FIRST OBSERVATION OF THE INCLUSIVE DECAY $B \rightarrow X_S\eta$ WITH THE BELLE DETECTOR

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAIʻI IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PHYSICS

DECEMBER 2010

By
Kurtis Alan Nishimura

Dissertation Committee:

Thomas Browder, Chairperson
Gary Varner
Jason Kumar
Pui Lam
Thomas Ernst

Acknowledgements

I owe a tremendous debt of gratitude to a number of people for their professional and personal assistance. This dissertation would not be possible without them.

Though I was completely new to the subject, Tom Browder welcomed me into the world of B decays without hesitation. His enthusiasm for beautiful physics is contagious, and I am grateful for his continuous guidance and wisdom. Behind every analysis is a rich world of instrumentation that makes it all possible, and I thank Gary Varner for introducing me to it and allowing me to contribute to current and future detectors.

This work is truly a collaborative effort of countless Belle and KEKB members. I appreciate the excellent insights and suggestions of my Belle analysis referees: Mikihiko Nakao, Shohei Nishida, and C.H. Wang. Thanks also to all the current and past members of the UH Belle group; you've endured many talks on the same topic, and always provided new and helpful feedback. I entered Belle without a clue as to how to analyze data, much less find somewhere to eat around KEK. Thanks to Jeremy, Minori-chan, David, Fujikawa-san, Nakahama-san, and Nakayama-san, for both technical help and your friendship. To these and all the Belle and KEKB members, it has been a privilege to work with you all.

I extend a special thanks to Andres and Sandip. Though our research has been vastly different, our friendship and informal collaboration has improved my own work and broadened my understanding of the field. Here's to our past conversations about life and physics over beers, and all those to come. A number of other friends have helped me stay sane over the years. Chynna, Matt, and Jeff, thanks for making me laugh, letting me vent occasionally, and even proofreading papers for me that I'm sure were not exactly page-turners. Thanks also to my officemate Himansu, who has both offered help on analysis issues and suffered through a plethora of noise and distractions.

To all of my friends all over the world, thank you for your continuing support. Despite the fact that I disappear for months at a time, you always welcome me back as if we saw one another yesterday. I owe a special debt to Erin, who has endured late nights of work, meetings at all hours, months-long separations, and that glazed-over look that I get when my mind is clearly still on work. Thanks for your infinite patience and for pushing me toward the finish line.

My parents have given endless support to me throughout my many, many years of education, and so I am especially grateful to them. They have always encouraged me to find something that I enjoy doing, and they'll be happy to know that I've found it.

ABSTRACT

The decay of a B meson into an η' meson and an inclusive charmless state with unit strangeness, $B \to X_s\eta'$, has stimulated significant theoretical interest since it was first observed in 1998. A number of models were proposed to explain the unexpectedly large branching fraction and the spectrum of X_s mass, which peaks above $2.0~\mathrm{GeV}/c^2$. One of the most initially appealing explanations was the QCD anomaly model, in which the η' couples strongly to two gluons. However, despite a number of theoretical calculations and improved measurements of the decay, no explanation has been universally accepted as accounting for available experimental data. The unique relationship between the η' and the η suggest that the complimentary decay, $B \to X_s\eta$, could elucidate the nature of the $X_s\eta'$ process.

We report the first measurement of inclusive $B \to X_s\eta$ decays, based on a pseudo-inclusive reconstruction technique using a sample of 657×10^6 $B\bar{B}$ pairs accumulated with the Belle detector at the KEKB e^+e^- collider. For $M_{X_s} < 2.6~\mathrm{GeV}/c^2$, we measure a branching fraction of $(26.1 \pm 3.0(\mathrm{stat})^{+1.9}_{-2.1}(\mathrm{syst})^{+4.0}_{-7.1}(\mathrm{model})) \times 10^{-5}$ and a direct CP asymmetry of $\mathcal{A}_{CP} = -0.13 \pm 0.04^{+0.02}_{-0.03}$. Over half of the signal occurs in the range $M_{X_s} > 1.8~\mathrm{GeV}/c^2$, above all currently known exclusive contributions to $B \to X_s\eta$.

The lack of significant suppression of this decay relative to its η' counterpart indicates that explanations for the $B \to X_s\eta'$ signal that invoke mechanisms specific to the η' are disfavored.

Contents

List of Tables

List of Figures

List of Acronyms

ADC Analog-to-Digital Converter

CDC Central Drift Chamber

CKM Cabibbo, Kobayashi, and Maskawa

CM Center-of-Mass

CP Charge-conjugation and Parity

DAQ Data Acquisition

DSSD Double-sided Silicon Strip Detector

ECL Electromagnetic Calorimeter

EFC Extreme Forward Calorimeter

GDL Global Decision Logic

HER High Energy Ring

IP Interaction Point

KEK High Energy Accelerator Research Organization, located in Tsukuba, Japan

KEKB An Asymmetric Electron-Positron Collider for B Physics at KEK

KLM K-long (K_L^0) and Muon (μ) Detector

KM Kobayashi and Maskawa

L1 Level-1 Trigger

L3 Level-3 Trigger

L4 Level-4 Trigger

LER Low Energy Ring

LHC Large Hadron Collider

LR Likelihood Ratio

MC Monte Carlo

PDF Probability Density Function

PDG Particle Data Group

PID Particle Identification

PMT Photomultiplier Tube

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

RPC Resistive Plate Counter

SM Standard Model of Particle Physics

SVD Silicon Vertex Detector

TDC Time-to-Digital Converter

TOF Time-of-Flight Counter

TSC Trigger Scintillator

GSW Glashow, Salam, and Weinberg

Chapter 1

Introduction

The Standard Model (SM) of particle physics is currently the best accepted theory explaining interactions of fundamental particles. It was developed and refined throughout the 1960s and 1970s, and has since remained relatively unaltered. The SM can account for nearly all observations of experimental high energy physics [1]. However, despite its impressive track record, certain observations and theoretical motivations imply that the SM is incomplete. There are a large number of theories and models proposing to supplement or complete the SM, but ultimately these proposals must be verified experimentally. The goal of experimental high energy physics, then, is to investigate the SM from an observational point of view. This includes looking for new phenomena that cannot be explained in the SM, as well as measuring known or expected SM processes and searching for inconsistencies with theoretical predictions.

