detector

The K_L^0 and muon (KLM) detector was the outermost of the Belle sub-detectors. Its purpose was the efficient detection of K_L^0 and muons with momenta greater than 600 MeV/c. It was comprised of an octagonal barrel region, and two endcaps: forward and backward. The KLM covered a polar angle range of $20^\circ < \theta < 155^\circ$ [12].

The KLM was made up of alternating layers of charged particle detectors and iron. The layers of charged particle detectors were made up of glass electrode resistive plate counters (RPCs). There were 15 layers of RPCs and 14 layers of iron. The iron layers were 4.7 cm thick and provided approximately 3.9 interaction lengths. The K_L^0 particles produced showers in the iron layers of the KLM and the CsI(Tl) layers of the ECL. Muons traveled further in the detector and were discriminated from charged hadrons by penetrating through the alternating layers of RPC and iron. Figure 1.18 shows the muon detection efficiency versus momentum [12].

The RPCs were constructed from two parallel plate electrodes. The plates had a resistivity of $\geq 10^{10}$ Ω-cm. Between the plates was a gas-filled layer. As a charged
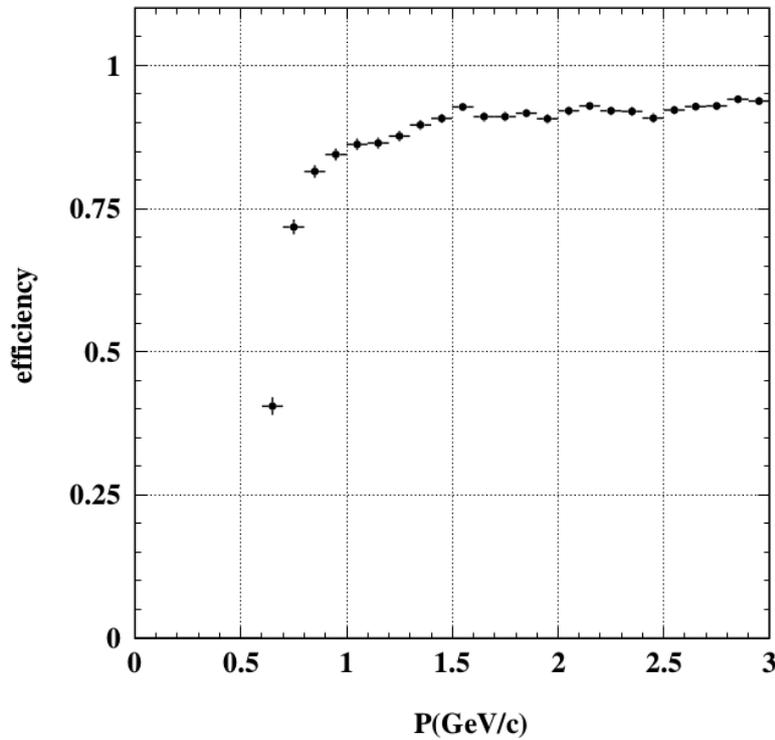
Figure 1.18.: Muon detection efficiency vs momentum.

particle passed through this layer it could create an electric discharge. This discharge was contained by the plate resistivity and gas, and was picked up by the detector pickup-strips that provided event position and time [12].

1.2.8. Solenoid and Iron Yoke

The previously mentioned 1.5 T magnetic field was generated by a cylindrical super-conducting solenoid. It generated the magnetic field in a volume that was 4.4 m in length and 3.4 m in diameter. The components of the Belle detector structure contained iron and other materials. The iron of the Belle detector was used as a return path for the magnetic flux and as an experimental support structure [12]. Figures 1.19 and 1.20 show diagrams of the physical structure while Fig. 1.21 shows magnetic field strength measurements as a function of z for several values of r [12].
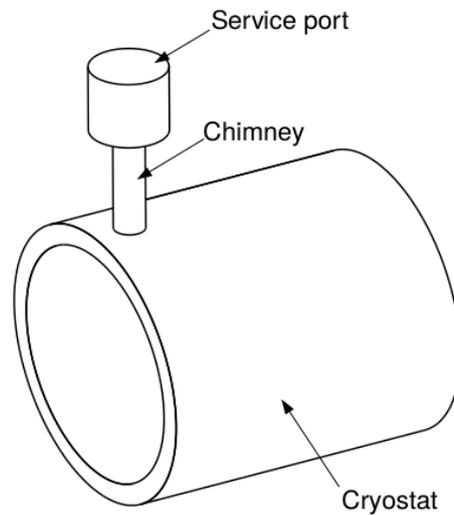
Figure 1.19.: Diagram of the solenoid with cryostat
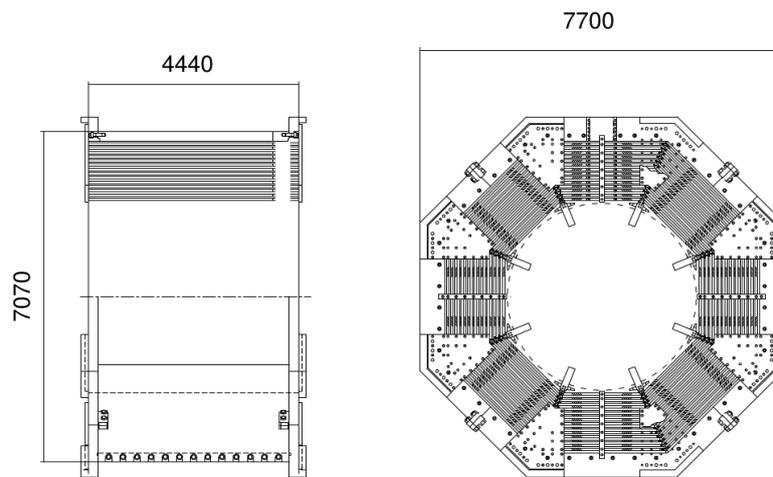
Figure 1.20.: View of the Belle barrel iron yoke structure

1.2.9. Trigger System

The general purpose of a trigger system was to identify potential physics events of interest. It was required to work in the high event rate environment of the KEKB collider and it was required be efficient in identifying physics events. See Table 1.1 for a list of physics processes at the $\Upsilon(4S)$ resonance [12, 19].

The Belle trigger system was primarily composed of a hardware-level trigger and a software-level trigger. The former was dubbed the Level-1 trigger and the
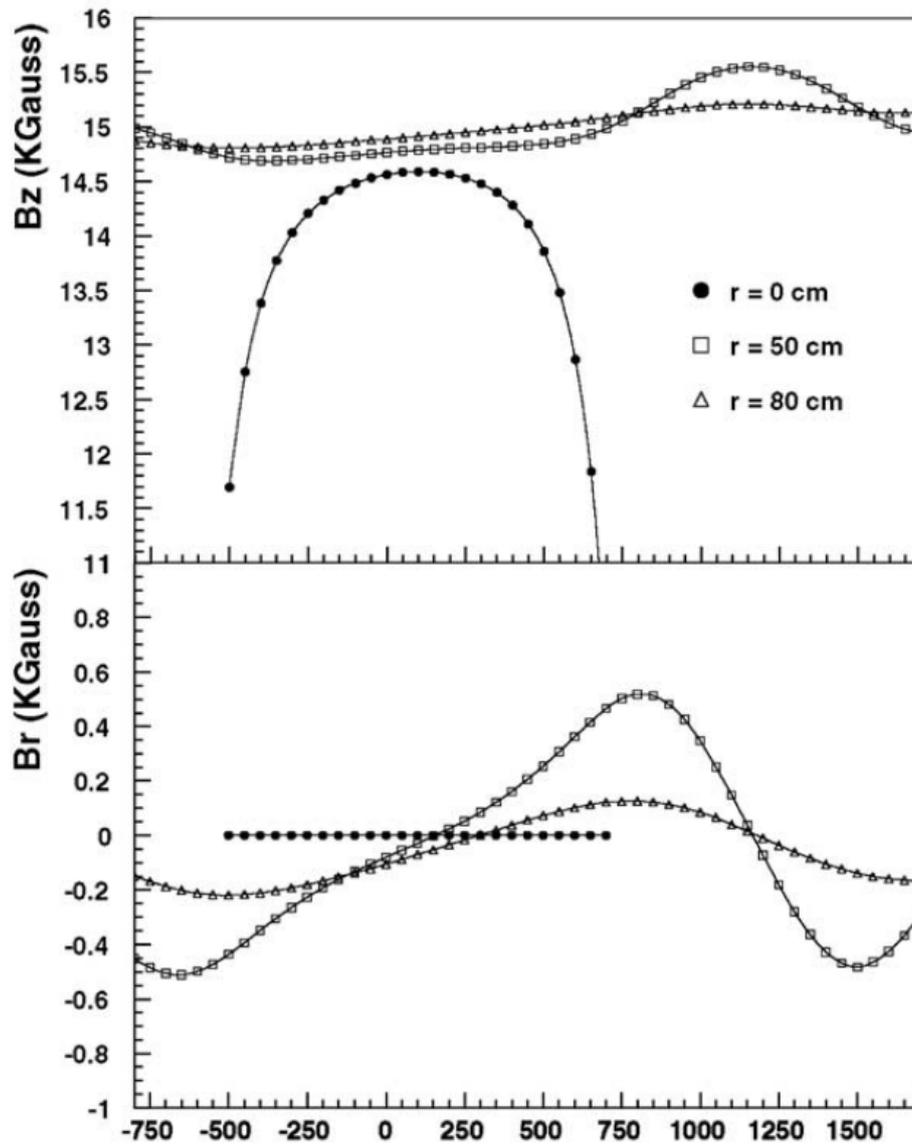
Figure 1.21.: Magnetic field strength measurements

latter the Level-3 trigger. The Level-3 trigger was implemented in a system of online computers. The Level-1 trigger consisted of all sub-detector trigger systems and the Global Decision Logic (GDL), as seen in Fig. 1.22. The GDL utilized information from sub-detector systems in the form of information on tracks and deposited energy [12].

The Level-3 trigger implemented various algorithms that reduced the data from the Level-1 trigger. A final Level-4 trigger was implemented offline. Events that pass the Level-4 trigger were used as data for physics analyses [19].

Physics Process	Cross Section (nb)
$\Upsilon(4S) \to B\bar{B}$	1.15
Hadron production from continuum	2.8
$\mu^+\mu^- + \tau^+\tau^-$	1.6
Bhabha: $e^+e^- \to e^+e^-$ ($\theta_{lab} \geq 17^\circ$)	44
$\gamma\gamma$ ($\theta_{\mathrm{lab}} \geq 17^\circ$)	2.4
2γ ($\theta_{\mathrm{lab}} \geq 17^\circ$), $p_t \geq 0.1$ GeV/c)	15

Table 1.1.: Physics processes at the $\Upsilon(4S)$ resonance
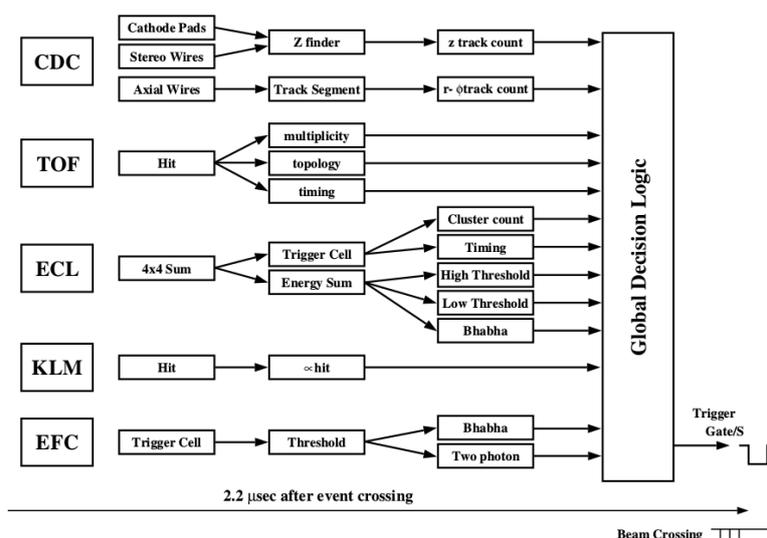
Figure 1.22.: Chart of the Level-1 Trigger at Belle

1.2.10. Data Acquisition

Initial Data Acquisition System

The Belle data acquisition (DAQ) system acquired information in parallel from the SVD, CDC, ACC, TOF, ECL, KLM, and EFC sub-detectors. The latter six were read out by charge-to-time and time-to-digital systems. The KLM used only time-to-digital systems while the SVD used flash ADCs [12, 20].

DAQ Upgrade

Pre-upgrade, the Belle DAQ was comprised of the sub-detector readouts, event builder, online computer system, and storage. In the middle of 2001 all but the readout systems were merged into a "switchless event building farm". This upgrade also constituted part of the newer Level-2 trigger system that improved data reduction [21].
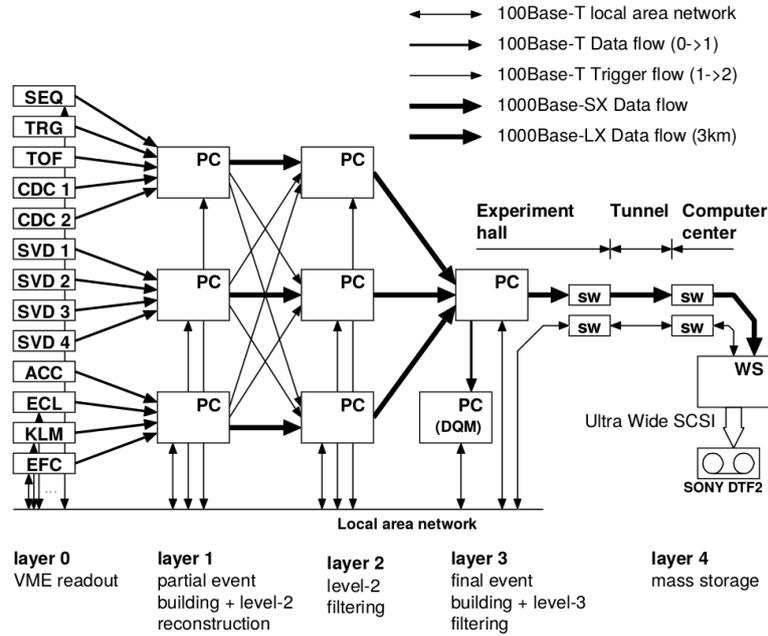
Figure 1.23.: Chart of upgraded Belle DAQ system with streamlined event builder and Level-2 trigger

1.3. Particle Identification (PID)

Track information from the ACC, TOF, and CDC can be combined to form a likelihood ratio for particle hypotheses. For the kaon ($K^\pm$)/ pion ($\pi^\pm$) separation, for example, the likelihood ratio is formed as [22]

$$\mathcal{L}(K^\pm) = \frac{P_\text{ACC}(K^\pm)P_\text{TOF}(K^\pm)P(K^\pm)_\text{CDC}}{P_\text{ACC}(K^\pm)P_\text{TOF}(K^\pm)P(K^\pm)_\text{CDC} + P_\text{ACC}(\pi^\pm)P_\text{TOF}(\pi^\pm)P(\pi^\pm)_\text{CDC}} \quad (1.7)$$

where P are probability distributions determined by the sub-detectors.

The likelihood ratio for the pion is then simply

$$\mathcal{L}(\pi^{\pm}) = 1 - \mathcal{L}(K^{\pm}) \tag{1.8}$$

These quantities can then be used for physics analyses that involve the need to separate different species of particles.

Chapter 2.

SuperKEKB and Belle II

2.1. SuperKEKB

SuperKEKB is the successor accelerator to KEKB. It is designed to have a peak luminosity of 8×10^{35} cm^{-2}s^{-1}, versus KEKB's achieved luminosity of 2.11×10^{34} cm^{-2}s^{-1} [23]. As shown in Fig. 2.1, SuperKEKB's luminosity will be higher than that of the Large Hadron Collider (LHC). Table 2.1 gives a comparison between the KEKB and SuperKEKB accelerator parameters.
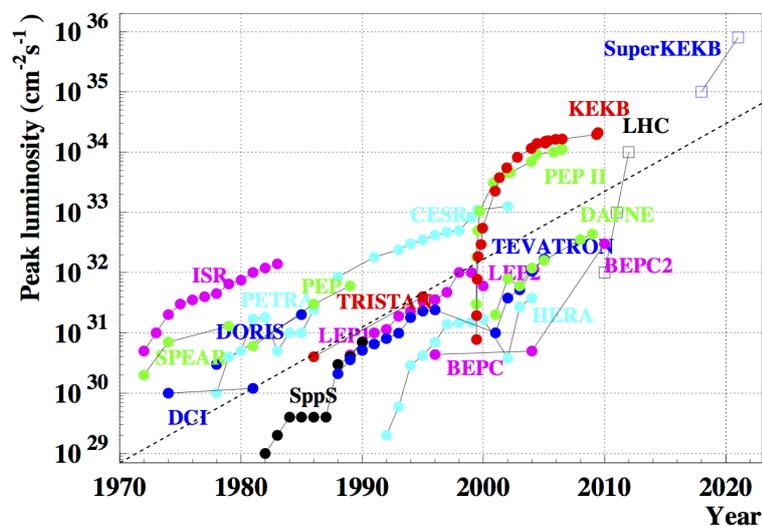
Figure 2.1.: Accelerator luminosities

2.1.1. Nano-Beams

SuperKEKB achieves this increase in luminosity through the use of "nano-beams". This is done by modifying the vertical β-function at the interaction point (IP), β^*, by squeezing it. The parameter β_y^* characterizes the vertical dimension of the beam [24] and is squeezed, minimizing the overlap of the longitudinal component of the two colliding beams [23, 25].

If d is the region of overlap between the two beams, then the value of d is given by $d \cong \frac{\sigma_x^*}{\phi}$, where σ_x^* is the horizontal beam size at the IP and ϕ is the horizontal half crossing angle. This overlap region is smaller than the beam bunch length, σ_z, and can be considered as effective bunch length. When this parameter is minimized, it sets an effective lower limit on the β_y^* function $\beta_y^* > d$ [23]. Figure 2.2 is a diagram of this strategy.
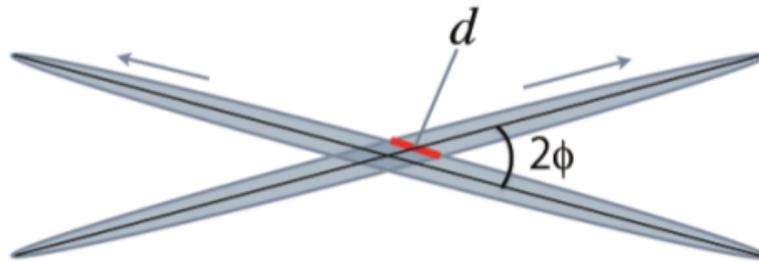
Figure 2.2.: Nano-Beam collision diagram

For detailed technical information about luminosity, beam optics, RF design, magnets, linac, and other items relating to the nano-beam scheme, readers are directed to the Belle II Technical Design Report (TDR), from which the preceding information was obtained [23].

2.2. Belle II

2.2.1. Belle II Overview

The Belle II experiment is the successor to the Belle experiment. They share the same design philosophy and Belle II's sub-detectors largely share the same purpose as

Parameter	KEKB Achieved	SuperKEKB
Energy (GeV) LER/HER	3.5/8.0	4.0/7.0
ξ_y LER/HER	0.129/0.090	0.090/0.088
β_y^* (mm) LER/HER	5.9/5.9	0.27/0.41
I (A) LER/HER	1.64/1.19	3.60/2.62
Luminosity ($\times$ 10^{34} cm^{-2}s^{-1})	2.11	80

Table 2.1.: Accelerator parameters compared between KEKB and SuperKEKB, from the Belle II TDR. Here, I is the beam current and ξ_y is the beam-beam parameter that characterizes the strength of beam-beam interactions [24]

.

those of Belle. Belle II is a general purpose magnetic spectrometer that has upgraded sub-detectors, data analysis software, and computing systems. Figure 2.3 shows a schematic of the full Belle II experiment [25].

Belle II will collect a 50 ab^{-1} integrated luminosity data sample, compared to Belle's 1 ab^{-1} integrated luminosity sample, and will hopefully provide more insight into fundamental interactions by studying the physics of B-mesons at the high intensity frontier [25]. Its sub-detectors are summarized in the following sections.
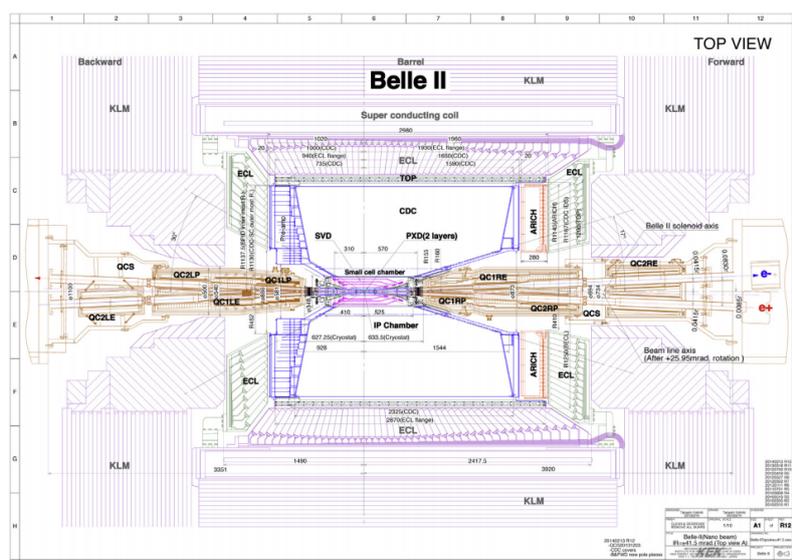
Figure 2.3.: Belle II schematic

2.2.2. Vertex Detector

The vertex detector is comprised of two components, the silicon vertex detector (SVD) and the pixel detector (PXD), described below. Together these structures total six layers: four from the SVD and two from the PXD [25]. These two detectors, coupled with the smaller beam pipe radius at the IP, compensate for SuperKEKB's boost of $\beta\gamma = 0.28$, which is smaller than that of KEKB.

Silicon Vertex Detector

The Belle II SVD is used to measure decay vertices of B-mesons. It is designed to be radiation hard and made with silicon strip sensors [23].

The Belle II SVD works in conjunction with the CDC and the Pixel Detector (PXD) to reconstruct tracks and vertices. It has an acceptance of $17° < \theta < 150°$, with an inner radius of 38 mm and outer radius of 140 mm [23, 25].

To meet all the performance requirements the Belle II SVD is composed of four layers of doubled sided silicon strip detectors (DSSDs) and is read out by Application Specific Integrated Circuits (ASICs), the APV25, originally developed for CMS [23, 25]. Figure 2.4 shows the four layers of the SVD with the sensors slanted in the forward region.
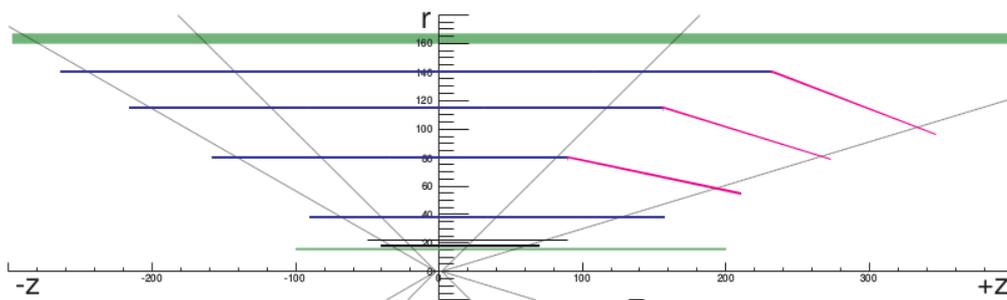
Figure 2.4.: Belle II SVD; units are in millimeters

Pixel Detector

Due to the high luminosity of the SuperKEKB accelerator and the narrower beam pipe, backgrounds increase and the occupancy of the vertex detectors increase beyond usefulness. To overcome this issue, the innermost sub-detector is made up of pixel

detectors. The pixel detectors have lower occupancy and a larger number of channels. Occupancy is defined as the fraction of channels hit in a triggered event [23]. The pixel detectors are comprised of DEPFETs (DEpleted Field Effect Transistors) [23, 25].

A pixel sensor of the PXD has current-digitizing electronics at its end. These electronics are situated outside of the acceptance domain. The PXD is made up of two layers of sensors, with inner radius of 14 mm and outer radius of 22 mm. The inner radius is comprised of eight sensors that are 15 mm in width and have an active sensor length of 90 mm [23, 25]. The outer radius is comprised of twelve sensors that are 15 mm in width with an active sensor length of 123 mm. The sensors are attached to a mechanical support structure with cooling [23]. Figure 2.26 shows the layout of the PXD sensors.
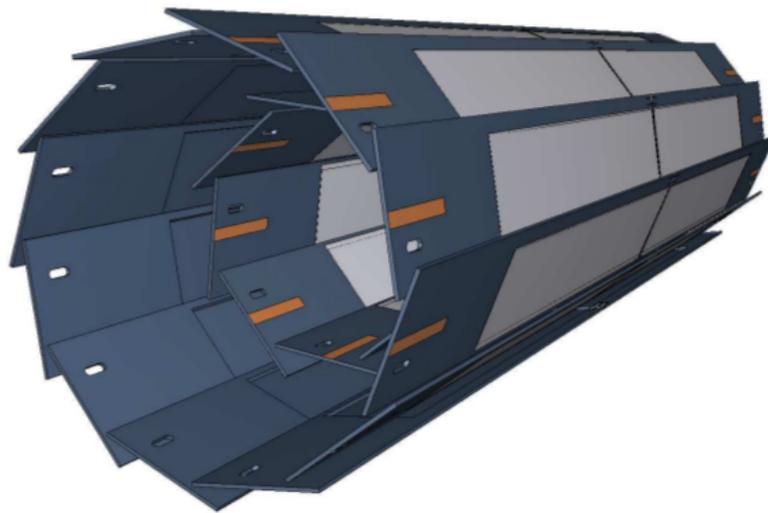
Figure 2.5.: Layout of the PXD sensors

2.2.3. imaging Time of Propagation Detector

Overview

The imaging Time of Propagation (iTOP) detector is part of the particle identification (PID) system of Belle II. It is used for charged kaon and pion separation [23, 25, 26] in the higher-background environment of Belle II.

Good separation from the iTOP for $K^{\pm}/\pi^{\pm}$ transverse momentum in the range of 1-4 GeV, with 85-90% efficiency and less than 5% misidentification, is required to achieve Belle II's flavor physics goals [23, 27].
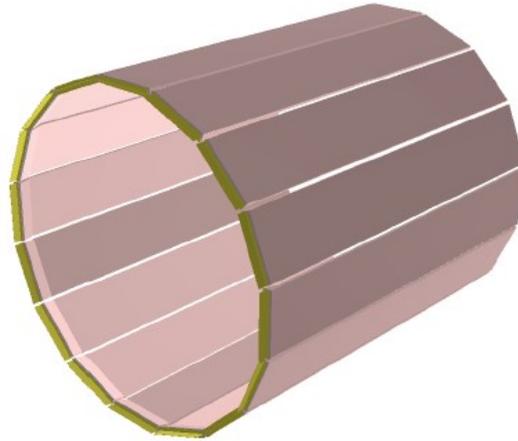
Figure 2.6.: A drawing of the arrangement of the 16 iTOP modules

The iTOP is an array of Cherenkov photon generators and measurement devices. It is is segmented into 16 modules. Figure 2.6 shows that arrangement. A single iTOP module consists of a quartz (fused silica) bar with total dimensions 2600 x 450 x 20 mm^3 (2x 1250 mm quartz radiator + 100 mm length spherical mirror), prism, micro-channel plate photomultiplier tubes (MCP-PMTs), readout electronics, cooling, and a QBB (and its associated cables and power supply) that houses the quartz and serves as a support structure.

At the end of the quartz radiator the prism is glued. It serves the function of an expansion wedge that allows the Cherenkov photons to reach the PMTs mounted on the readout electronics, as the height of the readout electronics and MCP-PMT windows are greater than that of the quartz radiator, which is 20 mm [26]. The prism at its maximum height is 51 mm, with a length of 100 mm, and a width of 456 mm. It is sloped at an angle of 18.07°. This is shown in Fig. 2.7. Between the readout electronics and the prism there is a silicone "optical cookie" that optically couples the two components. Between the optical cookie and prism there is placed a small amount of oil to make sure there are no voids in between the optical contacts [26].

When charged particles, particularly charged kaons and pions, pass through a quartz bar faster than the speed of light in quartz, Cherenkov radiation is produced. By total internal reflection the light is directed toward the readout electronics that digitize arrival the time of the photons. Some Cherenkov photons travel directly to the end of

the quartz bar where the readout electronics are located. Others travel to the opposite end and are reflected back to the electronics end by a spherical mirror [23, 25–27].

The Cherenkov angle for the produced photons is defined by the relation $\cos \theta_c = 1/n\beta$, where n is the index of refraction and β is the speed of the particle. For fused silica, $n = 1.47$ at a wavelength of 410 nm [28]. For $K^\pm$ and $\pi^\pm$, θ_c is different and leads to different propagation times in the quartz [29]. The timing and hit position information obtained from the readout electronics, along with other sub-detector information, can produce particle identification information [27]. Time-of-flight information is also included due to the use of the SuperKEKB radio-frequency clock as a reference for particle collisions [27]. Figures 2.8 and 2.9 show the basic principle of the iTOP [26].
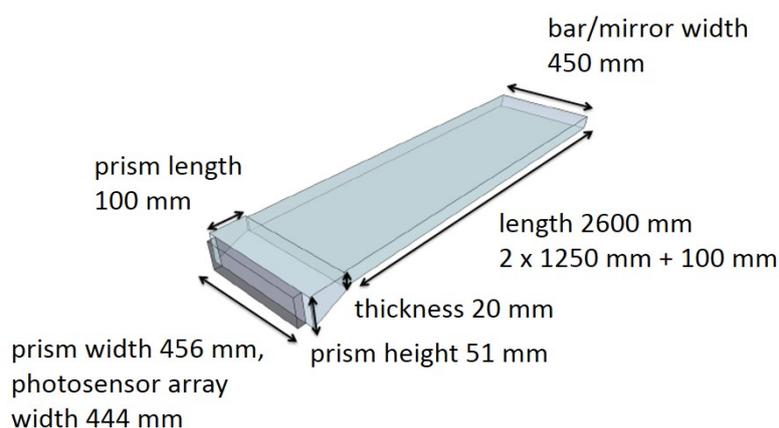
Figure 2.7.: iTOP module with parts labeled, without QBB
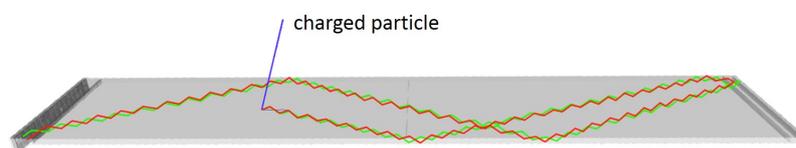
Figure 2.8.: iTOP principle of operation

The readout electronics consist of four individual readout modules (colloquially termed "boardstacks") for every iTOP module. Each readout module is connected to eight micro-channel plate photomultiplier tubes (MCP-PMTs), for a total of 32 MCP-PMTs per iTOP module. The Cherenkov photons are observed by the MCP-PMT pixels. The signals from the MCP-PMTs are read by the readout modules (boardstacks). Each readout module is comprised of four carrier boards and a Standard Control Read-Out
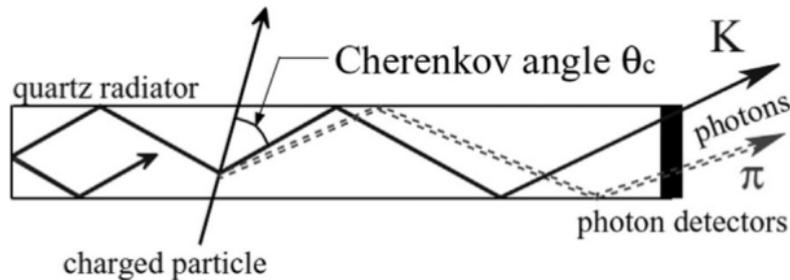
Figure 2.9.: iTOP principle of operation

Data ("SCROD") board. The carrier boards ("carriers") each contain four waveform-sampling Ice Ray Sampler, version X (IRSX) ASICs, developed at the University of Hawaii Manoa Department of Physics Instrumentation Development Laboratory (IDL). The carriers pass the waveform information to the SCROD that extracts the timing information of the photon pulses. Photon hit times and positions are used to generate particle hypothesis [26, 27].

MCP-PMTs

Micro-channel plate photomultiplier tubes (MCP-PMTs) are different from the familiar large tube PMTs. A photon impinges on the photocathode and launches a cascade in the micro-channel plates internal to the PMT. They drift to the anode and a signal is output on the pins on the back of the MCP-PMT. Figure 2.10 shows this concept.
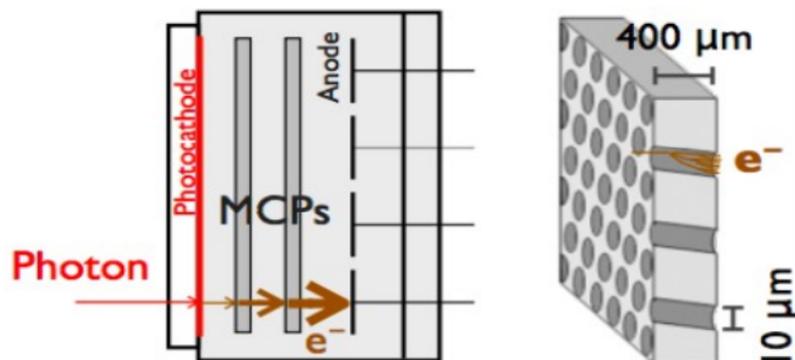
Figure 2.10.: Diagram of the internal working of an MCP-PMT

The MCP-PMTs are produced by HPK. Four PMTs are placed together into an assembly. Two of these assemblies are used for each electronic readout module. The

(a)   **(b)**
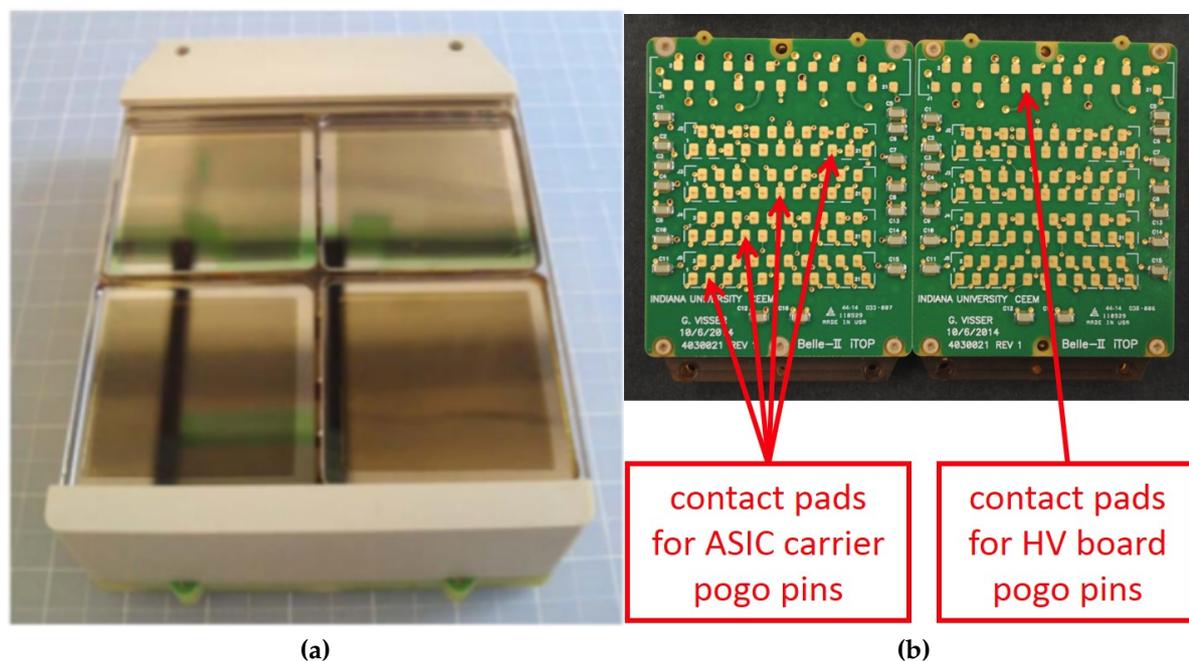
Figure 2.11.: MCP-PMT module assembly with four MCP-PMTs; two of these assemblies are used per a single electronic readout module (left); MCP-PMT PCBs showing pogo pin contact pads for HV and signal routing(right)

MCP-PMT windows have a wavelength cutoff filter at 340 nm to limit chromatic dispersion. The quantum efficiency of this design of PMT is approximately 28% at $\sim$ 400 nm [30, 31]. Characterization and verification of MCP-PMTs was carried out at Nagoya University [26, 29].

The MCP-PMTs are coupled to a printed circuit board (PCB) that provides high voltage (HV) for MCP-PMT operation, and signal routing from the output of the MCP-PMTs to the readout electronics. These specific PCBs are known as "frontboards". The signals are transferred from the frontboards to the carriers via metallic Mill-Max pogo pins that are in physical electrical contact with the frontboard [26, 27]. Polyether ether ketone (PEEK - $C_{21}H_{18}O_3$) structures are used to keep the MCP-PMTs and frontboards aligned to ensure good electrical contact, as shown in Fig. 2.11. The PEEK structures are attached to the PCBs with PEEK screws. These PMT assemblies are optically coupled to an iTOP module prism with an intervening optical cookie.
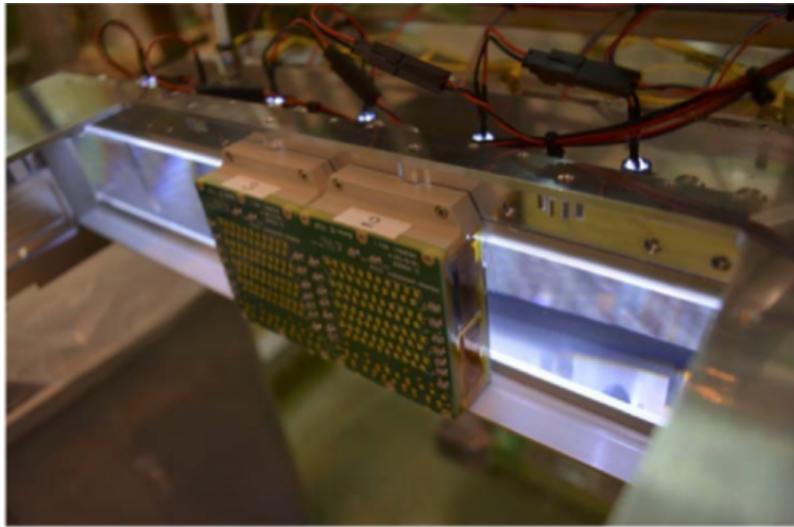
Figure 2.12.: Two MCP-PMT assemblies coupled to an iTOP module prism

Heat Removal

The IRSX ASICs and FPGAs can heat up considerably. While they are rated to operate at temperatures up to 125°C (398.2 K), heat removal is required for efficient operation.

Each of the carriers has an Alodine-coated 6101 aluminum (Al) heat sink screwed to it to provide head dissipation. On each heat sink a copper (Cu) heat spreader for the FPGA is attached (Fig. 2.13). The copper spreader is situated on top of the carrier FPGAs with an intervening layer of Bergquist Liqui-Bond SA 3505 two-part epoxy that is applied to the heat spreader.

Figure 2.13.: carrier Al heat sink with Cu head spreader

Thermal grease (T-Global S606C thermal grease) is then applied to the thermal vias on the carrier PCB edges. This helps to further efficiently remove heat. In addition, a

gap filler (Bergquist Gap Filler 4000) is injected into the final assembly on top of the SCROD FPGA, between it and the SCROD heat sink, for improved heat removal [27].

Readout Electronics

The readout module's four carrier boards and one SCROD board are connected by mezzanine connectors that allow for data transfer rates at the gigabit/second level. This module is electronically coupled to the MCP-PMT assembly [27].

Each carrier contains four IRSX ASICs, two on each side. These ASICs are fast waveform samplers, implemented using a switched capacitor array that digitizes the MCP-PMT signals. Each ASIC contains 8 channels and is capable of 2.7 giga-samples per second (GS/s). Signals from the MCP-PMT frontboards are routed through the pogo pins and go through a two-stage amplifier process before reaching the ASICs [26]. The amplifier chains allows for amplification by a factor of greater than 2×10^5 for single photon signals. The efficiency for these signals is greater than 80%. The ASICs themselves are controlled and provided a clock by an FPGA, one per carrier. The FPGA is the Xilinx Zynq Z-7030 System-on-Chip (SoC) [27].

The pogo pins are put together in assemblies and are attached to the front a carrier. There are two of these assemblies per carrier, corresponding to two frontboards. The pogo pins make physical contact with the MCP-PMT assembly frontboard metal contacts (the contacts are shown in Fig. 2.11). There is a pin for every PMT anode. The pogo pins themselves have a stroke provided by a spring-loaded mechanism that is engaged when the pins make contact with the MCP-PMT PCB [27]. The pin stroke is only 1.44 mm.

The SCROD board's primary component is the Xilinx Zynq Z-7045 SoC FPGA. The SCROD takes data from the carriers and performs what is known as "feature extraction", on the data. Feature extraction analyzes the waveform data from each channel of the carriers (that had been collected from photons impinging on the MCP-PMTs) [27].

The SCROD uses the Belle II timing system to send trigger and clock signals to the carrier FPGAs via a signal fan-out. The SCROD board itself has two fiber optic transceivers. The transceivers' operating rate is 2.544 GB/s. One transceiver is responsible for sending trigger bits from the IRSX ASICs to the Belle II global decision

(a)   **(b)**

Figure 2.14.: SCROD with Xilinx Zynq FPGA, three mezzanine and one power connector
shown (left); Carrier with two of four IRSX ASICs shown, as well as three mezza-
nine connectors and top edges of POGO pin assemblies to the right (right); One
SCROD and four carriers constitute one electronic readout module ("boardstack")

logic and the other sends readout data to the Belle II DAQ. The Joint Test Action Group
(JTAG) protocol is used to download firmware. The firmware is written in VHDL [27].

Figures 2.14, 2.15, and 2.16 show individual SCROD and carrier circuit boards and
their assembly into individual iTOP electronic readout modules.
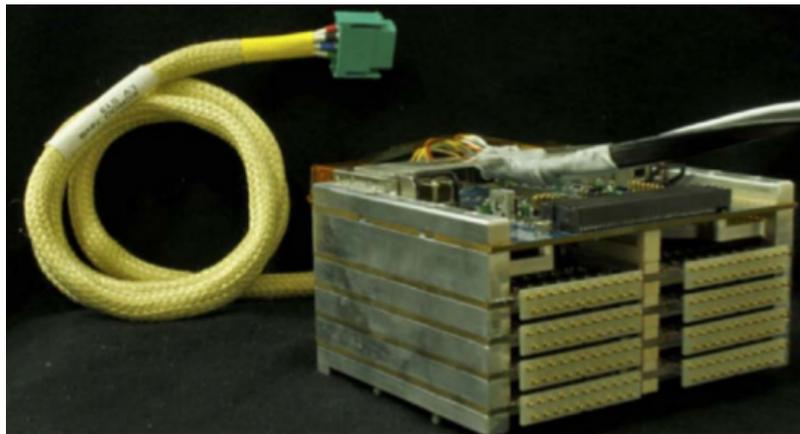
Figure 2.15.: A single electronic readout module, colloquially known as a "boardstack"; the
SCROD board in this picture is on top with four carrier boards under, with POGO
pin assemblies (two per carrier, eight total); yellow cable is power

Testing of the SCROD and carriers was extensive. The carriers were tested indi-
vidually, injecting electrically generated double pulses directly into the pogo pins to
measure the timing resolution. The timing resolution from electrical pulses alone was
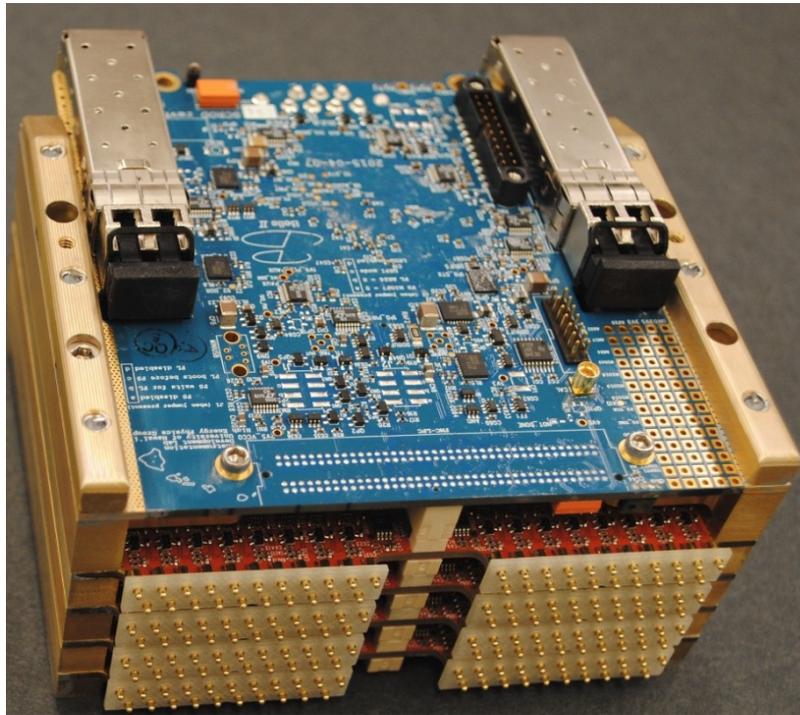
Figure 2.16.: A single electronic readout module, colloquially known as a "boardstack"; the SCROD board in this picture is on top (blue) with two optical fiber transceivers; four carrier boards (red) are located under, with pogo pin assemblies (two per carrier, eight total);

demonstrated to be approximately in the range of 20-30 ps. Fig. 2.17 shows a sample measurement [27].