We will begin in this chapter by summarizing briefly the current state of the SM, including some of its shortcomings. In Chapter 2, we will focus specifically on the η and η' mesons and their two associated B meson decays: $B \to X_s\eta'$ and $B \to X_s\eta$. Measurements of the former have stimulated significant theoretical activity, suggesting possible contributions from new physics. These possibilities can be further scrutinized and probed by the latter, which we have recently measured for the first time. In Chapter 3, we will introduce the basic methods of particle detection and the apparatus for this measurement: the KEKB accelerator and the Belle detector. Specific analysis procedures and systematic errors for this measurement will be discussed in Chapters 4 and 5, respectively. Chapter 6 summarizes the final measurement results and gives some discussion of their implications.

1.1 The Standard Model

The SM includes a set of fundamental fields, whose quanta we will henceforth refer to as particles. These include the fundamental fermions, the constituents of matter, and a set of gauge bosons that account for three of the four known fundamental forces: the electromagnetic, strong, and weak forces. The SM also posits the existence of a special particle known as the Higgs boson. These fundamental particles and interactions are described briefly here. For a more formal and complete description see, for example, Reference [2].

1.1.1 Standard Model Fermions

The fundamental fermions can be organized into two basic categories: quarks and leptons.

Quarks come in six varieties, known as flavors: up, down, charm, strange, top, and bottom. The quarks are arranged into three generations as follows:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.1.1}$$

The "up-type" quarks, u, c, and t, have electric charge +2/3, while the "down-type" quarks, d, s, and b, have electric charge -1/3. Each quark flavor comes in three "colors," usually referred to as red, green, and blue, though this naming convention is arbitrary.

Similar to the quarks, there are six types of leptons: the electron, muon, tau, electron neutrino, muon neutrino, and tau neutrino. As before, they are organized into three generations:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \tag{1.1.2}$$

The neutrinos, ν_e, ν_μ, and ν_τ, have no electric charge while the electron, muon, and tau each have charge -1.

Masses of all the quarks and leptons are given in Table 1.1. All quarks and leptons have anti-particle partners with equal mass and opposite electric charge.

The number and identity of the fundamental fermions is not an imposed constraint of the SM. Rather, the listed particles are a reflection of experimental observations. The SM framework does, however, impose constraints on possible new particle content. For example, if there were a fourth generation of quarks, it should be accompanied by a

Particle name	Symbol	Electric charge ($e = 1.60 \times 10^{-19}$ C)	Mass (GeV/c^2)
quarks			
up	u	+2/3	2.49×10^{-3}
down	d	-1/3	5.05×10^{-3}
strange	s	-1/3	1.01×10^{-1}
charm	c	+2/3	1.27
bottom	b	-1/3	4.19
top	t	+2/3	1.72×10^{2}
leptons			
electron	e	-1	5.11×10^{-6}
muon	μ	-1	1.06×10^{-3}
tau	τ	-1	1.78
electron neutrino	ν_e	0	$< 2 \times 10^{-9}$
muon neutrino	ν_μ	0	$< 2 \times 10^{-9}$
tau neutrino	ν_τ	0	$< 2 \times 10^{-9}$
gauge bosons			
gluon	g	0	0
photon	γ	0	0
W boson	W	+1	8.04×10^{1}
Z boson	Z^0	0	9.12×10^{1}
Higgs boson	H	0	$> 1.14 \times 10^{2}$

Table 1.1: A list of the particles of the Standard Model, including standard abbreviation, electric charge, and mass [1]. Note that each listed particle has a corresponding anti-particle with the opposite charge. Because free quarks are not observed, quark masses are estimates based on experimental observations (e.g., measured meson masses) coupled with theoretical calculations.

fourth generation of leptons. Experimental data indicates that the number of neutrinos with $m_\nu \lesssim 45$ GeV$/c^2$ is 2.9840 ± 0.0082 [3], supporting the notion that there are only three light quark and lepton families.[1]

1.1.2 Standard Model Interactions and Gauge Bosons

The three interactions incorporated into the SM are, in order of decreasing strength, the strong force, the electromagnetic force, and the weak force. Coupling constants for each force are given in Table 1.2.[2] These forces are natural consequences of the imposition of

[1]A new family of heavy quarks and leptons is not ruled out by experimental data. Current limits on possible fourth generation quark and lepton masses are given in Reference [1].

[2]The exact couplings depend on the energy scale of interest and in some cases, the exact particles involved, so the values in Table 1.2 should be treated as approximate.

Force	Coupling constant	Value
strong	α_s	≤ 1
electromagnetic	$\alpha = \frac{e^2}{4\pi\hbar c}$	$\approx 1/137 = 7.3 \times 10^{-3}$
weak	$\frac{G(Mc^2)^2}{(\hbar c)^3}$	1.17×10^{-5}
gravity	$\frac{G_N M^2}{4\pi\hbar c}$	5×10^{-40}

Table 1.2: Approximate coupling constants of the fundamental forces [4]. Although it is not currently included in the SM, gravity is also listed for comparison.

a gauge-symmetry into the quantum theoretical framework of the SM. These symmetries also lead to particles called "gauge-bosons," which can be thought of as the spin-one particles that mediate the force of interest.[3]

The Strong Force

The strong force is carried by eight gauge bosons, called gluons. Each is massless and couples only to particles that carry color charge (i.e., the quarks and other gluons). The theory describing strong interactions is referred to as quantum chromodynamics (QCD), and is based on the symmetry group SU(3). As seen in Table 1.2, the strong force has a large coupling constant relative to the electromagnetic and weak forces. This can complicate theoretical calculations that are usually based on perturbation theory. In fact, the strong force becomes notably non-perturbative at relatively low energy scales, $\mathcal{O}(500\mathrm{MeV})$. As a result, the spectrum of particles we observe experimentally is not the set of quarks previously listed. Rather, we observe color-neutral bound states of quarks and gluons known as hadrons.

Hadrons can primarily be classified into baryons, combinations of three quarks, and mesons, combinations of a quark and an anti-quark. The most well-known baryons are the proton and neutron, which are bound states of uud and udd, respectively. Mesons are less familiar from our every-day experience, as even the longest lived mesons have lifetimes of only $\mathcal{O}(10^{-8}\mathrm{s})$. We mention here only a few specific examples: the lightest mesons are made from combinations of up and down quarks and anti-quarks, and are known as pions, designated π; mesons containing a strange quark (anti-quark) with an up or down anti-quark (quark) are known as kaons, designated K; and mesons containing a

[3]Unlike the strong, weak, and electromagnetic forces, gravity is expected to be mediated by a spin-two particle known as the graviton.

bottom quark (anti-quark) with an up or down anti-quark (quark) are known as B mesons. We will discuss some of these mesons in more detail later.

Other more exotic hadronic states may exist, such as combinations of two gluons (glueballs), sets of two quarks and two anti-quarks (tetraquarks), or four quarks and an anti-quark (pentaquarks). However, none of these states have been conclusively identified. More information on the status of these exotic hadrons can be found in the literature, for example in Reference [5].

The Electromagnetic Force

The electromagnetic force is carried by a single massless gauge boson, the photon. It acts on particles that carry electric charge: the quarks, the electron, muon, and tau leptons, and the W^+ and W^- (see below). The field theory describing the interaction of the photon with charged particles is known as quantum electrodynamics (QED), and is based on the U(1) symmetry group.

The Weak Force

The weak force is carried by three gauge bosons, the W^+, the W^-, and the Z^0. As its name suggests, it is the weakest of the three forces in the SM. It is the only known force that can change the flavor of a lepton or quark. For example, it is responsible for the radioactive decay of the neutron:

$$n \to p + e + \bar{\nu}_e \tag{1.1.3}$$

At the quark level, this process changes the flavor of one of the neutron's constituent quarks from down to up:

$$d \to u + e + \bar{\nu}_e \tag{1.1.4}$$

The weak force couples to all the fundamental fermions, and is the only SM force that couples to neutrinos, making them both unique and notoriously difficult to detect experimentally. In fact the electromagnetic and weak forces have a unified description known as electroweak theory, or the Glashow, Weinberg, and Salam (GSW) theory (see, among others, Refs. [6, 7, 8]), which is based on a broken SU(2) $\times$ U(1) symmetry.

The weak force couples the up-type quarks to the down-type quarks via the W boson. Unlike the gauge bosons from the electromagnetic and strong forces, the W and Z bosons acquire masses of order $\sim 100\mathrm{MeV}/c^2$ through the Higgs mechanism, which we describe in the next section. However, the quark eigenstates of the weak interaction are not the same as the quark mass eigenstates. Rather, they are related by a matrix transformation:

$$\begin{pmatrix} \tilde{d} \\ \tilde{s} \\ \tilde{b} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1.5}$$

where the states $\tilde{d}$, $\tilde{s}$, and $\tilde{b}$ are the weak interaction eigenstates and d, s, and b are the mass eigenstates. The matrix $\mathbf{V}$ is known as the Cabibbo, Kobayashi, and Maskawa (CKM) matrix. This effective rotation between the weak interaction eigenstates and the mass eigenstates means that the magnitude of the effective coupling for flavor changing processes in the quark sector depends on the transition involved. Experimentally, the magnitudes of these elements are found to be:

$$\begin{pmatrix} |V_{ud}| \approx 0.974 & |V_{us}| \approx 0.225 & |V_{ub}| \approx 0.004 \\ |V_{cd}| \approx 0.225 & |V_{cs}| \approx 0.973 & |V_{cb}| \approx 0.041 \\ |V_{td}| \approx 0.009 & |V_{ts}| \approx 0.041 & |V_{tb}| \approx 0.999 \end{pmatrix} \tag{1.1.6}$$

where the values are based on a combination of various measurements, as tabulated by Reference [9]. From the values, it is clear that transitions within a quark family, (i.e., $t \to b$, $c \to s$, and $d \to u$), are favored over transitions between families.

$\mathbf{V}$ is a 3×3 complex matrix, defined in principle by nine magnitudes and nine phases. However, in order to preserve probability, the matrix must be unitary, thus reducing the degrees of freedom to three rotation angles and six complex phases. Five of the six phases can be absorbed into redefinitions of the quark fields, leaving one irreducible complex phase. The presence of this complex phase results in a potential asymmetry between the behavior of matter and antimatter. In the underlying theory, the substitution of matter into antimatter is a combination of two operations, known as charge-conjugation and parity (C and P), so this asymmetry is known as CP violation. This mechanism of CP violation can only occur if there are three (or more) families of quarks, though it was postulated by Kobayashi and Maskawa (KM) in 1973, when only two families were known. Ultimately, experimental discovery of the third quark family and subsequent confirmation of the KM mechanism of CP violation led Kobayashi and Maskawa to receive the Nobel Prize in Physics in 2008.

1.1.3 The Higgs Boson

The Higgs particle is the only remaining unobserved particle in the SM. The field associated with the Higgs couples to every massive particle in the SM (both the fundamental fermions and the massive $W^{\pm}$ and Z^0 gauge bosons). It is through these interactions, known as the Higgs mechanism, that these particles obtain a mass. Observation of the Higgs boson is a primary goal of many current experiments, including those taking place at the Tevatron [10, 11] and the Large Hadron Collider (LHC) [12]. However, the details of the Higgs mechanism are beyond the scope of this document, as they are not necessary to explore the B meson decay channels of interest.