An emulator board was designed and constructed to act as a mock up of the PMT frontboard signals that would be received by the pogo pins. This emulation was used to test the amplifier chain on the carriers by injecting simulated (electrically generated) signals into the pogo pins and measuring the gain [27].

The final test done before transferring iTOP readout modules to KEK was the laser qualification test. A special laser test setup was built at the University of Hawaii Manoa Department of Physics Detector Development Laboratory for the purpose. The laser was attenuated by optical filters and fiber optic cables so as to collect single photon events. A full boardstack was placed on a stationary test stand, fully instrumented with eight MCP-PMTs. All items were fully powered. The stationary stand and laser fiber head were all contained in a light-tight dark box [27].

The laser was fired into each individual MCP-PMT channel and the signals were recorded. After pedestal subtraction, timing resolution, gain, and laser efficiency were
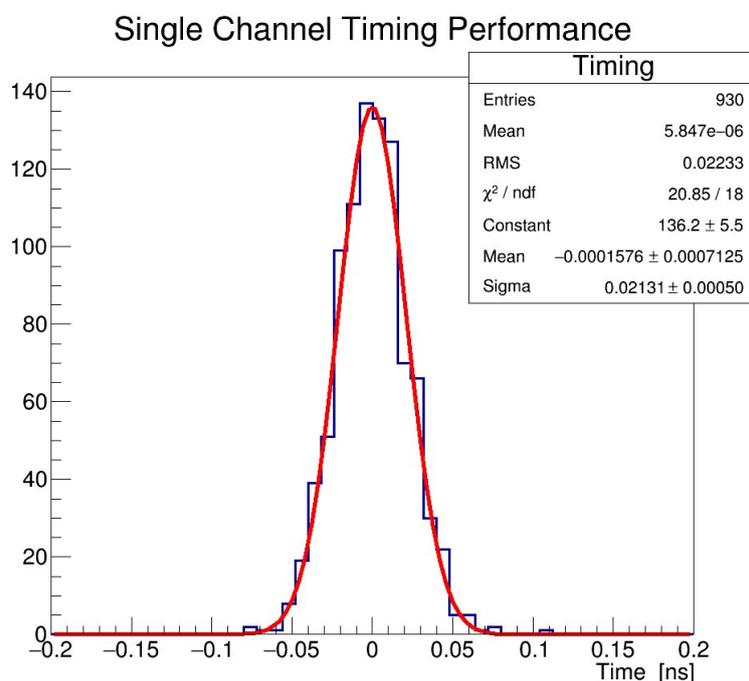
Figure 2.17.: Timing resolution from electrical double pulses, directly injected into the carrier POGO pins; carriers were tested individually using the SCROD board for each carrier test; The RMS shown here is approximately 22 ps.

determined. An 800 Hz signal was used to trigger the laser. An independent calibration signal was fed into the SCROD and used for calibration and timing measurements. A 127.2 MHz clock from a front-end timing switch (FTSW) system was provided to the boardstack, emulating that of the SuperKEKB accelerator RF timing system and FTSW clock distribution [27].

Using these signals, the data were reconstructed and timing information was obtained. The timing information was derived from the leading edge of the photon pulse and the calibration pulse.

Figure 2.18 shows a plot of timing resolution for one channel. Timing resolution for a channel is the sum in quadrature of IRSX resolution, MCP-PMT transit time spread (TTS), FTSW clock jitter, and the TDC resolution that was used at the laser test stand. Longer tails of the data distribution are a consequence of the well-known kinematic recoil of photoelectrons from the surface of the MCP web structure [27]. Figure 2.19 from [27] shows the distribution for all iTOP channels for the laser tests. Channels that had a resolution greater than 100 ps were usually recovered after a retest.
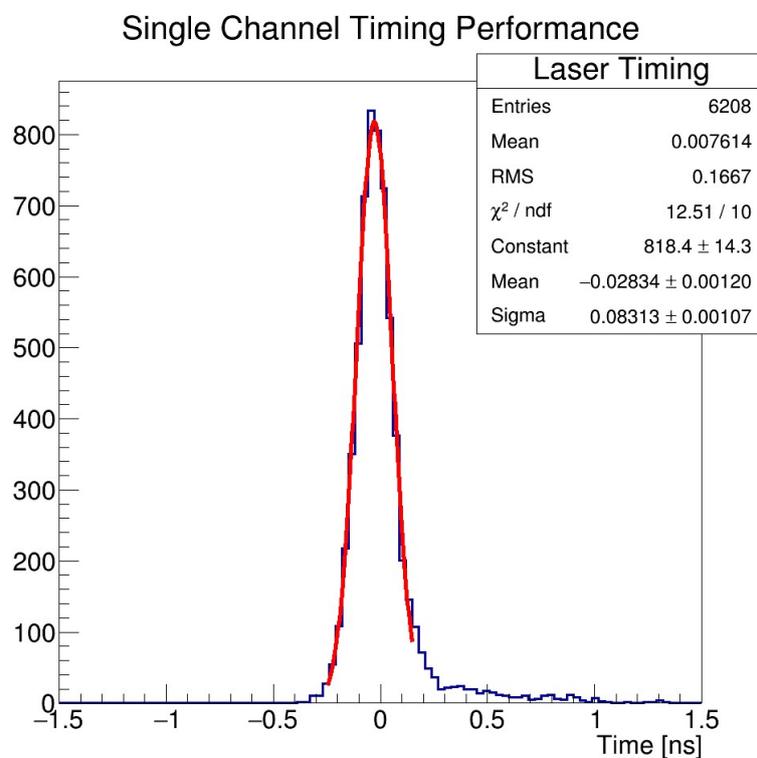
Figure 2.18.: Timing resolution for one channel, demonstrating the timing resolution.

Assembly

After boardstack assembly and electrical and laser testing was completed, boardstacks were packed with their cables and components and shipped to KEK in Japan. Figure 2.20 (left) shows a CAD rendering of a finally assembled boardstack, ready to be shipped to Japan. Four of these, complete with MCP-PMT assemblies, were mated to the QBB.

Quartz bar assembly was done on-site at KEK. Construction was performed in a special clean room that contains tools and jigs for precision assembly of iTOP modules (Fig. 2.23). After coupling with the MCP-PMTs and electronics, completed iTOP modules were tested for several days at KEK with cosmic rays, then installed in the Belle II detector. The process was completed by a three-man team from Futaba Corp (Fig. 2.24), supervised by a KEK engineer.

For further information on the iTOP see Refs. [26] and [27].
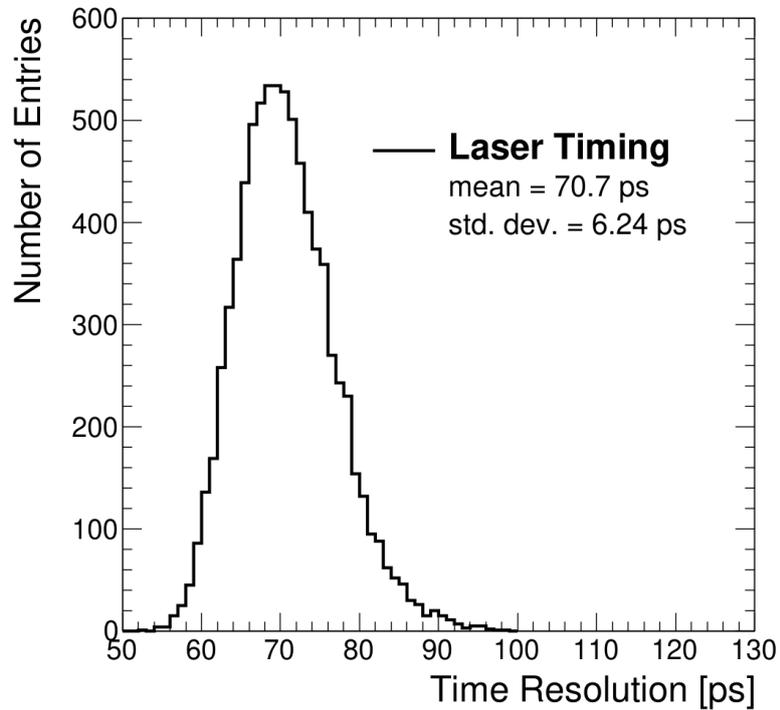
Figure 2.19.: Timing resolution distribution for all channels, demonstrating timing resolution; the resolution includes MCP-PMT TTS; data was collected using the laser test stand.

(a)

(b)
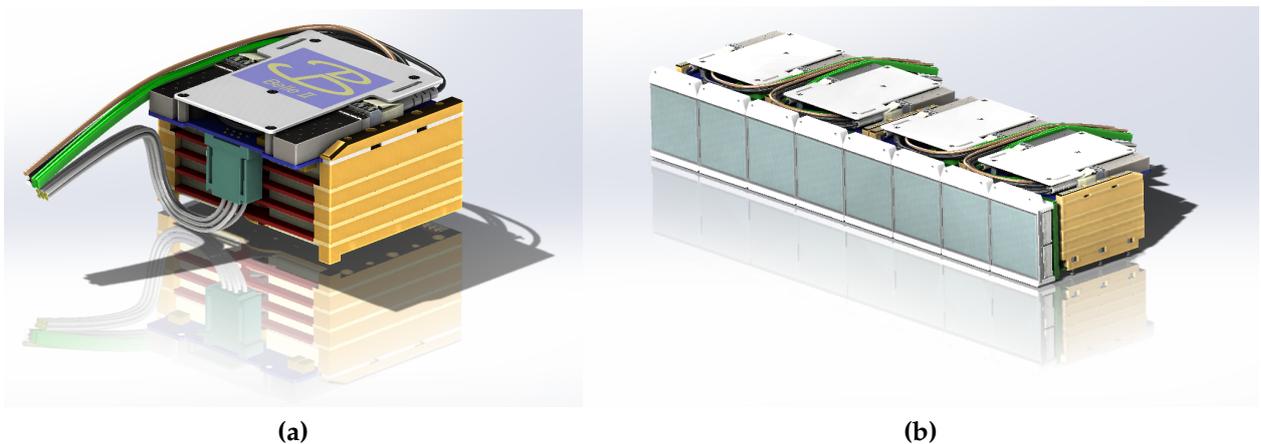
Figure 2.20.: CAD renderings of fully assembled boardstack (left) with its placement among three others in an iTOP module (right)
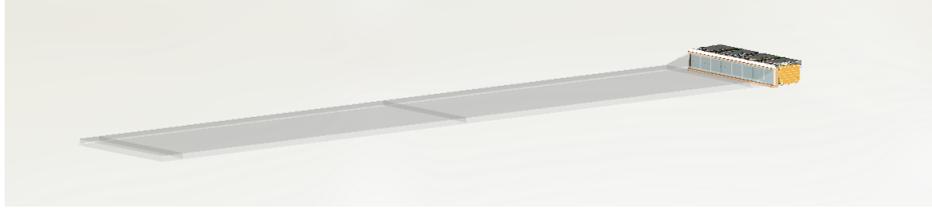
Figure 2.21.: An iTOP module with quartz, MCP-PMTs, and electronics, no QBB

Figure 2.22.: An iTOP module with quartz, MCP-PMTs, and electronics, with QBB

2.2.4. Aerogel Ring-Imaging Cherenkov Detector

The Aerogel Ring-Imaging Cherenkov (ARICH) detector is part of the PID system, in the forward endcap of the detector. It is designed to discriminate between kaons and pions over their momentum range. The ARICH also can separate pions, muons, and electrons with momentum below 1 GeV/c [6].

The ARICH components consist of an aerogel Cherenkov radiator, an expansion volume that allows the Cherenkov light to form Cherenkov rings, photodetectors, and an electronic readout system [23, 25]. Figure 2.25, from the Belle II TDR, shows a schematic of the ARICH.

The silica aerogel used as the Cherenkov radiator is structured in two layers, with refractive indices of 1.055 and 1.065. The indices are chosen so that the Cherenkov rings overlap on the photodetector plane [23]. The Cherenkov photons impinge on the photocathode of the PMTs that generate photoelectrons. This signal is then acquired by the readout system [23].

Figure 2.23.: Quartz bar assembly, with QBB, in a clean room at KEK

Figure 2.24.: Installation of the first iTOP module in Belle II
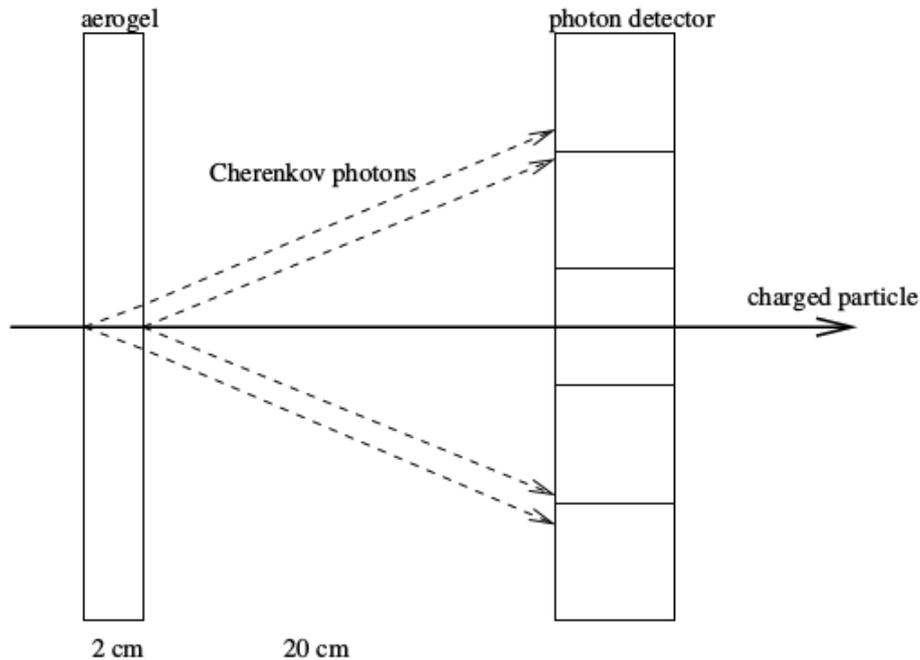
Figure 2.25.: Belle II ARICH schematic

2.2.5. Central Drift Chamber

The Belle II CDC is a larger, upgraded version of the Belle CDC, designed to provide more precise measurements of charged tracks and their momenta. The CDC is part of the PID system and provides dE/dx information needed for particle identification [6].

Due to the reliability of the earlier Belle CDC the Belle II CDC follows the same general configuration. Table 2.2 summarizes the main physical differences between the two. Further information on the CDC and its structure and electronics can be found in the Belle II TDR [6].

2.2.6. K_L^0 and Muon Detector

The KLM is the outermost sub-detector, located outside of the solenoid and is similar in design to the original Belle KLM. The barrel KLM is constructed of alternating plates of iron and RPC detector elements, with the first two layers being scintillators. All endcap RPCs were replaced with scintillators. [23, 25].

Parameter	Belle CDC	Belle II CDC
Inner cylinder radius (mm)	77	160
Outer cylinder radius (mm)	880	1130
Radius of innermost sense wire (mm)	88	168
Radius of outermost sense wire (mm)	863	1111.4
Number of sense wires	8400	14336
Diameter of sense wire (μm)	30	30
Number of layers	50	56
Gas	He-C_2H_6	He-C_2H_6

Table 2.2.: Parameters compared between the Belle CDC and the Belle II CDC from the Belle II TDR
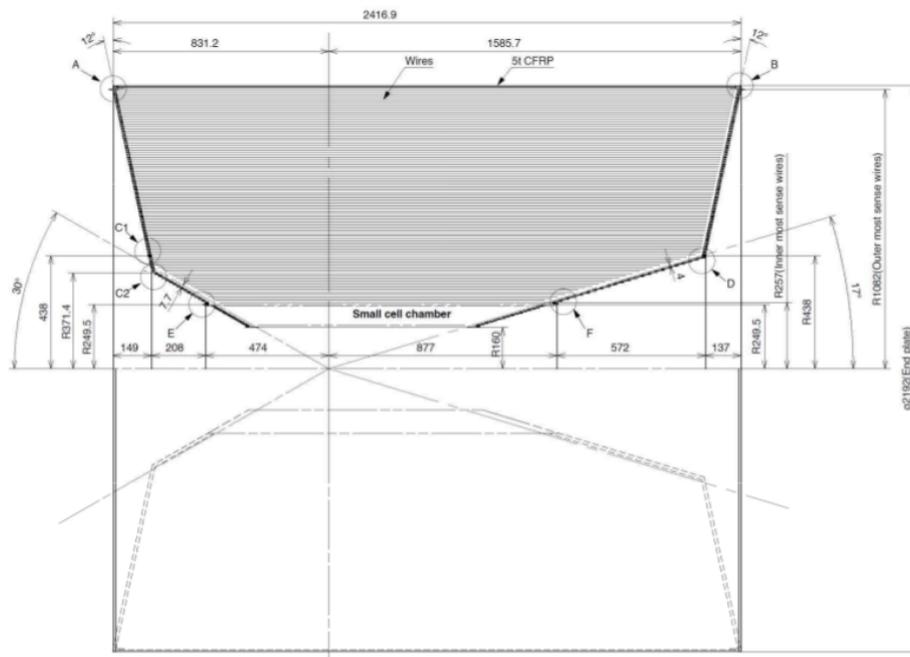
Figure 2.26.: Diagram of the Belle II CDC

The KLM barrel has an octagonal shape and covers the polar angle region of 40° $< \theta < 129^{\circ}$. The KLM endcaps covers $25^{\circ} < \theta < 40^{\circ}$ (forward) and $129^{\circ} < \theta < 155^{\circ}$ (backward) [25]. A side view of the KLM is shown in Fig. 2.27.
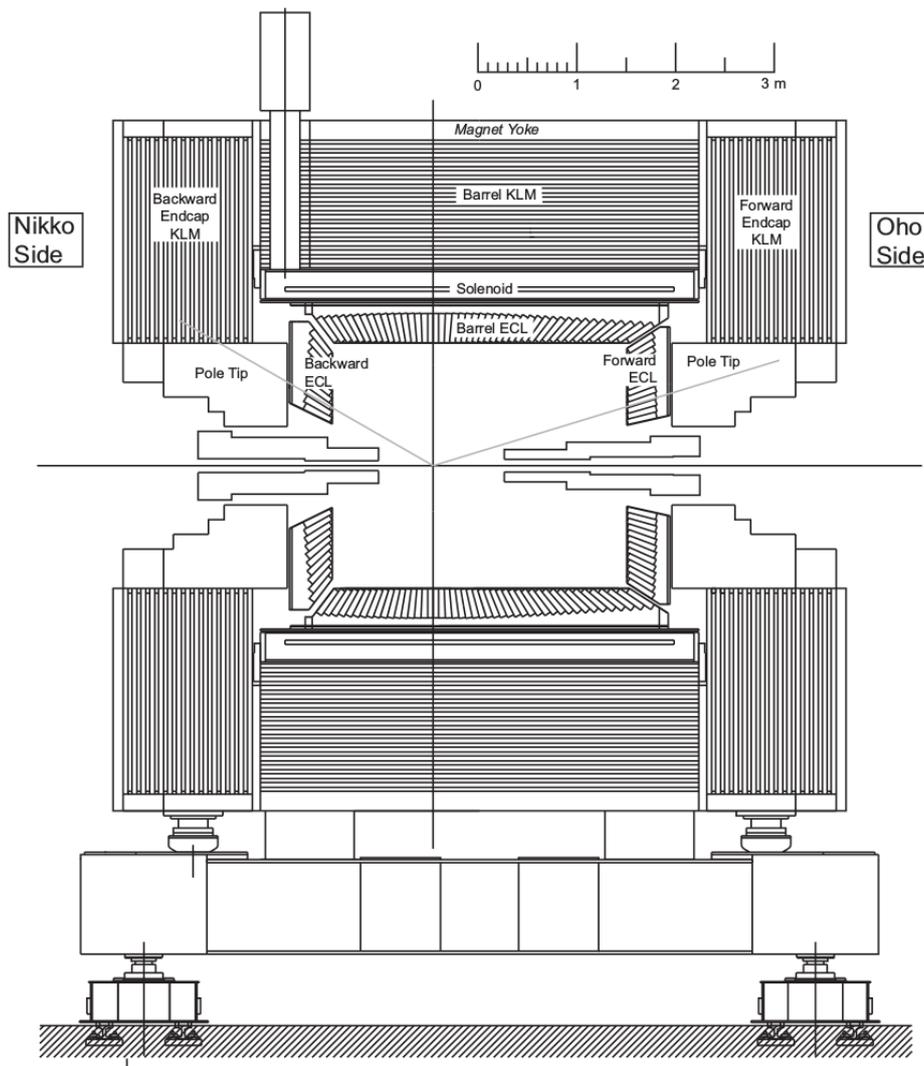
Figure 2.27.: Side view of the Belle II KLM

2.2.7. Iron Yoke and Solenoid

Yoke

The Belle II iron yoke is comprised of a barrel and movable endcaps. As mentioned previously, part of the KLM structure is partially composed of iron. This iron serves as a return path for the solenoid magnetic flux. A profile of the KLM is shown in Fig. 2.28 [6, 25].
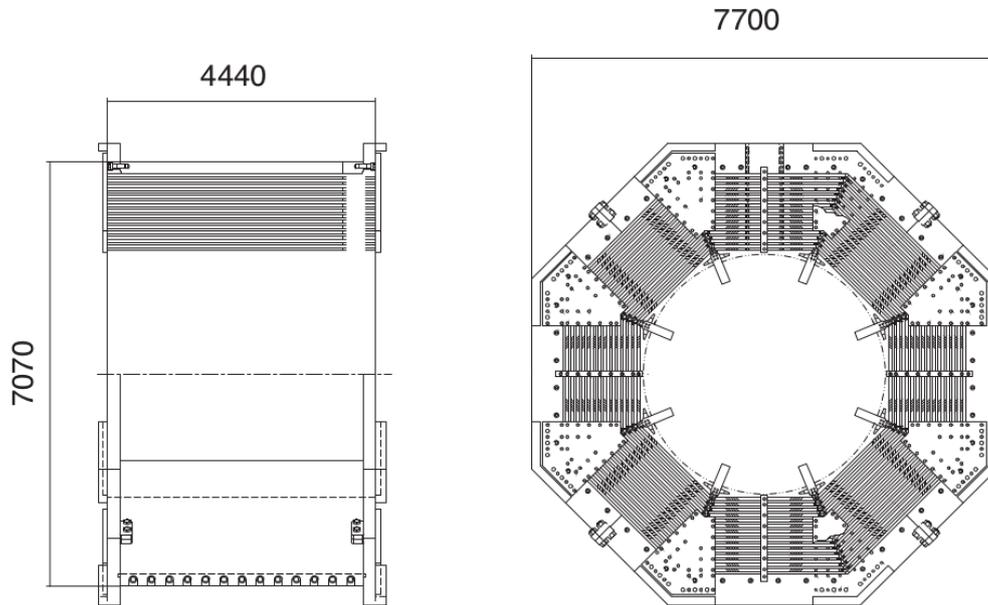
Figure 2.28.: Schematic of the Belle II Iron Yoke, from the Belle II TDR

Solenoid Magnet

The Belle II cylindrical superconducting solenoid provides a magnetic field in a volume 3.4 m in diameter and 4.4 m in length. It produces a 1.5 T magnetic field. More detailed information, including details on the cryogenics, can be found in the Belle II TDR [6,25].

2.2.8. Trigger System

The higher luminosity at Belle II provided by the SuperKEKB accelerator requires an upgraded trigger system and DAQ. From [23, 25], requirements for this upgraded trigger system are an average trigger rate of 30 kHz, approximately 5μs latency, 10 ns or less timing resolution, minimum two-event separation of 200 ns, and it must be efficient for hadronic events from $\Upsilon(4S) \rightarrow B\bar{B}$. To achieve this, the Belle trigger system has been upgraded with new electronics and other, new technologies [23]. Figure 2.29 shows a schematic of the Belle II trigger system.
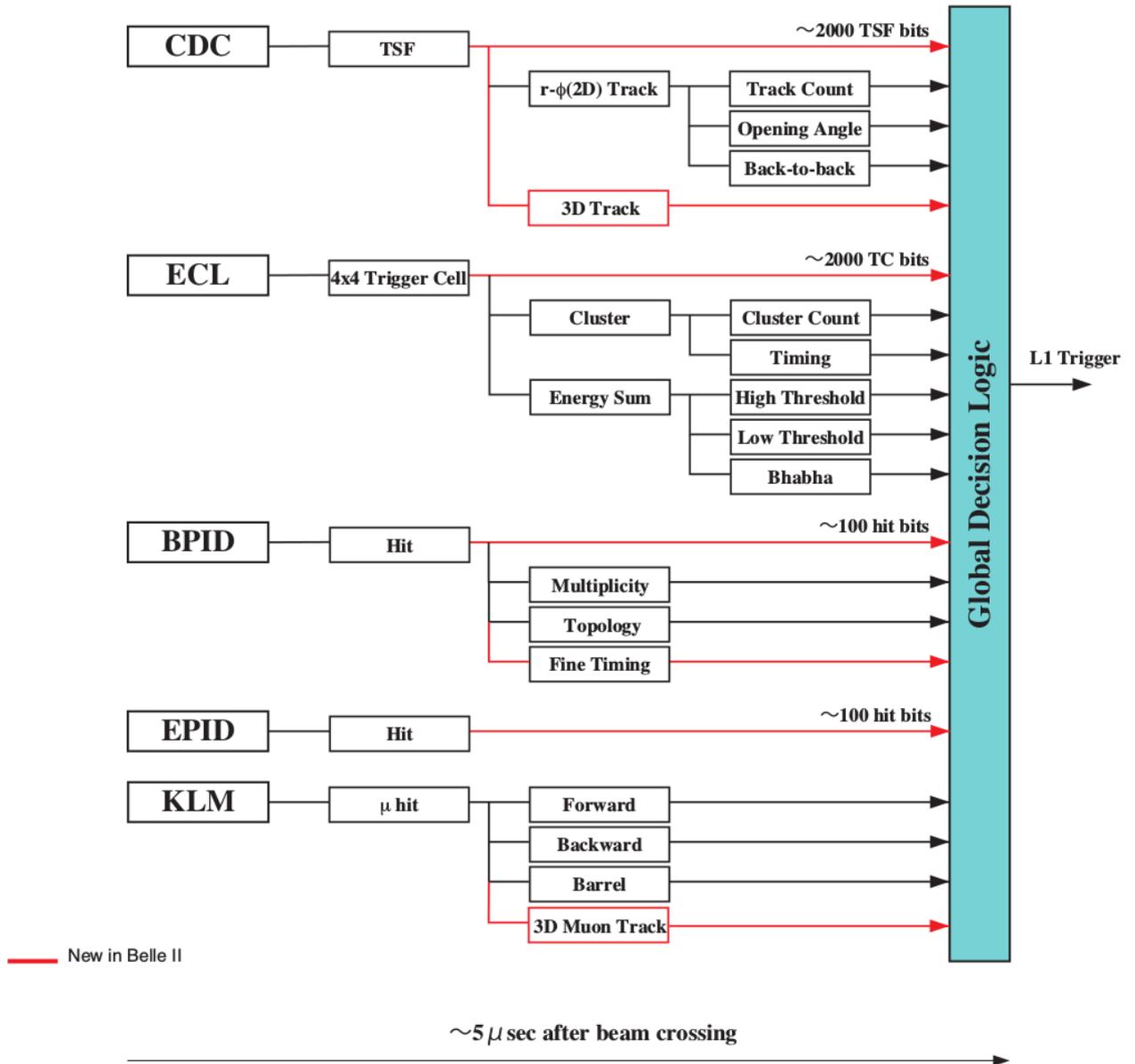
Figure 2.29.: Schematic of the Belle II Trigger System

2.2.9. Data Acquisition

The DAQ system for Belle II is a significant upgrade from Belle. It is upgraded to handle the higher luminosity environment that SuperKEKB provides.

The Level-1 trigger provides information to the DAQ on when to read out detector signals. All outer detectors are read out via optical fibers. In the DAQ there is a new data stream system that manages this throughout the entire experiment via fiber optic

connectors to the front end electronics: Belle2Link. Information is sent to the readout system: COmmon Pipeline Platform for Electronics Readout (COPPER), from the Belle2Link through optical fiber cables, and then to an event builder. A high level diagram of the DAQ can be found in Fig. 2.30 and more detailed information can be found in [23, 25].
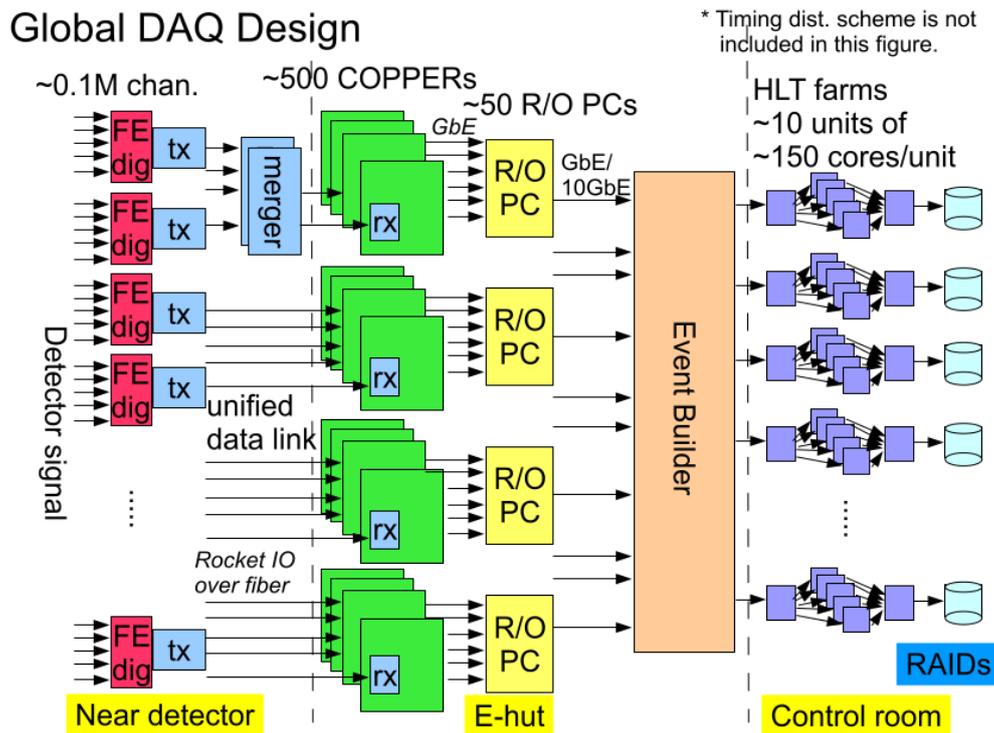
Figure 2.30.: Schematic of the Belle II DAQ

Chapter 3.

Standard Model of Particle Physics

The Standard Model of particle physics (SM) has been a successful description of nature. For most phenomena and interactions in high energy elementary particle physics, its predictions have shown to be accurate. However it does not include descriptions of several phenomena: neutrino mass, gravity, dark matter, and dark energy, being the major areas. The fundamental particle content of the SM - leptons, quarks, gauge bosons, and the Higgs boson - are shown in Fig 3.1. As of this writing, the Higgs boson is the most recently observed of the fundamental particles [32,33].
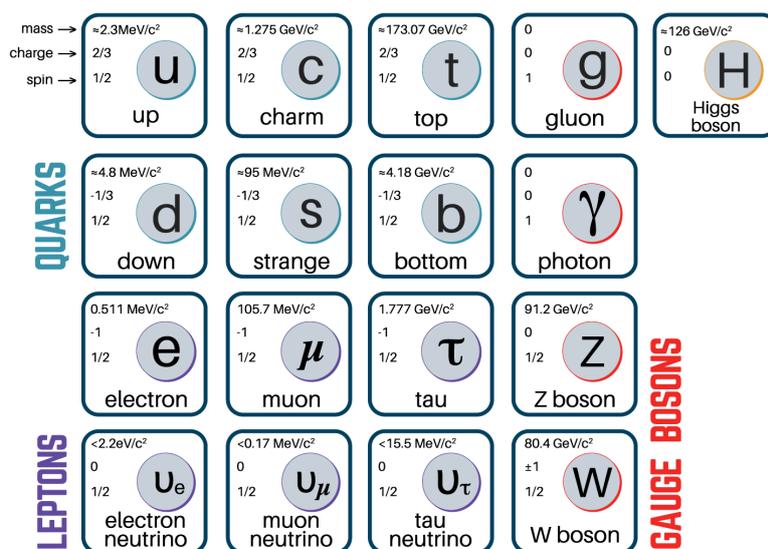
Figure 3.1.: Fundamental particle content of the Standard Model of particle physics.

3.1. Pre-Standard Model Symmetry

3.1.1. SU(3) Flavor Symmetry

Before the introduction of quantum chromodynamics (QCD), the quantum field theory that describes the strong force between quarks and gluons, Gell-Mann and others developed a method that grouped mesons (quark anti-quark bound states) and baryons (three-quark bound states) according to their charge and strangeness quantum number. Multiplets involving bound states of the u, d, and s quarks are representations of the $SU(3)$ group. For instance, the nonet, seen in Fig. 3.2, breaks up into an octet and singlet. This implies an approximate flavor symmetry in the multiplet spectrum.

These bound states can be characterized by their isospin and strangeness, in addition to baryon number. Following the convention in [34], states can be labeled by their hypercharge Y and the third component T_3 of isospin. Here $Y = S + B$, where S and B are the strangeness and baryon numbers, respectively. From these the relationship $Q = T_3 + \frac{1}{2}Y$ is obtained. This relationship between charge (Q) and isospin holds for hadrons containing the u, d, and s quarks and is called the Gell-Mann-Nishijima formula.
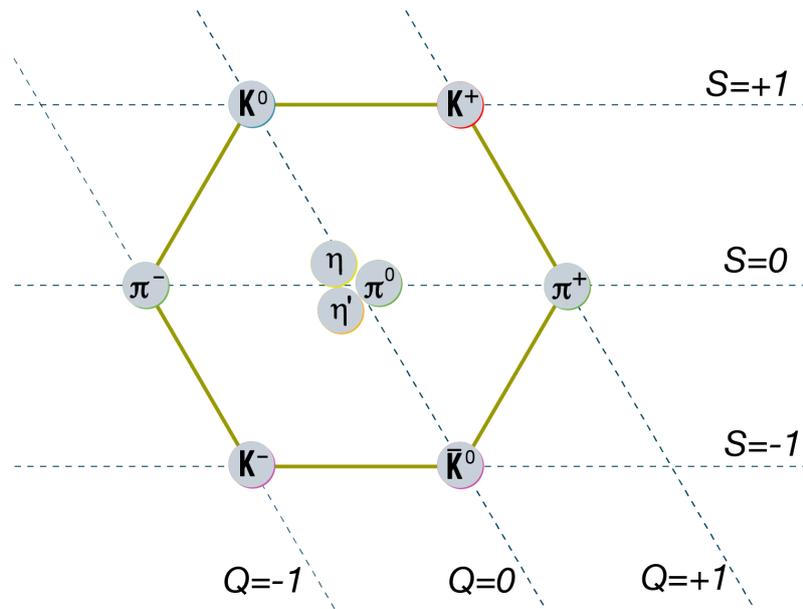
Figure 3.2.: Meson nonet. Grouping is arranged according to electric charge and strangeness

η' Meson

The η' meson, of interest in this thesis, is a mixture of the η_1 singlet and η_8 octet states:

$$\begin{bmatrix} \eta \\ \eta' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \eta_8 \\ \eta_1 \end{bmatrix} \tag{3.1}$$

$$\eta_8 = \frac{1}{\sqrt{6}} \left[u\bar{u} + d\bar{d} - 2s\bar{s} \right] \tag{3.2}$$

$$\eta_1 = \frac{1}{\sqrt{3}} \left[u\bar{u} + d\bar{d} + s\bar{s} \right] \tag{3.3}$$

Further discussion of this is given in subsequent sections.

3.2. Symmetry and Noether's Theorem

For continuous systems the object of interest is the Lagrangian density $\mathcal{L}$, that has units of energy per volume. The relation between the Lagrangian (L) and Lagrangian density is

$$L = \int d^3x \mathcal{L} \tag{3.4}$$

The Lagrangian formalism provides descriptions of continuous entities as well as non-mechanical entities such as electromagnetic fields. A field is defined as a function of space and time and is the dynamical object of a theory. In the SM, particles are described as quantum excitations of the quantized fields, which, being the dynamical variables, are promoted to operators .

Consider a local classical field $\phi(\vec{x}, t)$. The action S is defined as

$$S = \int dt L = \int d^4x \mathcal{L} \tag{3.5}$$

If only considering the field ϕ, then the Lagrangian density is a function of ϕ and its partial derivatives with respect to space and time: $\mathcal{L}(\phi, \partial_\mu \phi)$, where $\mu = 0, 1, 2, 3$, for one time and three spatial dimensions [35]. The Langrangian density can also be a function of multiple fields of different types.

The Euler-Lagrange equations in field-theoretic form are

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0 \tag{3.6}$$

where the index i denotes the field type. These equations can be understood as the equations for the operators in the Heisenberg picture, quantum mechanically.

3.3. The Standard Model

The Standard Model (SM) describes processes involving the electromagnetic force, weak nuclear force, and strong nuclear force.

3.3.1. Lagrangian Density

The Lagrangian density (for brevity, henceforth simply "Lagrangian") for the SM is [36]

$$
\begin{aligned}
\mathcal{L}_{SM} \;=\;& i\bar{L}_{Li}\,\slashed{D}L_{Li} + i\bar{e}_{Ri}\,\slashed{D}e_{Ri} + i\bar{q}_{Li}\,\slashed{D}q_{Li} + i\bar{u}_{Ri}\,\slashed{D}u_{Ri} + i\bar{d}_{Ri}\,\slashed{D}d_{Ri} \\
&- Y^e_{ij}\bar{L}_{Li}\phi e_{Rj} - Y^{ij}_d\bar{q}_{Li}\phi d_{Rj} - Y^{ij}_u\bar{q}_{Li}\tilde{\phi}u_{Rj} \\
&+ |D_\mu\phi|^2 - m^2|\phi|^2 - \frac{\lambda}{4}|\phi|^2 \\
&- \frac{1}{4}\left(B_{\mu\nu}B^{\mu\nu} + W_{a\mu\nu}W^{a\mu\nu} + G_{b\mu\nu}G^{b\mu\nu}\right) \\
&+ H.c.
\end{aligned}
\tag{3.7}
$$

where $i, j = 1, 2, 3$ stand for the generations.

Here $\tilde{\phi} = i\sigma_2\phi^*$. The notation $\slashed{D}$ means $\gamma^\mu D_\mu$, where $D_\mu = \partial_\mu - ig\frac{\tau}{2}\cdot W_\mu - ig'\frac{Y}{2}B_\mu - ig_s\frac{\lambda}{2}\cdot G_\mu$ is the covariant derivative, with W_μ and B_μ, G_μ being the electroweak and strong gauge fields, respectively. In the covariant derivative Y is the weak hypercharge and τ (Pauli matrices) and λ (Gell-Mann matrices) are the generators of the Lie groups $SU(2)$ and $SU(3)$, respectively.

Here, ϕ is the Higgs doublet and the Y are the matrices of Yukawa coupling constants. The L_L is a left-handed (L) $SU(2)$ doublet given by

$$
L_{Li} = \begin{bmatrix} \nu_{(e,\mu,\tau)L} \\ (e, \mu, \tau)_L \end{bmatrix}
\tag{3.8}
$$

For quark fields q_L is a left-handed quark doublet given by

$$q_{Li} = \begin{bmatrix} (u,c,t)_L \\ (d,s,b)_L \end{bmatrix} \tag{3.9}$$

Fields with a subscript R denote the right-handed components that come as $SU(2)$ singlets. They are denoted by the first generation labels [36]:

$$e_{Ri} = (e,\mu,\tau)_R \tag{3.10}$$

$$u_{Ri} = (u,c,t)_R \tag{3.11}$$

$$d_{Ri} = (d,s,b)_R \tag{3.12}$$

Right-handed neutrinos are absent in the SM. The last line, $H.c.$, means the Hermitian conjugate.

The SM Lagrangian has the Yang-Mills gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$.

CKM Matrix

After spontaneous symmetry breaking the Higgs field obtains an non-zero vacuum expectation value (vev) $\langle\phi\rangle \neq 0$ and the quarks obtain masses. The $Y_{u,d}$ are diagonalized as $M^f = vV_L^f Y^f V_R^{f\dagger}$, with $f = u,d$ and where v is proportional to the Higgs vev. As a result, the Lagrangian (in matrix form) proportional to $\bar{u}_L \gamma^\mu W_\mu^+ V_L^u V_L^{d\dagger} d_L + H.c.$ obtains [36,37]. Here u and d are in the mass basis. The matrix $V_L^u V_L^{d\dagger}$ is the Cabbibo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \tag{3.13}$$

The CKM matrix is unitary with 1 complex, CP-violating phase. It relates the quark mass eigenstates with the quark weak eigenstates and gives the magnitude of flavor-changing (FC) processes, such as in the loop-order penguin processes.

3.4. Symmetry and Noether's Theorem

If a field undergoes a continuous transformation then Noether's Theorem states [35]

> *For every **continuous symmetry** of the Lagrangian there exists a conserved current.*

Noether's Theorem is a fundamental theorem in physics and carries implications for the Standard Model.

3.5. Anomalies in the Standard Model

If a theory is quantized and the associated classical symmetries are no longer symmetries of the quantum theory, then the symmetry is said to be *anomalous*.

3.5.1. U(1) Problem

The Lagrangian of QCD in the massless, three-flavor limit is $\mathcal{L}_{QCD} = i \sum_{q=u,d,s} (\bar{q}_L \slashed{D} q_L + \bar{q}_R \slashed{D} q_R) - \frac{1}{4} G_{\mu\nu} G^{\mu\nu}$. It exhibits a global $U(3)_L \times U(3)_R$ symmetry. Given that $\langle \bar{q}_L q_R \rangle \approx \Lambda_{QCD}^3 \neq 0$, this symmetry is spontaneously broken. This broken symmetry implies the existence nine Goldstone modes. In the massive QCD case, these modes are pseudo-Goldstone modes (mesons) [36]. A problem arises due to the fact that the meson corresponding to the $U(1)_A$ symmetry is not observed. This is historically known as the $U(1)$ problem [38].