1.1.4 Standard Model Processes: Trees and Loops

Having now summarized the fundamental fields and interactions described by the SM, we will briefly discuss the Lagrangian formalism, insofar as it is helpful to understand why certain decay channels are of inherent interest. Particle interactions are described by a Lagrangian:

$$L_{int} = \int d^3x(\mathcal{L}_{int} + \mathcal{L}^{\dagger}{}_{int}) \tag{1.1.7}$$

where $\mathcal{L}$ and $\mathcal{L}^{\dagger}$ are the Lagrangian density and its Hermitian conjugate.[4] The matrix element, $\mathcal{M}_{if}$ for a typical process can then be represented as the expectation values of initial and final states i and f, connected by the interaction Lagrangian, and higher orders thereof:

$$\mathcal{M}_{if} = \langle f \,|\, \mathcal{L}_{int} + \mathcal{L}^{\dagger}_{int} \,|\, i \,\rangle + \frac{(-i)^2}{2!} \int \mathrm{d}^4x \, \langle f \,|\, (\mathcal{L}_{int} + \mathcal{L}^{\dagger}_{int})^2 \,|\, i \,\rangle + \ldots \tag{1.1.8}$$

For example, the Lagrangian density for the interaction between the light quarks, u and d, their antiquarks, $\bar{u}$ and $\bar{d}$, and the charged W bosons is given by:

$$\begin{aligned} \mathcal{L} &= i\tfrac{g}{2\sqrt{2}}[W^{+}_{\mu} V_{ud} \bar{u} \gamma^{\mu}(1+\gamma^5) d] \\ \mathcal{L}^{\dagger} &= i\tfrac{g}{2\sqrt{2}}[W^{-}_{\mu} (V^{\dagger})_{ud} \bar{d} \gamma^{\mu}(1+\gamma^5) u] \end{aligned} \tag{1.1.9}$$

where g is the weak coupling constant, and V_{ud} is an element of the CKM matrix.

This Lagrangian describes the interaction, $d \rightarrow u + W^-$.[5] A graphical representation of this process, known as a Feynman diagram, is shown in Figure 1.1. We can

[4] Including the Hermitian conjugate of $\mathcal{L}$ ensures that the Lagrangian is Hermitian, a requirement necessary for the theory to preserve unitarity, and thus probability.

[5] This Lagrangian actually describes any d, u, W process. For example, it also can be used to describe the annihilation of a quark-antiquark, as in $d + \bar{u} \rightarrow W^-$.

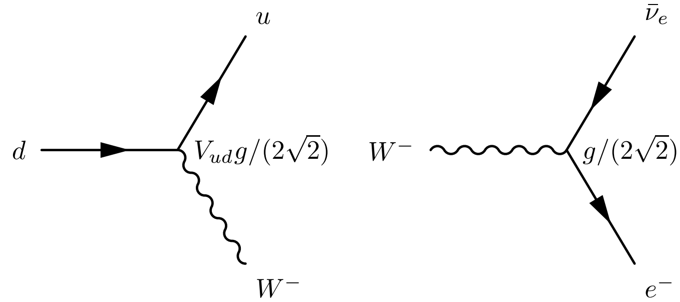

Figure 1.1: Feynman diagrams of the coupling between the down and up quarks and the W boson (left) and the coupling between the e, $\bar{\nu}_e$, and the W (right). The factors $gV_{ud}/(2\sqrt{2})$ and $g/2\sqrt{2}$ represent the coupling coefficients for the respective interactions.

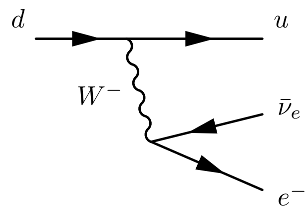

Figure 1.2: Tree-level diagram for the process $d \to u + e^- + \bar{\nu}_e$.

describe more complicated processes by combining such diagrams. For example, the other interaction shown in Figure 1.1, $W^- \to e^- + \bar{\nu}_e$, corresponds to the following piece of the electroweak Lagrangian:

$$\begin{aligned} \mathcal{L} &= i\tfrac{g}{2\sqrt{2}}[W^+_\mu e^+ \gamma^\mu (1+\gamma^5)\nu_e] \\ \mathcal{L}^\dagger &= i\tfrac{g}{2\sqrt{2}}[W^-_\mu \bar{\nu}_e \gamma^\mu (1+\gamma^5) e^-] \end{aligned} \tag{1.1.10}$$

We can then combine these two diagrams, as in Figure 1.2, to describe the β decay of the neutron ($d \to ue^- + \bar{\nu}_e$).

In the language of Eq. 1.1.8, this diagram is the lowest non-zero contribution to the matrix element for the process $d \to u + e^- + \bar{\nu}_e$ (in this case, it is of order $\mathcal{L}^2$, one factor of $\mathcal{L}$ for the $d \to u + W^-$ process from Eq. 1.1.9 and another for $W^- \to e^- \bar{\nu}_e$ from Eq. 1.1.10).

Contributions of this nature, in which the diagrams do not contain closed particle loops, are called tree-level diagrams. Although many processes can be described in this

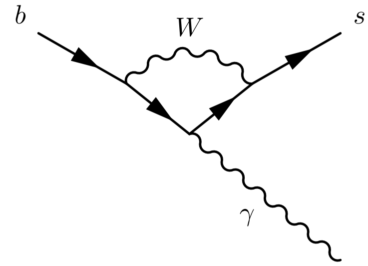

Figure 1.3: A loop diagram, known as a radiative penguin, contributing to the process $b \to s\gamma$.

fashion, some processes can only occur in the presence of closed loops. Diagrams of such processes are called loop diagrams. As an example, the weak interaction does not allow the process $b \to s\gamma$ to occur at tree-level.[6] Instead, this process occurs via loop diagrams known as radiative penguins. Figure 1.3 shows an example of such a diagram. The particles in the loop are not observable, and are instead known as "virtual." Virtual particles are not required to have the mass of their real counterparts, so they are able to contribute at energy scales below their nominal mass. Furthermore, the total amplitude for the process is the sum of all such loop contributions. If there were non-SM particles with couplings such that they could take the place of any of the virtual particles in such a loop diagram, the rate of this decay would be sensitive to them, even if they are too heavy to be produced directly. Significant differences between SM expectations and experimental data for such decays would indicate the presence of new physics. For this reason, decays involving loops are an excellent testbed to search for physics beyond the SM.