The group $U(1)_L \times U(1)_R$ is isomorphic to $U(1)_V \times U(1)_A$, where $U(1)_V$ rotates quark field phases in the same direction and $U(1)_A$ rotates them in the opposite direction. The relation $U(3)_L \times U(3)_R = SU(3)_L \times SU(3)_R \times U(1)_V \times U(1)_A$ then obtains. If the meson corresponding to the $U(1)_A$ symmetry missing, then it can be said that $U(1)_A$ is not actually a symmetry of the quantum QCD Lagrangian that can

be spontaneously broken to produce Goldstone bosons. It is anomalous [35, 36]. This means that the η' is not a Goldstone mode corresponding to an approximate symmetry, as it is heavier than would be expected from an approximate symmetry.

The divergence of the current corresponding to the $U(1)_A$ anomaly is:

$$\partial_\mu j^{\mu 5} = -\frac{3g_s}{4\pi} G_{\mu\nu} \tilde{G}^{\mu\nu} + 2i \sum_{q=u,d,s} m_q \bar{q} \gamma^5 q \qquad (3.14)$$

where g_s is strong coupling constant, q are the quark spinor fields, m_q is the quark mass, and i is the imaginary unit. Here, G is the gluonic field strength tensor and $\tilde{G}$ is its dual. The current associated with a global $U(1)_A$ symmetry is therefore not conserved. This implies an anomalous coupling between the η' and gluons [39, 40].

η' Mass

The $U(1)$ problem and the existence of the η' meson mass with a large mass implies that non-perturbative effect might play a role. A way to understand these non-perturbative effects is through lattice QCD (LQCD). Recent LQCD calculations give the η' mass as $M(\eta') = 911 \pm 64$ (stat.) MeV/c^2. This is in good agreement with the world average of $M(\eta') = 957.78 \pm 0.06$ MeV/c^2 [37]. Some theoretical models attempt to explain the experimental measurements of the η' mass in terms of gluonic couplings to the η' [39–47]. Other methods are available to understand the η' mass but are not discussed here [36].

3.6. Amplitudes and the S-matrix

If the interaction Lagrangian is known then the probability amplitude for a transition from one state to another may be determined, if perturbation theory is assumed to be valid. Taking the following derivation from [48], a state Φ is defined as $\Phi(t) = U(t, t_0)\Phi(t_0)$, where t_0 is some fixed time. The operator $U(t, t_0)$ satisfies $U(t_0, t_0) = 1$ and, in the interaction picture, evolves according to

$$i\frac{\partial U(t,t_0)}{\partial t} = H_I U(t,t_0) \tag{3.15}$$

where H_I is the interaction Hamiltonian. [1] Solving by iterations yields the Dyson series:

$$U(t,t_0) = 1 - i \int_{t_0}^{t} dt_1 H_I(t_1) + (-i)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2) +$$
$$\cdots + (-i)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H_I(t_1) H_I(t_2) \ldots H_I(t_n) + \ldots \tag{3.16}$$

If the state Φ is in some initial state Φ_i at time t_0 and evolves to some final state Φ_f at some time $t > t_0$, then the probability for this transition (process) to occur is given by $|\langle \Phi_f | U(t,t_0) | \Phi_i \rangle|^2 = |U_{fi}(t,t_0)|^2$, i.e. some matrix element of the operator $U(t,t_0)$. The average transition probability per unit time has a limit provided that $t_0 \to -\infty$ and $t \to \infty$. This prompts the definition of a matrix S, defined as

$$S = U(\infty, -\infty) \tag{3.17}$$

From equation 3.16, the expansion of the S-matrix is given as

$$S = 1 - i \int_{-\infty}^{\infty} dt_1 H_I(t_1) + (-i)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1) H_I(t_2) +$$
$$\cdots + (-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \cdots \int_{-\infty}^{t_{n-1}} dt_n H_I(t_1) H_I(t_2) \ldots H_I(t_n) + \ldots \tag{3.18}$$

The S-matrix is unitary and therefore satisfies the relation $SS^\dagger = S^\dagger S = 1$. The transition amplitudes S_{fi} satisfy

---

[1] The Lagrangian and Hamiltonian are related by a Legendre transformation.

$$\sum_f |S_{fi}|^2 = 1 \tag{3.19}$$

for all final states, i.e. all the transition probabilities sum to one.

One type of process that can occur in the SM is known as the "penguin" process. These processes can be used as probes of SM physics as well as physics beyond the SM (BSM). Decays of the B-meson that involve penguin processes provide probes of the CKM matrix elements as processes such as $b \to s$ do not occur at tree-level. They proceed through loops, e.g. as seen in Fig. 3.3, and act as flavor-changing neutral currents (FCNC), which are forbidden at first-order in the SM [35, 36, 49]. Penguin diagrams are sensitive to BSM physics if a species other than the W is involved in the loop.
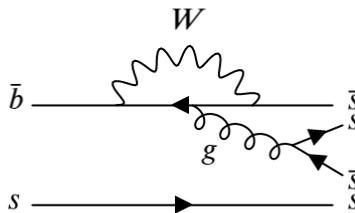
Figure 3.3.: Feynman diagram for a penguin process

3.7. The Strange B Meson B_s^0

3.7.1. B_s^0 Properties

There are two electrically neutral B mesons: the B_d^0 and the B_s^0 mesons. The B_s^0 meson is a hadron consisting of a $\bar{b}$ quark and and s quark ($\bar{b}s$). Its invariant mass is 5366.89 $\pm$ 0.19 MeV/c^2 with a mean lifetime $\tau = (1.509 \pm 0.004) \times 10^{-12}$ s. Its first excited state B_s^{0*} has an invariant mass of $5415.4^{+1.8}_{-1.5}$ MeV/c^2 and decays to $B_s\gamma$ [50].

3.7.2. $B - \bar{B}$ Mixing

The B_s^0 and B_d^0, can mix with their anti-particles. Mass eigenstates can be formed by the particles and anti-particles [50]:

$$|B_{L,H}\rangle = p|B_q\rangle \pm q|\bar{B}_q\rangle \tag{3.20}$$

where $q = d, s$ and L and H denote "light" and "heavy", respectively.

The time evolution of the flavor states is given by

$$|B_q(t)\rangle = g_+(t)|B_q\rangle + \frac{p}{q}g_-(t)|\bar{B}_q\rangle \tag{3.21}$$

$$|\bar{B}_q(t)\rangle = g_+(t)|\bar{B}_q\rangle + \frac{q}{p}g_-(t)|B_q\rangle \tag{3.22}$$

The mixing oscillation occurs with probabilities given by

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma_q t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_q t}{2}\right) \pm \cos(\Delta m_q t)\right] \tag{3.23}$$

where $\Delta\Gamma_q = \Gamma_L - \Gamma_H$ is the total width difference, $m_q = m_H - m_L$ is the mass difference, and $\Gamma_q = \frac{\Gamma_H + \Gamma_L}{2}$ [37]. The preceding assumes CPT conservation.

The study of neutral B-meson mixing can help to constrain the CKM matrix and measure CP violation in B-meson systems. The dominant Feynman diagrams for this process, obtained from [50] are shown in Fig. 3.4.
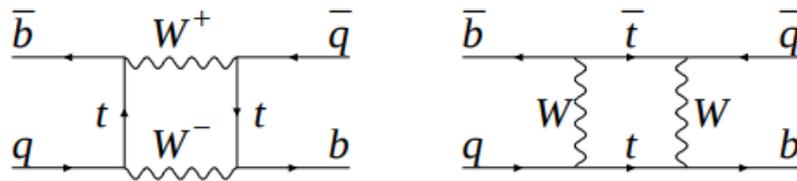
Figure 3.4.: Box diagrams for neutral B-meson mixing

Chapter 4.

Analysis

4.1. Motivation

This analysis seeks to perform a semi-inclusive measurement of charmless B_s^0 meson decays to a final state with an η' meson. Similar measurements have been published by both the CLEO and BaBar collaborations but never using the B_s^0 meson. Inclusive methods have also been used for the analysis of other modes such as $B \to K^* l^+ l^-$ and $B \to X_s \gamma$ [51–53]. The semi-inclusive branching fractions do not contain hadronic form factors so they are theoretically preferred.

CLEO measured an excess of events and obtained the branching fraction (BF) for $B \to \eta' X$: $\mathcal{B}(B \to \eta' X) = [4.6 \pm 1.1 \text{ (stat.)} \pm 0.4 \text{ (syst.)} \pm 0.5 \text{ (bkg)}] \times 10^{-4}$, with $2.0 < \left| \mathbf{p}_{\text{cm}}^{\eta'} \right| < 2.7 \text{ GeV}/c$ [54,55]. BaBar similarly measured the BF as $\mathcal{B}(B \to \eta' X_s) = [3.9 \pm 0.8 \text{ (stat.)} \pm 0.5 \text{ (syst.)} \pm 0.8 \text{ (model)}] \times 10^{-4}$, for X_s mass $M(X_s)$ less than $2.35 \text{ GeV}/c^2$ [56]. These measurements generated some interest as the measured BFs were larger than expected. The predicted BF for $B \to \eta' X$ is 1.3 $\times 10^{-4}$ in [41], for instance. The larger-than-expected BF prompted a theoretical discussion of explanations. One possible idea of interest to explain the branching fraction enhancement is to characterize the process in terms of a $b \to sg^*$ transition, where the off-shell g^*, via a $\eta' gg$ coupling, decays to $g^* \to g\eta'$ [39,40].

Other similar analyses have been performed by both the Belle and BaBar experiments [57–59], using semi-inclusive methods, but no analogous inclusive measurement in the B_s^0 sector has been done. There is a relative dearth of measurements in general of the B_s^0 compared to the B_u and B_d^0 [37].

Belle's 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ sample provides opportunities to study such decays of B_s^0 mesons, in particular the partial BF for the semi-inclusive decay $B_s^0 \to \eta' X_{s\bar{s}}$. However, as no theoretical prediction of the BF for this mode is known to be published as of this writing, we must look to previous analyses to estimate what might be the expected BF, assuming an $SU(3)$ symmetry [56–58].

The purpose of this analysis is then to attempt to the measure the partial branching fraction for the decay $B_s^0 \to \eta' X_{s\bar{s}}$, where $X_{s\bar{s}}$ is reconstructed as system of kaons and pions. Here, "partial" means the mass range of the $X_{s\bar{s}}$ is limited to below 2.4 GeV/c^2. A blind analysis is performed whereby cuts are optimized and fixed on MC before the $\Upsilon(5S)$ data is examined.

As of this writing, some relevant analysis material is given in [2].

4.2. Amplitude

From [60] the effective Hamiltonian for this process has the terms

$$H \sim \lambda_u M + (P_t - P_c)\lambda_t \tag{4.1}$$

and the amplitude is given by

$$A = \lambda_u C' + \lambda_t (P' - \frac{1}{2}P'_{EW} + Q' + P'_a) \tag{4.2}$$

The parameters $\lambda_i = V_{is}^* V_{ib}$ with $i = u, c, b$ are the CKM parameters. The CKM unitarity condition $\lambda_c = -(\lambda_t + \lambda_u)$ is utilized. The $P_{t,c}$ are penguins with an internal top or charm quark; P'_a is the anomaly contribution for the $b \to sg^*(g^* \to g\eta')$ process; Q' is the $b \to s\bar{s}s$ contribution; P'_{EW} is the contribution from the electroweak diagrams; P' is the contribution from the QCD penguins; and C' is the contribution from the color-suppressed diagram. The lowest-order diagrams contributing to the amplitude are shown in Fig. 4.1.

(a) QCD Penguin Diagram

(b) Electroweak Penguin Diagram

(c) Electroweak Penguin Diagram

(d) $b \to s\bar{s}s$ Diagram

(e) Color-suppressed diagram

(f) Anomaly diagram
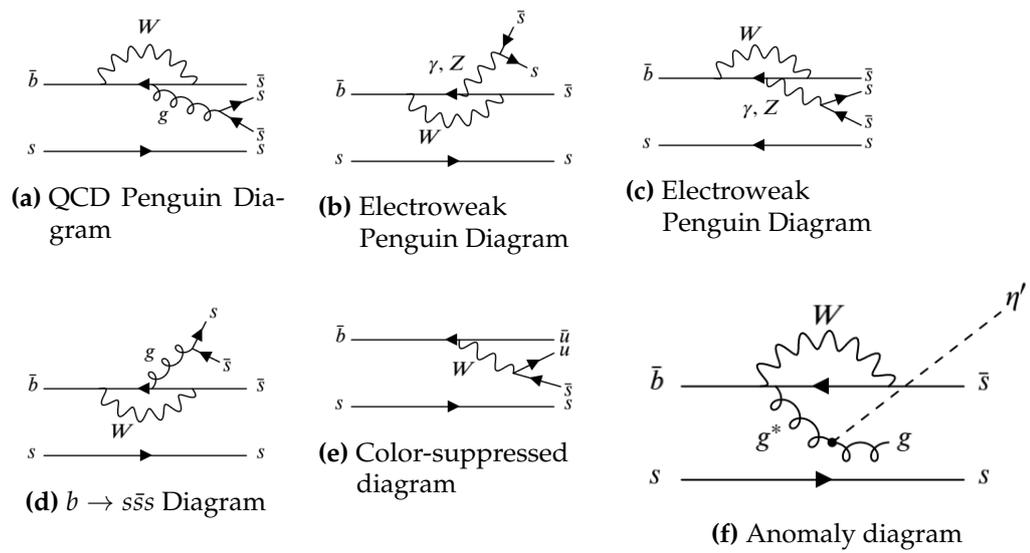
Figure 4.1.: Lowest-order Feynman diagrams contributing to the amplitude for $B_s^0 \to \eta' X_{s\bar{s}}$

4.3. Analysis Method

4.3.1. Overview

The measurement of the partial branching fraction is undertaken for the inclusive process of $B_s^0 \to \eta' X_{s\bar{s}}$. This is the first attempt to measure the BF of this mode. The $X_{s\bar{s}}$ system - which includes one s-quark and one s-anti-quark - is reconstructed as a system of two kaons (either two charged K, or one charged K and one K_S^0) with up to four pions, with at most one neutral pion (π^0). The η' is reconstructed in the channel $\eta' \to \eta(\eta \to \gamma\gamma)\pi^+\pi^-$. Unless otherwise stated explicitly, the charge conjugate modes are automatically included by implication.

The partial branching fraction (branching fraction for $M(X_{s\bar{s}}) \leq 2.4\,\text{GeV}/c^2$) for $B_s^0 \to \eta' X_{s\bar{s}}$ is measured, and the measurement of the individual sub-modes is left for a possible future study. For this study a shorthand may sometimes be used. Modes of the type $B_s^0 \to \eta' K^+ K^- + n\pi$ are called "non-K_S^0 modes" or "modes without (a) K_S^0". Modes of the type $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ are called "K_S^0 modes" or "modes with (a) K_S^0". Here, n is an integer such that $n \in [0, 4]$. These are the experimental signatures in this analysis, for $B_s^0 \to \eta' X_{s\bar{s}}$.

For the final result, the weighted average of the central values of the branching fractions for $B_s^0 \to \eta' K^+ K^- + n\pi$ and $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ is taken. Both of these classes of modes are reconstructed independently. An unbinned maximum likelihood fit is used to extract the signal and background yields.

Some misreconstruction is expected and allowed, as the analysis method is not used to reconstruct any particular exclusive mode. Background can come from $B\bar{B}$ decays, continuum ($e^+ e^- \to q\bar{q}, q = u, d, s, c$), and combinatorial background.

The two variables used to study reconstructed events are the beam-energy-constrained mass, M_{bc}

$$M_{\text{bc}} = \sqrt{\frac{E_{\text{beam}}^2}{c^4} - \frac{|\mathbf{P}_{B_s^0}|^2}{c^2}} \qquad (4.3)$$

where $E_{\text{beam}} = \sqrt{s}/2$ and $|\mathbf{p}_{B_s^0}|$ is the magnitude of the B_s^0 three-momentum in the center-of-mass (CM) frame, and ΔE

$$\Delta E = E_{B_s^{0(*)}} - E_{\text{beam}} \tag{4.4}$$

where $E_{B_s^{(*)}}$ is the energy of the $B_s^{(*)}$ measured from reconstruction in the CM frame. For the $\Upsilon(5S)$ data the CM energy is $\sqrt{s} = 10.866$ GeV.

4.3.2. Analysis Flow Summary

The general progression of the data analysis follows the path: apply initial cuts to reduce the data, perform best candidate selection, apply cuts to define a signal region, fit the data to extract signal and background yields, and determine statistical significance and final BF results.

If there are multiple B_s^0 candidates reconstructed in an event, a best candidate must be selected. This is what is referred to as best candidate selection (BCS).

4.3.3. Decay Mode

In $B_s^0 \to \eta' X_{s\bar{s}}$, the B_s^0 comes from the decay of the $\Upsilon(5S)$: $e^+e^- \to \Upsilon(5S) \to B_s^{0(*)}\bar{B}_s^{0(*)}$. The proportion $\Upsilon(5S)$ decays to $B_s^{0(*)}$ is 0.172 ± 0.030, specifically by the processes $\Upsilon(5S) \to B_s^*\bar{B}_s^*$, $\Upsilon(5S) \to B_s\bar{B}_s^*/B_s^*\bar{B}_s$, and $\Upsilon(5S) \to B_s\bar{B}_s$. These proceed at a rate of 87.0%, 7.3%, and 100.0% - (87.0 + 7.3)% = 5.7%, respectively [8]. Here, B_s^{0*} decays as $B_s^{0*} \to B_s^0\gamma$. The soft photon is not reconstructed in this analysis.

The modes searched for in this analysis and their shorthands are given in Table 4.1. All of these are searched for in an event, and the best B_s^0 candidate reconstructed as one of these modes is kept. Modes and their charge conjugates are counted and reconstructed as the same type of mode [1].

---

[1]They are reconstructed using the same function in the reconstruction code.

Mode Shorthand	Mode
0π	$B_s^0 \to \eta' K^+ K^-$
1π	$B_s^0 \to \eta' K^+ K^- \pi^0$
2π	$B_s^0 \to \eta' K^+ K^- \pi^+ \pi^-$
3π	$B_s^0 \to \eta' K^+ K^- \pi^+ \pi^- \pi^0$
4π	$B_s^0 \to \eta' K^+ K^- \pi^+ \pi^- \pi^+ \pi^-$
$1\pi 1K_S^0$	$B_s^0 \to \eta' K_S^0 K^+ \pi^-$
$2\pi 1K_S^0$	$B_s^0 \to \eta' K_S^0 K^+ \pi^- \pi^0$
$3\pi 1K_S^0$	$B_s^0 \to \eta' K_S^0 K^+ \pi^- \pi^+ \pi^-$
$4\pi 1K_S^0$	$B_s^0 \to \eta' K_S^0 K^+ \pi^- \pi^+ \pi^- \pi^0$

Table 4.1.: Modes of interest and their shorthands

Experiment	Runs	Integrated Luminosity (fb^{-1})
43	1013-1034	1.857
53	1-272	21.513
67	98-696	27.222
69	12-819, 892-1309	47.830
71	27-221, 2001-2244	22.938

Table 4.2.: $\Upsilon(5S)$ experiments at Belle and their corresponding run ranges

4.4. Data Sets

4.4.1. Experimental Data

For the final analysis Belle's $\Upsilon(5S)$ on-resonance data is used. Belle collected a 121.4 fb^{-1} integrated luminosity data sample at the $\Upsilon(5S)$ energy. A pre-selection is performed on the $\Upsilon(5S)$ data. This reduces background from $e^+e^- \rightarrow$ leptons and reduces the data size for faster analysis. A similar skim is also applied to Monte Carlo (MC).

Belle's $\Upsilon(5S)$ data is broken down into several running periods known in Belle as "experiments", with each experiment being divided into runs with specific run conditions. The experiments are listed in Table 4.2.

4.4.2. Signal Monte Carlo

A signal MC sample of 1×10^6 events is generated for $B_s^0 \rightarrow \eta' X_{s\bar{s}}$. Events are generated in proportion to experimental luminosity. The method of generation uses a quasi-two-body decay mode in which the B_s^0 is decayed to η' and $X_{s\bar{s}}$. The $X_{s\bar{s}}$ system is then fragmented via PYTHIA 6, which hadronizes the s s-anti-quark system in the $X_{s\bar{s}}$ particle [61].

The mass of the $X_{s\bar{s}}$ object is constrained to lie in the range 0.986 GeV/c^2 $\leq$ $M(X_{s\bar{s}}) \leq$ 3.0 GeV/c^2, and is generated as a flat distribution. The lower bound corresponds to slightly less than twice the mass of the $K^{\pm}$: 493.677 $\pm$ 0.013 GeV/c^2 [50]. EvtGen [62] is used to generate the particle decay chain and GEANT 3 [63] is used to simulate the Belle detector response, with PHOTOS [64] describing final state radiation. With the exception of $B_s^{*0} \rightarrow B_s^0 \gamma$, which is decayed according to the VSP_PWAVE EvtGen model, all other non-PYTHIA signal MC decays are generated according to the phase space model (PHSP) in EvtGen.

To obtain MC that approximates the effect of the QCD anomaly, the flat mass MC mentioned previously can scaled (reweighted/skimmed) according to the Atwood-Soni parameterization, i.e. the flat mass-generated MC is skimmed to a distribution approximating that in [39]. This reweights the generated flat mass MC data to form a new dataset. The generated mass distributions are shown in Fig. 4.2. This is done for demonstration purposes and the flat-mass MC is used for all practical purposes in this analysis.

To quickly verify the signal modes were correctly generated, the MC was manually scanned, event-by-event, for a few events, to determine if at least some of the modes of interest were generated. Further, by scanning the MC at generator level, a histogram of the number of daughter particles of the $X_{s\bar{s}}$ species can be obtained. This is shown in Fig. 4.3.
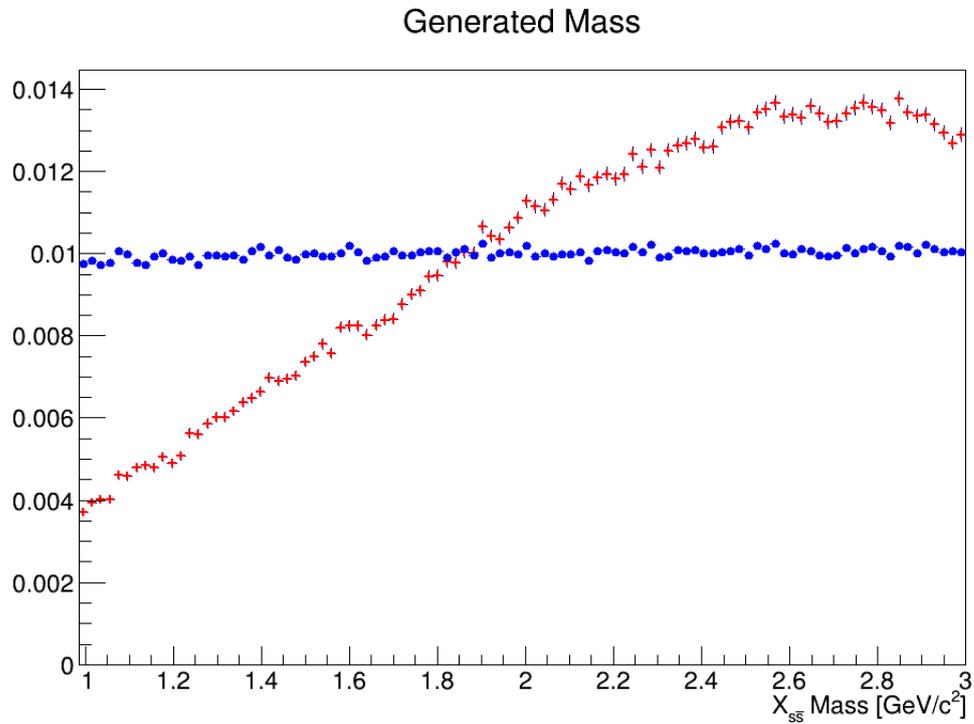
Generated Mass

Figure 4.2.: Generated $X_{s\bar{s}}$ distribution from MC; blue circles indicate the flat mass distributions; red crosses indicate the anomaly (Atwood-Soni) mass distribution, generated by reweighting the flat mass distribution
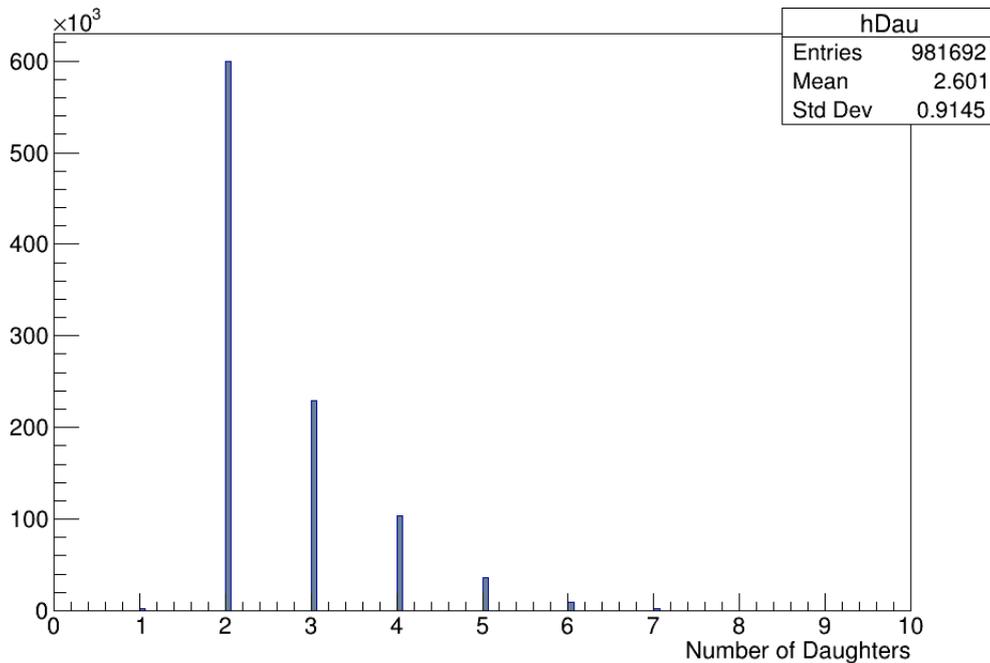
Figure 4.3.: Number of $X_{s\bar{s}}$ daughters in the generated signal MC

4.4.3. Control Monte Carlo

Control MC samples are generated in two modes that are topologically similar to the signal mode [65]. These modes are used to verify the analysis method and to study discrepancies between data and MC. The control modes studied are $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \rho^+$. The D_s^- is forced to three decay modes. The PHSP model in EvtGen is used for all decays in the $B_s^0 \to D_s^- \pi^+$ control mode and the SVS model for the $B_s^0 \to D_s^- \rho^+$ control mode. The decay submodes of the D_s^- are (with their EvtGen decay models):

- $D_s^- \to \phi \pi^-$ (SVS):

- $D_s^- \to \bar{K}^{*0} K^-$ (SVS):

- $D_s^- \to K^0 K^-$ (PHSP)

A MC sample of 1×10^5 events is generated for each D_s^- mode in each of the two control modes.

4.4.4. Background Monte Carlo

To understand backgrounds and their sources three types of background MC are used. Five streams (a stream is an approximately data-equivalent size) of $\Upsilon(5S)$ on-resonance $B\bar{B}$ generic MC are used to study non-continuum background. Four streams of $\Upsilon(5S)$ on-resonance continuum $e^+ e^- \to q\bar{q}$ MC are used to study non-peaking background. Five streams of generic $B_s^{0(*)} B_s^{0(*)}$ MC are used to study additional sources of background. Here, generic means that the MC includes most known (at the time of MC generation) decay modes, rather than just a single modes of interest.

4.5. Event Selection and Signal Reconstruction

4.5.1. Event Selection (Cuts)

Event reconstruction and handling is done using the Belle AnalysiS Framework (BASF) [66,67]. A BASF module is run on the KEK Central Computer System (KEKCC). The

following describes the selection cuts for this analysis. These are applied "online", i.e. they are written into the BASF analysis module that is used to process MC and data.

- $K^{\pm}$ and $\pi^{\pm}$

Charged tracks are separated using particle likelihood ratios, where the $K^{\pm}$ must satisfy the condition $\mathcal{L}(K : \pi) > 0.6$, for the likelihood ratio $\mathcal{L}(K : \pi) = \frac{\mathcal{L}(K^{\pm})}{\mathcal{L}(K^{\pm}) + \mathcal{L}(\pi^{\pm})}$, determined using information from the TOF, ACC, and CDC. Tracks that do not pass this requirement are considered to be charged pions, i.e. $\mathcal{L}(K : \pi) < 0.6$. The impact parameters requirements for charged tracks are $|dz| < 10.0$ cm and $|dr| < 1.0$ cm. The transverse momentum, p_t, must be greater than 0.05 GeV/c.

- γ

All photons are required to have energies greater than 0.05 MeV.

- π^0

Neutral pions (π^0) are mass-constrained to the world average mass [37]. They are required to have a $\chi^2 > 50$ from the fit, corresponding to the range $M(\gamma\gamma) \in [0.089, 0.180]$ GeV/c^2 ($\pm 5\sigma$ window). The two photons from the π^0 each must satisfy $\frac{E_9}{E_{25}} > 0.9$. This is the requirement that the ratio of a photon's deposited energy in a 3×3 ECL crystal array to that in a 5×5 crystal array must be greater than 0.9 around the central crystal. The condition $p_{\pi^0} > 0.2$ GeV/c is utilized to reduce the combinatorial background. Photons in the barrel and endcaps are required to have $E_\gamma^{\text{barrel}} > 0.05$ GeV and $E_\gamma^{\text{endcaps}} > 0.10$ GeV.

- K_S^0

K_S^0's are reconstructed in the channel $K_S^0 \to \pi^+\pi^-$. For both control modes the mass is required to be in the range $M(\pi^+\pi^-) \in [0.488, 0.508]$ GeV/c^2 and for the signal modes the range is $M(\pi^+\pi^-) \in [0.487, 0.509]$ GeV/c^2 ($\pm 3\sigma$ window).

- η and η'

The η' is reconstructed in the channel $\eta' \to \eta\pi^+\pi^-$, in the mass range $M(\eta\pi^+\pi^-) \in [0.933, 0.982]$ GeV/c^2 ($\pm 3.0\sigma$). The η is reconstructed from $\gamma\gamma$, with each photon having energy greater than 0.1 GeV. The required reconstructed mass range is $M(\gamma\gamma) \in [0.476, 0.617]$ GeV/c^2 ($\pm 2.5\sigma$). In signal MC, for B_s^0 candidates, this corresponds to a mass range of $[4.5\sigma_L, 9.2\sigma_R]$ GeV/c^2 for the η in the signal region, for the $B_s^0 \to \eta' K^+ K^- + n\pi$ signal class, and $[4.0\sigma_L, 10.2\sigma_R]$ GeV/c^2 for the

$B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ signal class. For the η', the corresponding signal region mass range of $[6.8\sigma_L, 7.3\sigma_R]$ GeV/c^2 is used for the $B_s^0 \to \eta' K^+ K^- + n\pi$ signal class and $[6.4\sigma_L, 6.9\sigma_R]$ GeV/c^2 for the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ signal class. Here, the subscripts L and R refer to the left and right sides of the mean of the mass distributions for the reconstructed η and η', in signal MC. To reduce background contributions from photons, photons used in η reconstruction are required to satisfy the ratio $A = \left| \frac{E_{\gamma 1} - E_{\gamma 2}}{E_{\gamma 1} + E_{\gamma 2}} \right| < 0.6$. Both the η and η' candidates are mass-constrained to their world average masses [37].

- ρ

  A mass range cut of $M(\pi^+ \pi^0) \in [0.696, 0.843]$ GeV/c^2 is used.

- B_s^0

  Due to the relatively high daughter multiplicity of some of the B_s^0 signal final states, there can be many random combinations that pass initial selection cuts. When this happens there can be many B_s^0 candidates per event. Too many candidates per event can cause the usage limit of computer resources to be exceeded. To get around this problem, a pre-skim is performed before best candidate selection. Candidates must pass the initial skimming criteria $M(X_{s\bar{s}}) \leq 3.0$ GeV/c^2, $M_{bc} > 5.30$ GeV/c^2 and $\Delta E < 0.35$ GeV before they are passed to the best candidate selection.

  A summary of all the preceding selection cuts can be found in Table 4.3. These cuts are applied before best candidate selection.

Cut	Value
$K^{\pm}$ ($\pi^{\pm}$)	$\mathcal{L}(K:\pi) > (<)\, 0.6$
$|dz|$	< 10.0 cm
$|dr|$	< 1.0 cm
p_{t}	> 50 MeV/c
E_{γ}	> 50 MeV
$\chi^2_{\pi^0}$	< 50
$|\mathbf{p}_{\pi^0}|$	> 200 MeV/c
$E_{\gamma}^{\mathrm{barrel}}(\pi^0)$	> 0.05 GeV
$E_{\gamma}^{\mathrm{endcaps}}(\pi^0)$	> 0.10 GeV
$\frac{E_9}{E_{25}}$ (π^0 photons)	> 0.9
$M(\pi^+\pi^-)$	$[0.487, 0.509]$ GeV/c^2
E_{γ}^{η}	> 100 MeV
$A = \left|\frac{E_{\gamma 1}-E_{\gamma 2}}{E_{\gamma 1}+E_{\gamma 2}}\right|$ (η)	< 0.6
$M(\gamma\gamma)$	$[0.476, 0.617]$ GeV/c^2
$M(\eta\pi^+\pi^-)$	$[0.933, 0.982]$ GeV/c^2
$|\mathbf{p}_{\mathrm{cm}}^{\eta'}|$	> 1.5 GeV/c
M_{bc}	> 5.30 GeV/c^2

Table 4.3.: Table of event selection criteria (cuts), applied before best candidate selection.

4.5.2. Signal Region

This subsection describes the cuts imposed after best candidate selection. They define the signal region.

- $|\mathbf{p}_{\mathrm{CM}}^{\eta'}| > 1.5$ GeV/c (daughter η' of the best B_s^0 candidate)

- -0.12 GeV $\leq \Delta E \leq 0.05$ GeV

- $R_2 \leq 0.6$ where $R_2 \equiv \frac{H_2}{H_0}$, the ratio of the second to zeroth order Fox-Wolfram moments [68, 69]. This variable discriminates between events with a spherical (signal) topology and jet-like (continuum background) topology. R_2 distributions from MC can be seen in Figs. 4.4 and 4.5.

- Neural Network output variable cut. A neural network is trained on variables that also describe event shape. It outputs a variable that may be cut on to reduce some background. This will be discussed in a later section.

- $M(X_{s\bar{s}}) \leq 2.4 \, \text{GeV}/c^2$, in bins of $0.2 \, \text{GeV}/c^2$
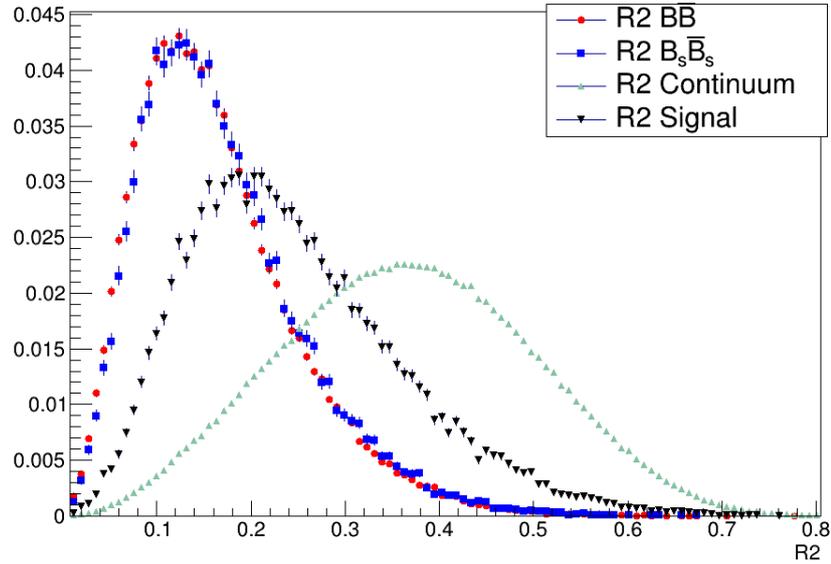
Figure 4.4.: MC R2 distributions for signal (black curve), $B_s^{0(*)} \bar{B}_s^{0(*)}$ (red curve), $B\bar{B}$ (blue curve), $q\bar{q}$ (mint green curve); $B_s^0 \to \eta' K^+ K^- + n\pi$ modes
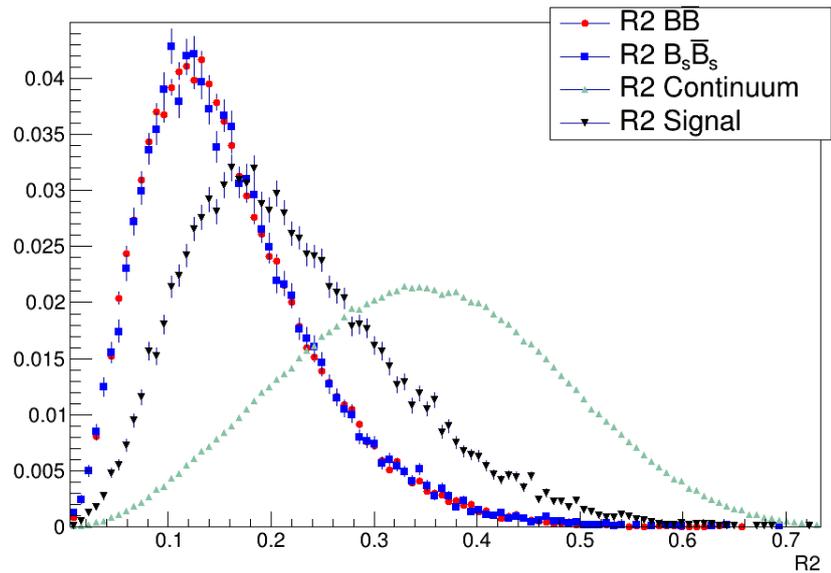
Figure 4.5.: MC R2 distributions for signal (black curve), $B_s^{0(*)} \bar{B}_s^{0(*)}$ (red curve), $B\bar{B}$ (blue curve), $q\bar{q}$ (mint green curve); $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes

(a) $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes
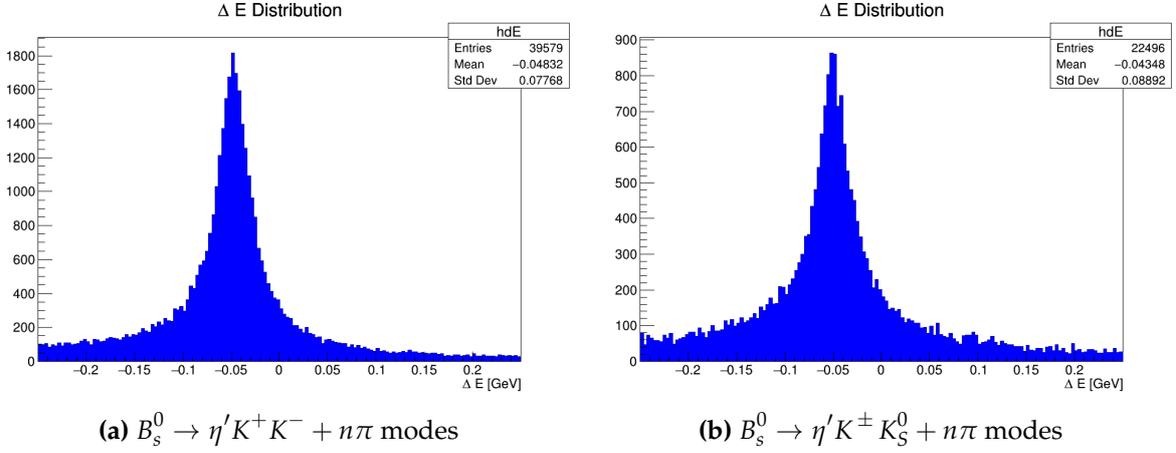
Figure 4.6.: ΔE (signal MC), only an R2 cut is applied; this variable is biased due to its use in best candidate selection

4.5.3. Multiple Candidates Per Event - Best Candidate Selection

It is possible that in a single event multiple B_s^0 candidates are reconstructed. After cuts are applied, more than one candidate per event may survive and it becomes necessary to perform a best candidate selection (BCS). To select the best candidate the χ^2 value

$$\chi^2 = \frac{\chi_{\text{vtx}}^2}{ndf} + \frac{(\Delta E - \Delta E_{\text{mean}})^2}{\sigma_{\Delta E}^2} \tag{4.5}$$

is determined for every B_s^0 candidate in an event. The candidate with the lowest χ^2 is taken to be the best candidate in that event. To determine the χ^2 parameters, the exclusive signal modes given in Table 4.1 are generated in MC and ΔE distributions are obtained from their reconstruction. From these distributions, values ΔE_{mean} and $\sigma_{\Delta E}$ are determined. The values of $\sigma_{\Delta E}^2$ and ΔE_{mean} are given in Table 4.4. These values are determined from fitting to the exclusive decay mode ΔE distributions, as shown in Figs. 4.7, 4.8, 4.9, and 4.10. The quantity ΔE is determined on an candidate-by-candidate basis, in each event.

The quantity $\chi_{\text{vtx}}^2 / ndf$ is the reduced χ^2 of the vertex fit, obtained using the charged primary daughters of the reconstructed B_s^0. All candidates whose vertex fit is successful are considered in the best candidate selection and no cut on fit quality is made. Figures 4.11 and 4.12 show the number of candidate per event before and after best candidate

Mode	$\Delta E_{\mathbf{mean}}$ (GeV)	$\sigma_{\Delta E}$ (GeV)
0π	-0.0453573 $\pm$ 0.000047	0.017109 $\pm$ 0.000036
1π	-0.049658 $\pm$ 0.00017	0.02328 $\pm$ 0.00015
2π	-0.046177 $\pm$ 0.00011	0.018615 $\pm$ 0.000087
3π	-0.049477 $\pm$ 0.0007	0.0257 $\pm$ 0.00066
4π	-0.047075 $\pm$ 0.00073	0.02266 $\pm$ 0.00064
$1\pi 1K_S^0$	-0.0458125 $\pm$ 0.000086	0.017414 $\pm$ 0.000066
$2\pi 1K_S^0$	-0.050096 $\pm$ 0.00043	0.02324 $\pm$ 0.00038
$3\pi 1K_S^0$	-0.046244 $\pm$ 0.00038	0.02075 $\pm$ 0.00032
$4\pi 1K_S^0$	-0.05496 $\pm$ 0.0037	0.0272 $\pm$ 0.0033

Table 4.4.: ΔE_{mean} and $\sigma_{\Delta E}$ values from MC for the flat mass model, using only correctly reconstructed events
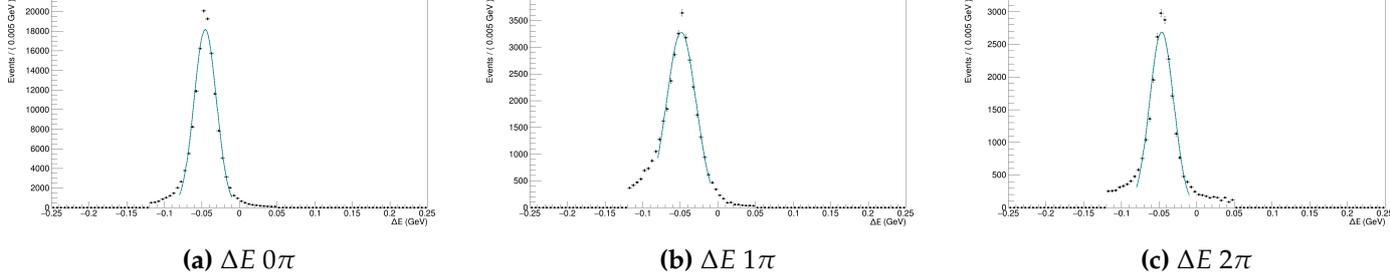
(a) $\Delta E\ 0\pi$ **(b)** $\Delta E\ 1\pi$ **(c)** $\Delta E\ 2\pi$

Figure 4.7.: Fits to ΔE distributions in MC for the individual exclusive modes given in Table 4.1. Results from these fits are used in the best candidate selection (BCS) χ^2 and are tabulated in Table 4.4.

selection, respectively. Best candidate selection is done prior to the application of the final signal region cuts.

From signal MC, by counting all events in the signal region with and without BCS (and without MC truth-matching), the BCS efficiency can be determined. For the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes it is approximately 85.5 % and for the $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes it is 43.2%. The fraction of best candidates that are correctly reconstructed is 94.0% for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes and 60.4% for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes.

(a) $\Delta E\ 3\pi$

(b) $\Delta E\ 4\pi$

Figure 4.8.: Fits to ΔE distributions in MC for the individual exclusive modes given in Table 4.1. Results from these fits are used in the best candidate selection (BCS) χ^2 and are tabulated in Table 4.4.

(a) $\Delta E\ 1\pi 1K_S^0$

(b) $\Delta E\ 2\pi 1K_S^0$

(c) $\Delta E\ 3\pi 1K_S^0$

Figure 4.9.: Fits to ΔE distributions in MC for the individual exclusive modes given in Table 4.1. Results from these fits are used in the best candidate selection (BCS) χ^2 and are tabulated in Table 4.4.

(a) $\Delta E\ 4\pi 1K_S^0$

Figure 4.10.: Fits to ΔE distributions in MC for the individual exclusive modes given in Table 4.1. Results from these fits are used in the best candidate selection (BCS) χ^2 and are tabulated in Table 4.4.

(a) Number of candidates per event; $B_s^0 \rightarrow$ $\eta' K^+ K^- + n\pi$

(b) Number of candidates per event; $B_s^0 \rightarrow$ $\eta' K^\pm K_S^0 + n\pi$

Figure 4.11.: Number of candidates per event before best candidate selection in signal MC, inside the signal region, without MC truth-matching applied.
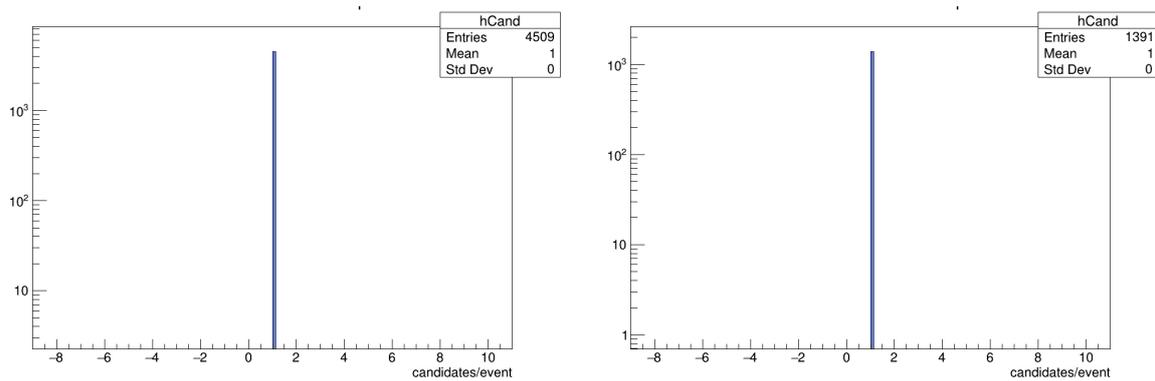
(a) Number of candidates per event; $B_s^0 \rightarrow$ $\eta' K^+ K^- + n\pi$

(b) Number of candidates per event; $B_s^0 \rightarrow$ $\eta' K^\pm K_S^0 + n\pi$

Figure 4.12.: Number of candidates per event after best candidate selection in the signal region with MC truth-matching applied.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$
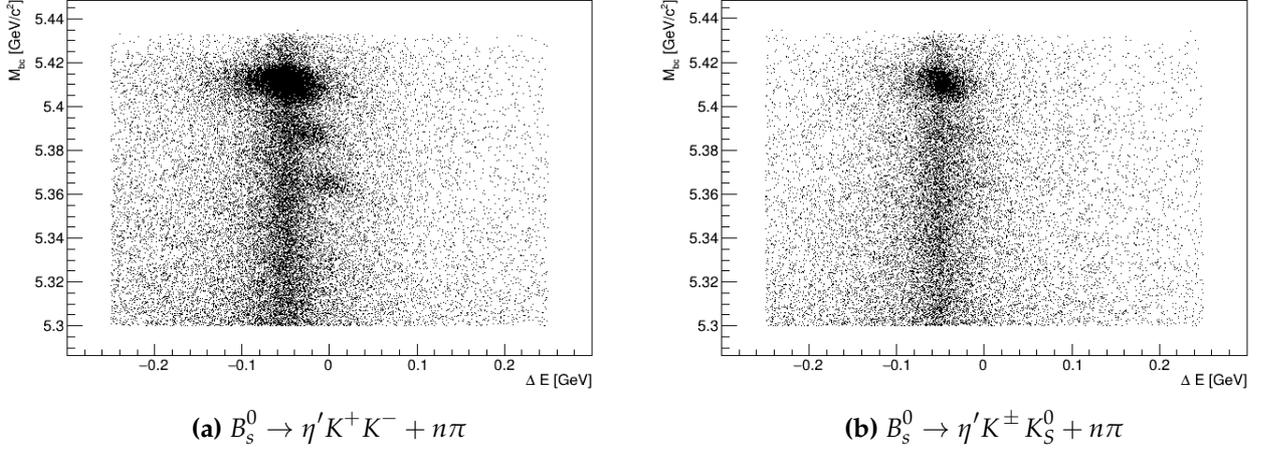
Figure 4.13.: M_{bc} vs ΔE; no signal region cuts are applied; the band at around ΔE = -0.05 GeV indicates bias in ΔE because the variable ΔE is used in the best candidate selection method.

4.6. Fitting and PDFs

The fitting is performed using unbinned maximum likelihood fits to M_{bc} distributions in the signal region [37,70]. The PDF(s) utilized in this analysis is the sum of three Gaussian PDFs with fixed normalizations given by the fractions shown in Fig. 1.3, for the signal peaks. An ARGUS PDF models the non-peaking background [71]. The full fitting model is the sum of a triple Gaussian and ARGUS PDFs. The parameters for the signal Gaussian PDFs are determined from the control mode $B_s^0 \to D_s^- \rho^+$. As ΔE is used for best candidate selection, it is assumed to be biased and is therefore not appropriate for fitting. This biasing may be seen in the band in Fig 4.13. For this reason only M_{bc} is used for fitting. To extract the signal yield one-dimensional fits to M_{bc} are done in bins of $X_{s\bar{s}}$ mass. Final signal region selection and fitting are done using RooFit [72].

The parameters for fitting to $\Upsilon(5S)$ data for the ARGUS PDF are determined by fitting to data neural network sidebands, done using a looser and non-overlapping (with the signal region) neural network cut. This is done due to the dearth of statistics in data, leading to an unphysical shape in some mass bins if the ARGUS parameter is allowed to float. These parameters are given in Tables 4.6 and 4.7. The ARGUS endpoint is fixed at 5.434 GeV/c^2, the kinematic limit of the M_{bc} distribution. Table 4.5 gives a summary of the PDFs used in fitting and signal extraction.

Component	Function	Parameters Value Source
Signal	Gaussian Sum	fixed: control mode
Non-peaking $q\bar{q}$ + Other Comb.	ARGUS	data - fixed, MC - floated
Other Backgrounds	N/A	Negligible amount survives cuts

Table 4.5.: PDF Summary Table

Mass Bin (GeV/c^2)	ARGUS Parameter
0.8 - 1.0	-19.8 $\pm$ 21.7
1.0 - 1.2	-18.4 $\pm$ 3.0
1.2 - 1.4	-21.3 $\pm$ 3.0
1.4 - 1.6	-21.4 $\pm$ 3.1
1.6 - 1.8	-13.5 $\pm$ 2.9
1.8 - 2.0	-5.0 $\pm$ 3.0
2.0 - 2.2	-7.8 $\pm$ 2.8
2.2 - 2.4	-7.4 $\pm$ 2.4

Table 4.6.: ARGUS parameters in $X_{s\bar{s}}$ mass bins determined from data neural network sidebands, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

Mass Bin (GeV/c^2)	ARGUS Parameter
0.8 - 1.0	-
1.0 - 1.2	-
1.2 - 1.4	-22.0 $\pm$ 11.2
1.4 - 1.6	-8.3 $\pm$ 6.5
1.6 - 1.8	-13.5 $\pm$ 5.1
1.8 - 2.0	-12.2 $\pm$ 4.1
2.0 - 2.2	-6.5 $\pm$ 3.6
2.2 - 2.4	3.1 $\pm$ 3.0

Table 4.7.: ARGUS parameters in $X_{s\bar{s}}$ mass bins determined from data neural network sidebands, for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes

4.7. Background Suppression

4.7.1. Neural Network and Background Suppression Variables

Backgrounds relevant to this study can come from continuum $q\bar{q}$, $B\bar{B}$, and $B_s^{0(*)}\bar{B}_s^{0(*)}$ processes.

A neural network (NN), NeuroBayes (NB), is used to reduce the dominant background source: continuum [73]. NeuroBayes has two phases: training and expert. The training phase determines all the weights that are used to discriminate between signal and background. The expert phase passes the data through the trained network and uses the weights from the training phase to separate background and signal events. It generates a network output variable - called "nnout" in this analysis - that is cut on to reject background events.

Event shape information is used to train the NN to discriminate against continuum events, which are jet-like, from non-continuum events, which tend to have a more spherical topology. To do this, a set of modified Fox-Wolfram moments is used [68, 69, 74, 75]. Training of the NN is done with these moments, the transverse energy, missing mass squared, and the B_s^0 flight direction.

4.7.2. NeuroBayes Training

To train the NeuroBayes neural network, correctly truth-matched signal events are used for the signal training sample. For background training events, on-resonance *uds* and charm continuum MC are used.

The NN training is accomplished by inputting the distributions of these variables, obtained from MC, into the training phase of the NN. These variables are shown in Figs. 4.14, 4.15, 4.16, 4.17, 4.18, 4.19. The blue curves are signal and the red curves are the background.

Figure 4.20 shows the network training results. The network purity plots - purity vs the network output - should be almost perfectly on-diagonal. This effectively describes the probability that a B_s^0 candidate came from an event came that is spherical (network output closer to 1) or jet-like (network output closer to -1). The output of the training phase demonstrates a reasonable separation between signal and background.
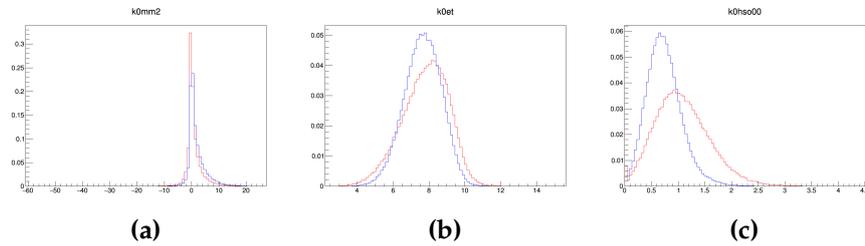
Figure 4.14.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.
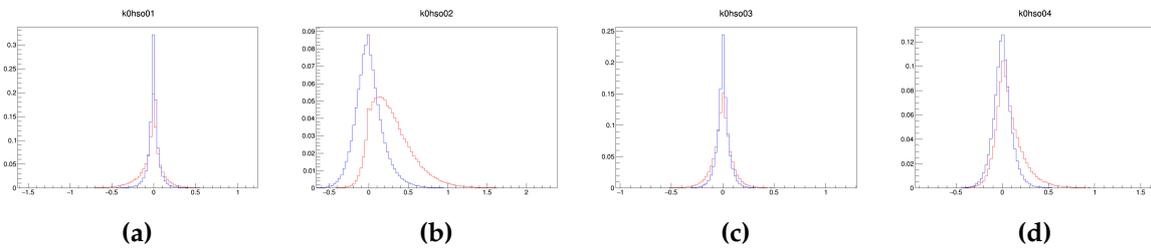
Figure 4.15.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.
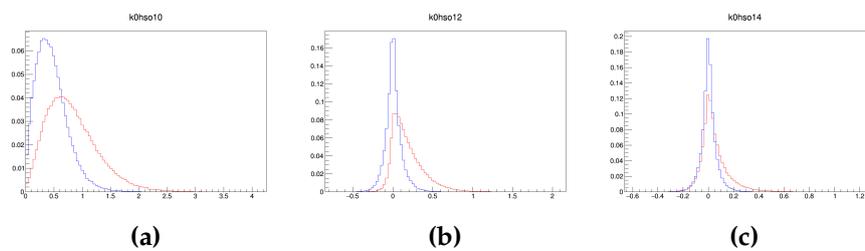
Figure 4.16.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.
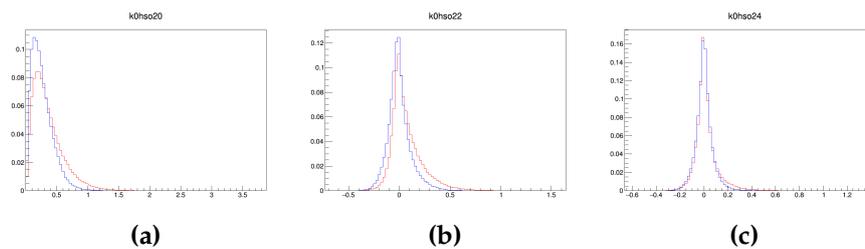
Figure 4.17.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.
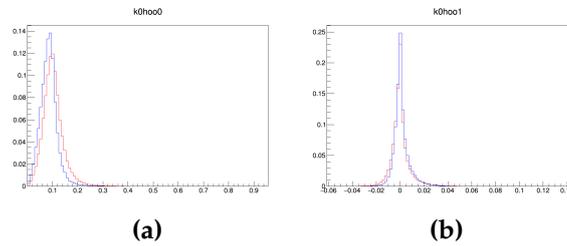
Figure 4.18.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.
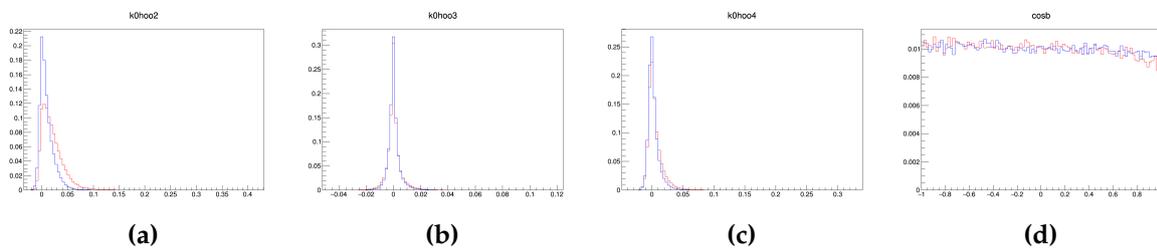
Figure 4.19.: Variables used for NeuroBayes training from MC. The red curve are background ($q\bar{q}$ continuum). The blue curve is signal MC.

To check for overtraining, the training sample and an independent validation samples are used. The validation sample and training samples are passed through the NN to obtain the nnout distributions. The normalized nnout distributions are compared to see if they are similar. If they are, overtraining is unlikely and the network generalizes. This is seen in Fig. 4.21.

4.7.3. Figure-of-Merit Determination for the Neural Network Output Cut

To determine the cut on the NN output variable nnout, a figure-of-merit (FOM) is optimized. This is accomplished using MC. A number of signal events are combined with one stream of background MC. The resulting MC sample is run through the expert phase of the NN to obtain the nnout distribution.

The FOM used in this analysis is defined as
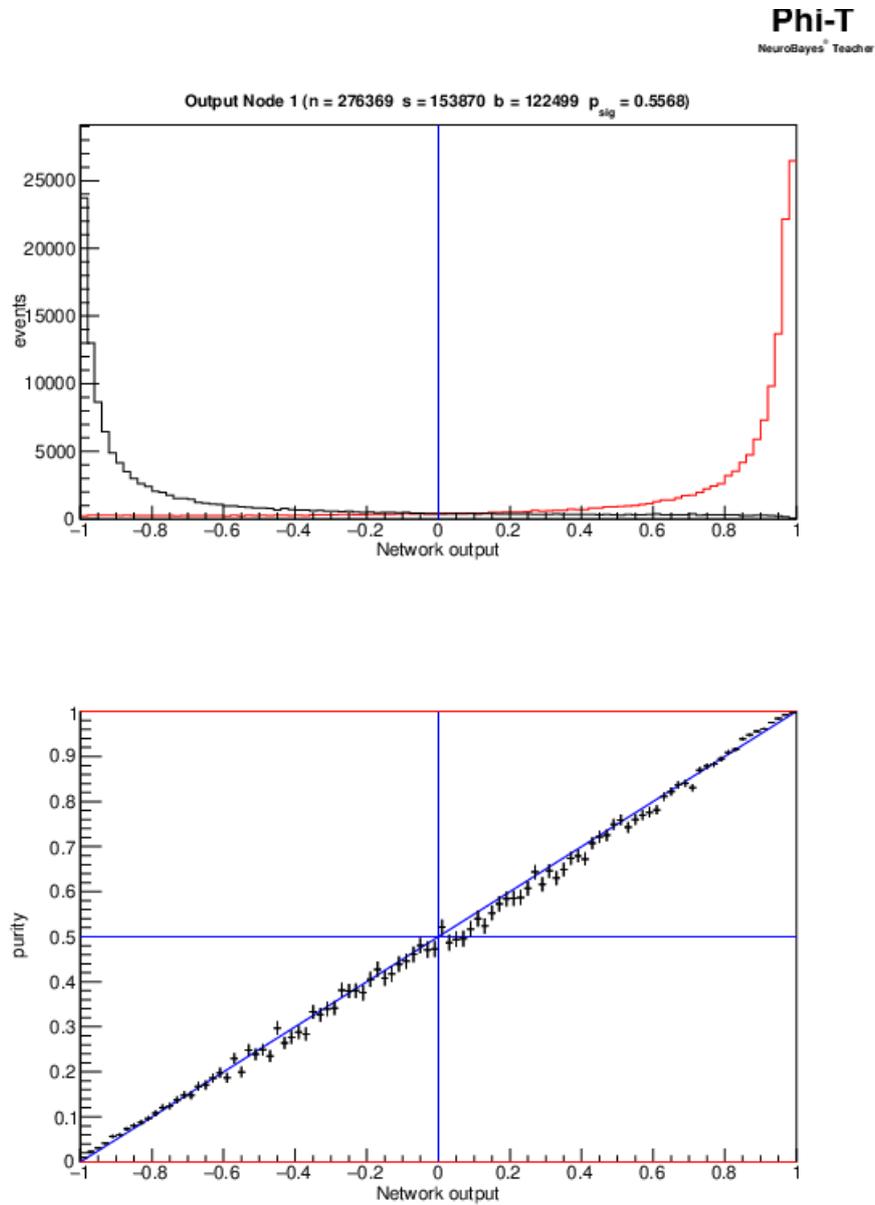
Figure 4.20.: NeuroBayes training phase network output from MC

$$FOM = \frac{S}{\sqrt{S+B}} \tag{4.6}$$

where S is the number of signal events and B is the number of background events.

Optimization of the FOM uses, from signal MC, on the order of 100 times the expected signal events in data assuming a BF of 2×10^{-4}. These are combined with
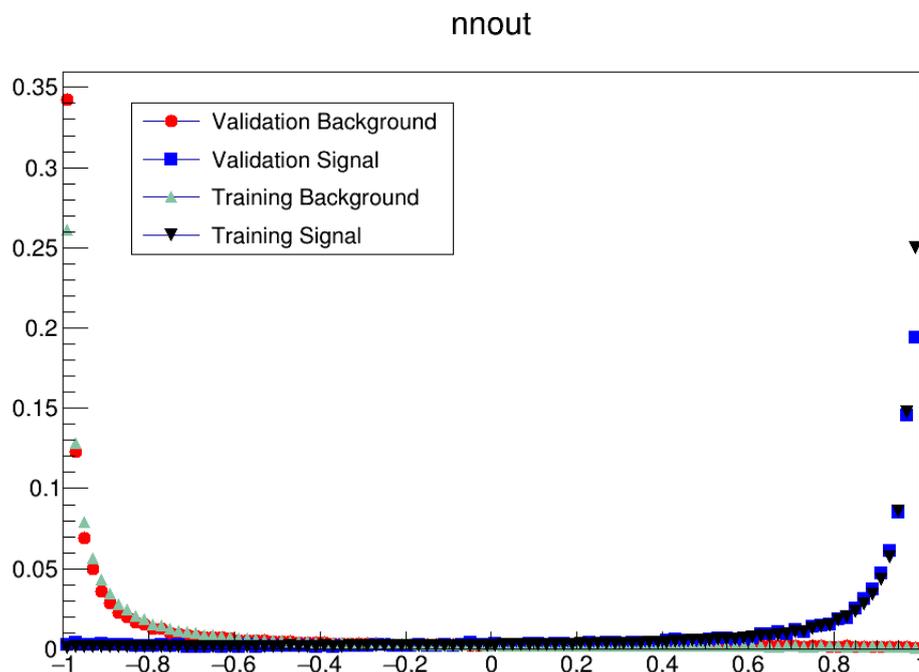
nnout

Figure 4.21.: Overtraining Test - normalized nnout distributions for MC - with validation background sample (red circles), validation signal sample (blue squares), training background sample (green upward triangles), and training signal sample (black downward triangles)

one stream of on-resonance background MC. To compensate for the enhanced number of signal events, when calculating the FOM, the signal yield, S, is scaled by a factor of 10^{-2}, in both the numerator and denominator, in equation 4.6. By fitting the M_{bc} distribution in bins of nnout, with the PDFs in Table 4.5, signal (S) and background (B) yields can be ascertained. The value of nnout at which the FOM maximizes or plateaus is taken as the cut value on nnout. This is done for the whole range $M(X_{s\bar{s}}) \leq 2.4 \, \text{GeV}/c^2$, not in individual $X_{s\bar{s}}$ mass bins.

The FOM maximizes around the value nnout = 0.95 for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. For $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes, it maximizes at around nnout = 0.6. The optimization is shown in Figs. 4.22 and 4.23.
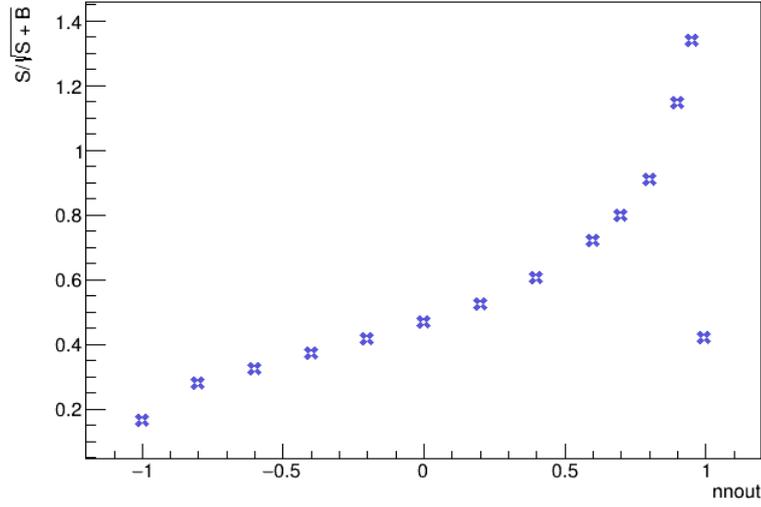
Figure 4.22.: FOM vs nnout; the nnout cut value of 0.95 is chosen for this analysis. Each point is obtained from the fitted signal and background yields in M_{bc} for the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes
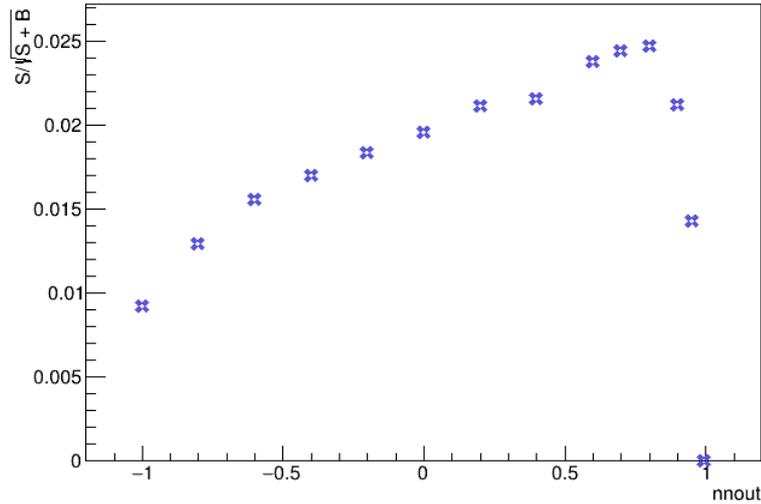
Figure 4.23.: FOM vs nnout; the nnout cut value of 0.6 is chosen for this analysis. Each point is obtained from the fitted signal and background yields in M_{bc} for the $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes

4.8. Signal Reconstruction Efficiency

The signal reconstruction efficiencies are determined by analysis of the signal MC. The signal events are reconstructed as the modes given in Table 4.1. For example, if the mode $B_s^0 \to \eta' K^+ K^- \pi^0$ is reconstructed, then it is a mode in the $B_s^0 \to \eta' K^+ K^- + n\pi$ class that is reconstructed. The sum of the four-momenta of the constituents of the

$K^+K^-\pi^0$ system (in this example) is the four-momentum of the $X_{s\bar{s}}$ particle. The magnitude of this is the $X_{s\bar{s}}$ mass. Concretely, the four-momentum for the reconstructed B_s^0 is

$$P_{B_s^0} = P_{\eta'} + \sum_j P_j \qquad (4.7)$$

where P_j is the four-momentum of a $K^\pm$, K_S^0, $\pi^\pm$, or π^0, coming from the $X_{s\bar{s}}$. The mass of the $X_{s\bar{s}}$ system is the determined from the daughters of the reconstructed B_s^0, in the sum $\sum_j P_j$.

The reconstruction efficiency per mass bin, for both the $B_s^0 \to \eta' K^+ K^- + n\pi$ and $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ classes of modes, is defined as

$$\epsilon_i = N_i^{\text{rec}} / N_i^{\text{gen}} \qquad (4.8)$$

where N_i^{rec} is the number of signal events reconstructed after all signal region cuts are applied, in the i-th $X_{s\bar{s}}$ mass bin, determined from fitting the M_{bc} distribution.

The quantity N_i^{gen} is defined as

$$N_i^{\text{gen}} = N_i^{B_s^0 \to \eta' K^+ K^- + n\pi} + N_i^{B_s^0 \to \eta' K^\pm K_S^0 + n\pi} + N_i^{\text{other}} \qquad (4.9)$$

where N_i^{other} is the number of generated B_s^0 events that are both unreconstructed signal events and non-signal events.

The reconstruction efficiency is then fully written as

$$\epsilon_i = N_i^{\text{rec}} / (N_i^{B_s^0 \to \eta' K^+ K^- + n\pi} + N_i^{B_s^0 \to \eta' K^\pm K_S^0 + n\pi} + N_i^{\text{other}}) \qquad (4.10)$$

The quantity N_i^{gen} is determined by counting the number of $X_{s\bar{s}}$ particles generated in each $X_{s\bar{s}}$ mass bin, in signal MC.

The error on the efficiency is determined by binomial statistics where

Bin Number	Bin (GeV/c^2)	Eff.(%)	Xfeed. Eff. (%)
1	0.8-1.0	5.59 $\pm$ 0.41	-0.03 $\pm$ 0.017
2	1.0-1.2	3.76 $\pm$ 0.09	0.004 $\pm$ 0.008
3	1.2-1.4	2.96 $\pm$ 0.08	0.030 $\pm$ 0.014
4	1.4-1.6	0.96 $\pm$ 0.05	0.0002 $\pm$ 0.010
5	1.6-1.8	0.58 $\pm$ 0.04	0.02 $\pm$ 0.010
6	1.8-2.0	0.36 $\pm$ 0.03	-0.002 $\pm$ 0.02
7	2.0-2.2	0.24 $\pm$ 0.02	-0.002 $\pm$ 0.008
8	2.2-2.4	0.15 $\pm$ 0.02	0.006 $\pm$ 0.007

Table 4.8.: Reconstruction efficiency in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes, for the flat $X_{s\bar{s}}$ mass model, determined from signal MC. Backgrounds from signal cross-feed are largely negligible, but their efficiency can still be used as a systematic uncertainty. Branching fractions for the η and η' decays are not included.

$$\sigma_{\epsilon_i} = \frac{\sqrt{(N_i^{\text{rec}})^2(\sigma_i^{\text{fail}})^2 + (N_i^{\text{fail}})^2(\sigma_i^{\text{rec}})^2}}{(N_i^{\text{rec}} + N_i^{\text{fail}})^2} \qquad (4.11)$$

where σ_i^{rec} is the uncertainty from fit on N_i^{rec}, $N_i^{\text{fail}} = N_i^{\text{gen}} - N_i^{\text{rec}}$ and $\sigma_i^{\text{fail}} = \sqrt{N_i^{\text{fail}}}$. Here, N_i^{fail} are the events that fail to pass final selection criteria.

The efficiency is determined in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins. Calculating bin-by-bin efficiencies in this way allows for an analysis that is nearly independent of $X_{s\bar{s}}$ models. Efficiency determinations are given in Tables 4.8 and 4.9. They are represented graphically in Figs. 4.24 and 4.25. Fits to signal MC M_{bc} distributions are shown in Figs. 4.26, 4.27, 4.27, and 4.29.

Bin Number	Bin (GeV/c^2)	Eff.(%)	Xfeed. Eff. (%)
1	0.8-1.0	0.0	0.0
2	1.0-1.2	0.02 ± 0.001	-0.002 ± 0.002
3	1.2-1.4	0.25 ± 0.02	-0.002 ± 0.007
4	1.4-1.6	0.9 ± 0.05	0.04 ± 0.014
5	1.6-1.8	0.69 ± 0.04	0.025 ± 0.015
6	1.8-2.0	0.48 ± 0.04	0.04 ± 0.017
7	2.0-2.2	0.38 ± 0.03	0.02 ± 0.017
8	2.2-2.4	0.18 ± 0.03	-0.002 ± 0.008

Table 4.9.: Reconstruction efficiency in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes, for the $X_{s\bar{s}}$ flat mass model, determined from signal MC. Backgrounds from peaking cross-feed are largely negligible, but their efficiency can still be used as a systematic uncertainty. Branching fractions for the η and η' decays are not included. The BF for $K^0 \rightarrow K_S^0$ is implicitly included in MC generation.
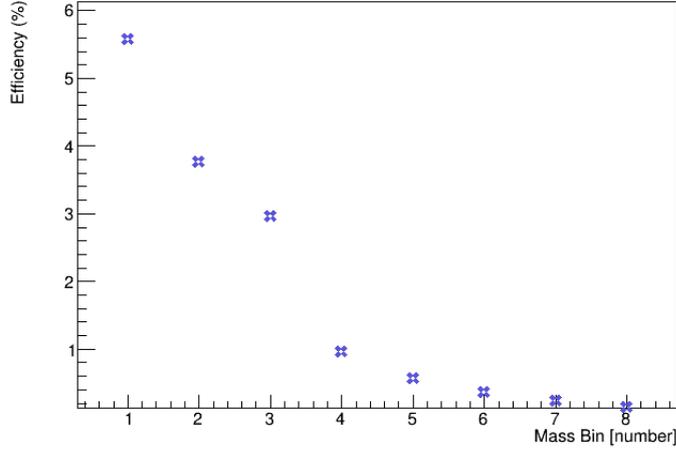
Figure 4.24.: Efficiency vs Mass Bin for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes
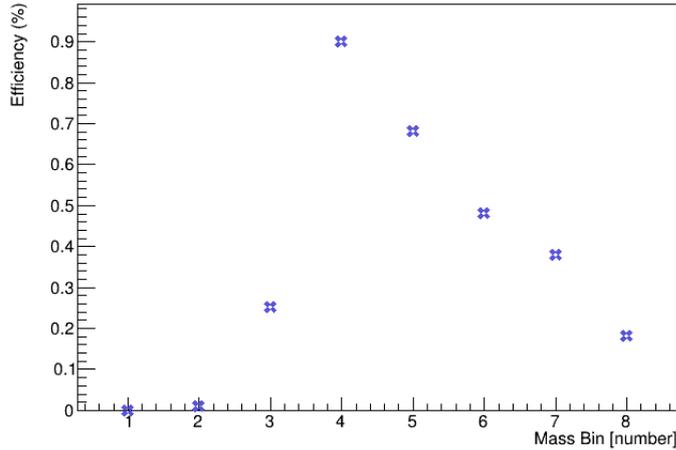
Figure 4.25.: Efficiency vs Mass Bin for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes

The order of the application of the final signal region cuts, after BCS, is checked in signal MC. All permutations of cuts are implemented and the signal reconstruction efficiency is re-determined and checked for discrepancies that could arise from the different orderings. None were apparent.

4.8.1. Mode Migration

It is possible that the $X_{s\bar{s}}$ candidate is misreconstructed. To understand this, all of the previously-mentioned exclusive signal modes are generated individually.

Each exclusive submode had a MC sample of 5×10^5 events generated using the EvtGen phase space (PHSP) model. Using the PHSP EvtGen model to generate the
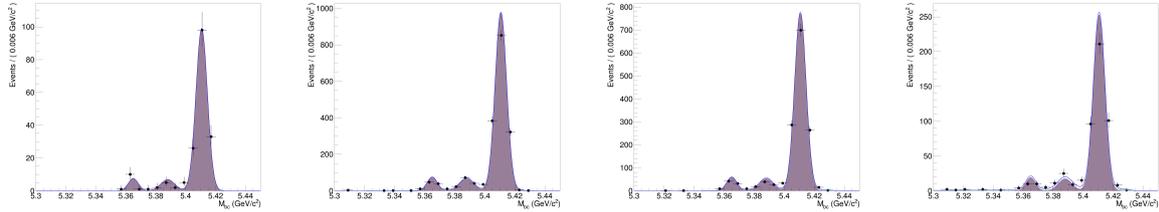
(a) Bin 1: $0.8 \leq M(X_{s\bar{s}}) \leq$ **(b)** Bin 2: $1.0 \leq M(X_{s\bar{s}}) \leq$ **(c)** Bin 3: $1.2 \leq M(X_{s\bar{s}}) \leq$ **(d)** Bin 4: $1.4 \leq M(X_{s\bar{s}}) \leq$
1.0 GeV/c^2 1.2 GeV/c^2 1.4 GeV/c^2 1.6 GeV/c^2

Figure 4.26.: M_{bc} distributions in individual $X_{s\bar{s}}$ mass bins; $B_s^0 \rightarrow K^+ K^- + n\pi$ modes. The dark shaded curve is the signal Gaussian portion of the fit, the light shaded curve is the ARGUS portion of the fit, and the blue curve is the sum of the two.
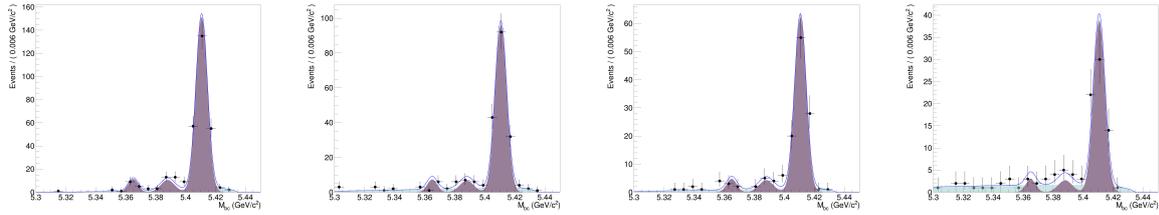
(a) Bin 5: $1.6 \leq M(X_{s\bar{s}}) \leq$ **(b)** Bin 6: $1.8 \leq M(X_{s\bar{s}}) \leq$ **(c)** Bin 7: $2.0 \leq M(X_{s\bar{s}}) \leq$ **(d)** Bin 8: $2.2 \leq M(X_{s\bar{s}}) \leq$
1.8 GeV/c^2 2.0 GeV/c^2 2.2 GeV/c^2 2.4 GeV/c^2

Figure 4.27.: M_{bc} distributions in individual $X_{s\bar{s}}$ mass bins; $B_s^0 \rightarrow K^+ K^- + n\pi$ modes. The dark shaded curve is the signal Gaussian portion of the fit, the light shaded curve is the ARGUS portion of the fit, and the blue curve is the sum of the two.
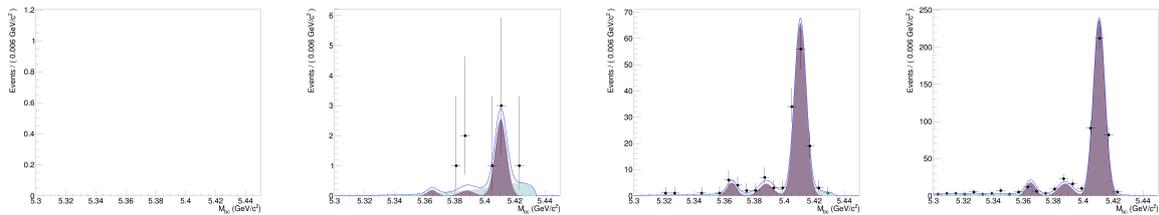
(a) Bin 1: $0.8 \leq M(X_{s\bar{s}}) \leq$ **(b)** Bin 2: $1.0 \leq M(X_{s\bar{s}}) \leq$ **(c)** Bin 3: $1.2 \leq M(X_{s\bar{s}}) \leq$ **(d)** Bin 4: $1.4 \leq M(X_{s\bar{s}}) \leq$
1.0 GeV/c^2 1.2 GeV/c^2 1.4 GeV/c^2 1.6 GeV/c^2

Figure 4.28.: M_{bc} distributions in individual $X_{s\bar{s}}$ mass bins; $B_s^0 \rightarrow K^\pm K_S^0 + n\pi$ modes. The dark shaded curve is the signal Gaussian portion of the fit, the light shaded curve is the ARGUS portion of the fit, and the blue curve is the sum of the two.
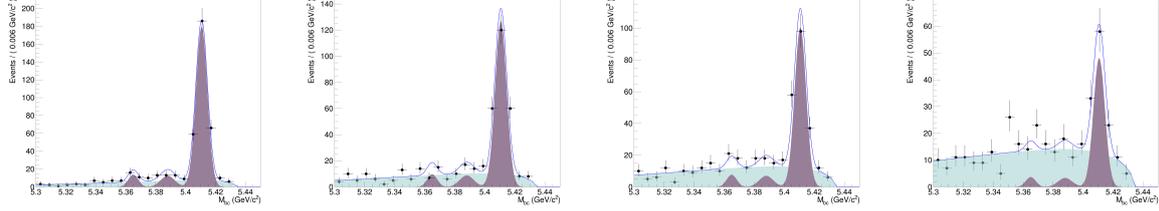
(a) Bin 5: $1.6 \leq M(X_{s\bar{s}}) \leq$ **(b)** Bin 6: $1.8 \leq M(X_{s\bar{s}}) \leq$ **(c)** Bin 7: $2.0 \leq M(X_{s\bar{s}}) \leq$ **(d)** Bin 8: $2.2 \leq M(X_{s\bar{s}}) \leq$
$1.8\,\mathrm{GeV}/c^2$ $2.0\,\mathrm{GeV}/c^2$ $2.2\,\mathrm{GeV}/c^2$ $2.4\,\mathrm{GeV}/c^2$

Figure 4.29.: M_{bc} distributions in individual $X_{s\bar{s}}$ mass bins; $B_s^0 \to K^{\pm}K_S^0 + n\pi$ modes. The dark shaded curve is the signal Gaussian portion of the fit, the light shaded curve is the ARGUS portion of the fit, and the blue curve is the sum of the two.

signal samples for this study was done to ensure the generation of a pure sample without hadronic background that would come from using PYTHIA to fragment $X_{s\bar{s}}$ generically.

Each mode was reconstructed as all other modes (e.g. the zero pion MC sample was reconstructed not only as the zero pion submode but also two, three, etc. pion submodes), to determine the level of misreconstruction. The results are shown graphically in the form of a mode migration matrix, shown in Fig. 4.30. The amount of misreconstruction from this source is negligible. The off-diagonal elements between the $B_s^0 \to \eta'K^+K^- + n\pi$ and $B_s^0 \to \eta'K^{\pm}K_S^0 + n\pi$ modes are likely due to generic B_s^0 decays, as feed-across is physically unlikely.

The value of each individual mode migration matrix element is obtained by taking the Gaussian yield of the fit model in each individual mode MC sample, and dividing by the number of generated events. The off-diagonal elements quantify the misreconstruction of each mode as another mode. Peaking can occur due to events that may be only partially misreconstructed, e.g. if an event is reconstructed as K^+K^- but one of the kaons is from a different B_s^0 meson or mode.
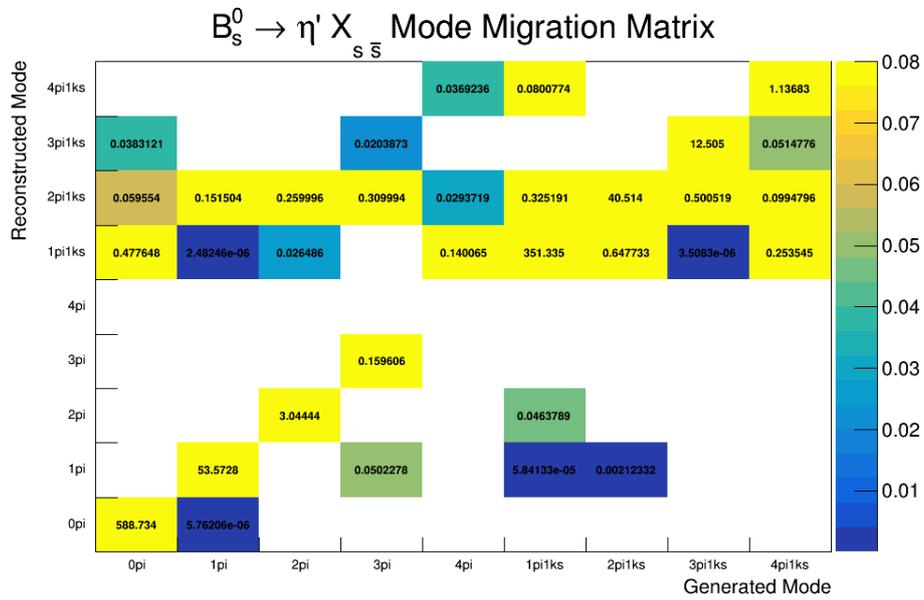
Figure 4.30.: Mode migration matrix from MC; elements are calculated by taking the Gaussian fit yield of the particular mode and dividing by the total number of events generated, and multiplying by a factor of 10000 (efficiency in percent multiplied by 100). The on-diagonal elements are the proportion of correctly reconstructed modes and the off-diagonal elements are the proportion of misreconstructed modes.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$s
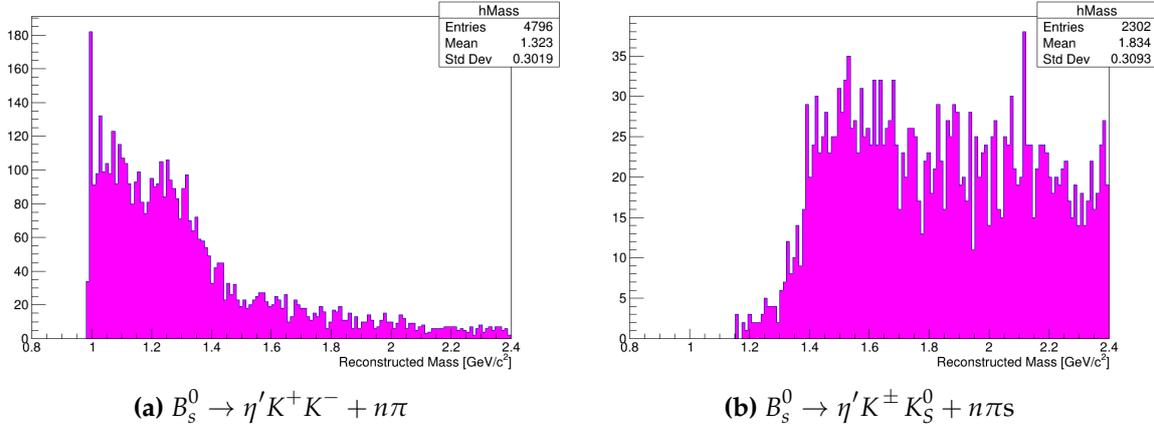
Figure 4.31.: Reconstructed $X_{s\bar{s}}$ mass distribution from signal MC

4.8.2. Mass Reconstruction

$X_{s\bar{s}}$ Mass Spectrum

Figure 4.31 shows the reconstructed $X_{s\bar{s}}$ mass from signal MC. Checks can be done to determine if the correct mass, and thus correct B_s^0 candidates, are being correctly reconstructed. In Fig. 4.32 the solid diagonal band depicts the correctly reconstructed $X_{s\bar{s}}$ mass in signal MC. Points outside of this band show fraction of incorrect reconstruction. As another check on reconstruction effectiveness, signal MC is used to analyze the difference between reconstructed and generated $X_{s\bar{s}}$ mass in each event, as shown in Figs. 4.33 and 4.34. These show that after all cuts are applied, the correct $X_{s\bar{s}}$ mass is being reconstructed and a strong signal is visible.