1.1.5 Inclusive and Exclusive Decays

Although the fundamental SM processes describe interactions with quarks, they are not directly observable in nature. For example, although we described the quark-level transition $b \to s\gamma$ in the diagram in Figure 1.3, in reality the initial b and final s quarks are bound in hadronic states by the strong force. The b quark could be bound in a B meson, and the s quark in a kaon (K), a combination of s and $\bar{u}$ or $\bar{d}$. The $b \to s\gamma$ process thus encompasses many possible specific hadronic decays, only one of which is $B \to K^*\gamma$. For

[6]For convenience, from here on we will drop the addition symbol when describing processes. So $b \to s + \gamma$ and $b \to s\gamma$ refer to the same process.

this reason, we call the underlying quark-level process an inclusive decay. In contrast, when specific hadronic states are specified, we call it an exclusive decay.

Because theoretical calculations involving mesons are complicated by the underlying dynamics of hadronic systems, predictions for exclusive processes are significantly more difficult to calculate, and suffer from large uncertainties. However, because they involve very specific initial and final states that can be searched for experimentally, they can be measured with high relative accuracy. In contrast, if calculations for part or all of the process are restricted to the quark level, the theoretical uncertainties are reduced. To compare such predictions to experiment, one must measure all possible realizations of the inclusive process. For example, an inclusive measurement of $b \to s\gamma$ should include $B \to K^*\gamma$, $B \to K\pi\gamma$, etc. Although such measurements are difficult, they can be compared to more precise theoretical predictions, and create an excellent laboratory to search for new physics.

1.2 Beyond the Standard Model

We have mentioned two possible windows for testing the SM and searching for new physics: loop processes and inclusive processes. However, we have not yet addressed why we expect physics that extends beyond the SM. Although the SM accurately describes a large number of experimental results, some theoretical arguments imply that it is an incomplete description of the fundamental laws of nature. Further, some observations cannot be accounted for within the framework of the SM.

One of the most obvious deficits of the SM is that it only includes three of the four known fundamental forces; it fails to incorporate gravity. In the same vein, though the SM can describe the electromagnetic and weak forces with one underlying theory, the electroweak force, it does not provide a similar unification for the strong force. The SM also provides no prediction for the number or masses of the fundamental fermions, instead describing them in an ad-hoc manner to match experimental observations.

On the experimental side, one of the strongest pieces of evidence for physics beyond the SM comes from astrophysical observations. A number of these observations suggest a large amount of gravitational matter that cannot be attributed to SM particles. In fact, SM particles account for less than 5% of the energy in the universe [4].

There are many theories and models that propose to supplement the SM to solve these and other theoretical and observational problems. One of the most popular theories is known as supersymmetry, in which each fundamental fermion has a bosonic partner, and vice versa [13]. The presence of these additional particles and their couplings could change the evolution of the SM forces with energy, resulting in their unification at some higher scale. Some supersymmetric theories also propose to describe gravity with the other fundamental interactions. The additional supersymmetric particles could also account for the dark matter whose gravitational influence has been observed by astronomers.

Of course, supersymmetry is by no means the only possible extension of the SM, but it provides an illustrative example of what new particles may exist. One approach to reveal new physics is to search for these new particles, supersymmetric or otherwise, by direct detection. Passive searches include dark matter detection experiments, while active attempts to produce new particles can be conducted in collider environments, such as previous experiments at the Large Electron Positron collider (LEP) and existing experiments at the Tevatron and LHC. New particles could also be detected indirectly, as they may contribute to loop diagrams, as discussed in Section 1.1.4. Their presence could influence the observed rates for certain processes, such as decays of the B meson. Thus, precision measurements of such decays are a valuable tool to test the SM and search for physics beyond.

1.3 B Decays

Conservation of energy dictates that a particle can only decay into a state with lower overall mass. As such, more massive particles tend to have a richer variety of possible decays. In the quark sector, the b quark is the most massive quark that forms bound states.[7] As can be seen from the CKM matrix magnitudes, the strongest hadronic b decays at tree level are expected to be via the $b \to c$ transition, since V_{cb} is an order of magnitude larger than V_{ub}. However, loops allow for other transitions, albeit with smaller matrix elements, such as the $b \to s$ transition we have previously described, and even $b \to d$ decays. Other decays with leptons in the final state are also allowed.

[7]Although the t quark is the most massive of the known quarks, its large mass makes it so unstable that it does not survive long enough to form hadronic bound states.

This rich variety of underlying b decays coupled with the relative ease of production in accelerators (see Section 3.1) makes the B meson an excellent testbed to study the b quark and the predictions of the SM. For our purposes, the primary observables of interest for a B meson decay are its branching fraction and its direct CP asymmetry.

1.3.1 Branching Fraction

A particle or state with a finite lifetime has a corresponding spread in its rest mass. The lifetime, τ, is related to this spread, or width, Γ, by $\Gamma = \hbar/\tau$. The width for a given decay is related to the square of the matrix element for that process. The total width is the sum over each possible decay of the state or particle:

$$\Gamma = \sum_i \Gamma_i \tag{1.3.1}$$

or in terms of the lifetime:

$$\frac{1}{\tau} = \hbar \sum_i \frac{1}{\tau_i} \tag{1.3.2}$$

In some cases it is practical to measure the width or lifetime directly. In our case, where we will have a large source of B and $\bar{B}$ mesons, it is more convenient to measure the fraction of decays that proceed via a particular channel, known as a branching fraction, $\mathcal{B}$. For example, for a B^0 meson decay, $B^0 \to K^+\pi^-$, we define the experimentally measured branching ratio as:

$$\mathcal{B}(B^0 \to K^+\pi^-) = \frac{N(B^0 \to K^+\pi^-)}{N(B^0 \to X)} \tag{1.3.3}$$

where the numerator is the total number of $B^0 \to K^+\pi^-$ decays that occured and the denominator is the total number of B^0 decays to any final state.[8] In terms of the decay width, this corresponds to:

$$\mathcal{B}(B^0 \to K^+\pi^-) = \frac{\Gamma(B^0 \to K^+\pi^-)}{\sum_i \Gamma_i} \tag{1.3.4}$$

where the summation is over all possible decays.