Further, if the reconstruction efficiencies in section 4.8 are used to correct the number of reconstructed events obtained from another, independent signal MC sample, then the corrected distributions should be similar to that of the distribution of generated events in $X_{s\bar{s}}$ mass bins, in overall shape. The structure of the generated mass distributions is approximately recovered when this correction is done, as can be seen in Figs. 4.35, 4.36, and 4.37. The shape is generally that of the generated flat $X_{s\bar{s}}$ mass distribution.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes
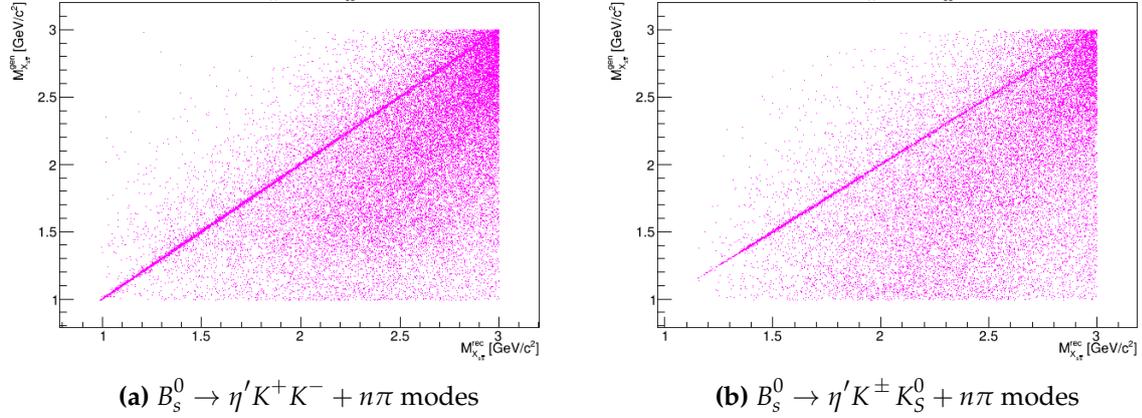
Figure 4.32.: Signal MC generated $X_{s\bar{s}}$ mass versus reconstructed $X_{s\bar{s}}$ mass; no signal region cuts applied

(a) Before neural network cut

(b) After neural network cut
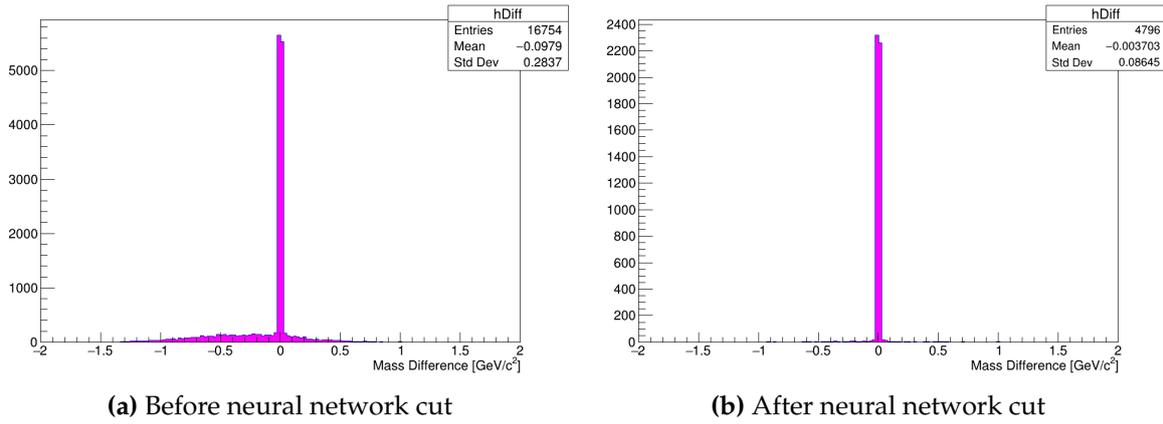
Figure 4.33.: Difference between generated and reconstructed $X_{s\bar{s}}$ mass in signal MC in the range $M(X_{s\bar{s}}) \leq 2.4\,\mathrm{GeV}/c^2$, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(a) Before neural network cut

(b) After neural network cut
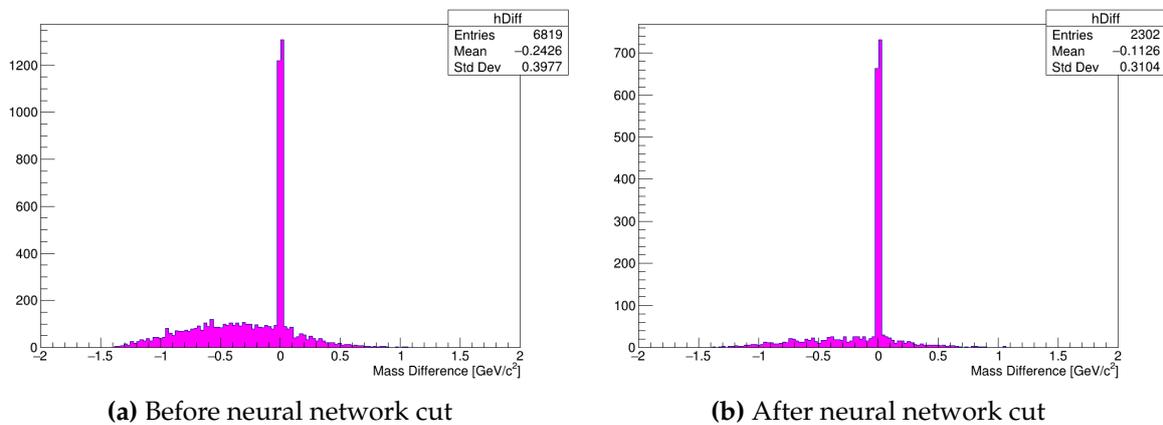
Figure 4.34.: Difference between generated and reconstructed $X_{s\bar{s}}$ mass in signal MC for $M(X_{s\bar{s}}) \leq 2.4\,\mathrm{GeV}/c^2$, for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes
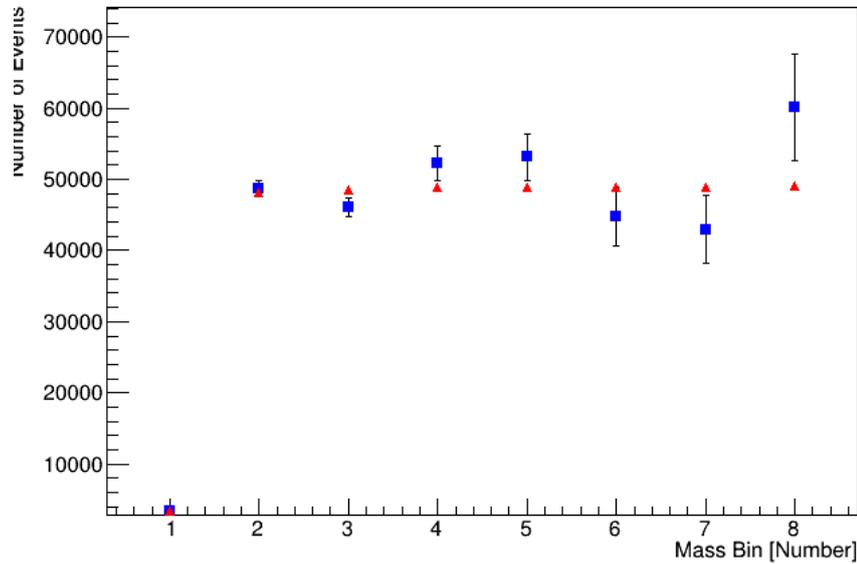
Figure 4.35.: Number of efficiency-corrected events in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The reconstructed distributions (blue) have approximately the same general shape and yields the generated (red) flat mass distribution
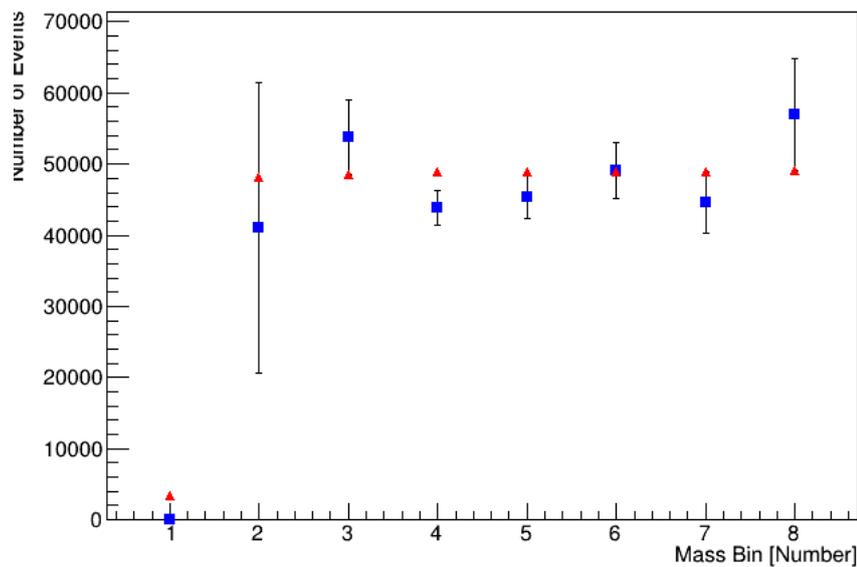
Figure 4.36.: Number of efficiency-corrected events in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The reconstructed distributions (blue) have approximately the same general shape and yields the generated (red) flat mass distribution
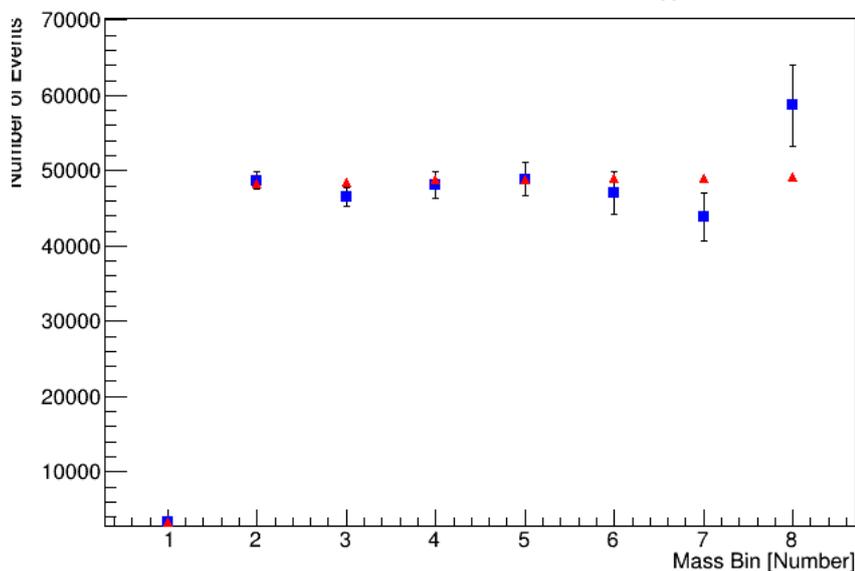
Figure 4.37.: Number of efficiency-corrected events in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for the weighted average of the $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ and $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The reconstructed distributions (blue) have approximately the same general shape and number of the generated (red) flat mass distribution
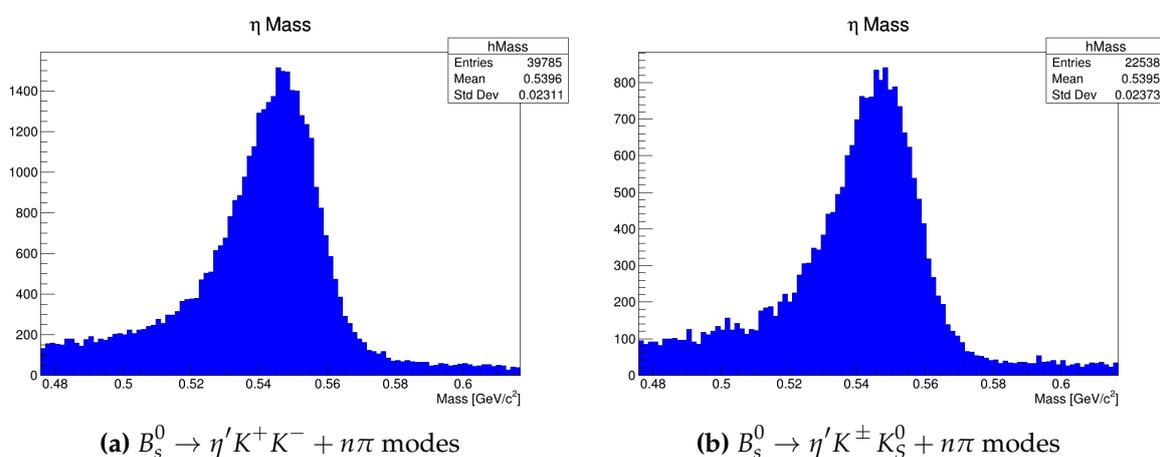
(a) $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes

Figure 4.38.: η mass reconstructed from signal MC without kinematic fitting, from B_s^0 candidates

4.8.3. η and η' Masses From Signal Monte Carlo

A check on the reconstruction of the η and η' masses in signal MC is also done. The masses, before the signal region selections are applied, of the η and η', without kinematic fitting, are shown in Figs. 4.38 and 4.39, respectively. Further, figures 4.40 and 4.41 show an appropriate correlation between the $X_{s\bar{s}}$ mass and the η' momentum.
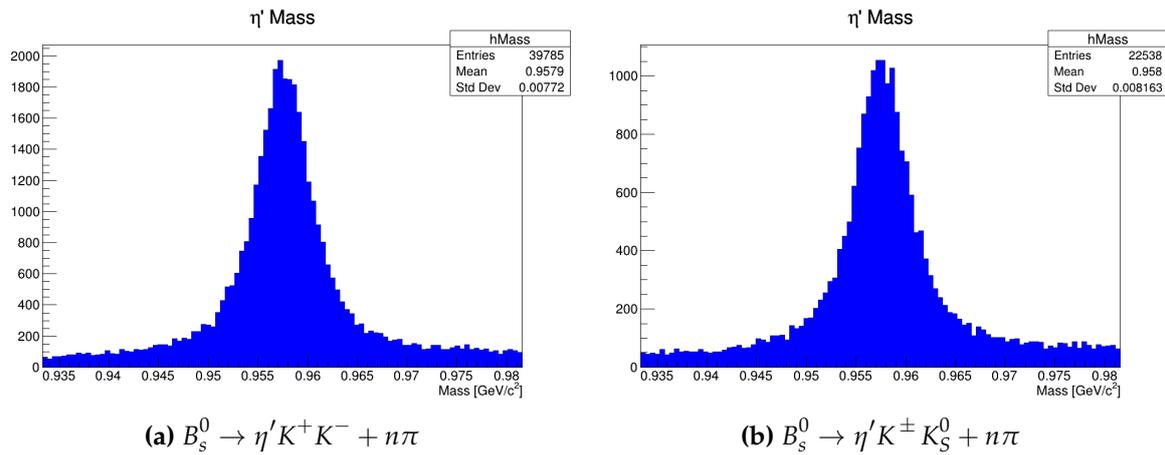
(a) $B_s^0 \to \eta' K^+ K^- + n\pi$

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$

Figure 4.39.: η' mass reconstructed from signal MC without kinematic fitting, from B_s^0 candidates

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes
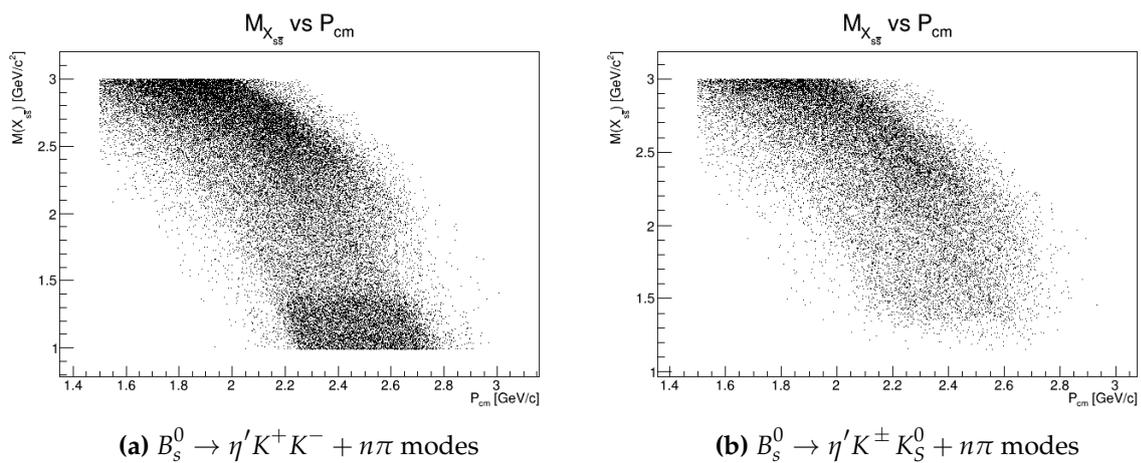
Figure 4.40.: $M(X_{s\bar{s}})$ vs $p_{cm}^{\eta'}$ - Showing the correlation between the two variables, with no signal region cuts applied.

(a) $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ **(b)** $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$
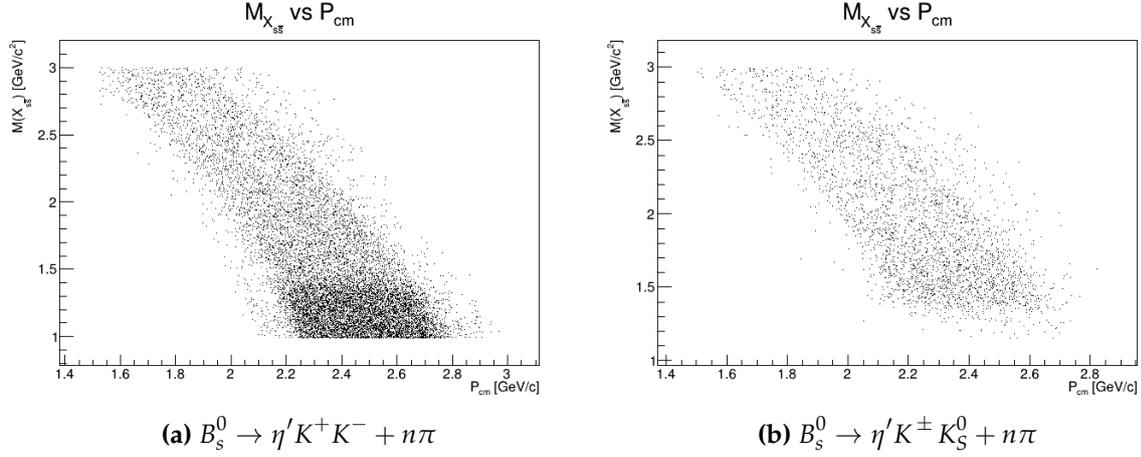
Figure 4.41.: $M(X_{s\bar{s}})$ vs $p_{cm}^{\eta'}$ - Showing the correlation between the two variables, with MC truth-matching applied but no signal region cuts.

4.9. Background Estimation

Background levels are estimated by fitting M_{bc} in a high statistics MC sample (multiple streams of background MC) in 0.2 GeV/c^2 bins of $X_{s\bar{s}}$ mass. For non-peaking backgrounds fitting is done with the ARGUS portion of the previously-discussed fit model. For peaking backgrounds (backgrounds that can mimic the signal), the full Gaussian + ARGUS model is used.

4.9.1. Peaking Backgrounds

Peaking Background Studied Using Generic MC

Background from generic $B\bar{B}$ and $B_s^{0(*)}\bar{B}_s^{0(*)}$ decays are considered as potential sources of peaking background. These backgrounds are studied by applying the same reconstruction and fitting procedure on high statistics $B\bar{B}$ and $B_s^{0(*)}\bar{B}_s^{0(*)}$ MC samples as the signal MC. Yields from this procedure are then scaled down to be concordant with experimental integrated luminosity.

It is determined that there is negligible peaking background in individual $X_{s\bar{s}}$ mass bins from generic $B\bar{B}$ decays, after all signal region cuts are applied. For generic $B_s^0\bar{B}_s^0$ decays there is similarly negligible peaking background. Signal decays in the $B_s^0\bar{B}_s^0$

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$
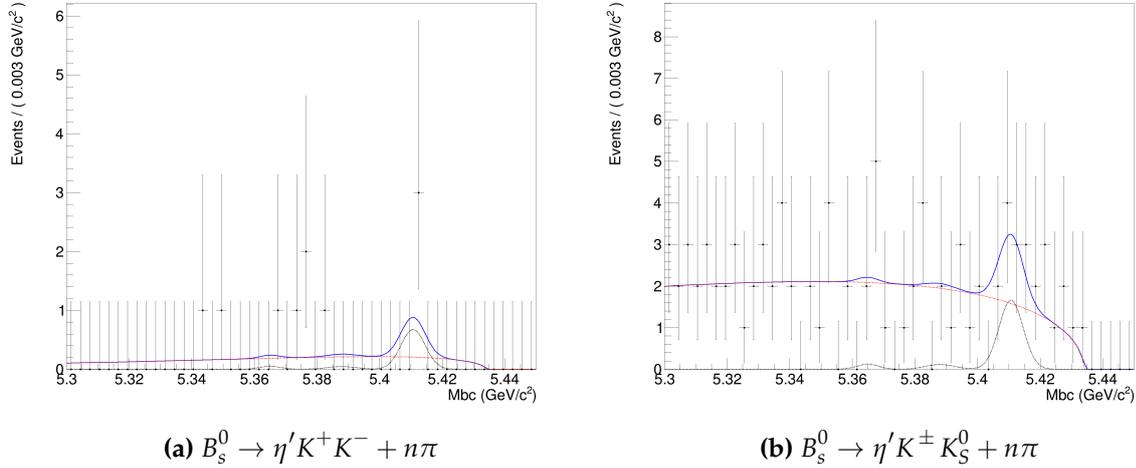
Figure 4.42.: Gaussian + ARGUS fit to five streams of $B_s^{0(*)} \bar{B}_s^{0(*)}$ MC for $M(X_{s\bar{s}}) \le 2.4\,\mathrm{GeV}/c^2$; when scaled to be concordant with experimental integrated luminosity, the number of peaking events is less than one. The red curve is the ARGUS portion of the fit, the black curve is the Gaussian portion of the fit, and the blue curve is the sum of the two.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0 $\pm$ 0.1
1.2-1.4	0.0
1.4-1.6	0.2 $\pm$ 0.3
1.6-1.8	0.0
1.8-2.0	0.4 $\pm$ 0.3
2.0-2.2	0.2 $\pm$ 0.4
2.2-2.4	0.0

Table 4.10.: Number of peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \to \eta' K^+ K^- + n\pi$. A high statistics MC sample is used. When scaled to be concordant with experimental integrated luminosity the number of peaking events is negligible.

MC sample are rejected using MC truth-matching. The results of this study are given in Tables 4.10, 4.11, 4.12, and 4.13.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0
1.2-1.4	0.0
1.4-1.6	0.1 ± 0.3
1.6-1.8	0.7 ± 0.5
1.8-2.0	0.0
2.0-2.2	1.2 ± 1.0
2.2-2.4	-0.7 ± 0.6

Table 4.11.: Number of peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins, for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes. A high statistics MC sample is used. When scaled to be concordant with experimental integrated luminosity the number of peaking events is negligible.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.2 ± 0.3
1.2-1.4	0.0
1.4-1.6	0.2 ± 0.2
1.6-1.8	0.0
1.8-2.0	0.2 ± 0.2
2.0-2.2	0.0 ± 0.1
2.2-2.4	0.0

Table 4.12.: Number of peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC (signal-like events are filtered out through MC truth-matching) in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes. A high statistics MC sample is used. When scaled to be concordant with experimental integrated luminosity the number of peaking events is negligible.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0
1.2-1.4	0.0
1.4-1.6	0.0
1.6-1.8	0.0
1.8-2.0	0.2 $\pm$ 0.4
2.0-2.2	0.6 $\pm$ 0.5
2.2-2.4	0.5 $\pm$ 0.7

Table 4.13.: Number of peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC (signal-like events are filtered out through MC truth-matching) in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins, for $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes. A high statistics MC sample is used. When scaled to be concordant with experimental integrated luminosity the number of peaking events is negligible.

Other Potential Peaking Backgrounds

Although the CM momentum cut on the η' helps to reduce some $b \to c$ background, there could be some residual peaking background due to other sources. The two background sources studied in this section are $B_s^0 \to \eta' D_s^- \pi^+$ and $B_s^0 \to \eta' \eta_c$, neither of which have been measured and published by any experiment to date. Both are Cabbibo-favored. To understand these backgrounds 1×10^5 MC events were generated. The resulting MC samples were analyzed with the analysis procedure to determine the efficiency and the expected number of peaking events.

As with the $B_s^0 \to \eta' X_{s\bar{s}}$ signal analysis, these backgrounds were evaluated in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins, with the η_c and $D_s^- \pi^+$ being the analogues of $X_{s\bar{s}}$ in their respective modes. If there were reconstructed events seen in bins where none were generated in MC (e.g. outside outside of the mass range of the η_c), then they were assumed to be a signal anyway and the efficiency and corresponding expected number of background events were determined. The number of expected peaking events was determined by assuming a BF of [3.3 $\pm$ 0.4] $\times 10^{-4}$ for $B_s \to \eta' \eta_c$, as that is the measured BF for $B_s \to \eta' J/\psi(1S)$ [37]. Similarly for $B_s^0 \to \eta' D_s^- \pi^+$ it is assumed that the BF is similar to $B \to D^- \pi^+ \rho^0$, which has a BF of [1.1 $\pm$ 1.0] $\times 10^{-3}$ [37].

There is also the tree-level process $B_s^0 \to \bar{D}^0 \eta'$, with $D^0 \to K^+ K^-$, that could contribute to the peaking background. From [37], the color-suppressed decay $B^0 \to \bar{D}^0 \eta'$ has a measured branching fraction of $\mathcal{B}(B^0 \to \bar{D}^0 \eta') =$ [1.38 $\pm$ 0.16] $\times 10^{-4}$. The Cabbibo-suppressed channel $D^0 \to K^+ K^-$ has a branching fraction measured to be $\mathcal{B}(D^0 \to K^+ K^-) =$ [4.08 $\pm$ 0.06] $\times 10^{-3}$. Multiplication of these with a CKM factor of approximately 0.05, gives 2.8×10^{-8}. Factoring in any reconstruction efficiency less than unity and the number of $B_s^{0(*)} \bar{B}_s^{0(*)}$ pairs produced at Belle, the number of expected events from this channel is negligible.

Other backgrounds sources are possible but due to the strangeness requirements of the final state, along with potential suppression, there are few modes that could possibly contribute an appreciable background. One of these is the decay $B \to \eta' K^{*0}$, where K^{*0} is reconstructed as a system of kaons and pions. This system can also proceed via a $b \to sg$ transition. It can contribute to the background if a pion is misidentified as a kaon.

To study this, 1×10^5 $B \to \eta' K^{*0}$ events were generated in MC. After all cuts were made few events remain, giving a negligible efficiency and no expected peaking

events. In each $X_{s\bar{s}}$ mass bin, the total number of expected events is less than one. This makes sense as BaBar measured the branching fraction for the mode $B \to \eta' K^*$ to be $\mathcal{O}\left(10^{-6}\right)$ [76].

Other modes of this type are thought to be similarly negligible, given previous BF measurements made in Belle and BaBar. These include charmless B meson decays with an η or η', $B^0 \to (\pi^0/\eta/\eta')K_S^0 K_S^0$, $B \to \eta'\eta' K$, $B^0 \to \eta K^0, \eta\eta, \eta\phi, \eta'\phi$, and $B \to \eta'\eta' K$ [59, 77–79].

4.9.2. Non-Peaking Background Estimates

Non-peaking continuum background from $e^+ e^- \to q\bar{q}$ ($q = u, d, s, c$) is the dominant background source in this analysis, as in many rare B-decay analyses. High statistics continuum MC samples generated at the $\Upsilon(5S)$ energy are used to study continuum background. Fours streams of continuum MC, five streams of $B\bar{B}$ MC, and five streams of $B_s^0 \bar{B}_s^0$ MC (with signal candidates removed by truth-matching), are analyzed with all selections are applied, to estimate the level of all non-peaking background. Results in subsequent tables are results from fitting, scaled to be concordant with experimental integrated luminosity, to account for the high statistics MC samples used for estimation. All non-peaking backgrounds are combined and fitted with a single ARGUS function (in individual $X_{s\bar{s}}$ mass bins). Fitting the full $\Upsilon(5S)$ data uses the combined model given in Table 4.5.

Plots of peaking and non-peaking backgrounds are shown in Figures 4.42, 4.43, 4.44, and 4.45. They include the $X_{s\bar{s}}$ mass cut off at 2.4 GeV/c^2. Numerical estimates of background yields are given in Tables 4.14, 4.15, 4.16, 4.17, 4.18, and 4.19.
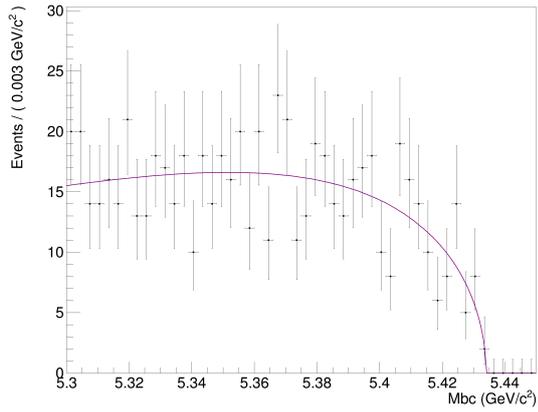
(a) $B_s^0 \to \eta' K^+ K^- + n\pi$
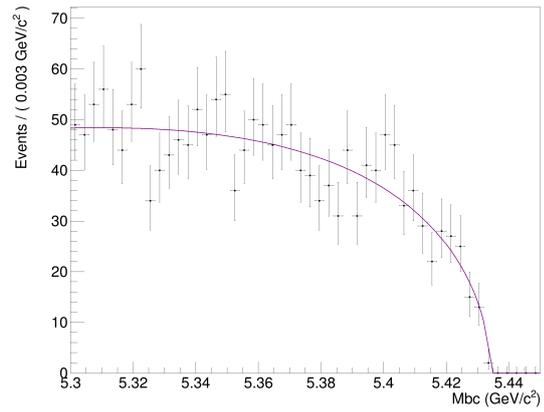
(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$

Figure 4.43.: ARGUS fit to four streams of continuum MC
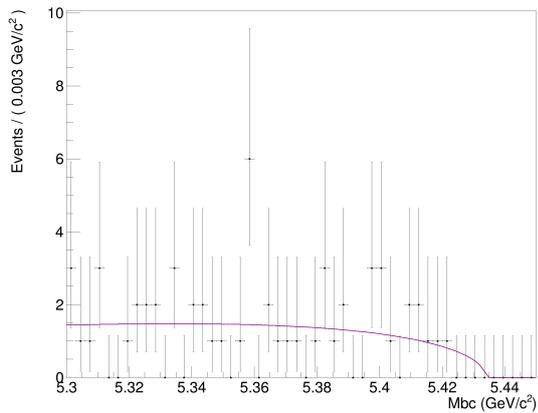
(a) $B_s^0 \to \eta' K^+ K^- + n\pi$
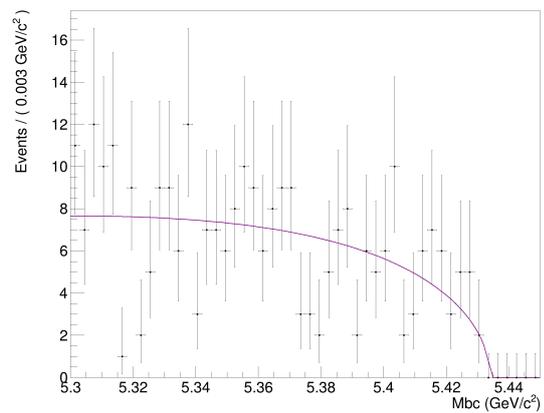
(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$

Figure 4.44.: ARGUS fit to five streams of generic $B\bar{B}$ MC

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$
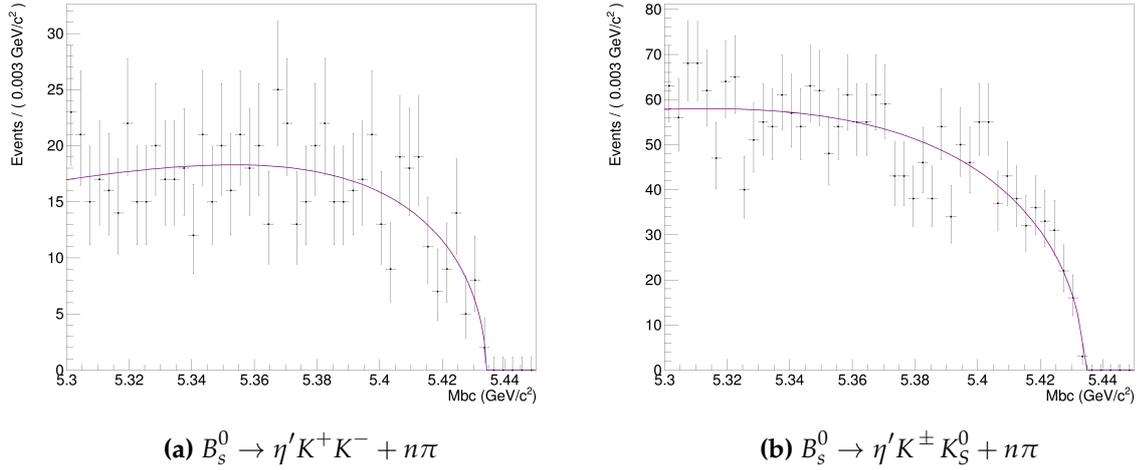
Figure 4.45.: ARGUS fit to five streams of generic $B\bar{B}$ MC + five streams of $B_s^0 \bar{B}_s^0$ MC + four streams of continuum MC; all non-peaking backgrounds are fit a single ARGUS function; continuum dominates the other background sources

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.2 $\pm$ 0.5
1.0-1.2	37.8 $\pm$ 3.0
1.2-1.4	30.5 $\pm$ 2.8
1.4-1.6	27.2 $\pm$ 2.5
1.6-1.8	21.0 $\pm$ 2.3
1.8-2.0	20.5 $\pm$ 2.3
2.0-2.2	12.2 $\pm$ 1.8
2.2-2.4	14.0 $\pm$ 1.9

Table 4.14.: Number of estimated non-peaking background events estimated from $q\bar{q}$ MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0
1.2-1.4	13.0 $\pm$ 1.8
1.4-1.6	43.8 $\pm$ 3.2
1.6-1.8	57.8 $\pm$ 3.8
1.8-2.0	85.0 $\pm$ 4.5
2.0-2.2	103.0 $\pm$ 5.0
2.2-2.4	148.8 $\pm$ 6.0

Table 4.15.: Number of estimated non-peaking background events estimated from $q\bar{q}$ MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.2 $\pm$ 0.2
1.2-1.4	1.7 $\pm$ 0.6
1.4-1.6	0.6 $\pm$ 0.4
1.6-1.8	2.4 $\pm$ 0.7
1.8-2.0	2.2 $\pm$ 0.7
2.0-2.2	1.8 $\pm$ 0.7
2.2-2.4	1.9 $\pm$ 0.6

Table 4.16.: Number of estimated non-peaking background events estimated from $B\bar{B}$ (non-$B_s^0 \bar{B}_s^0$) MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0
1.2-1.4	0.0
1.4-1.6	0.5 $\pm$ 0.4
1.6-1.8	3.3 $\pm$ 0.9
1.8-2.0	8.3 $\pm$ 1.3
2.0-2.2	15.8 $\pm$ 2.0
2.2-2.4	27.2 $\pm$ 2.4

Table 4.17.: Number of estimated non-peaking background events estimated from $B\bar{B}$ (non-$B_s^0\bar{B}_s^0$) MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.2 $\pm$ 0.2
1.2-1.4	0.0
1.4-1.6	0.0
1.6-1.8	0.2 $\pm$ 0.2
1.8-2.0	0.2 $\pm$ 0.2
2.0-2.2	0.2 $\pm$ 0.2
2.2-2.4	0.7 $\pm$ 0.3

Table 4.18.: Number of estimated non-peaking background events estimated from $B_s^0\bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

Mass Bin (GeV/c^2)	Est. Events
0.8-1.0	0.0
1.0-1.2	0.0
1.2-1.4	0.0
1.4-1.6	0.3 $\pm$ 0.2
1.6-1.8	0.4 $\pm$ 0.3
1.8-2.0	2.0 $\pm$ 0.7
2.0-2.2	5.4 $\pm$ 1.1
2.2-2.4	8.5 $\pm$ 1.4

Table 4.19.: Number of estimated non-peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, for $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes. A high statistics MC sample is used. Results are obtained after scaling the fitted MC results to be concordant with experiment integrated luminosity.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes
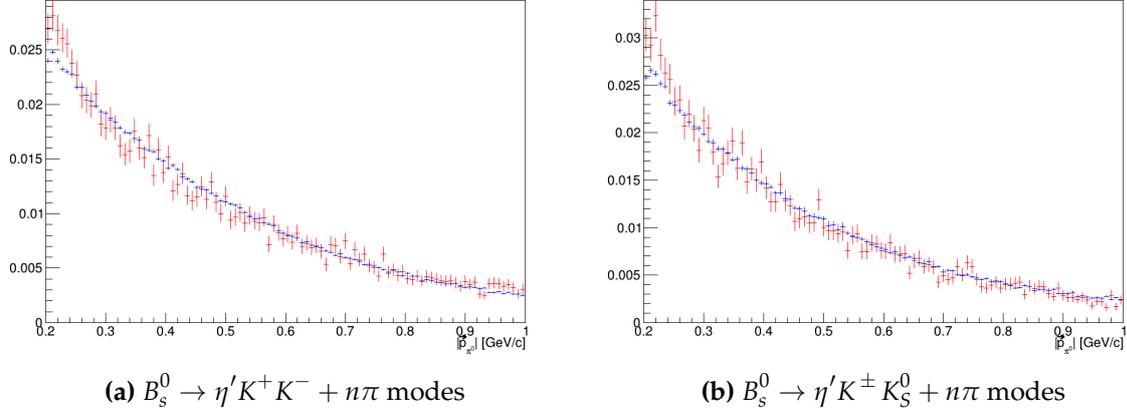
Figure 4.46.: $|\mathbf{p}_{\pi^0}|$ for signal MC and continuum MC, in the range $0.2 < |\mathbf{p}_{\pi^0}| < 1.0\,\text{GeV}/c$; the blue curve is continuum and the red curve is the signal; no additional cuts to be applied are apparent from these plots

4.9.3. Additional Check on the NN

As a further check, the efficiencies of the signal and background separately are plotted versus nnout. For the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes the area around the nnout cut (nnout = 0.95), from nnout = 0.855 to nnout = 0.97 in the signal efficiency is not varying rapidly. Similarly, the background efficiency is not varying rapidly. The same can be seen for the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes. See Figures 4.47 and 4.48. These figures also demonstrate the functionality of the neural network: the background efficiency drops faster than the signal efficiency.

4.9.4. Possible Additional Cut

Further, Fig. 4.46 demonstrates that no additional cut is apparent based on the π^0 momentum comparison between signal and continuum MC.

4.9.5. Cut-flow

The cut efficiencies are given in Tables 4.20 and 4.21.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes
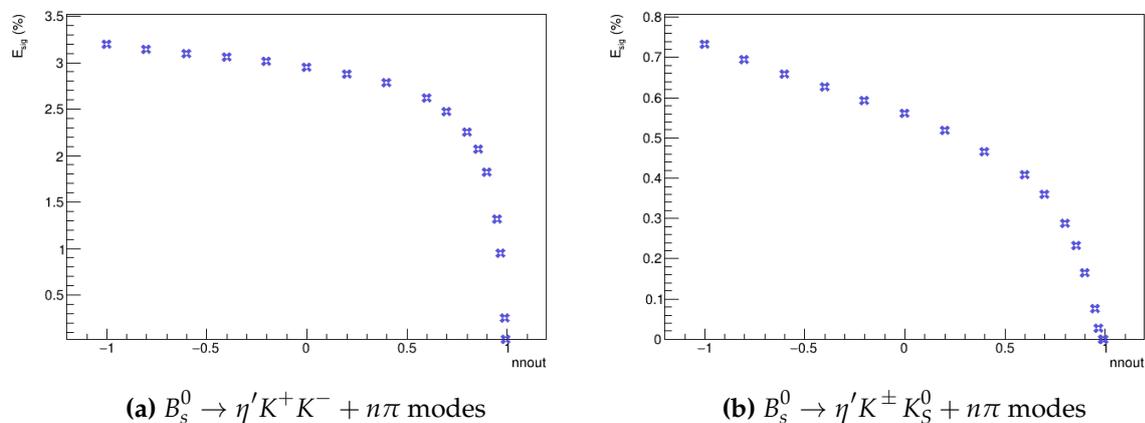
Figure 4.47.: Efficiency vs nnout in signal MC for $M(X_{s\bar{s}}) \leq 2.4\,\mathrm{GeV}/c^2$

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes
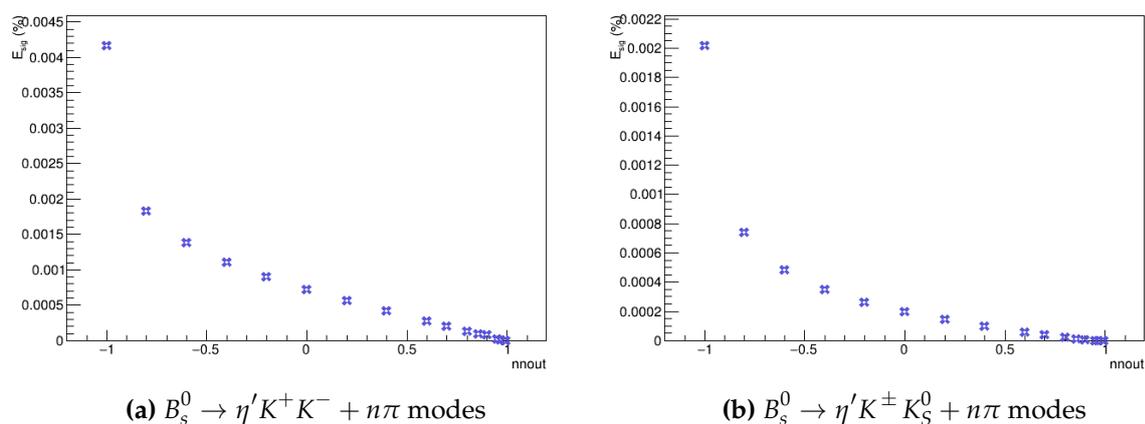
Figure 4.48.: Efficiency vs nnout (all generic background MC sources) in signal MC for $M(X_{s\bar{s}}) \leq 2.4\,\mathrm{GeV}/c^2$

Cut	Sig. (%)	$q\bar{q}$ (%)	$B_s^0 \bar{B}_s^0$ (%)	$B\bar{B}$ (%)
Loose M_{bc} + ΔE cuts + BCS	3.7	0.03	0.08	0.03
+ tighter ΔE	3.6	0.02	0.05	0.02
+ R2	3.5	0.02	0.05	0.02
+ nnout	1.4	2.6×10^{-5}	3.9×10^{-5}	3.9×10^{-5}

Table 4.20.: Cut-flow table for analysis for the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes, for in the flat mass model for $M(X_{s\bar{s}}) \leq 2.4\,\mathrm{GeV}/c^2$; the "+" sign indicates a cut that were added to the cut listed in the above row.

Cut	Sig. (%)	$q\bar{q}$ (%)	$B_s^0 \bar{B}_s^0$ (%)	$B\bar{B}$ (%)
Loose $M_{bc} + \Delta E$ cuts + BCS	0.96	0.02	0.05	0.02
+ tighter ΔE	0.9	0.01	0.03	0.01
+ R2	0.9	0.01	0.03	0.01
+ nnout	0.5	1.5×10^{-5}	1.2×10^{-3}	5.3×10^{-4}

Table 4.21.: Cut-flow table for analysis for the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes, for in the flat mass model for $M(X_{s\bar{s}}) \le 2.4 \, \text{GeV}/c^2$; the "+" sign indicates a cut that were added to the cut listed in the above row.
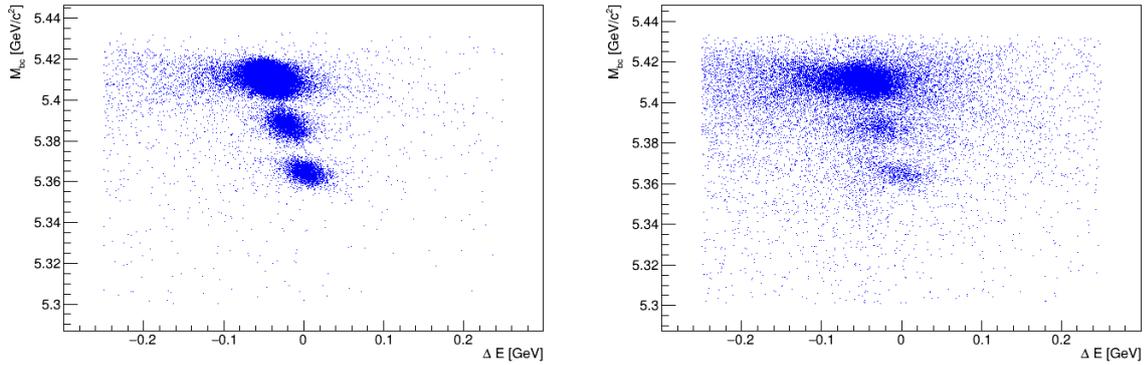
(a) $B_s \to D_s^- \pi^+$ MC; no signal region cuts are applied

(b) $B_s \to D_s^- \rho$ MC; no signal region cuts are applied

Figure 4.49.: M_{bc} vs ΔE for control MC

4.10. Control

4.10.1. Monte Carlo

Two control modes are studied to determine discrepancies between the data and MC fitting, as well as to determine the soundness of the analysis method. The modes are $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \rho^+$. In practice, $B_s^0 \to D_s^- \rho^+$ alone is sufficient to act as a control for this analysis but for historical reasons both are analyzed.

(a) $B_s \to D_s^- (\to \phi \pi^-) \rho^+$ MC   **(b)** $B_s \to D_s^- (\to K^0 K^-) \rho^+$ MC   **(c)** $B_s \to D_s^- (\to K^{*0} K^-) \rho^+$ MC
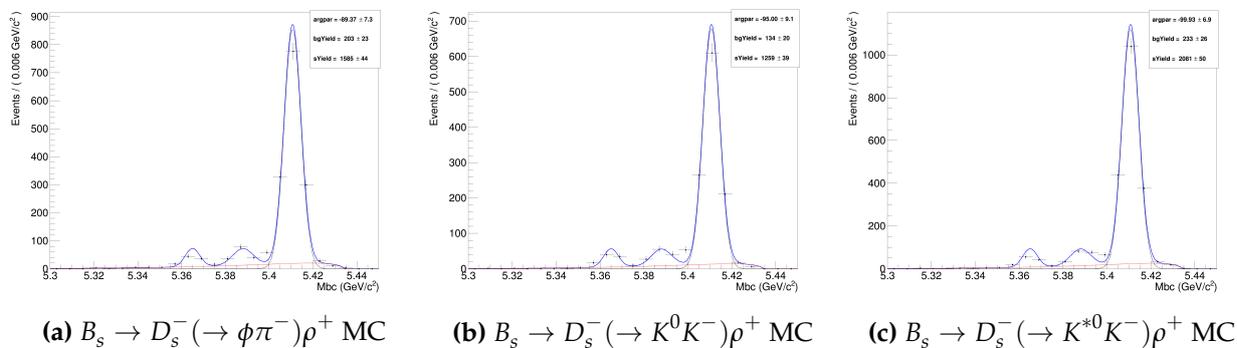
Figure 4.50.: M_{bc} fits with nnout≥ 0.95

(a) $B_s \to D_s^- (\to \phi \pi^-) \rho^+$ MC   **(b)** $B_s \to D_s^- (\to K^0 K^-) \rho^+$ MC   **(c)** $B_s \to D_s^- (\to K^{*0} K^-) \rho^+$ MC
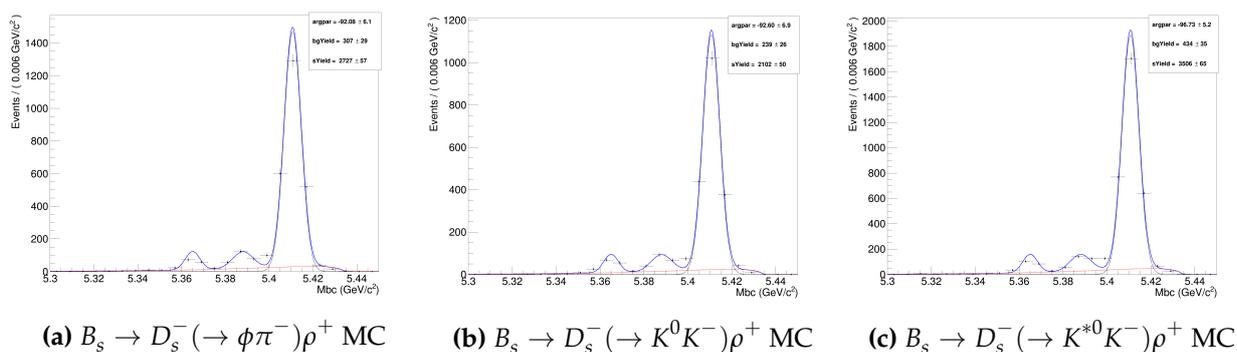
Figure 4.51.: M_{bc} fits with nnout≥ 0.6

(a) $B_s \to D_s^- (\to \phi \pi^-) \pi^+$ MC   **(b)** $B_s \to D_s^- (\to K^0 K^-) \pi^+$ MC   **(c)** $B_s \to D_s^- (\to K^{*0} K^-) \pi^+$ MC
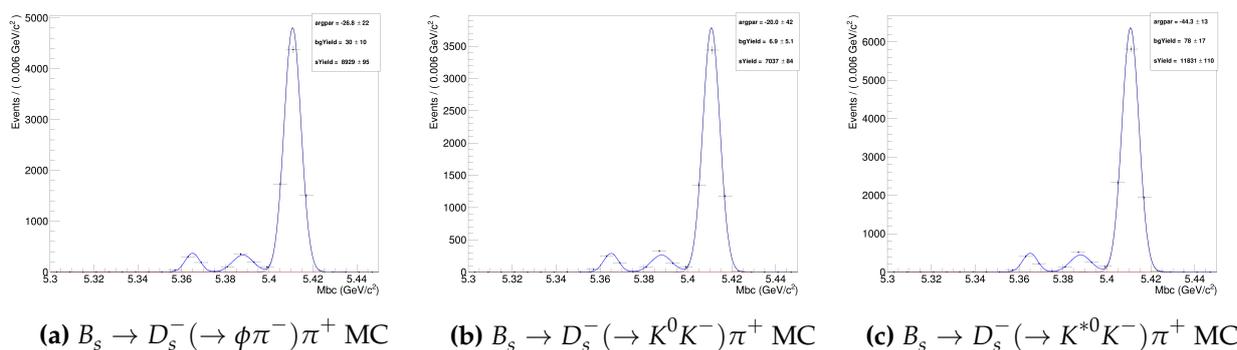
Figure 4.52.: M_{bc} fits with nnout≥ 0.95

Mode	ΔE_{mean} **(GeV)**
$B_s^0 \to D_s^- \pi^+$	-0.042881 $\pm$ 0.00011
$B_s^0 \to D_s^- \rho^+$	-0.043815 $\pm$ 0.00013

Table 4.22.: BCS χ^2 ΔE_{mean} values from MC for control modes

(a) $B_s \to D_s^- (\to \phi\pi^-)\pi^+$ MC   **(b)** $B_s \to D_s^- (\to K^0K^-)\pi^+$ MC   **(c)** $B_s \to D_s^- (\to K^{*0}K^-)\pi^+$ MC
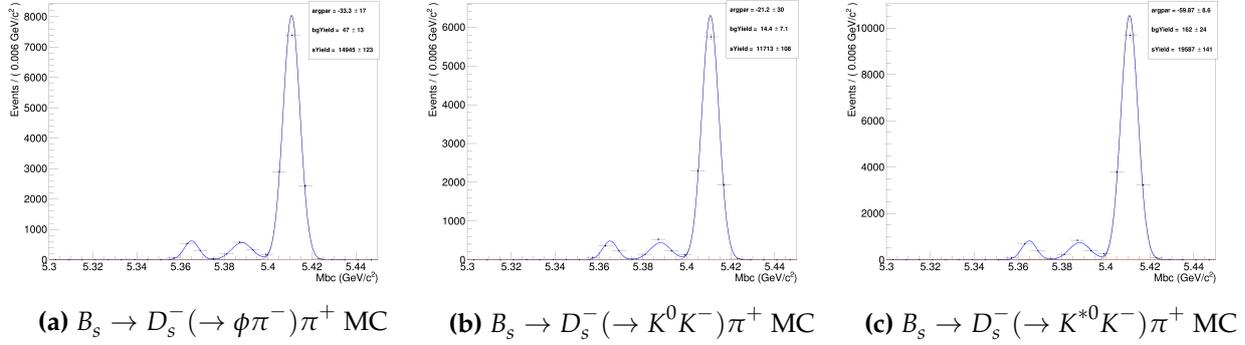
Figure 4.53.: M_{bc} fits with nnout≥ 0.6

Mode	$\sigma_{\Delta E}$ (GeV)
$B_s^0 \to D_s^- \pi^+$	0.01278 ± 0.000076
$B_s^0 \to D_s^- \rho^+$	0.02036 ± 0.00011

Table 4.23.: BCS χ^2 $\sigma_{\Delta E}$ values from MC for the control modes

4.10.2. Control Branching Fractions

Branching fractions for the control modes are calculated to check for consistency [65]. For the control modes the entire Belle $\Upsilon(5S)$ data sample was analyzed, 121.4 fb^{-1}. The discrepancies between data and MC are listed in Tables 4.29 and 4.32.

The branching fractions are calculated as:

$$\mathcal{B}(B_s^0 \to D_s^- \pi^+/\rho^+) = \frac{N_{\mathrm{sig}}}{2 \times N_{B_s^{0(*)}\bar{B}_s^{0(*)}}[\Sigma_{i=1,2,3}\mathcal{B}_i(D_s^- \to X_i^-)\epsilon_i^{MC}]} \qquad (4.12)$$

where i runs over the three modes listed in Table 5.1 and where $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ is the measured value for the number of $B_s^{0(*)}\bar{B}_s^{0(*)}$ pairs, obtained from [8]. Efficiencies (ϵ_i^{MC}) are determined by generating 1×10^5 events of each control mode in MC, one sample for each D_s^- sub-decay mode. Figures 4.49, 4.50, 4.51, 4.52, and 4.53 show the results from MC. Tables 4.22 and 4.23 give the BCS χ^2 values.

The previous Belle measurements are:

Mode	Branching Fraction(%)
$D_s^- \to \phi\pi^-$	4.50 ± 0.4
$D_s^- \to K^0K^-$	2.95 ± 0.14
$D_s^- \to K^{*0}K^-$	3.92 ± 0.12

Table 4.24.: D_s^- Branching Fractions from [37]

D_s Sub-mode	$B_s^0 \to D_s^-\pi^+$ MC Efficiency(%)	$B_s^0 \to D_s^-\rho^+$ MC Efficiency(%)
$D_s^- \to \phi\pi^-$	8.9 ± 0.03	1.6 ± 0.01
$D_s^- \to K^0K^-$	7.0 ± 0.08	1.2 ± 0.01
$D_s^- \to K^{0*}K^-$	11.8 ± 0.1	2.1 ± 0.02

Table 4.25.: Efficiencies ϵ^{MC} estimated using the signal NN and nnout≥ 0.95

$\mathcal{B}(B_s^0 \to D_s^-\pi^+) = [3.67^{+0.35}_{-0.33}(\text{stat.})^{+0.43}_{-0.42}(\text{sys.}) \pm 0.49(f_s)] \times 10^{-3}$, and $\mathcal{B}(B_s^0 \to D_s^-\rho^+) = [8.5^{+1.3}_{-1.2}(\text{stat.}) \pm 1.1(\text{sys.}) \pm 1.3(f_s)] \times 10^{-3}$ [65, 80]. The PDG values are $\mathcal{B}(B_s^0 \to D_s^-\pi^+) = [3.00 \pm 0.21] \times 10^{-3}$ and $\mathcal{B}(B_s^0 \to D_s^-\rho^+) = [6.9 \pm 1.4] \times 10^{-3}$ [37].

Control studies utilized relevant $B_s^0 \to \eta' X_{s\bar{s}}$ analysis cuts, including the neural network cuts . The nnout cuts determined for the $B_s^0 \to \eta' K^+K^- + n\pi$ and and $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes are each applied to control MC and $\Upsilon(5S)$ data samples.

Using the $B_s^0 \to \eta' K^+K^- + n\pi$ NN cut, the control BFs are determined as:

$$\mathcal{B}(B_s^0 \to D_s^-\pi^+) = (4.0 \pm 0.2(\text{stat.})) \times 10^{-3} \tag{4.13}$$

D_s^- Sub-mode	$B_s^0 \to D_s^-\pi^+$ MC Efficiency(%)	$B_s^0 \to D_s^-\rho^+$ MC Efficiency(%)
$D_s^- \to \phi\pi^- (\phi \to K^+K^-)$	14.9 ± 0.1	2.7 ± 0.06
$D_s^- \to K^0K^-$	11.7 ± 0.1	2.1 ± 0.05
$D_s^- \to K^{0*}K^-$	19.6 ± 0.1	3.5 ± 0.06

Table 4.26.: Efficiencies ϵ^{MC} estimated using the signal NN and nnout≥ 0.6

(a) No NN cut   **(b)** nnout$\geq$ 0.95   **(c)** nnout$\geq$ 0.6
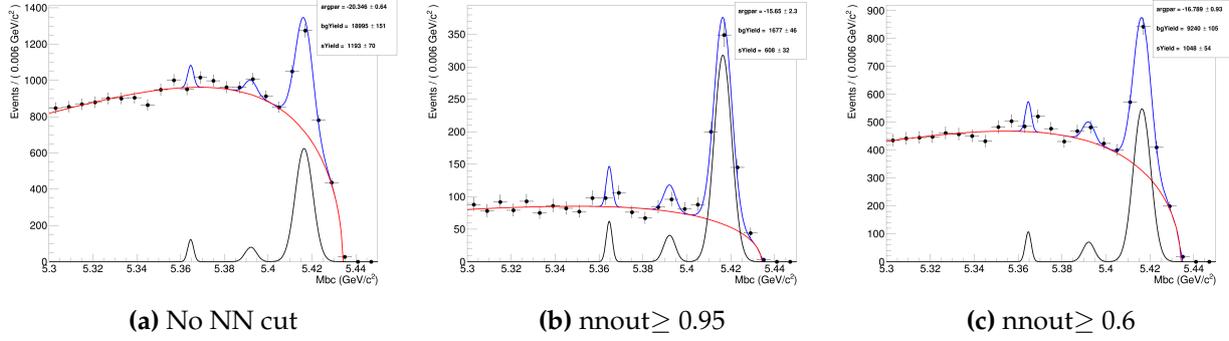
Figure 4.54.: $B_s^0 \to D_s^- \pi^+$ - Results from fitting to the full 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ data sample

$$\mathcal{B}(B_s^0 \to D_s^- \rho^+) = (9.7 \pm 0.9(\text{stat.})) \times 10^{-3} \qquad (4.14)$$

Using the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ NN cut, the control BFs are determined as:

$$\mathcal{B}(B_s^0 \to D_s^- \pi^+) = (4.1 \pm 0.5(\text{stat.})) \times 10^{-3} \qquad (4.15)$$

$$\mathcal{B}(B_s^0 \to D_s^- \rho^+) = (10.1 \pm 0.9(\text{stat.})) \times 10^{-3} \qquad (4.16)$$

The results from the full $\Upsilon(5S)$ data sample are shown in Figs. 4.54 and 4.55.

4.10.3. Comparison Numbers

The difference in the means of the Gaussians and the ratio of their standard deviations, between data and MC, can be checked for any disparities. The results are listed in Tables 4.29 and 4.32.

(a) No NN cut   **(b)** nnout$\geq$ 0.95   **(c)** nnout$\geq$ 0.6
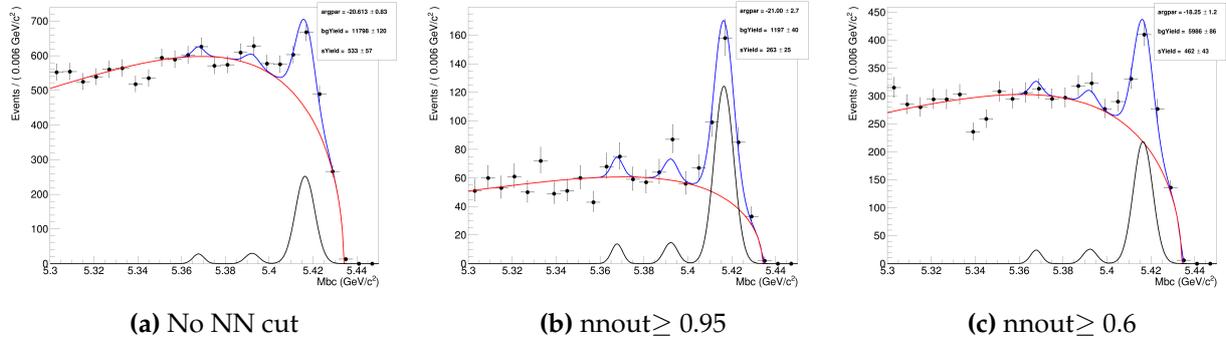
Figure 4.55.: $B_s^0 \to D_s^- \rho^+$ - Results from fitting to the full 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ data sample

μ (GeV/c^2)	σ (GeV/c^2)
5.4102 $\pm$ 0.00030	0.0040 $\pm$ 0.00027
5.3869 $\pm$ 0.0018	0.0037 $\pm$ 0.0013
5.3634 $\pm$ 0.0025	0.0035 $\pm$ 0.0020

Table 4.27.: Values for means and widths for $B_s^0 \to D_s^- \pi^-$ control fit to MC; values obtained after all cuts are applied

μ (GeV/c^2)	σ (GeV/c^2)
5.4164 $\pm$ 0.00020	0.00398 $\pm$ 0.00033
5.3921 $\pm$ 0.00130	0.00260 $\pm$ 0.00100
5.3646 $\pm$ 0.00039	0.00133 $\pm$ 0.00058

Table 4.28.: Values for means and widths for $B_s^0 \to D_s^- \pi^+$ control fit to data; values obtained after all cuts are applied

$\delta\mu$ (GeV/c^2)	R
0.0062 $\pm$ 0.00040	0.995 $\pm$ 0.106
0.0052 $\pm$ 0.00220	0.703 $\pm$ 0.366
0.0012 $\pm$ 0.00250	0.380 $\pm$ 0.273

Table 4.29.: Comparison values from $B_s^0 \to D_s^- \pi^+$ control study

μ (GeV/c^2)	σ (GeV/c^2)
5.4108 $\pm$ 0.000032	0.00426 $\pm$ 0.000028
5.3886 $\pm$ 0.00017	0.00545 $\pm$ 0.00015
5.3647 $\pm$ 0.00016	0.00435 $\pm$ 0.00015

Table 4.30.: Values for means and widths for $B_s^0 \rightarrow D_s^- \rho^+$ control fit to MC; values obtained after all cuts are applied

μ (GeV/c^2)	σ (GeV/c^2)
5.4164 $\pm$ 0.0004	0.00440 $\pm$ 0.0003
5.3922 $\pm$ 0.0019	0.00310 $\pm$ 0.0016
5.3677 $\pm$ 0.0015	0.00259 $\pm$ 0.0009

Table 4.31.: Values for means and widths for $B_s^0 \rightarrow D_s^- \rho^+$ control fit to data; values obtained after all cuts are applied

Here $\delta\mu$ is the difference in the means between the data and MC Gaussian means, and R is the ratio of their widths. Errors are obtained from addition in quadrature.

4.11. Correction Factors

The data/MC efficiency correction factors are determined and given as follows.

- Neural Network Output cut: nnout cut

A correction factor associated with the cut on nnout is determined. The neural network cut is removed and the signal (Gaussian) yield (N) of the fit is re-determined. This is done for both data and MC. The ratio of the data and MC results is then calculated, i.e a double ratio - r, and is used as the efficiency correction factor.

$\delta\mu$ (GeV/c^2)	R
0.0056 $\pm$ 0.0004	1.033 $\pm$ 0.0707
0.0036 $\pm$ 0.0019	0.569 $\pm$ 0.2940
0.0030 $\pm$ 0.0002	0.595 $\pm$ 0.2079

Table 4.32.: Comparison values from $B_s^0 \rightarrow D_s^- \rho^+$ control study

Ratio	Value
$r_{Data} = N_{standard}/N_{loose}$	0.504 ± 0.028
$r_{MC} = N_{standard}/N_{loose}$	0.539 ± 0.005
$r = r_{Data}/r_{MC}$	0.935 ± 0.054

Table 4.33.: nnout correction factor for $B_s^0 \to D_s^- \rho^+$, using the $B_s^0 \to \eta' K^+ K^- + n\pi$ NN cut

Ratio	Value
$r_{Data} = N_{standard}/N_{loose}$	0.893 ± 0.016
$r_{MC} = N_{standard}/N_{loose}$	0.912 ± 0.003
$r = r_{Data}/r_{MC}$	0.979 ± 0.017

Table 4.34.: nnout correction factor for $B_s^0 \to D_s^- \rho^+$, using the $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ NN cut

The signal yield with the standard cut is $N_{standard}$. The signal yield with a looser (removed) cut is N_{loose}. The ratios $r_{Data/MC}$ are calculated as $N_{standard}/N_{loose}$. The errors for $r_{Data/MC}$ are calculated using the binomial formula given in section 4.8. Here $N^{rec} = N_{standard}$, $\sigma_{rec} = \delta N^{rec}$, $N^{gen} = N_{loose}$, $\sigma_{gen} = \delta N^{gen}$, $N^{fail} = N^{gen}$ - N^{rec}, $\sigma_{fail} = \sqrt{N_{fail}}$, where δN^{rec} and δN^{gen} are the errors on the signal yield from fitting. These errors on r_{Data} and r_{MC} are then added in quadrature to obtain the error on r.

Correction factors that are ≤ 1.0 by less than 3σ are not applied but the deviation from 1.0 is taken as a conservative method of determining the corresponding systematic uncertainties.

- Best Candidate Selection (BCS)

The correction factor associated with BCS is obtained in a similar way to the correction factor for the neural network cut, by comparing signal yields with and without BCS.

Ratio	Value
$r_{Data} = N_{standard}/N_{No\_BCS}$	0.854 ± 0.022
$r_{MC} = N_{standard}/N_{No\_BCS}$	0.862 ± 0.005
$r = r_{Data}/r_{MC}$	0.990 ± 0.026

Table 4.35.: BCS correction factor for $B_s^0 \to D_s^- \rho^+$, using the $B_s^0 \to \eta' K^+ K^- + n\pi$ NN cut.

Ratio	Value
$r_{Data} = N_{standard}/N_{No\_BCS}$	0.901 ± 0.015
$r_{MC} = N_{standard}/N_{No\_BCS}$	0.863 ± 0.004
$r = r_{Data}/r_{MC}$	1.044 ± 0.018

Table 4.36.: BCS correction factor for $B_s^0 \to D_s^- \rho^+$, using the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ NN cut.

	Factor
$K^{\pm}$	1.0083 ± 0.0096
$\pi^{\pm}$	0.9809 ± 0.0127

Table 4.37.: PID correction factors

Only the correction factor for the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes is applied as it has a statistical significance of $\geq 3\sigma$ from unity.

- PID

Correction factors associated with PID were determined by the information provided by the Belle PID group, and use the momentum and polar angles of tracks to calculate correction factors and systematic uncertainties [81].

- $\pi^{\pm}$ Tracking Efficiency

From studies of $D^* \to \pi D$, $D \to \pi\pi K_S^0$, and $K_S^0 \to \pi^+ \pi^+$ in [82], the data/MC correction factor is 0.9987. The average charged pion multiplicity per mass bin from signal MC is used to calculate an overall correction factor in each bin as (correction factor)^{multiplicity}.

- $\eta \to \gamma\gamma$ Reconstruction

From table 9 in [83], a correction factor of 0.979 is applied for η reconstruction.

- $\pi^0 \to \gamma\gamma$ Reconstruction

From Table 18 in [83] a correction factor of 0.924 ± 0.009 (stat.) ± 0.011 (syst.), corresponding to the appropriate energy and momentum cuts used, is applied for π^0 reconstruction.

- K_S^0

From [84], the applied correction factor is 0.9857, for K_S^0 reconstruction.

Bin (GeV/c^2)	Correction/Bin
0.8 - 1.0	0.955
1.0 - 1.2	0.955
1.2 - 1.4	0.953
1.4 - 1.6	0.939
1.6 - 1.8	0.936
1.8 - 2.0	0.933
2.0 - 2.2	0.930
2.2 - 2.4	0.928

Table 4.38.: Efficiency correction factor per mass bin for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

Bin (GeV/c^2)	Correction/Bin
0.8 - 1.0	1.008
1.0 - 1.2	0.955
1.2 - 1.4	0.955
1.4 - 1.6	0.955
1.6 - 1.8	0.951
1.8 - 2.0	0.947
2.0 - 2.2	0.947
2.2 - 2.4	0.945

Table 4.39.: Efficiency correction factor per mass bin for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes

To incorporate these correction factors (CF) into the final branching fraction calculations, the average charged $K^\pm$, $\pi^\pm$, and π^0 multiplicities per $X_{s\bar{s}}$ mass bin in signal MC are determined. The total correction factor, calculated bin-by-bin, is the multiplication of all relevant correction factors, with those associated with track multiplicities being raised to the average track multiplicity, e.g. ($K^\pm$ correction factor)$^{\text{avg.mult.}}$. Final correction factors per $X_{s\bar{s}}$ mass bin are given in Tables 4.38 and 4.39.

4.12. Systematic Uncertainties

4.12.1. Multiplicative Systematic Uncertainties

The multiplicative systematic uncertainties are dealt with separately from the additive systematic uncertainties. Within the multiplicative systematic uncertainties, the uncertainties related to PYTHIA are treated separately in the final BF calculations. The non-PYTHIA uncertainties in each $X_{s\bar{s}}$ mass bin are added in quadrature to obtain a bin-by-bin systematic uncertainty. This does not include the $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ uncertainty, as it is included only in the final BF weighted average.

- Neural Network cut (cut on nnout)

The systematic uncertainty associated with the cut on the neural network variable, nnout, is determined as the difference from unity of the double ratio, r, in Tables 4.33 and 4.34. The uncertainty is 6.5% for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes and 2.1% for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes.

- Best Candidate Selection (BCS)

The BCS uncertainty is determined as the difference from unity of the double ratio, r, (BCS correction factor) in Tables 4.35 and 4.36. The uncertainty is approximately 0.96% for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes and 4.4% for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes.

- $\eta \to \gamma\gamma$ Reconstruction

From [85], a systematic uncertainty of 3.0% is assigned.

- $\pi^0 \to \gamma\gamma$ Reconstruction

From [85], a systematic uncertainty of 3.0% can be assigned. The average π^0 multiplicity per $X_{s\bar{s}}$ mass bin in signal MC is used to calculate the total bin systematic uncertainty as (0.03)*M, where M is the average π^0 multiplicity in signal MC.

- $\pi^\pm$ Tracking Efficiency

From studies of $D^* \to \pi D$, $D \to \pi\pi K_S^0$, and $K_S^0 \to \pi^+\pi^+$ in [82, 86], an uncertainty of 0.35% is used. The average charged pion multiplicity per mass bin is used to calculate an overall uncertainty in each bin as: total 0.0035 * (average multiplicity).

Mode	πID/track (%)	KID/track (%)	πID Total (%)	KID Total (%)
0π	1.28	0.95	2.56	1.90
1π	1.28	0.95	2.56	1.90
2π	1.28	0.95	5.12	1.90
3π	1.28	0.95	5.12	1.90
4π	1.28	0.95	7.68	1.90
$1\pi1K_S^0$	1.28	0.95	3.84	0.95
$2\pi1K_S^0$	1.28	0.95	3.84	0.95
$3\pi1K_S^0$	1.28	0.95	5.12	0.95
$4\pi1K_S^0$	1.28	0.95	5.12	0.95

Table 4.40.: Charged track PID efficiency uncertainties; $B_s^0 \to \eta'(\to \eta\pi^+\pi^-)X_{s\bar{s}}$

• Particle Identification (PID)

The uncertainty in the efficiency to identify kaons and pions is obtained by using tables prepared by the Belle PID group. They are calculated using charged kaon and pion momenta and polar angles [81]. The average number of charged kaons and pions per $X_{s\bar{s}}$ mass bin in signal MC is used to assign this systematic uncertainty (it is under the "CF" column header in the subsequent tables, CF meaning "correction factor", the uncertainty on the correction factor associated with kaon and pion identification) and the systematic uncertainty pertaining to tracking. This is calculated using the uncertainties given in Table 4.37.

The uncertainty on PID correction factors of 0.0096 for kaons and 0.0179 for pions are used. The relative uncertainty for a charged kaon is then (0.0096/1.0083)100% = 0.95%. For a charged pion it is (0.0127/0.9809)100% = 1.28%. The common daughter particles of a single B_s^0 are correlated so the uncertainties are added linearly. If M is the average multiplicity of charged kaons per $X_{s\bar{s}}$ mass bin in signal MC, and N is the average charged pion multiplicity per $X_{s\bar{s}}$ mass bin, then the total systematic uncertainty in percent per $X_{s\bar{s}}$ mass bin is M*(0.95)% for kaons and N*(1.28)% for pions.

• K_S^0 Reconstruction

A systematic uncertainty of 1.57% is assigned for each K_S^0 [84, 87].

Mode	K_S^0 Uncertainty (%)
0π	0
1π	0
2π	0
3π	0
4π	0
$1\pi 1K_S^0$	1.57
$2\pi 1K_S^0$	1.57
$3\pi 1K_S^0$	1.57
$4\pi 1K_S^0$	1.57

Table 4.41.: K_S^0 per mode; uncertainty is obtained by counting the number of modes with K_S^0

- Production Model

The uncertainty associated with the $\Upsilon(5S)$ production model (PM) is approximately 0.2% for $B_s^0 \to \eta' K^+ K^- + n\pi$ and 1.1% for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$. It is the relative change in the reconstruction efficiency of the $B_s^{0*}\bar{B}_s^{0*}$ signal peak with and without the model in [88] applied.

- $N_{B_s^{(*)}\bar{B}_s^{(*)}}$

From [8] the number of $N_{B_s^{(*)}\bar{B}_s^{(*)}}$ pairs is $(7.11 \pm 1.30) \times 10^6$ for the 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ data sample. The uncertainty on the number of pairs is used as the systematic uncertainty. It is $(1.30/7.11)100 \% \approx 18.3\%$. This is applied only to the final branching fraction average central value, as it is completely correlated between the $B_s^0 \to \eta' K^+ K^- + n\pi$ and $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes.

- PYTHIA Fragmentation of $X_{s\bar{s}}$

Due to a lack of statistics in this study, performing a data-driven evaluation as in [57] and [58] is not feasible. Performing a systematic uncertainty analysis based on published theoretical results relating to semi-inclusive B_s^0 decays is also not feasible as there is a paucity of results in this sector. Instead, sets of JETSET parameters JetSetPar PARJ(1, 2, 3, 4, 11, 12, 13, 25, and 26) are varied to increase and decrease the reconstruction efficiency in $X_{s\bar{s}}$ mass bins. The varied parameter sets are called the "alternative tunings" (AT). These varied sets differ from the Belle standard set. Descriptions of these parameters and their values are given in Tables 4.42 and 4.43,

Parameter	Description
PARJ(1)	baryon suppression
PARJ(2)	s vs u, d quark suppression
PARJ(3)	s quark further suppression
PARJ(4)	spin-1 diquark suppression vs spin-0 diquarks
PARJ(11)	probability of spin-1 light mesons
PARJ(12)	probability of spin-1 strange meson
PARJ(13)	probability of spin-1 meson with c or heavier quark
PARJ(25)	η suppression factor
PARJ(26)	η' suppression factor

Table 4.42.: JETSET parameter descriptions

respectively. The models described in the appendix of [57] are used as a starting reference.

The difference between unity and the ratio of the re-tuned reconstruction efficiency and nominal efficiency is used as the systematic uncertainty for the PYTHIA fragmentation model. These are given in Tables 4.44 and 4.45. This uncertainty can be positive if the reconstruction efficiency of the alternative tuning is smaller than that of the standard tuning, and negative if the reconstruction efficiency due to the alternative tuning is larger than that of the standard PYTHIA tuning. A negative error therefore does not indicate a negative efficiency. At no point is any efficiency negative, either in the standard or alternative tuning. This method includes the effect of changing the proportion of unreconstructed modes[2].

If a tuning tends to decrease (increase) reconstruction efficiency, but if a bin nevertheless has an increase (decrease), then in those bins an uncertainty of 0.0 is assigned.

This uncertainty is alternatively called the fragmentation model (FM) uncertainty.

- Sub-mode branching fraction uncertainties

Branching fractions and their uncertainties from [37] are found in Table 5.1.

---

[2]In the appendices, a study of the proportion of unreconstructed modes in the standard signal MC is given.

Parameter	Standard	Buckley et al.	Nishimura	AT1	AT2
PARJ(1)	0.1	0.073	0.073	0.2	0.1
PARJ(2)	0.3	0.2	1	0.2	0.4
PARJ(3)	0.4	0.94	0.94	0.4	0.4
PARJ(4)	0.05	0.032	0.032	0.264	0.008
PARJ(11)	0.5	0.31	0.01	0.9	0.1
PARJ(12)	0.6	0.4	0.01	0.6	0.6
PARJ(13)	0.75	0.54	0.54	0.75	0.75
PARJ(25)	1	0.63	1	0.1	1
PARJ(26)	0.4	0.12	0.12	0.4	0.12

Table 4.43.: JETSET parameters used to tune the fragmentation of the $X_{s\bar{s}}$ system in PYTHIA. AT1 and AT2 are used to tune PYTHIA to obtain the systematic uncertainties due to fragmentation.

Bin Number	Bin (GeV/c^2)	Standard	AT1	AT2	Syst.
1	0.8-1.0	5.59 ± 0.41	4.91 ± 0.4	5.03 ± 0.4	$^{+10.0}_{-0.0}$
2	1.0-1.2	3.76 ± 0.09	3.99 ± 0.09	3.75 ± 0.09	$^{+0.4}_{-5.9}$
3	1.2-1.4	2.96 ± 0.08	3.04 ± 0.08	2.77 ± 0.08	$^{+6.4}_{-2.8}$
4	1.4-1.6	0.96 ± 0.05	1.04 ± 0.05	0.89 ± 0.04	$^{+8.0}_{-8.3}$
5	1.6-1.8	0.58 ± 0.04	0.78 ± 0.04	0.49 ± 0.03	$^{+14.7}_{-35.3}$
6	1.8-2.0	0.36 ± 0.03	0.48 ± 0.03	0.29 ± 0.03	$^{+21.1}_{-33.6}$
7	2.0-2.2	0.24 ± 0.02	0.32 ± 0.03	0.17 ± 0.02	$^{+28.7}_{-37.4}$
8	2.2-2.4	0.15 ± 0.02	0.32 ± 0.02	0.11 ± 0.02	$^{+23.7}_{-58.2}$

Table 4.44.: Comparison of reconstruction efficiencies and their associated relative systematic uncertainties (%) between PYTHIA tunings (Standard, AT1, and AT2) given in Table 4.43, used in systematic uncertainty estimation; tuning is done in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes.

Bin Number	Bin (GeV/c^2)	Standard	AT1	AT2	Syst.
1	0.8-1.0	0.0	0.0	0.0	$^{+0.0}_{-0.0}$
2	1.0-1.2	0.02 ± 0.001	0.001 ± 0.004	0.01 ± 0.006	$^{+23.7}_{-0.0}$
3	1.2-1.4	0.2 ± 0.02	0.2 ± 0.03	0.2 ± 0.02	$^{+18.3}_{-2.3}$
4	1.4-1.6	0.9 ± 0.05	0.8 ± 0.04	0.8 ± 0.05	$^{+6.6}_{-0.0}$
5	1.6-1.8	0.7 ± 0.04	0.8 ± 0.04	0.6 ± 0.04	$^{+12.5}_{-10.5}$
6	1.8-2.0	0.5 ± 0.04	0.5 ± 0.04	0.4 ± 0.03	$^{+20.2}_{-14.4}$
7	2.0-2.2	0.4 ± 0.03	0.5 ± 0.04	0.3 ± 0.03	$^{+30.7}_{-23.2}$
8	2.2-2.4	0.2 ± 0.03	0.3 ± 0.03	0.2 ± 0.03	$^{+0.0}_{-74.5}$

Table 4.45.: Comparison of reconstruction efficiencies and their associated relative systematic uncertainties (%) between PYTHIA tunings (Standard, AT1, and AT2) given in Table 4.43, used in systematic uncertainty estimation; tuning is done in 0.2 GeV/c^2 $X_{s\bar{s}}$ mass bins for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes.

Mode	Branching Fraction(%)
$\eta \rightarrow \gamma\gamma$	39.41 ± 0.2
$\eta' \rightarrow \pi^+\pi^-\eta$	42.6 ± 0.7

Table 4.46.: η, η' branching fractions from Ref. [37]

$\delta\mathcal{B}(\eta' \rightarrow \eta\pi\pi) = 0.7$

$\delta\mathcal{B}(\eta \rightarrow \gamma\gamma) = 0.20$

- Cross-feed

The cross-feed values in Tables 4.8 and 4.9 are used to estimate this uncertainty.

4.12.2. Additive Systematics Uncertainties

- PDF Parameters - Signal and Background

Mass Bin (GeV/c^2)	$\mathcal{B}(\eta' \to \eta\pi\pi)$	$\mathcal{B}(\eta \to \gamma\gamma)$	Track	K CF	π CF
0.8 - 1.0	0.7	0.2	0.7	1.90	2.56
1.0 - 1.2	0.7	0.2	0.7	1.90	2.56
1.2 - 1.4	0.7	0.2	0.7	1.90	2.56
1.4 - 1.6	0.7	0.2	0.7	1.90	2.56
1.6 - 1.8	0.7	0.2	0.7	1.90	2.56
1.8 - 2.0	0.7	0.2	0.7	1.90	2.56
2.0 - 2.2	0.7	0.2	0.7	1.90	2.56
2.2 - 2.4	0.7	0.2	0.7	1.90	2.56

Table 4.47.: Relative systematics summary table I; all values are in percent (%); for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

Mass Bin (GeV/c^2)	$\mathcal{B}(\eta' \to \eta\pi\pi)$	$\mathcal{B}(\eta \to \gamma\gamma)$	Track	K CF	π CF
0.8 - 1.0	0.7	0.2	1.05	0.0	0.0
1.0 - 1.2	0.7	0.2	1.05	0.95	3.84
1.2 - 1.4	0.7	0.2	1.05	0.95	3.84
1.4 - 1.6	0.7	0.2	1.05	0.95	3.84
1.6 - 1.8	0.7	0.2	1.05	0.95	3.84
1.8 - 2.0	0.7	0.2	1.05	0.95	3.84
2.0 - 2.2	0.7	0.2	1.05	0.95	3.84
2.2 - 2.4	0.7	0.2	1.05	0.95	3.84

Table 4.48.: Relative systematics summary table I; all values are in percent (%); for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes

Mass Bin (GeV/c^2)	nnout	FM	K_S^0	η	π^0
0.8 - 1.0	8.1	$+10.0$ -0.0	0.0	3.0	0.0
1.0 - 1.2	8.1	$+0.4$ -5.9	0.0	3.0	0.003
1.2 - 1.4	8.1	$+6.4$ -2.8	0.0	3.0	0.07
1.4 - 1.6	8.1	$+8.0$ -8.3	0.0	3.0	0.65
1.6 - 1.8	8.1	$+15.0$ -35.0	0.0	3.0	0.76
1.8 - 2.0	8.1	$+21.0$ -34.0	0.0	3.0	0.88
2.0 - 2.2	8.1	$+29.0$ -37.0	0.0	3.0	1.02
2.2 - 2.4	8.1	$+24.0$ -58.0	0.0	3.0	1.06

Table 4.49.: Relative systematics summary table II; all values are in percent (%); for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

Mass Bin (GeV/c^2)	nnout	FM	K_S^0	η	π^0
0.8 - 1.0	4.8	$^{+0.0}_{-0.0}$	1.57	3.0	0.0
1.0 - 1.2	4.8	$^{+23.7}_{-0.0}$	1.57	3.0	0.0
1.2 - 1.4	4.8	$^{+18.3}_{-2.3}$	1.57	3.0	0.0
1.4 - 1.6	4.8	$^{+6.6}_{-11.6}$	1.57	3.0	0.02
1.6 - 1.8	4.8	$^{+12.5}_{-10.5}$	1.57	3.0	0.18
1.8 - 2.0	4.8	$^{+20.2}_{-14.4}$	1.57	3.0	0.34
2.0 - 2.2	4.8	$^{+30.7}_{-23.2}$	1.57	3.0	0.32
2.2 - 2.4	4.8	$^{+0.0}_{-74.5}$	1.57	3.0	0.39

Table 4.50.: Relative systematics summary table II; all values are in percent (%); for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes

Mass Bin (GeV/c^2)	BCS	xfeed	PM
0.8 - 1.0	0.8	0.03	0.2
1.0 - 1.2	0.8	0.004	0.2
1.2 - 1.4	0.8	0.03	0.2
1.4 - 1.6	0.8	0.002	0.2
1.6 - 1.8	0.8	0.02	0.2
1.8 - 2.0	0.8	0.002	0.2
2.0 - 2.2	0.8	0.002	0.2
2.2 - 2.4	0.8	0.008	0.2

Table 4.51.: Relative systematics summary table III; all values are in percent (%);for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$

Mass Bin (GeV/c^2)	BCS	xfeed	PM
0.8 - 1.0	5.2	0.0	1.1
1.0 - 1.2	5.2	0.002	1.1
1.2 - 1.4	5.2	0.007	1.1
1.4 - 1.6	5.2	0.04	1.1
1.6 - 1.8	5.2	0.03	1.1
1.8 - 2.0	5.2	0.04	1.1
2.0 - 2.2	5.2	0.02	1.1
2.2 - 2.4	5.2	0.01	1.1

Table 4.52.: Relative systematics summary table III; all values are in percent (%); for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$

The fit parameters for the final fit to data have a statistical uncertainty associated with them, as determined from the fit results from the $B_s^0 \to D_s^- \rho^+$ control. The parameters for data fitting for the signal Gaussian in Table 4.31 are used to determine a systematic uncertainty. When fitting to data for the signal yield, the parameters are allowed to float within their 1σ errors. The absolute difference in fitted signal yield from the two types of fits is used as an additive systematic uncertainty.

For the background ARGUS function, the ARGUS parameters in each $X_{s\bar{s}}$ mass bin are allowed to float between their 1σ errors (Tables 4.6 and 4.7) that are determined from $\Upsilon(5S)$ data NN sidebands. The absolute difference in fitted signal yield is used as an additive systematic uncertainty.