[8]Note that for the remainder of this work, charge conjugate modes are implied unless otherwise stated. So in this example, the branching fraction $B^0 \to K^+\pi^-$ is the sum of the number of $B^0 \to K^+\pi^-$ and $\bar{B}^0 \to K^-\pi^+$ decays, divided by the total number of B^0 and $\bar{B}^0$ decays.

1.3.2 Direct CP Asymmetry

As alluded to in Section 1.1.2, the complex phase in the CKM matrix allows for an asymmetry between matter and antimatter, known as CP asymmetry. In B decays, CP asymmetry can be observed indirectly, via mixing between B^0 and $\bar{B}^0$, or directly, via an asymmetry between the rate of a B^0 decay and the conjugate $\bar{B}^0$ decay, or similarly a B^+ decay and the conjugate B^- decay. Experimentally, direct CP asymmetry is defined as:

$$\mathcal{A}_{CP} = \frac{\mathcal{B}(b) - \mathcal{B}(\bar{b})}{\mathcal{B}(b) + \mathcal{B}(\bar{b})} \tag{1.3.5}$$

where the states with a b correspond to B^- or $\bar{B}^0$ and those with $\bar{b}$ correspond to B^+ or B^0, and the decays are the appropriate charge conjugate states of one another. Theoretically, this can be expressed in terms of the decay widths,

$$\mathcal{A}_{CP} = \frac{\Gamma(b) - \Gamma(\bar{b})}{\Gamma(b) + \Gamma(\bar{b})} \tag{1.3.6}$$

In both the experimental and theoretical cases, the asymmetry has the benefit of being a ratio of similar terms. For example, in the experimental case, many uncertainties in measuring the branching fraction will be the same for both the b and $\bar{b}$ states, and will thus cancel out in a determination of $\mathcal{A}_{CP}$. Similarly, uncertainties due to, for example, hadronic form factors, may cancel out in the theoretical expression, resulting in more precise theory predictions.

Chapter 2

The η and η' Mesons

We mentioned a few hadrons specifically in the previous chapter. To motivate this decay analysis, we should discuss some of the propeties of the mesons in more detail. Specifically, we will discuss mesons made out of u, d, and s quarks, with a special focus on the η' and η, including the $b \to s$ transitions involving these mesons, $B \to X_s\eta'$ and $B \to X_s\eta$.

2.1 Light (u,d,s) Mesons

Although the underlying theory of QCD is based on an SU(3) symmetry of quark and gluon colors, the light hadron spectrum can be understood as an approximate symmetry between the light quark flavors (u, d, and s), known as flavor SU(3). This approach allows us to classify hadrons into groups based on representations of SU(3). Although it is obvious from the variations in the masses within these subgroups that this symmetry is not exact, it nonetheless provides some valuable insights into some properties of the hadrons. We focus here on the light mesons.

By combining quark anti-quark pairs, we can create a spin-0 pseudoscalar or spin-1 pseudovector meson ground state. If we restrict ourselves to the light quarks, we expect a total of nine possible quark anti-quark combinations:

$$\begin{pmatrix} u \\ d \\ s \end{pmatrix} \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix} = \begin{pmatrix} u\bar{u} & u\bar{d} & u\bar{s} \\ d\bar{u} & d\bar{d} & d\bar{s} \\ s\bar{u} & s\bar{d} & s\bar{s} \end{pmatrix} \tag{2.1.1}$$

Meson	Quark content	Mass (MeV/c^2)
π^0	$\frac{1}{2}(u\bar{u} - d\bar{d})$	135.0
π^+	$u\bar{d}$	139.6
π^-	$d\bar{u}$	
K^+	$u\bar{s}$	493.7
K^-	$s\bar{u}$	
K^0	$d\bar{s}$	497.6
$\bar{K}^0$	$s\bar{d}$	
η_8	$\frac{1}{2\sqrt{3}}(u\bar{u} + d\bar{d} - 2s\bar{s})$	(See Section 2.2)
η_0	$\frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} + s\bar{s})$	
η	(See Section 2.2)	547.9
η'		957.8

Table 2.1: Mesons of the pseudoscalar nonet, including their quark content and masses.

The underlying transformation rules of the theory, known as flavor SU(3), dictate that the pseudoscalar eigenstates are a combination of these quark anti-quark pairs.[1] These nine states are known as the pseudoscalar nonet. Their quark content and masses are given in Table 2.1. The first eight are known as the pseudoscalar octet, since they form an octet representation of SU(3). The remaining state, the η_0, is a singlet representation of SU(3), and is hence known as the pseudoscalar singlet.

2.2 The η and η'

The η_8 and η_0 are not experimentally observed states. Rather, we observe linear combinations of these two states known as the η and η'. This mixture can be characterized by a mixing angle θ:

$$\begin{pmatrix} \eta \\ \eta' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \eta_8 \\ \eta_0 \end{pmatrix} \tag{2.2.1}$$

The value of this η–η' mixing angle must be determined experimentally. One theoretical approximation relates the mixing angle to the masses of η and η':

$$\tan^2\theta = \frac{m_\eta^2 - \frac{1}{3}(4m_K^2 - m_\pi^2)}{\frac{1}{3}(4m_K^2 - m_\pi^2) - m_{\eta'}^2} \tag{2.2.2}$$

[1] Each of these eigenstates has a pseudovector equivalent, but we restrict our discussion to the pseudoscalar states.

This gives the value $|\theta| \approx 11^\circ$. Other estimations give a range of values: $-10^\circ < \theta < -20^\circ$. Though the exact values vary, the accepted range of θ indicates that the η is predominantly composed of the octet η_8 and the η' of the singlet η_0.