- Fit Bias

The fit bias is an additive systematic uncertainty and is determined by toy MC ensemble studies, using the RooMCStudy package [89]. Linearity tests are performed using various assumptions for the branching fraction, allowing for Poisson fluctuations. There are 5000 toy MCs produced. These toys are fit with the fit model described earlier, with floating yields; floating signal yields are allowed to fluctuate negatively. A plot of fitted signal events versus generated events from the toy MCs is obtained. This plot is fit with a first order polynomial of the form $f(x) = p0 + p1x$. If there is little or no fit bias then the slope of the line, $p1$, should be unity, and $p0$ should be zero. This test is performed over all $X_{s\bar{s}}$ mass bins. The signal yield from fits to toy MC is biased by 0.41 events for the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes and 0.0004 events for the $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes. The results of this study are shown in the linearity test plots in Fig. 4.56.

Inclusion of Additive Systematics in the Final Results

All additive systematic uncertainties are included in the final results by adding them in quadrature with both the high and low asymmetric errors from fits to data for the signal yield.
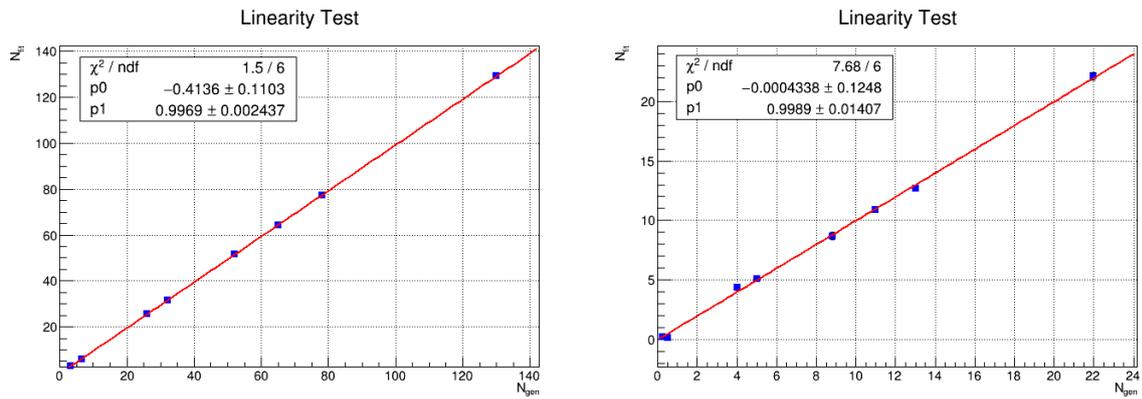
(a) Linearity test for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes   **(b)** Linearity test for $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes

Figure 4.56.: Linearity tests for each different signal class. Each data point corresponds to a different assumption on the branching fraction for $B_s^0 \to \eta' X_{s\bar{s}}$. The vertical axis is the fitted signal yield and the horizontal axis is the number of generated signal events.

Chapter 5.

Results

5.1. Final Signal Extraction

Results for the BF for $B_s^0 \to \eta'(\to \eta\pi^+\pi^-)X_{s\bar{s}}$ are extracted by fitting the full 121.4 integrated luminosity $Y(5S)$ data sample in bins of $X_{s\bar{s}}$ mass. These fit results are then used to calculate a BF in each $X_{s\bar{s}}$ mass bin using equation 5.1, in the mass range $M(X_{s\bar{s}}) \leq 2.4\,\text{GeV}/c^2$. The BF is calculated as

$$\mathcal{B}(B_s^0 \to \eta'(\to \eta\pi^+\pi^-)X_{s\bar{s}})_j = \frac{N_{\text{sig}}^j}{2 \times N_{B_s^{0(*)}\bar{B}_s^{0(*)}}\kappa_j\epsilon_{\text{MC}}^j[\mathcal{B}(\eta \to \gamma\gamma)\mathcal{B}(\eta' \to \pi^+\pi^-\eta)]} \tag{5.1}$$

where j denotes the mass bins of $X_{s\bar{s}}$, and κ_j are the bin-by-bin efficiency correction factors. The number of $B_s^0\bar{B}_s^0$ pairs $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ is $(7.11 \pm 1.30) \times 10^6$ [8].

Mode	Branching Fraction(%)
$\eta \to \gamma\gamma$	39.41 $\pm$ 0.20
$\eta' \to \pi^+\pi^-\eta$	42.6 $\pm$ 0.7

Table 5.1.: η, η' Branching Fractions from the PDG [37]

$M(X_{s\bar{s}})$ Bin (j) (GeV/c^2)	N_{sig}	Cumulative $\mathcal{B}(M(X_{s\bar{s}}) \leq j)$ (10^{-4})	$\mathcal{B}_j$ (10^{-4})
0.8 - 1.0	-	-	-
1.0 - 1.2	$0.4^{+2.6}_{-1.9}$	$0.05^{+0.3}_{-0.2}$ (stat.) $\pm$ 0.004 (syst.) $^{+0.0002}_{-0.003}$ (FM)	$0.05^{+0.3}_{-0.2}$ (stat.) $\pm$ 0.004 (syst.) $^{+0.0002}_{-0.003}$ (FM)
1.2 - 1.4	$0.08^{+2.4}_{-1.7}$	$0.06^{+0.5}_{-0.4}$ (stat.) $\pm$ 0.004 (syst.) $^{+0.001}_{-0.003}$ (FM)	$0.01^{+0.4}_{-0.3}$ (stat.) $\pm$ 0.001 (syst.) $^{+0.0008}_{-0.0003}$ (FM)
1.4 - 1.6	$0.7^{+2.5}_{-1.8}$	$0.4^{+1.2}_{-0.9}$ (stat.) $\pm$ 0.03 (syst.) $^{+0.03}_{-0.03}$ (FM)	$0.3^{+1.1}_{-0.8}$ (stat.) $\pm$ 0.03 (syst.) $^{+0.03}_{-0.03}$ (FM)
1.6 - 1.8	$0.4^{+2.1}_{-1.4}$	$0.7^{+2.0}_{-1.4}$ (stat.) $\pm$ 0.03 (syst.) $^{+0.07}_{-0.1}$ (FM)	$0.3^{+1.6}_{-1.1}$ (stat.) $\pm$ 0.02 (syst.) $^{+0.04}_{-0.1}$ (FM)
1.8 - 2.0	$1.4^{+2.6}_{-2.0}$	$2.4^{+3.8}_{-2.8}$ (stat.) $\pm$ 0.1 (syst.) $^{+0.4}_{-0.7}$ (FM)	$1.7^{+3.3}_{-2.5}$ (stat.) $\pm$ 0.1 (syst.) $^{+0.4}_{-0.6}$ (FM)
2.0 - 2.2	$0.3^{+3.7}_{-3.4}$	$3.0^{+8.1}_{-7.0}$ (stat.) $\pm$ 0.1 (syst.) $^{+0.6}_{-0.9}$ (FM)	$0.6^{+7.1}_{-6.4}$ (stat.) $\pm$ 0.05 (syst.) $^{+0.2}_{-0.2}$ (FM)
2.2 - 2.4	$-2.3^{+3.8}_{-3.4}$	$-4.0^{+14.1}_{-12.6}$ (stat.) $\pm$ 0.6 (syst.) $^{+2.2}_{-5.0}$ (FM)	$-7.0^{+11.6}_{-10.4}$ (stat.) $\pm$ 0.6 (syst.) $^{+1.6}_{-4.1}$ (FM)

Table 5.2.: Results from the 121.4 fb^{-1} $\Upsilon(5S)$ data sample for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes; rows with dashes indicate bins where no events, background or signal, are found; the asymmetric statistical uncertainties include the additive systematic uncertainties by addition in quadrature; it and its errors are calculated from the fit to data, given in the N_{sig} column; "FM" means "Fragmentation Model" uncertainty, the uncertainty due to PYTHIA fragmentation; the "syst." uncertainty is the sum in quadrature of all (relative) systematic uncertainties, not including the PYTHIA and $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ uncertainties; the "$\mathcal{B}_j$" column is the BF central value in each bin and $\mathcal{B}(M(X_{s\bar{s}}) \leq j)$ is the cumulative branching fraction up to that bin, with uncertainties added in quadrature

$M(X_{s\bar{s}})$ Bin (GeV/c^2)	N_{bkg}
0.8 - 1.0	1.0 ± 1.0
1.0 - 1.2	29.6 ± 5.8
1.2 - 1.4	29.9 ± 5.8
1.4 - 1.6	21.3 ± 4.9
1.6 - 1.8	13.7 ± 3.9
1.8 - 2.0	19.7 ± 4.7
2.0 - 2.2	21.7 ± 4.9
2.2 - 2.4	23.3 ± 5.4

Table 5.3.: Background yields for the 121.4 fb^{-1} $\Upsilon(5S)$ data sample; $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

5.1.1. $B_s^0 \to \eta' K^+ K^- + n\pi$

Table 5.2 gives the results for $B_s^0 \to \eta' K^+ K^- + n\pi$ modes. Table 5.3 gives the non-peaking background yield from data. Figures 5.1, 5.2, and 5.3 show the fits to data in $X_{s\bar{s}}$ mass bins.
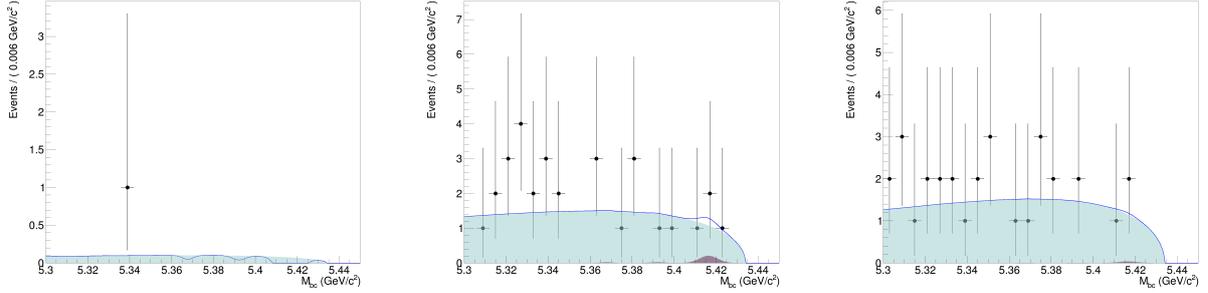
(a) Bin 1: $0.8 \leq M(X_{s\bar{s}}) \leq 1.0$ GeV/c^2
(b) Bin 2: $1.0 \leq M(X_{s\bar{s}}) \leq 1.2$ GeV/c^2
(c) Bin 3: $1.2 \leq M(X_{s\bar{s}}) \leq 1.4$ GeV/c^2

Figure 5.1.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^+ K^- + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.
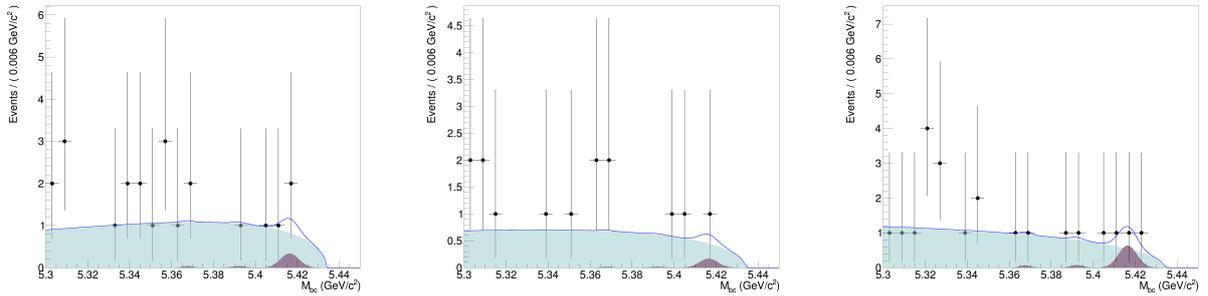
(a) Bin 4: $1.4 \leq M(X_{s\bar{s}}) \leq 1.6$ GeV/c^2
(b) Bin 5: $1.6 \leq M(X_{s\bar{s}}) \leq 1.8$ GeV/c^2
(c) Bin 6: $1.8 \leq M(X_{s\bar{s}}) \leq 2.0$ GeV/c^2

Figure 5.2.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^+ K^- + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.

(a) Bin 7: $2.0 \leq M(X_{s\bar{s}}) \leq 2.2$ GeV/c^2

(b) Bin 8: $2.2 \leq M(X_{s\bar{s}}) \leq 2.4$ GeV/c^2
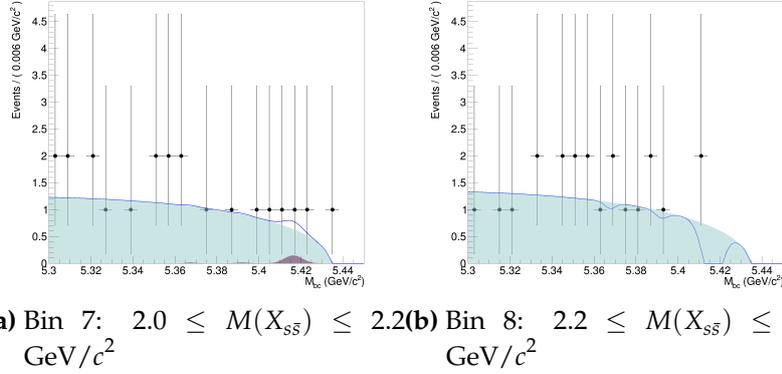
Figure 5.3.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^+ K^- + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.

$B_s^0 \to \eta' K^+ K^- + n\pi$ Combined Result

By summing all the bins, the combined result for the central value for the BF for the $B_s^0 \to \eta' K^+ K^- + n\pi$ modes is determined to be approximately [-4.0 $^{+14.1}_{-12.6}$ (stat.) $\pm$ 0.6 (syst.) $^{+2.2}_{-5.0}$ (FM)] $\times 10^{-4}$.

Mass Bin (j) (GeV/c^2)	N_{sig}	Cumulative $\mathcal{B}(\,M(X_{s\bar{s}}) \leq j)\ (10^{-4})$	$\mathcal{B}_j\ (10^{-4})$
0.8 - 1.0	-	-	-
1.0 - 1.2	-	-	-
1.2 - 1.4	$0.3^{+1.4}_{-0.8}$	$0.5^{+2.5}_{-1.5}$ (stat.) $\pm$ 0.04 (syst.)$^{+0.09}_{-0.01}$ (FM)	$0.5^{+2.5}_{-1.5}$ (stat.) $\pm$ 0.04 (syst.) $^{+0.09}_{-0.01}$ (FM)
1.4 - 1.6	$2.0^{+3.0}_{-2.2}$	$1.4^{+2.8}_{-1.8}$ (stat.) $\pm$ 0.08 (syst.)$^{+0.2}_{-0.01}$ (FM)	$1.0^{+1.4}_{-1.1}$ (stat.) $\pm$ 0.07 (syst.) $^{+0.06}_{-0.0}$ (FM)
1.6 - 1.8	$1.2^{+3.3}_{-2.6}$	$2.2^{+3.5}_{-2.4}$ (stat.) $\pm$ 0.1 (syst.)$^{+0.2}_{-0.09}$ (FM)	$0.8^{+2.1}_{-1.6}$ (stat.) $\pm$ 0.06 (syst.) $^{+0.1}_{-0.08}$ (FM)
1.8 - 2.0	$4.8^{+4.2}_{-3.4}$	$6.6^{+5.3}_{-4.0}$ (stat.) $\pm$ 0.3 (syst.)$^{+1.1}_{-0.7}$ (FM)	$4.4^{+3.9}_{-3.1}$ (stat.) $\pm$ 0.3 (syst.) $^{+0.9}_{-0.6}$ (FM)
2.0 - 2.2	$-2.4^{+3.9}_{-3.2}$	$3.8^{+7.0}_{-5.5}$ (stat.) $\pm$ 0.4 (syst.)$^{+2.0}_{-1.4}$ (FM)	$-2.8^{+4.6}_{-3.8}$ (stat.) $\pm$ 0.2 (syst.) $^{+0.8}_{-0.6}$ (FM)
2.2 - 2.4	$-1.1^{+3.6}_{-2.9}$	$1.2^{+11.3}_{-9.0}$ (stat.) $\pm$ 0.4 (syst.)$^{+2.0}_{-3.3}$ (FM)	$-2.6^{+8.9}_{-7.1}$ (stat.) $\pm$ 0.2 (syst.) $^{+0.0}_{-1.9}$ (FM)

Table 5.4.: Results from the 121.4 fb^{-1} $\Upsilon(5S)$ data sample for $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes; rows with dashes indicate bins where no events, background or signal, are found; the asymmetric statistical uncertainties include the absolute additive systematic uncertainties by addition in quadrature; it and its errors are calculated from the fit to data, given in the N_{sig} column; the "FM" means "Fragmentation Model" uncertainty, the uncertainty due to PYTHIA fragmentation; the "syst." uncertainty is the sum in quadrature of all (relative) systematic uncertainties, not including the PYTHIA and $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ uncertainties; the "$\mathcal{B}_j$" column is the BF central value in each bin and $\mathcal{B}(M(X_{s\bar{s}}) \leq j)$ is the cumulative branching fraction up to that point, with uncertainties added in quadrature

$M(X_{s\bar{s}})$ Bin (GeV/c^2)	N_{bkg}
0.8 - 1.0	-
1.0 - 1.2	-
1.2 - 1.4	6.7 $\pm$ 2.8
1.4 - 1.6	29.0 $\pm$ 5.8
1.6 - 1.8	42.8 $\pm$ 7.0
1.8 - 2.0	52.2 $\pm$ 7.8
2.0 - 2.2	102.0 $\pm$ 10.8
2.2 - 2.4	152.0 $\pm$ 12.8

Table 5.5.: Background yields for the 121.4 fb^{-1} $\Upsilon(5S)$ data sample; $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes. The points are data and the shaded region is the fit.

5.1.2. $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$

Table 5.4 gives the results for $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes. Table 5.5 gives the non-peaking background yield from data. Figures 5.4, 5.5, and 5.6 show the fits to data in $X_{s\bar{s}}$ mass bins.
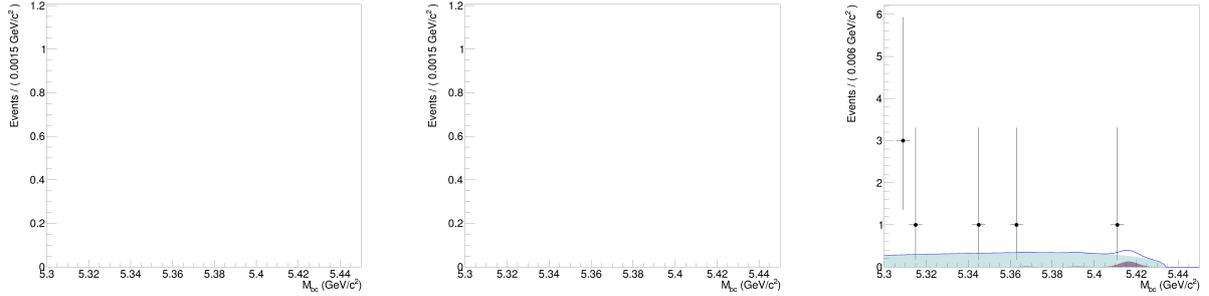
(a) Bin 1: $0.8 \leq M(X_{s\bar{s}}) \leq 1.0$ GeV/c^2 **(b)** Bin 2: $1.0 \leq M(X_{s\bar{s}}) \leq 1.2$ GeV/c^2 **(c)** Bin 3: $1.2 \leq M(X_{s\bar{s}}) \leq 1.4$ GeV/c^2

Figure 5.4.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.
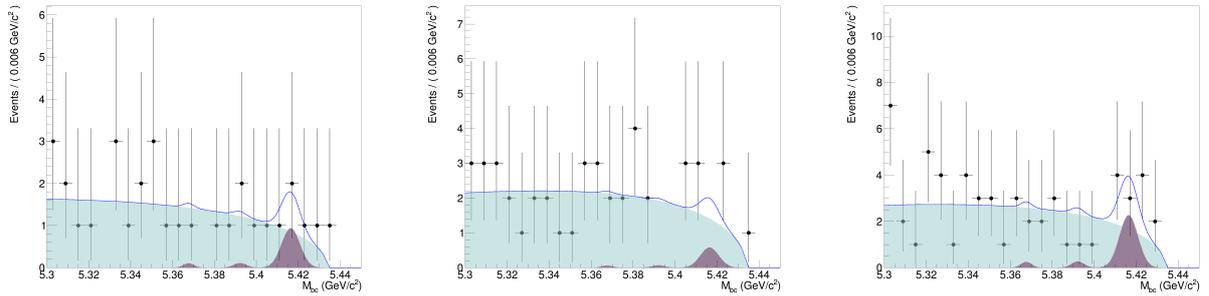
(a) Bin 4: $1.4 \leq M(X_{s\bar{s}}) \leq 1.6$ GeV/c^2 **(b)** Bin 5: $1.6 \leq M(X_{s\bar{s}}) \leq 1.8$ GeV/c^2 **(c)** Bin 6: $1.8 \leq M(X_{s\bar{s}}) \leq 2.0$ GeV/c^2

Figure 5.5.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.

(a) Bin 7: $2.0 \leq M(X_{s\bar{s}}) \leq 2.2$ GeV/c^2

(b) Bin 8: $2.2 \leq M(X_{s\bar{s}}) \leq 2.4$ GeV/c^2
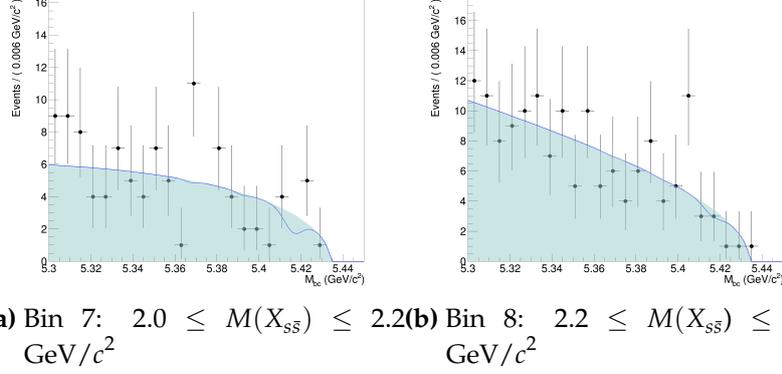
Figure 5.6.: Fits to the $\Upsilon(5S)$ data for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$. The points are data, the shaded regions are the fit components (ARGUS: light shaded region, Gaussian sum: dark shaded region), the solid blue curve is the full fit.

$B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ Combined Result

By summing all the bins, the combined result for the central value for the BF for the $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes is determined to be approximately: $[1.2 \, ^{+11.3}_{-9.0} \, (\text{stat.}) \pm 0.4 \, (\text{syst.}) \, ^{+2.0}_{-3.3} \, (\text{FM})] \times 10^{-4}$.

Initial Conclusion

Using the profile likelihood ratio statistic [37], it is determined that no statistically significant signal excess is observed or evidenced. This was done using the RooStats package [89]. As such, a 90% confidence level upper limit is evaluated.

5.1.3. Combined Results

Weighting Method

The final combined result for the upper limit at 90% confidence is obtained by the following method.

1. Add together all branching fraction central values from each $X_{s\bar{s}}$ mass bin, linearly.

2. Add the absolute additive systematic uncertainties in quadrature with both the high and low asymmetric uncertainties obtained from the signal yield fit results. For brevity, "statistical uncertainties" will now imply that the additive systematic uncertainties are included with the fit uncertainties through addition in quadrature with the asymmetric fit uncertainties.

3. Add together all high statistical uncertainties from each bin in quadrature; add together all low statistical errors from each bin in quadrature.

4. Add the other uncertainties from each bin together in quadrature; FM systematics are added linearly in a class; FM uncertainties of the combined results of the two classes are added in quadrature for the final weighted average BF central value.

5. There should now be a BF central value with (asymmetric) statistical uncertainties and symmetric systematic uncertainties, one each for the classes. See Table 5.6.

6. Symmetrize the statistical uncertainties for each class by taking their average and define it as σ, e.g. $[(18.4 + 13.2) \times 10^{-4}]/2 \equiv \sigma$.

7. Take the weighted average of the two classes of BFs, using $\mu' = \sum_j \frac{\mu_j w_j}{\sum_j w_j}$ where w_j are the weights of each class; the weight associated with each class is $w_j = 1/\sigma_j^2$, where j denotes one of the classes.

8. The statistical uncertainty on the final average BF is $\sigma_{\text{stat.}} = 1/\sqrt{\sum_j w_j}$.

9. Combine the systematic errors for the weighted average by addition in quadrature

10. As the $N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$ uncertainty is completely correlated between the two classes, it is included only after the weighted average above is calculated

11. The upper limit at 90% confidence level is calculated by integrating a Gaussian likelihood in the physically allowed region above zero, where the Gaussian width, σ, is the estimated by adding all positive uncertainties in quadrature.

Class	$\mathcal{B}$ (10^{-4})
$B_s^0 \to \eta' K^+ K^- + n\pi$	-4.0 $^{+14.1}_{-12.6}$ (stat.) $\pm$ 0.6 (syst.) $^{+2.2}_{-5.0}$ (FM)
$B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$	1.2 $^{+11.3}_{-9.0}$ (stat.) $\pm$ 0.4 (syst.) $^{+2.0}_{-3.3}$ (FM)
Weighted Average	-0.7 $\pm$ 8.1 (stat.) $\pm$ 0.7 (syst.) $^{+3.0}_{-6.0}$ (FM) $\pm$ 0.1 ($N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$)
$\mathcal{B}_{UL}^{90\%}$	13.8

Table 5.6.: Combined results for the 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ dataset.

$\leq$ **Mass Bin (GeV/c^2)**	**Cumulative $\mathcal{B}_{UL}^{90\%}$ (10^{-4})**
1.0	-
1.2	0.4
1.4	0.7
1.6	1.9
1.8	3.1
2.0	7.6
2.2	11.1
2.4	13.8

Table 5.7.: $\mathcal{B}_{UL}^{90\%} \leq M(X_{s\bar{s}})$, 90% upper limits. Result in each mass bin is the cumulative result of all bins up to and including that bin.

5.1.4. Subsidiary Result: Upper Limit on $B_s^0 \to \phi(\to K^+K^-)\eta'$

A subsidiary measurement can be made of the currently unobserved $B_s^0 \to \phi(\to K^+K^-)\eta'$ decay. To do this, 1×10^6 events of the exclusive $B_s^0 \to \phi\eta'$ decay were generated and reconstructed in MC.

A $\pm 3\sigma$ range about the reconstructed $\phi(1020)$ mass peak is determined, seen in Fig. 5.8, where σ is the Breit-Wigner distribution width obtained from fitting. This range is used to assign a range for $M(X_{s\bar{s}})$ in which to search for this decay, in the $\Upsilon(5S)$ data. The range determined from MC is $M(X_{s\bar{s}}) \in [1.006, 1.03]$ GeV/c^2. The efficiency correction factor for the corresponding parent 1.0-1.2 GeV/c^2 mass bin is therefore used when calculating the BF.

Events are reconstructed[1] in the same way as B_s^0 candidates in the $B_s^0 \to X_{s\bar{s}}\eta'$ analysis but reconstructing only the two-charged kaon final state with zero pions. The

---

[1]Reconstruction of $B_s^0 \to \phi(\to K^+K^-)\eta'$ and $B_s^0 \to X_{s\bar{s}}\eta'$ MC, and $\Upsilon(5S)$ data are performed with the same BASF module.

efficiency is estimated to be $(7.9 \pm 0.03)\%$, as determined from fitting to M_{bc} in MC, as seen in Fig. 5.9.

The fit to the $\Upsilon(5S)$ data in the given mass range is shown in Fig. 5.10. There is no statistically significant signal peak, with only $1.0^{+1.7}_{-0.9}$ events found. Factoring in $\mathcal{B}(\phi \to K^+K^-) = (49.2 \pm 0.5)\%$ [37], the branching fraction central value is $[11.3^{+18.9}_{-10.3}$ (stat.) $\pm$ 1.0 (syst.) $\pm$ 2.1 $(N_{B_s^{0(*)}\bar{B}_s^{0(*)}})] \times 10^{-6}$ and the corresponding 90% confidence level upper limit is approximately 3.6×10^{-5}. LHCb gives the 90% confidence level upper limit as 8.2×10^{-7} [90].
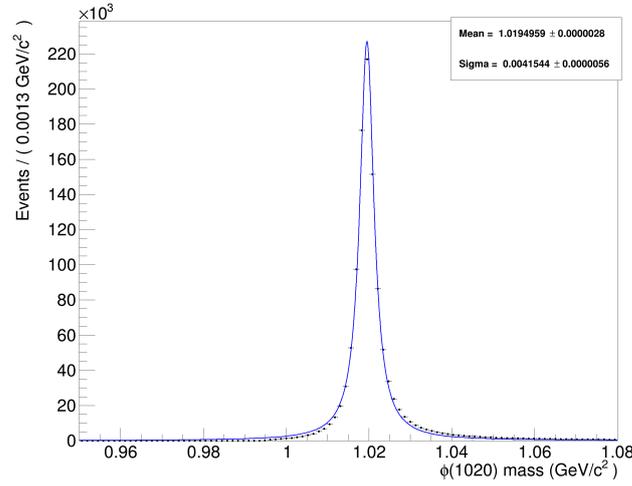
Figure 5.7.: Generated ϕ mass for the exclusive $B_s^0 \to \phi\eta'$ decay in MC. Black points are MC, the blue curve is the fit.
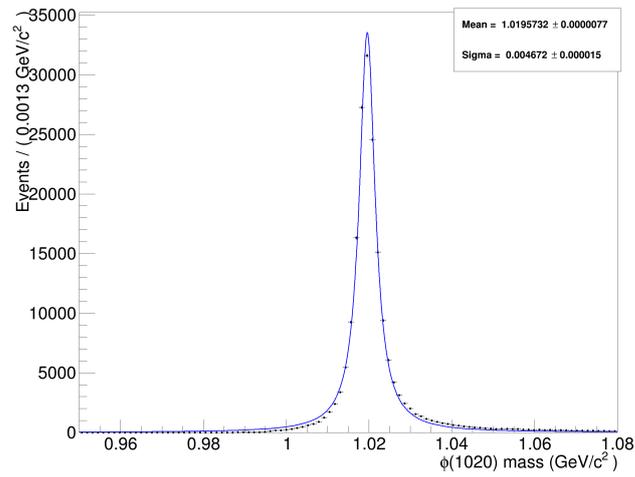
Figure 5.8.: Reconstructed ϕ mass for the exclusive $B_s^0 \rightarrow \phi\eta'$ decay in MC Black points are MC, the blue curve is the fit.
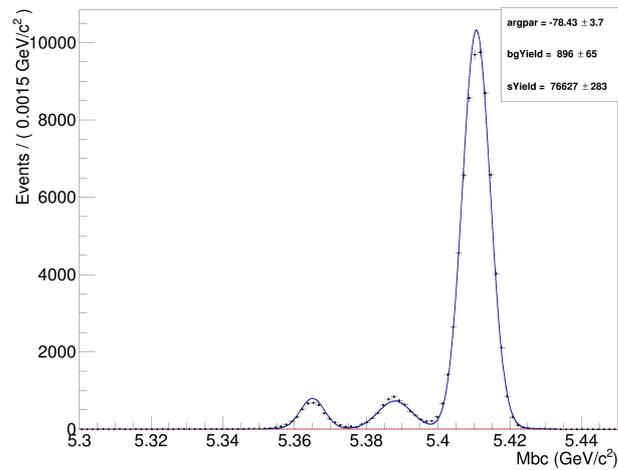
Figure 5.9.: Fit to M_{bc} for the $B_s^0 \rightarrow \phi\eta'$ MC. The points are MC, the solid curve is the fit.
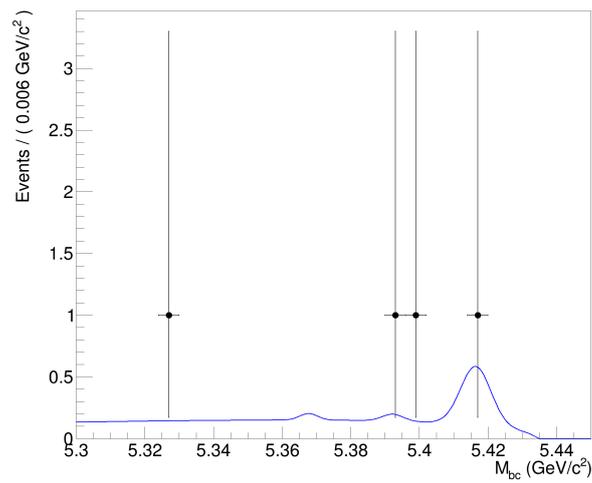
Figure 5.10.: $B_s^0 \to \phi(\to K^+K^-)\eta'$ decay results for the $\Upsilon(5S)$ data. The black points are data and the blue curve is the fit.

5.1.5. Discussion and Conclusion

Using Belle's 121.4 fb^{-1} integrated luminosity $\Upsilon(5S)$ data set, an upper limit of 13.8×10^{-4} is set at 90% confidence for the partial branching fraction of $B_s^0 \to \eta' X_{s\bar{s}}$, in the mass range $M(X_{s\bar{s}}) \leq 2.4\,\text{GeV}/c^2$. The corresponding BF central value is [-0.7 $\pm$ 8.0 (stat.) $\pm$ 0.9 (syst.) $^{+3.0}_{-6.0}$ (FM) $\pm$ 0.1 ($N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$)] $\times 10^{-4}$, where "FM" refers to the PYTHIA fragmentation model.

The BaBar branching fraction measurement for $B \to X_s\eta'$ is [3.9 $\pm$ 0.8 (stat.) $\pm$ 0.5 (syst.) $\pm$ 0.8 (model)] $\times 10^{-4}$. Using this and the BF central value given above for $B_s^0 \to \eta' X_{s\bar{s}}$, a ratio, R($\eta'$), between them can be determined. This allows for the comparison and possible ruling out of some theoretical models. The ratio is determined to be -0.2 ± 2.1(stat.)± 0.2(syst.)$^{+0.8}_{-1.5}$(FM)± 0.03($N_{B_s^{0(*)}\bar{B}_s^{0(*)}}$), and < 3.5 at 90% confidence level. The upper limit is determined using the Bayesian method previously described.

The conservative 90% upper limits for $B_s^0 \to \eta' X_{s\bar{s}}$ and R(η') may be due to one or more things. First, including the high $X_{s\bar{s}}$ mass bins that have low sensitivity, in the final total and average, could be an issue. This is because the methods described earlier, particularly the NN training and cut optimization, were biased toward lower $X_{s\bar{s}}$ mass bins, as they have higher efficiency. A potential future study could possibly compensate for that by optimizing selections using a fine-grained analysis method focusing on optimizations in individual $X_{s\bar{s}}$ mass bins.

Another potential reason for the lack of statistically significant signal instead of, or in addition to, the previously mentioned optimization issue, is a potentially badly broken $SU(3)_F$ symmetry. This may be hinted at in the upper limit for $B_s^0 \to \phi\eta'$ and the measured branching fraction for $B \to K^*\eta'$ of [2.8 $\pm$ 0.6] $\times 10^{-6}$ [37]. The BFs of these two modes should be approximately equivalent by an $SU(3)_F$ symmetry. However, the upper limit on the BF for $B_s^0 \to \phi\eta'$ from LHCb is lower than the BF measured for $B \to K^*\eta'$. This might indicate a broken flavor symmetry.

Whatever the reason, it is hoped that this analysis will spur further experimental and theoretical studies of the B_s^0 meson, particularly with respect to inclusive analyses. A potential future analysis might be to update this analysis with updated (and well maintained and documented) tools and methods, with an additional η' decay mode, to potentially uncover as-yet found issues. It is hoped that Belle II will collect a larger $\Upsilon(5S)$ data sample than Belle and that this study can be repeated using the Belle II data. For now it is hoped that this analysis entices the reader to study the B_s^0 meson.

Appendix A.

Supplemental NeuroBayes Summary Plots
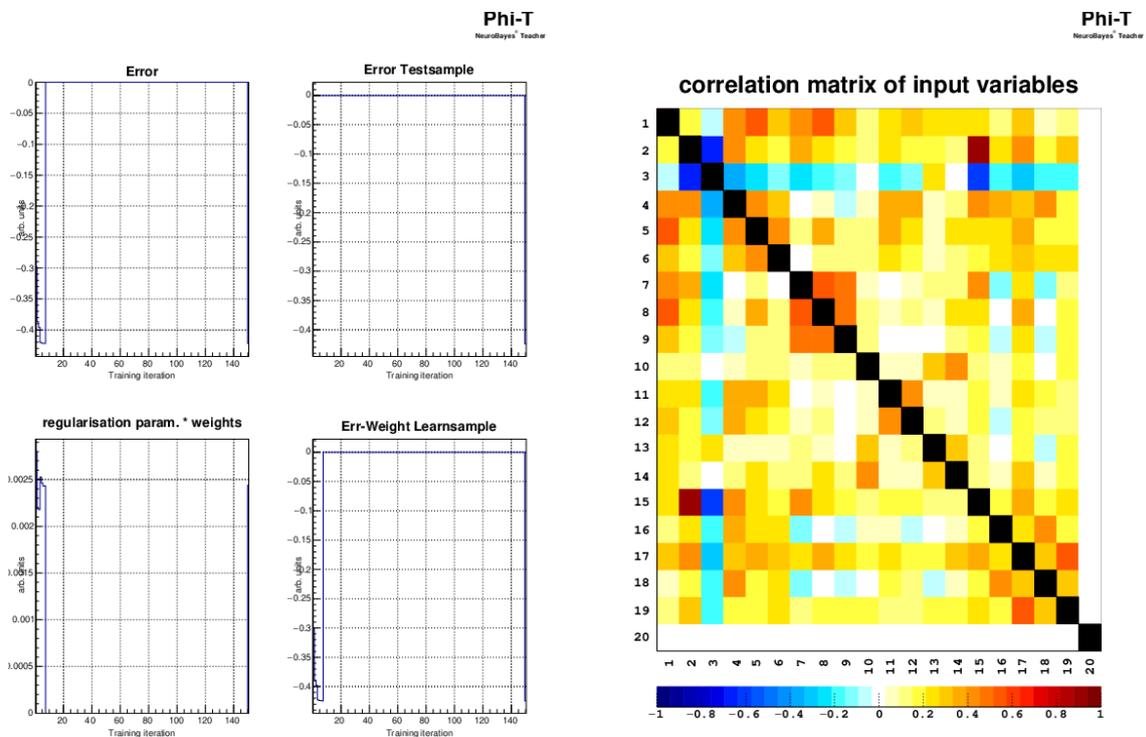
(a) Error per training iteration; error goes to approx-
imately zero after several training iterations

(b) Correlation between input variables; plot indi-
cates low correlation between variables

Figure A.1.: Supplemental NeuroBayes figures

Appendix B.

Additional Checks

B.0.1. Data Run Number Statistics

Run number statistics can also be checked. At the beginning of the BASF reconstruction module (code where all the reconstruction is done), the current run number of the event can be printed out to a log file. The number of events accessed at the beginning of the module can be counted and comparisons can be made to all known run numbers. The number of unique run numbers of the $\Upsilon(5S)$ data can be found internally to Belle using the Belle File Search Engine and specifying the data type as 5S_onresonance and the skim type as HadronB(J) [91].

(a) $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ Modes   **(b)** $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ Modes
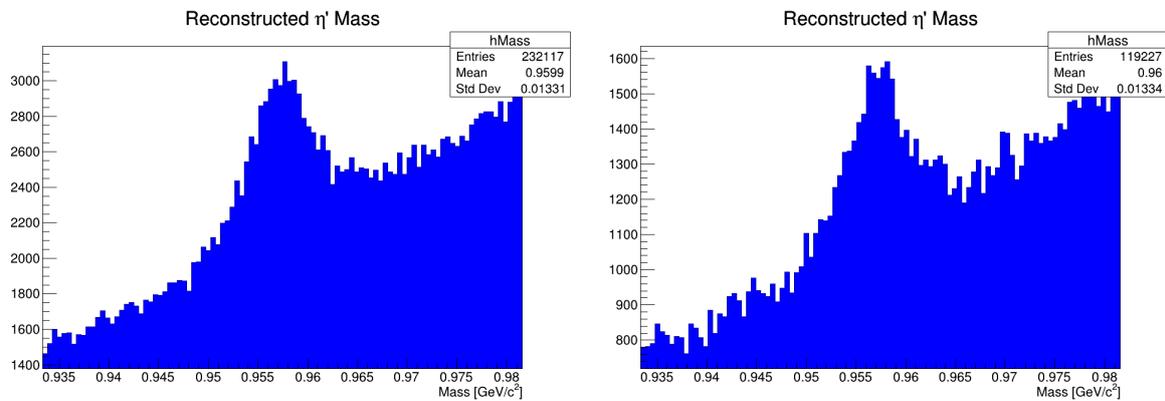
Figure B.1.: Raw η' mass from $\Upsilon(5S)$ data, corresponding to $M(X_{s\bar{s}}) \leq 2.4\,\text{GeV}/c^2$. No signal region cuts are applied.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ Modes

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ Modes
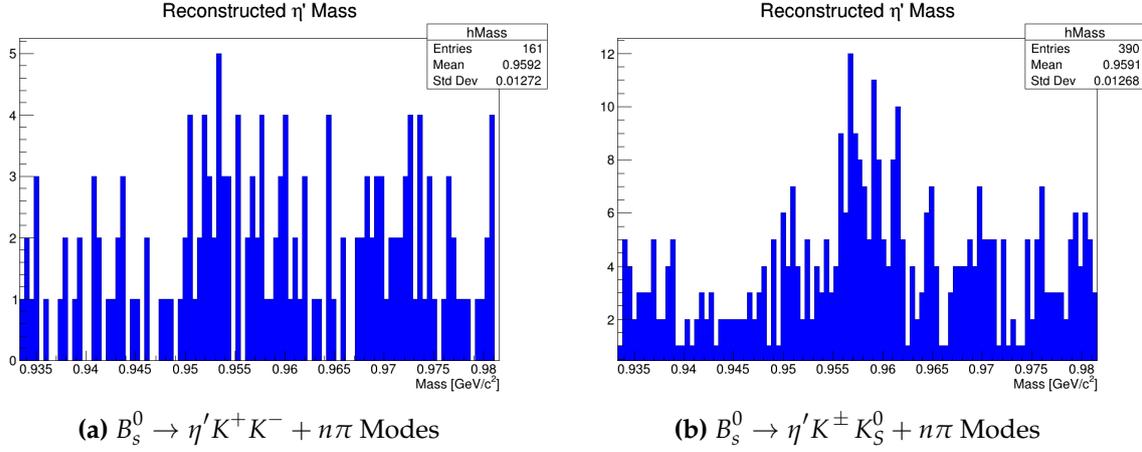
Figure B.2.: Raw η' mass from from data, corresponding to $M(X_{s\bar{s}}) \leq 2.4 \,\mathrm{GeV}/c^2$. All final signal region cuts are applied.