2.3 η' and the QCD Anomaly

For a classical field theory, we expect by Noether's theorem a conserved current, j^μ, associated with each continuous symmetry of the action. When such a current fails to survive quantization of the theory, it is called an anomaly. The QCD axial-vector current, $j^{\mu 5}$, is one such anomaly, with a divergence given by [14]:

$$\partial_\mu j^{\mu 5} = \sum_{q=u,d,s} 2im_k q\gamma_5 \bar{q} - \frac{3\alpha_S}{4\pi} G_{\alpha\beta}\tilde{G}^{\alpha\beta} \tag{2.3.1}$$

where q and $\bar{q}$ denote the light quark and anti-quark fields, $G_{\alpha\beta}$ is the gluonic field strength, and $\tilde{G}_{\alpha\beta}$ is its dual. The first term is a sum over the three flavors of light quarks and their anti-quarks, and can thus be interpreted as the η_0. Thus, this equation implies a coupling between the η_0 and a two-gluon (also known as gluonium, or gg) state.

Returning to the physical states, the η', being composed of mostly η_0, should also couple strongly to gluons. One interpretation is that the η' is actually a mixture of η_0,η_8, and gg, and it is this presence of a significant gg component that is responsible for the increased mass of the η' relative to the other members of the pseudoscalar nonet. Under this interpretation, as much as $\sim 25\%$ of the η' could be composed of gluonium [14]. The presence of η–η' mixing indicates that the η could also have some gluonium component, but since the η is made up primarily of the octet η_8, its gluonium content is usually assumed to be negligible.

2.4 $B \to X_s\eta'$

A possible candidate to observe the enhanced $\eta' gg$ coupling is in $b \to sg^*$ processes, where a b quark decays to an s quark and an intermediate unstable gluon state g^*, which in turn could couple via the $\eta' gg$ anomaly mechanism to produce an η'. The diagram for such a $b \to s\eta' g$ transition is shown in Figure 2.1. This class of diagram, where a gluon is emitted from the loop, is known as a strong penguin.

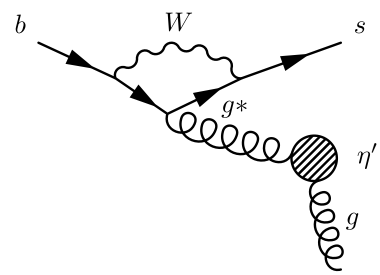

Figure 2.1: A $b \to s\eta' g$ QCD penguin diagram.

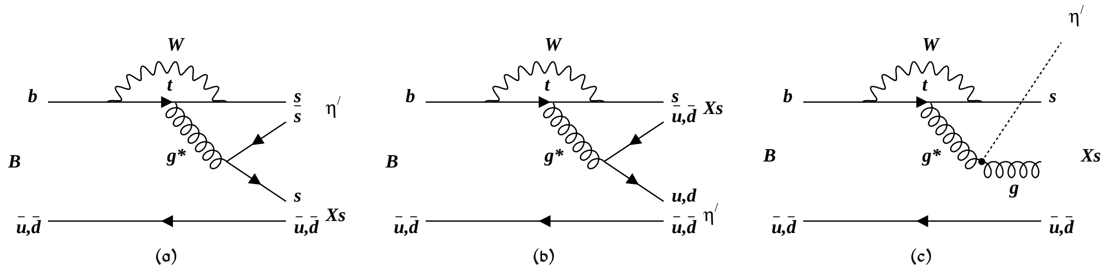

Figure 2.2: (Left, Middle) Two $B \to X_s\eta'$ diagrams that occur via "two-body" mechanisms that do not involve the QCD anomaly coupling. (Right) The QCD anomaly contribution to $B \to X_s\eta'$.

As discussed in Section 1.1.5, the quarks and gluons in Figure 2.1 cannot be seen directly. Instead, a candidate for experimental searches could contain the initial b quark bound with a $\bar{u}$ or $\bar{d}$ in a B meson. The final state gluon would create one or more $q\bar{q}$ pairs, giving an overall final state of an η' and an outgoing hadronic system with a net strangeness of one, which we denote as X_s. The experimentally observable decay is thus $B \to X_s\eta'$. Unfortunately, observation of this decay alone does not necessarily imply that the anomaly mechanism is responsible. For example, Figure 2.2 indicates three possible diagrams that could contribute to the decay. As the underlying processes are different, they will lead to different distributions of other observables, such as M_{X_s}. Thus, it is important to determine both the branching fraction as well as the mass spectrum of the X_s system.

The CLEO collaboration peformed the first such measurement in 1998 [15] using a data sample of 3.3×10^6 $B\bar{B}$ pairs and found a branching fraction of $\mathcal{B}(B \to X_s\eta'; p^*_{\eta'} > 2.0\mathrm{GeV}/c) = (6.2 \pm 1.6(\mathrm{stat}) \pm 1.3(\mathrm{syst})^{+0.0}_{-1.5}(\mathrm{bkg})) \times 10^{-4}$, where the restriction on the

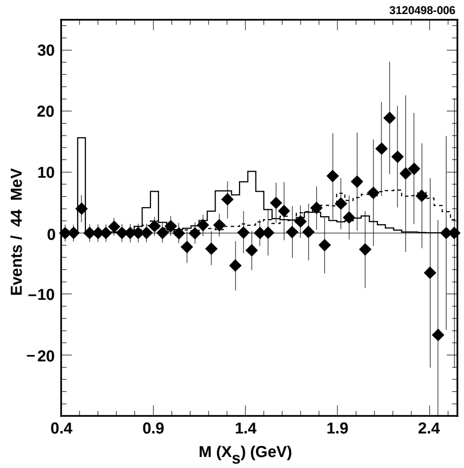

Figure 2.3: Spectrum of M_{X_s} observed by CLEO for $B \to X_s\eta'$ in Ref. [15]. Data points with error bars are the measurement result, corrected for detection efficiency and based on a subtraction of $q\bar{q}$ backgrounds. Histograms represent the expectation from $b \to sq\bar{q}$ (solid) and $b \to s\eta' g$ (dashed).

center-of-mass (CM) momentum of the η', $p^*_{\eta'}$, was chosen to reduce backgrounds from $b \to c$ decays. This branching fraction was considered anomalously large, and the mass spectrum of the inclusive X_s state exhibited a peak above 2.0 GeV/c^2, as shown in Figure 2.3.