Experiment	Total Number of Events	Skim calls	Analysis Module calls
43	6310259	6310259	6168818
53	76628683	75199284	73500310
67	96517366	96070370	93896334
69	167433009	167291661	163522044
71	80656738	80262406	78438098
All	427546055	425133980	415525604

Table B.1.: Number of events per run from the summary statistics that are printed at the end of every BASF log file, for all of the $\Upsilon(5S)$ data

Figures B.3 and B.4 compare the unique run numbers printed to the log files from the beginning of the BASF analysis module. Figures B.5 and B.6 compare the unique run numbers printed to the log files from the end of the BASF analysis module. Figures B.7 and B.8 compare the unique run numbers printed to the log files from the beginning and end of the BASF analysis module. Figures B.9 and B.10 compare the run numbers of all the events at the beginning and end of the BASF analysis module. All of the preceding figures are produced after the standard skim is applied. Table B.1 compares the total number of events, as determined from the summary statistics at the end of all the log files, to the number of BASF module calls on KEKCC.
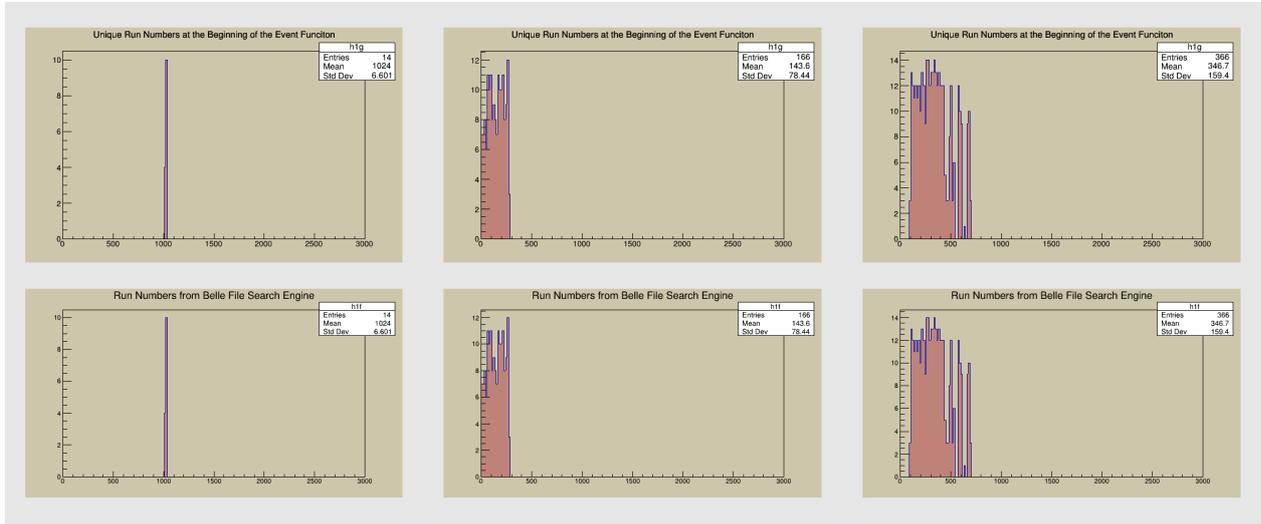
(a) Experiment 43   **(b)** Experiment 53   **(c)** Experiment 67

Figure B.3.: Comparison of the unique run numbers printed to a log file at the beginning of the BASF analysis module (upper figures) and the runs found on the Belle file search engine, for the $\Upsilon(5S)$ on-resonance, caseB, data (lower figures)
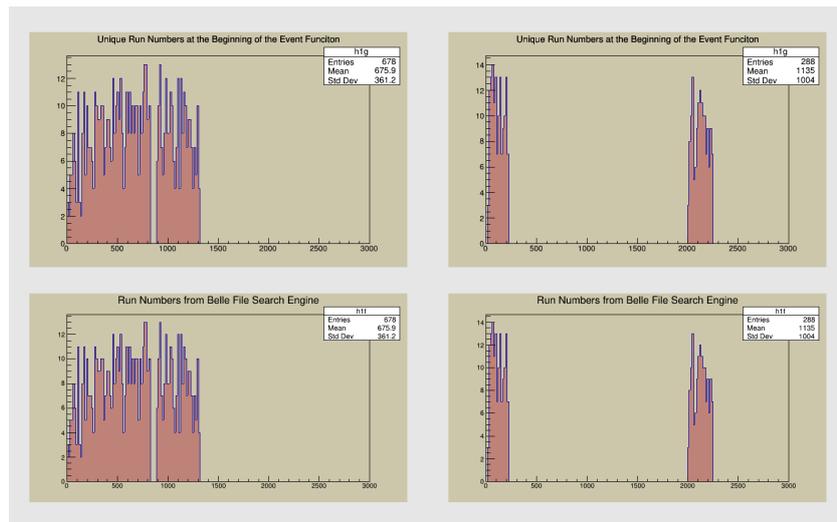
(a) Experiment 69   **(b)** Experiment 71

Figure B.4.: Comparison of the unique run numbers printed to a log file at the beginning of the BASF analysis module (upper figures) and the runs found on the Belle file search engine, for the $\Upsilon(5S)$ on-resonance, caseB, data (lower figures)

(a) Experiment 43   **(b)** Experiment 53   **(c)** Experiment 67

Figure B.5.: Comparison of the unique run numbers printed to a log file at the end of the BASF analysis module (upper figures) and the runs found on the Belle file search engine, for all of the $\Upsilon(5S)$ data (lower figures)

(a) Experiment 69   **(b)** Experiment 71
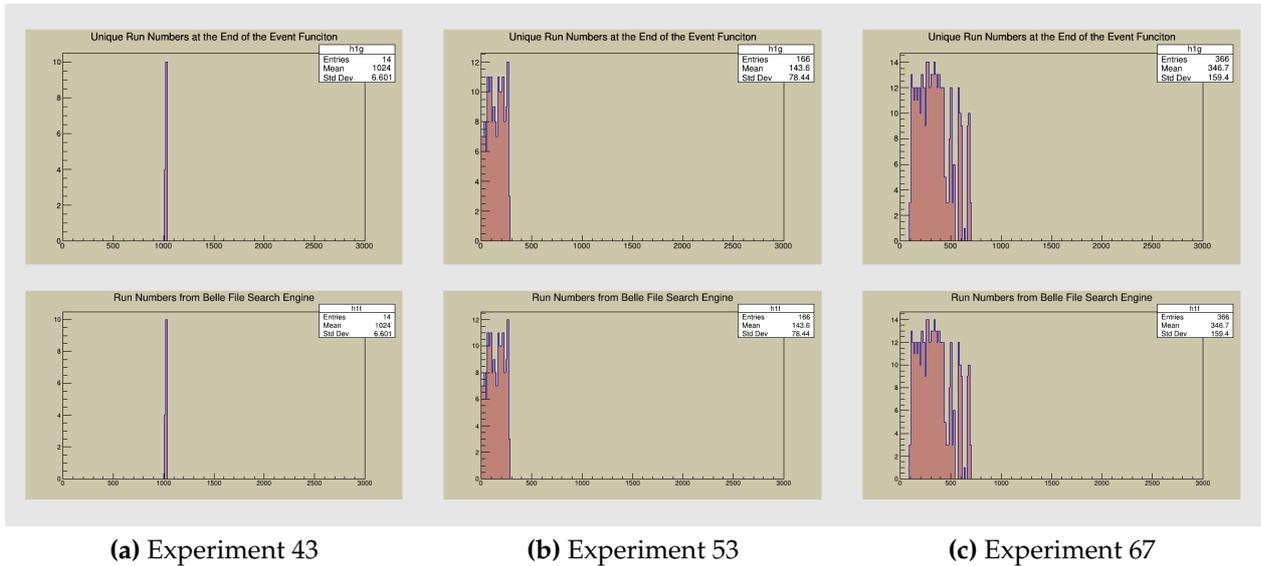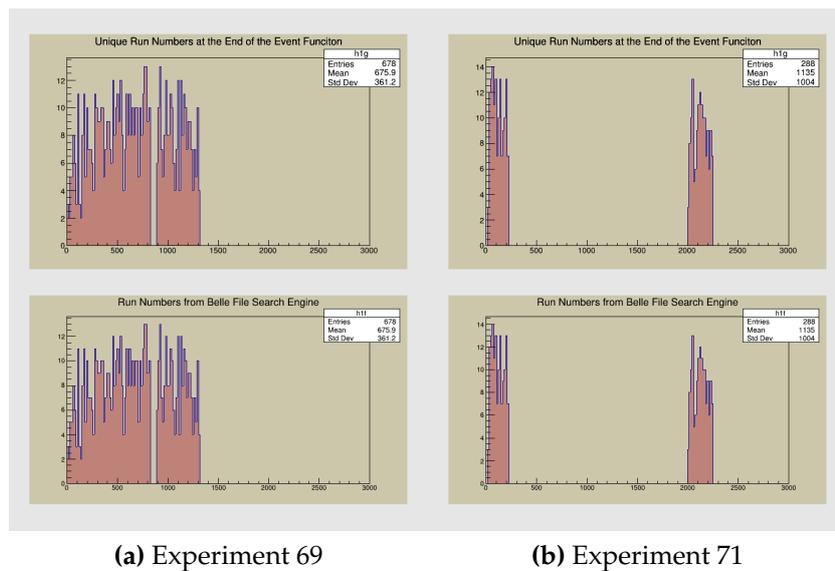
Figure B.6.: Comparison of the unique run numbers and printed to a log file at the end of the BASF analysis module (upper figures) and the runs found on the Belle file search engine, for all of the $\Upsilon(5S)$ data (lower figures)

(a) Experiment 43 **(b)** Experiment 53 **(c)** Experiment 67
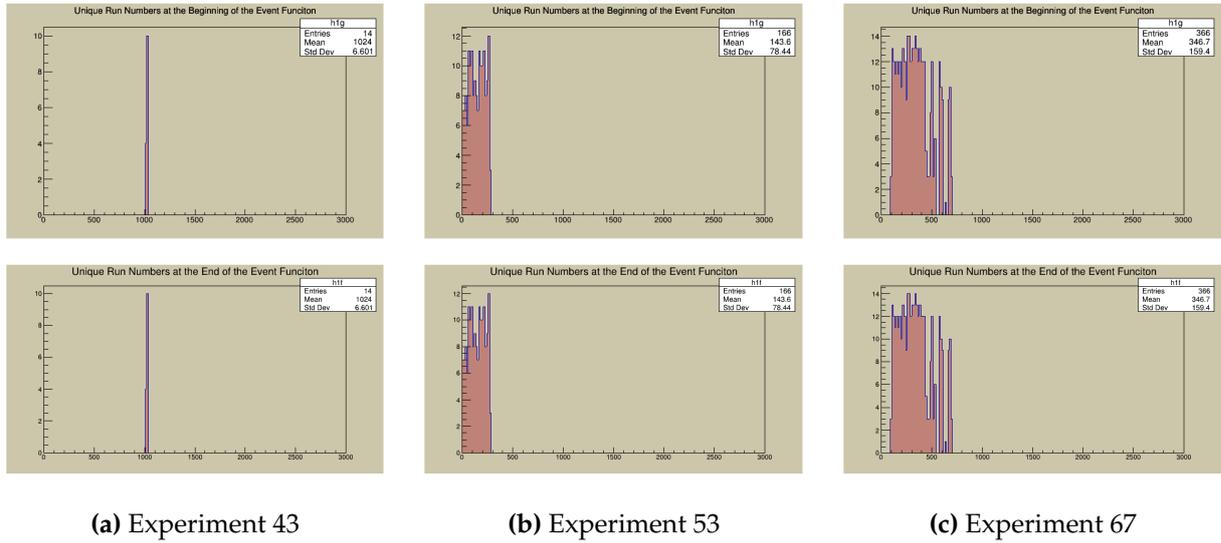
Figure B.7.: Comparison of the unique run numbers printed to a log file at the beginning of the BASF analysis module (upper figures) and the end of the BASF analysis module (lower figures), for all of the $\Upsilon(5S)$

(a) Experiment 69 **(b)** Experiment 71
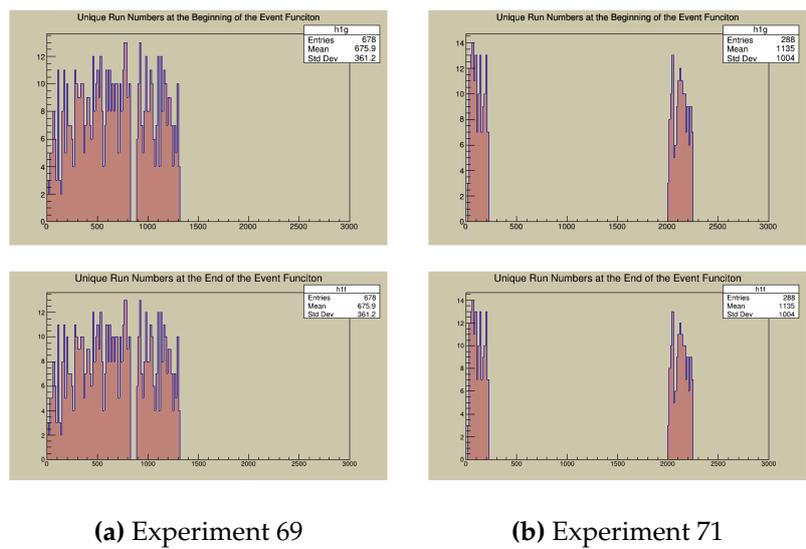
Figure B.8.: Comparison of the unique run numbers printed to a log file at the beginning of the BASF analysis module (upper figures) and the end of the BASF analysis module (lower figures), for all of the $\Upsilon(5S)$ data

(a) Experiment 43   **(b)** Experiment 53   **(c)** Experiment 67
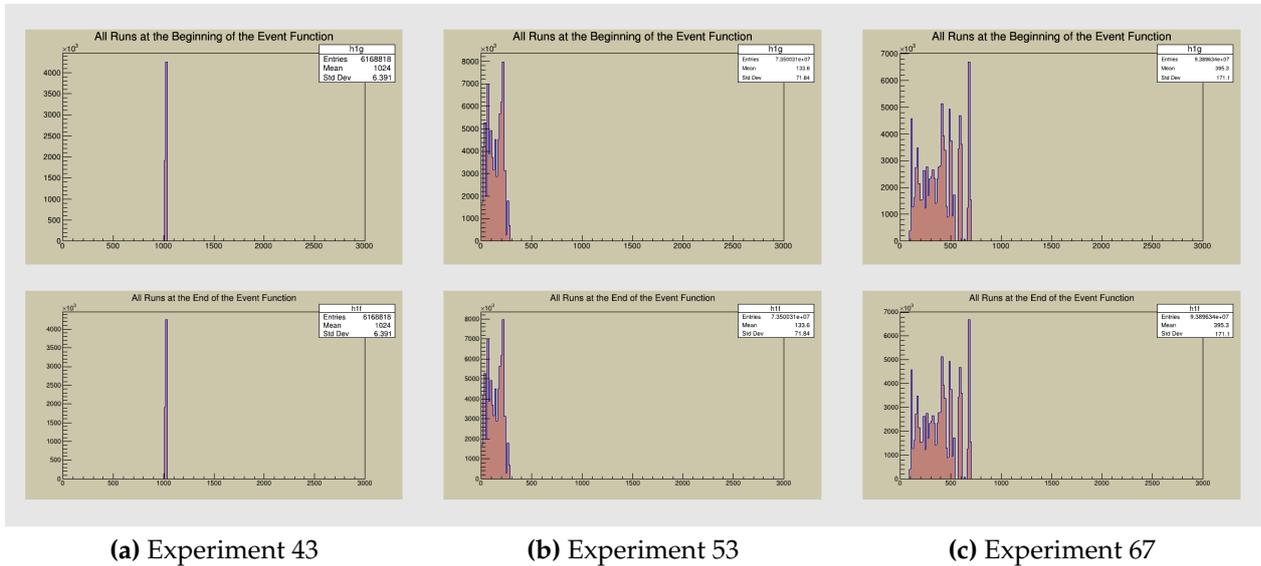
Figure B.9.: Comparison of the run numbers all events, printed to a log file at the beginning of the BASF analysis module (upper figures) and the end the BASF analysis module (lower figures), for all of the $\Upsilon(5S)$ data

(a) Experiment 69   **(b)** Experiment 71
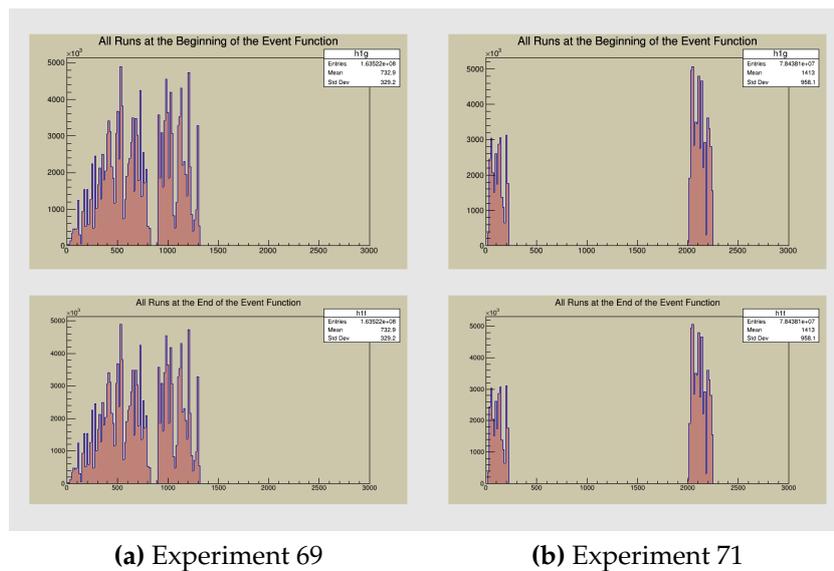
Figure B.10.: Comparison of the run numbers all events, printed to a log file at the beginning of the BASF analysis module (upper figures) and the end the BASF analysis module (lower figures), for all of the $\Upsilon(5S)$ data

(a) $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ Modes

(b) $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$
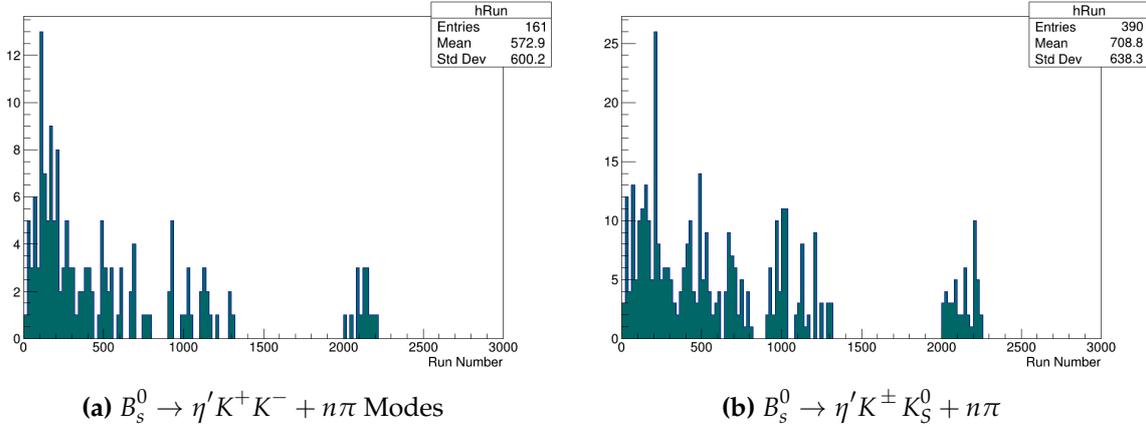
Figure B.11.: Histogram of the run numbers for the $\Upsilon(5S)$ data. All signal region cuts are applied, for $M(X_{s\bar{s}}) \leq 2.4\,\text{GeV}/c^2$. All experiments were processed.

B.0.2. Some BASF/BASF2 Comparisons

As an additional check, comparison between BASF and Belle AnalysiS Framework 2 (BASF2) can be done. The Belle-to-BelleII (B2BII) framework is used to analyzed Belle data in BASF2. The B2BII process coverts the Belle data format to the Belle II data format [92, 93].

There are some differences to note for this analysis, which uses the light-2020-ichep BAF2 release. The ECL acceptance in the BASF2 framework differs slightly than the one used in the BASF reconstruction module. There are also seems to be slight differences in the efficiency of the E9/E25 cut. These cause the number of π^0's to differ. With both discrepancies included, the BASF π^0 list is 0.3% larger than in BASF2. If both frameworks use the BASF2 ECL acceptance, this difference reduces to 0.03%. The K_S^0 list in BASF is 0.7% larger, and pre-mass-constrained fit, the η and η' lists in BASF are 8.1×10^{-4}% smaller and 0.03% larger, respectively. These discrepancies seem to matter less when all final selections are applied.

Given these (relatively minor) discrepancies, reconstruction can be compared at different stages. Cuts in BASF2 are the same as in BASF and in the same order. The following results are obtained using the same signal MC as is used to obtain the reconstruction efficiencies in Tables 4.8 and 4.9. The same NeuroBayes expertise used in the main analysis is also used.

Figure B.12 shows the comparison in η and η' reconstruction. The η' includes the mass-constrained fit on the η. Histograms are normalized to their respective integrals.

(a) η meson reconstruction

(b) η' meson reconstruction
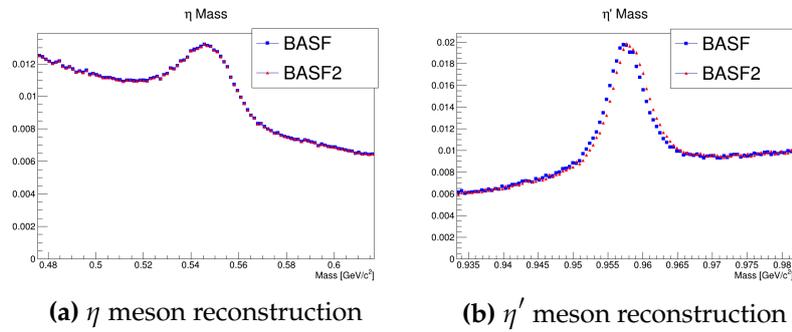
Figure B.12.: Comparison between η (left) η' (right) reconstructed in BASF (red circles) and BASF2 (blue squares)

These demonstrate little discrepancy between the two frameworks in intermediate meson reconstruction.

Bin Number	Bin (GeV/c^2)	BASF Eff.(%)	BASF2 Eff.(%)
1	0.8-1.0	5.6 ± 0.41	5.5 ± 0.42
2	1.0-1.2	3.8 ± 0.09	3.7 ± 0.089
3	1.2-1.4	3.0 ± 0.08	2.8 ± 0.079
4	1.4-1.6	0.95 ± 0.05	0.92 ± 0.048
5	1.6-1.8	0.58 ± 0.04	0.58 ± 0.038
6	1.8-2.0	0.35 ± 0.03	0.35 ± 0.030
7	2.0-2.2	0.24 ± 0.02	0.22 ± 0.024
8	2.2-2.4	0.15 ± 0.02	0.12 ± 0.019

Table B.2.: Reconstruction efficiency in 0.2 GeV/c^2 bins for $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ in signal MC.

Bin Number	Bin (GeV/c^2)	BASF Eff.(%)	BASF2 Eff.(%)
1	0.8-1.0	0.0	0
2	1.0-1.2	0.01 ± 0.006	0.0079 ± 0.0056
3	1.2-1.4	0.25 ± 0.02	0.24 ± 0.026
4	1.4-1.6	0.90 ± 0.05	0.98 ± 0.049
5	1.6-1.8	0.68 ± 0.04	0.69 ± 0.043
6	1.8-2.0	0.48 ± 0.04	0.46 ± 0.037
7	2.0-2.2	0.38 ± 0.03	0.35 ± 0.034
8	2.2-2.4	0.18 ± 0.03	0.25 ± 0.030

Table B.3.: Reconstruction efficiency in 0.2 GeV/c^2 bins for $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ in signal MC.

Signal Reconstruction Efficiency Comparison

Module Calls

BASF2 Result

After analyzing the $\Upsilon(5S)$ data with BASF2, not statistically significant signal is found. There are approximately 50% fewer skimmed events in data (before final signal region cuts are applied), possibly due to fewer continuum events having a successful vertex fit using BASF2 fitter. Using the systematic uncertainties and correction factors from

Experiment	BASF Skim Calls	BASF2 Skim Calls
43	6310259	6310259
53	75199284	75199284
67	96070370	96070370
69	167291661	167291661
71	80262406	80262406

Table B.4.: Comparison of the number of module calls between BASF and BASF2 for the full Y(5S) data set.

Experiment	BASF Module Calls	BASF2 Module Calls
43	6168818	6168818
53	73500310	73500310
67	93896334	93896334
69	163522044	163522044
71	78438098	78438098

Table B.5.: Comparison of the number of module calls between BASF and BASF2 for the full Y(5S) data set.

the main analysis, and the BASF2 signal reconstruction efficiencies in Tables B.2 and B.3, the 90% confidence level upper limit is approximately 1.7×10^{-3}, consistent with the BASF analysis.

Appendix C.

Monte Carlo Checks Using a Fixed ARGUS Parameter

As the the fitting method was modified with respect to the ARGUS parameter (ARG-PAR) after data unblinding, the method is checked on MC. After data unblinding it was found that there were not enough data in several of the $X_{s\bar{s}}$ mass bins for a floating ARGUS parameter to admit a physical shape. The solution was to look in the $\Upsilon(5S)$ data NN sidebands. This is done by looking in a region on the neural network output variable that is below the standard selection.

For $B_s^0 \to \eta' K^+ K^- + n\pi$ modes the nominal selection is nnout = 0.95 and for $B_s^0 \to \eta' K^{\pm} K_S^0 + n\pi$ modes nnout = 0.6. For the former, the sideband region for which the ARGUS parameter is determined is -0.25 $\leq$ nnout < 0.95. For the latter it is -0.6 $\leq$ nnout < 0.6, all other cuts being the same. The data is fitted in these sideband regions with a floating ARGUS parameter. The ARGUS parameter determined from this fit is used as the fixed ARGUS parameter when fitting to the signal region in data. This same parameter is used to re-fit the signal and background MC to see if there is a significant difference as compared to a floating ARGUS parameter. No significant difference is found except that there are fewer random combination background events in signal MC, in a couple of bins. Results from sections 4.8 and 4.9 are reproduced in the following tables for direct comparison.

C.0.1. Signal Reconstruction Efficiency

Bin Number	Bin (GeV/c^2)	Eff.(%)	Fixed ARGPAR
1	0.8-1.0	5.6 ± 0.4	5.6 ± 0.4
2	1.0-1.2	3.8 ± 0.09	3.8 ± 0.09
3	1.2-1.4	3.0 ± 0.08	3.0 ± 0.08
4	1.4-1.6	1.0 ± 0.05	1.0 ± 0.05
5	1.6-1.8	0.6 ± 0.04	0.6 ± 0.04
6	1.8-2.0	0.4 ± 0.03	0.4 ± 0.03
7	2.0-2.2	0.2 ± 0.02	0.2 ± 0.02
8	2.2-2.4	0.2 ± 0.02	0.2 ± 0.02

Table C.1.: Comparison of the effect of floating versus fixed ARGUS parameter for signal reconstruction efficiency, estimated from signal MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Bin Number	Bin (GeV/c^2)	Eff.(%)	Fixed ARGPAR
1	0.8-1.0	0.0	0.0
2	1.0-1.2	0.01 ± 0.01	0.01 ± 0.01
3	1.2-1.4	0.2 ± 0.03	0.3 ± 0.02
4	1.4-1.6	0.9 ± 0.06	0.9 ± 0.04
5	1.6-1.8	0.7 ± 0.05	0.7 ± 0.04
6	1.8-2.0	0.5 ± 0.04	0.5 ± 0.04
7	2.0-2.2	0.4 ± 0.04	0.4 ± 0.03
8	2.2-2.4	0.2 ± 0.03	0.2 ± 0.03

Table C.2.: Comparison of the effect of floating versus fixed ARGUS parameter for signal reconstruction efficiency, estimated from signal MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0 $\pm$ 0.1	0.0 $\pm$ 0.1
1.2-1.4	-0.2 $\pm$ 0.1	-0.2 $\pm$ 0.2
1.4-1.6	0.2 $\pm$ 0.3	0.2 $\pm$ 0.3
1.6-1.8	-0.2 $\pm$ 0.1	-0.2 $\pm$ 0.1
1.8-2.0	0.4 $\pm$ 0.3	0.3 $\pm$ 0.3
2.0-2.2	0.2 $\pm$ 0.4	0.2 $\pm$ 0.3
2.2-2.4	-0.2 $\pm$ 0.2	-0.1 $\pm$ 0.3

Table C.3.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0	0.0
1.2-1.4	-0.2 $\pm$ 0.3	-0.2 $\pm$ 0.3
1.4-1.6	0.1 $\pm$ 0.3	0.2 $\pm$ 0.2
1.6-1.8	0.7 $\pm$ 0.5	0.6 $\pm$ 0.5
1.8-2.0	-0.2 $\pm$ 0.1	-0.2 $\pm$ 0.1
2.0-2.2	1.2 $\pm$ 1.0	1.4 $\pm$ 0.9
2.2-2.4	-0.7 $\pm$ 0.6	-0.5 $\pm$ 0.6

Table C.4.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

C.0.2. Peaking Backgrounds

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.2 $\pm$ 0.3	0.3 $\pm$ 0.3
1.2-1.4	0.0	0.0
1.4-1.6	0.2 $\pm$ 0.2	0.2 $\pm$ 0.2
1.6-1.8	-0.2 $\pm$ 0.1	-0.2 $\pm$ 0.1
1.8-2.0	0.2 $\pm$ 0.2	0.2 $\pm$ 0.2
2.0-2.2	0.0 $\pm$ 0.1	0.0 $\pm$ 0.1
2.2-2.4	0.0	-0.2 $\pm$ 0.1

Table C.5.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0	0.0
1.2-1.4	0.0	0.0
1.4-1.6	-0.2 $\pm$ 0.2	0.2 $\pm$ 0.7
1.6-1.8	-0.2 $\pm$ 0.1	-0.2 $\pm$ 0.1
1.8-2.0	0.2 $\pm$ 0.4	0.3 $\pm$ 0.3
2.0-2.2	0.6 $\pm$ 0.5	0.4 $\pm$ 0.5
2.2-2.4	0.5 $\pm$ 0.7	1.1 $\pm$ 0.6

Table C.6.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.2 $\pm$ 0.5	0.2 $\pm$ 0.2
1.0-1.2	37.8 $\pm$ 3.0	37.7 $\pm$ 3.1
1.2-1.4	30.5 $\pm$ 2.8	30.4 $\pm$ 2.8
1.4-1.6	27.2 $\pm$ 2.5	27.2 $\pm$ 2.6
1.6-1.8	21.0 $\pm$ 2.3	21.0 $\pm$ 2.3
1.8-2.0	20.5 $\pm$ 2.3	20.5 $\pm$ 2.3
2.0-2.2	12.2 $\pm$ 1.8	12.2 $\pm$ 1.7
2.2-2.4	14.0 $\pm$ 1.9	14.0 $\pm$ 1.9

Table C.7.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $q\bar{q}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0	0.0
1.2-1.4	13.0 $\pm$ 1.8	13.0 $\pm$ 1.8
1.4-1.6	43.8 $\pm$ 3.2	43.7 $\pm$ 3.3
1.6-1.8	57.8 $\pm$ 3.8	57.7 $\pm$ 3.8
1.8-2.0	85.0 $\pm$ 4.5	85.0 $\pm$ 4.6
2.0-2.2	103.0 $\pm$ 5.0	103.0 $\pm$ 5.1
2.2-2.4	148.8 $\pm$ 6.0	148.6 $\pm$ 6.1

Table C.8.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $q\bar{q}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

C.0.3. Non-Peaking Background

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.2 $\pm$ 0.2	0.2 $\pm$ 0.2
1.2-1.4	1.7 $\pm$ 0.6	1.7 $\pm$ 0.6
1.4-1.6	0.6 $\pm$ 0.4	0.6 $\pm$ 0.4
1.6-1.8	2.4 $\pm$ 0.7	2.4 $\pm$ 0.7
1.8-2.0	2.2 $\pm$ 0.7	2.3 $\pm$ 0.7
2.0-2.2	1.8 $\pm$ 0.7	1.8 $\pm$ 0.7
2.2-2.4	1.9 $\pm$ 0.6	1.9 $\pm$ 0.7

Table C.9.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0	0.0
1.2-1.4	0.3 $\pm$ 0.2	0.3 $\pm$ 0.2
1.4-1.6	0.5 $\pm$ 0.4	0.4 $\pm$ 0.3
1.6-1.8	3.3 $\pm$ 0.9	3.4 $\pm$ 0.9
1.8-2.0	8.3 $\pm$ 1.3	8.3 $\pm$ 1.3
2.0-2.2	15.8 $\pm$ 2.0	15.6 $\pm$ 1.9
2.2-2.4	27.2 $\pm$ 2.4	26.9 $\pm$ 2.4

Table C.10.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $B\bar{B}$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^{\pm} K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.2 ± 0.2	0.1 ± 0.2
1.2-1.4	0.0	0.0
1.4-1.6	0.0	0.0
1.6-1.8	0.2 ± 0.2	0.2 ± 0.2
1.8-2.0	0.2 ± 0.2	0.2 ± 0.2
2.0-2.2	0.2 ± 0.2	0.2 ± 0.2
2.2-2.4	0.7 ± 0.3	0.7 ± 0.3

Table C.11.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^+ K^- + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Mass Bin (GeV/c^2)	Est. Events	Fixed ARGPAR
0.8-1.0	0.0	0.0
1.0-1.2	0.0	0.0
1.2-1.4	0.0	0.0
1.4-1.6	0.3 ± 0.2	0.4 ± 0.2
1.6-1.8	0.4 ± 0.3	0.4 ± 0.3
1.8-2.0	2.0 ± 0.7	1.9 ± 0.7
2.0-2.2	5.4 ± 1.1	5.6 ± 1.1
2.2-2.4	8.5 ± 1.4	7.9 ± 1.3

Table C.12.: Comparison of the effect of floating versus fixed ARGUS parameter on the number of non-peaking background events estimated from $B_s^0 \bar{B}_s^0$ MC in 0.2 GeV/c^2 mass bins, $B_s^0 \rightarrow \eta' K^\pm K_S^0 + n\pi$ modes. The fixed ARGUS parameters are determined from $\Upsilon(5S)$ data NN sidebands. The middle column gives the original estimates, the right column gives the estimates due to the fixed ARGUS parameter.

Appendix D.

Unreconstructed Modes

D.0.1. Proportion of Unreconstructed Modes

• Unreconstructed Modes

There are some signal modes that are not reconstructed in this analysis. For example, modes with more than one π^0, modes with a K_L^0 instead of a K_S^0, and modes with more than six B_s^0 daughter particles (excluding the η'). To estimate the unreconstructed fraction, a generator-level scan was conducted. This fraction is defined as $\frac{N_{UR}}{N_{UR}+N_R}$, where N_{UR} is the number of events in signal MC that come from unreconstructed modes and N_R is the number of events in signal MC that come from reconstructed modes (events that were used to determine reconstruction efficiency). The unreconstructed fraction is shown in Fig. D.1. BASF2 was used to count N_R and N_{UR}. For the class $B_s^0 \to K^\pm K_S^0$, the majority of unreconstructed modes come from unreconstructed modes with a K_L^0 that also satisfy the pion requirements.

(a) $B_s^0 \to \eta' K^+ K^- + n\pi$ modes

(b) $B_s^0 \to \eta' K^\pm K_S^0 + n\pi$ modes
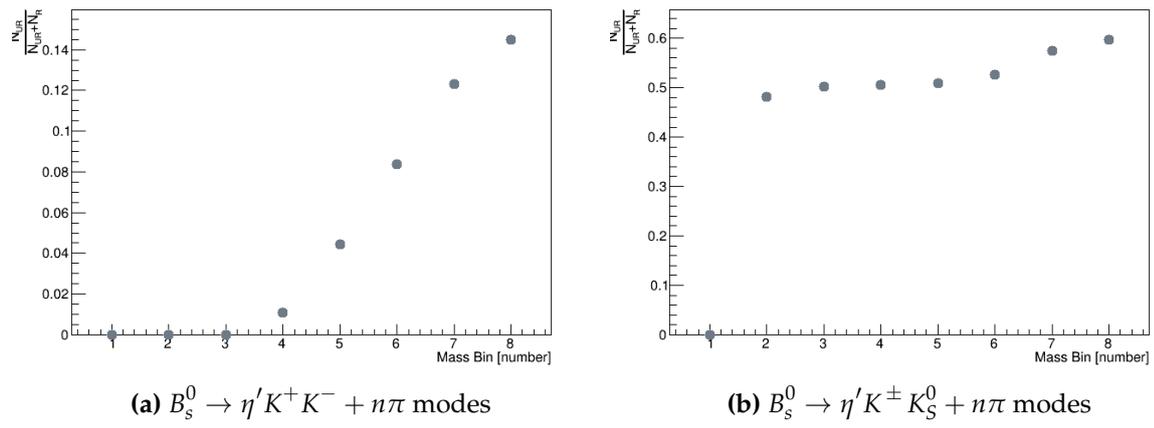
Figure D.1.: Fraction of unreconstructed events

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [94].

Bibliography

[1] S. Dubey, *Search for the Decay $B_s^0 \rightarrow \eta' X_{s\bar{s}}$ Using a Semi-Inclusive Method at the Belle Experiment*, PhD thesis, University of Hawaii Manoa, 2020.

[2] Analysis materials, https://stash.desy.de/projects/B2A/repos/bs2etaprimexss/browse, Bitbucket.

[3] Belle Collaboration, S. Dubey *et al.*, 2021, arvix:2102.10266.

[4] National Laboratory for High Energy Physics Report No. KEK–95-7, 1995 (unpublished).

[5] S. Kurokawa and E. Kikutani, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **499**, 1 (2003), KEK-B: The KEK B-factory.

[6] I. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **499**, 167 (2003), KEK-B: The KEK B-factory.

[7] K. Akai *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **499**, 45 (2003), KEK-B: The KEK B-factory.

[8] Belle Collaboration, S. Esen *et al.*, Phys. Rev. D **87**, 031101 (2013).

[9] J. Brodzicka *et al.*, Progress of Theoretical and Experimental Physics **2012** (2012), 04D001.

[10] A. J. Bevan *et al.*, The European Physical Journal C **74** (2014).

[11] T. Abe *et al.*, Progress of Theoretical and Experimental Physics **2013** (2013), 03A001.

[12] A. Abashian *et al.*, Nuclear Instruments and Methods in Physics Research Section

A: Accelerators, Spectrometers, Detectors and Associated Equipment **479**, 117 (2002), Detectors for Asymmetric B-factories.

[13] Y. Ushiroda, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **511**, 6 (2003), Proceedings of the 11th International Workshop on Vertex Detectors.

[14] H. Aihara *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **568**, 269 (2006), New Developments in Radiation Detectors.

[15] T. Iijima *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **453**, 321 (2000), Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[16] H. Kichimi *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **453**, 315 (2000), Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[17] H. Hirano *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **455**, 294 (2000).

[18] K. Miyabayashi, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **494**, 298 (2002), Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[19] Y. Ushiroda *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **438**, 460 (1999).

[20] S. Y. Suzuki, M. Yamauchi, M. Nakao, R. Itoh, and H. Fujii, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **453**, 440 (2000), Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[21] S. Y. Suzuki *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **494**, 535 (2002), Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[22] K. A. Nishimura, *FIRST OBSERVATION OF THE INCLUSIVE DECAY $B \to X_s \eta$ WITH THE BELLE DETECTOR*, PhD thesis, University of Hawaii Manoa, 2010.

[23] T. Abe *et al.*, (2010), 1011.0352.

[24] S. Y. Lee, *Accelerator Physics* (World Scientific, 2004).

[25] I. Adachi, T. Browder, P. Krizan, S. Tanaka, and Y. Ushiroda, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **907**, 46 (2018), Advances in Instrumentation and Experimental Methods (Special Issue in Honour of Kai Siegbahn).

[26] J. Fast, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **876**, 145 (2017), The 9th international workshop on Ring Imaging Cherenkov Detectors (RICH2016).

[27] D. Kotchetkov *et al.*, (2019), 1804.10782.

[28] M. Bessner, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **958**, 162318 (2020), Proceedings of the Vienna Conference on Instrumentation 2019.

[29] K. Inami, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **766**, 5 (2014), RICH2013 Proceedings of the Eighth International Workshop on Ring Imaging Cherenkov Detectors Shonan, Kanagawa, Japan, December 2-6, 2013.

[30] K. Nishimura, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **639**, 177 (2011), Proceedings of the Seventh International Workshop on Ring Imaging Cherenkov Detectors.

[31] K. Inami, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **936**, 556 (2019), Frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors.

[32] G. Aad *et al.*, Physics Letters B **716**, 1 (2012).

[33] S. Chatrchyan *et al.*, Physics Letters B **716**, 30 (2012).

[34] W. Greiner, S. Schramm, and E. Stein, *Quantum Chromodynamics Third Edition* (Springer, 2007).

[35] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory* (Westview Press Incorporated, 1995).

[36] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2014).

[37] Particle Data Group, P. Zyla *et al.*, Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).

[38] S. Weinberg, Phys. Rev. D **11**, 3583 (1975).

[39] D. Atwood and A. Soni, Physics Letters B **405**, 150 (1997).

[40] E. Kou, Phys. Rev. D **63**, 054027 (2001).

[41] Physics Letters B **419**, 369 (1998).

[42] I. Halperin and A. Zhitnitsky, Phys. Rev. D **56**, 7247 (1997).

[43] X.-G. He and G.-L. Lin, Physics Letters B **454**, 123 (1999).

[44] W.-S. Hou and B. Tseng, Phys. Rev. Lett. **80**, 434 (1998).

[45] A. Ali and A. Y. Parkhomenko, 2001, hep-ph/0112048.

[46] A. L. Kagan and A. Petrov, 1997, hep-ph/9707354.

[47] H. Fritzsch, Physics Letters B **415**, 83 (1997).

[48] J. J. Sakurai, *Advanced Quantum Mechanics* (Pearson Education, Inc., 2006).

[49] K. Lingel, T. Skwarnicki, and J. G. Smith, Annual Review of Nuclear and Particle Science **48**, 253âĂŞ306 (1998).

[50] Particle Data Group, M. Tanabashi *et al.*, Phys. Rev. D **98**, 030001 (2018).

[51] Belle Collaboration, S. Wehle *et al.*, Phys. Rev. Lett. **118**, 111801 (2017).

[52] Belle Collaboration, T. Saito *et al.*, Phys. Rev. D **91**, 052004 (2015).

[53] The Belle Collaboration, S. Watanuki *et al.*, Phys. Rev. D **99**, 032012 (2019).

[54] CLEO Collaboration, T. E. Browder *et al.*, Phys. Rev. Lett. **81**, 1786 (1998).

[55] CLEO Collaboration, G. Bonvicini *et al.*, Phys. Rev. D **68**, 011101 (2003).

[56] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **93**, 061801 (2004).

[57] Belle Collaboration, K. Nishimura *et al.*, Phys. Rev. Lett. **105**, 191803 (2010).

[58] K. Uchida, Belle Note 940, Internal Belle Note.

[59] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **80**, 112002 (2009).

[60] University of Mississippi, A. Datta, private communication, 2020.

[61] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **2006**, 026âĂŞ026 (2006).

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

[63] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, (CERN, Geneva, 1987).

[64] E. Barberio, B. van Eijk, and Z. WaÌǧs, Comp. Phys. Commun. **66**, 115 (1991).

[65] R. Louvot *et al.*, Belle Note 1036, Internal Belle Note.

[66] R. Itoh, in *9th International Conference on Computing in High-Energy and Nuclear Physics*, 1997.

[67] I. Adachi *et al.*, 2003, physics/0306120.

[68] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[69] G. C. Fox and S. Wolfram, Nuclear Physics B **149**, 413 (1979).

[70] R. J. Larsen and M. L. Marx, *An Introduction to Mathematical Statistics and its Applications* (Prentice Hall, 2012).

[71] ARGUS, H. Albrecht *et al.*, Phys. Lett. B **241**, 278 (1990).

[72] W. Verkerke and D. Kirkby, 2003, physics/0306116.

[73] M. Feindt and U. Kerzel, Nucl. Instrum. and Meth. A **559**, 190 (2006).

[74] Belle Collaboration, H. Kakuno *et al.*, Nucl. Instrum. Meth. A **533**, 516 (2004).

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

[76] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **98**, 051802 (2007).

[77] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **80**, 011101 (2009).

[78] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **74**, 051106 (2006).

[79] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **74**, 031105 (2006).

[80] R. Louvot *et al.*, Belle Note 1104, Internal Belle Note.

[81] Belle pid group website, https://belle.kek.jp/group/pid_joint/kid/kid.html, Accessed: 2019-2020.

[82] B. Bhuyan, Belle Note 1165 v1, Internal Belle Note.

[83] S. Lin, Belle Note 645, Internal Belle Note.

[84] Belle Collaboration, N. Dash *et al.*, Phys. Rev. Lett. **119**, 171801 (2017).

[85] Belle Collaboration, B. Pal *et al.*, Phys. Rev. D **92**, 011101 (2015).

[86] Belle Collaboration, S. Ryu *et al.*, Phys. Rev. D **89**, 072009 (2014).

[87] N. Dash *et al.*, Belle note 1472, Internal Belle Note.

[88] A. Abdesselam *et al.*, Study of two-body $e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}$ production in the energy range from 10.77 to 11.02 GeV, 2016, 1609.08749.

[89] L. Moneta *et al.*, 2010, 1009.1003.

[90] LHCb, R. Aaij *et al.*, JHEP **1705**, 158 (2017) (2017).

[91] Belle file search engine, http://bweb3.cc.kek.jp/, Accessed: 2020-09-30.

[92] T. Kuhr *et al.*, Computing and Software for Big Science **3**, 1 (2018).

[93] M. Gelb *et al.*, Computing and Software for Big Science **2** (2018).

[94] A. Buckley, The hepthesis LaTeX class.

List of figures

List of tables