This measurement inspired significant theoretical interest. An initial calculation was performed using conventional factorization approaches [16], but the resulting branching fraction fell short of the observed value by nearly a factor of five, and gave a mass spectrum which peaked at a significantly lower X_s mass. Other calculations included more exotic possibilites: the presence of a significant $c\bar{c}$ component within the η' [17], and possible contributions from the QCD anomaly [18]. However, the $c\bar{c}$ explanation predicts that the rate of $B \to \eta' K$ should be roughly half the rate of $B \to \eta' K^*$, but current experimental averages [1] indicate that the latter is actually suppressed relative to the former by more than an order of magnitude. Likewise, the anomaly explanation is similarly disfavored, as it requires a form factor for $\eta' gg$ that is incompatible with measurements of $\Upsilon(1S) \to \eta' X$ [19].

$B \to X_s\eta'$ was later measured in an improved analysis by CLEO with 9.7×10^6 $B\bar{B}$ pairs [20] and independently by the BaBar collaboration with 88.4×10^6 $B\bar{B}$ pairs [21], giving branching fractions lower than, but consistent with, the original measurement:

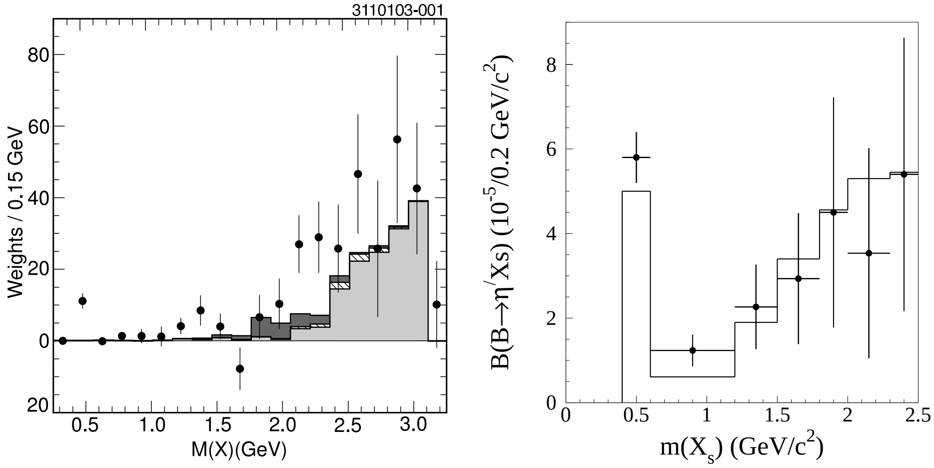

Figure 2.4: Observed X_s mass spectra in measurements of $B \to X_s\eta'$ from CLEO [20] (left), and BaBar [21] (right). For both plots, the points with error bars are the data. In the left plot, the histograms indicate expected background contributions from $b \to c \to s$ cascade decays (light gray), $B \to D^{(*)0}\eta'$ decays (dark gray), and $B \to D^{**0}\eta'$ (hatched). In the right plot, the histogram is the expectation from nonresonant $B \to X_s\eta'$.

$(4.6 \pm 1.1(\mathrm{stat}) \pm 0.4(\mathrm{syst}) \pm 0.5(\mathrm{bkg})) \times 10^{-4}$, and $(3.9 \pm 0.8(\mathrm{stat}) \pm 0.5(\mathrm{syst}) \pm 0.8(\mathrm{model})) \times 10^{-4}$, respectively. The mass spectra of these measurements, shown in Figure 2.4, confirm the peaking at high M_{X_s}. The currently accepted average for $B \to X_s\eta'$ is, as of this writing, [1]: $(4.2 \pm 0.9) \times 10^{-4}$. There is still no conclusive agreement within the community as to the underlying nature of the signal. Some theoretical treatments [22, 23] imply that the decay rate could be indicative of contributions from new physics.

2.5 $B \to X_s\eta$

The previously described η–η' mixing phenomenon suggests that a measurement of $B \to X_s\eta$ could help to clarify the underlying contributions to $B \to X_s\eta'$. For example, if the explanation is strongly linked to the QCD anomaly, then the rate for the η mode should be suppressed relative to the η' mode by the square of the tangent of the η–η' mixing angle, $\tan^2\theta \sim 0.1$. A more recent theoretical treatment of both the η' and η modes using soft-collinear effective theory includes possible contributions from nonperturbative diagrams involving charm quarks, though it is not possible to determine the extent to which these diagrams contribute without a measurement of the η mode [24].

The only previous search for $B \to X_s\eta$ was performed by CLEO with a data sample of 3.3×10^6 $B\bar{B}$ pairs, and placed an upper limit of $\mathcal{B}(B \to X_s\eta) < 4.4 \times 10^{-4}$ [15]. Though this is consistent with the QCD anomaly expectation, the limit is not restrictive enough to make any definite conclusions. The Belle detector, having collected hundreds of millions of $B\bar{B}$ pairs, is thus in an excellent position to measure $B \to X_s\eta$ and potentially clarify a piece of the η–η' puzzle.

Chapter 3

Experimental Methods and Apparatus

In order to measure the $B \to X_s\eta$ process, we require a source of B mesons and a method of detecting their decay products. In this chapter, we discuss production of B mesons with the $\Upsilon(4S)$ system at the KEKB collider, and the subsequent detection of B decays with the Belle detector.

3.1 The $\Upsilon(4S)$ and B Mesons

The Υ system is a mesonic state composed of b and $\bar{b}$, and was discovered at Fermilab in 1977 [25]. Because of the large mass of the b quark relative to the binding energy provided by the strong interaction, the $b\bar{b}$ pair in the Υ can be treated as nonrelativistic. This treatment allows the Υ system to be understood in analogy to the positronium system, a short lived e^+e^- bound state. As in positronium, the result is a spectrum of bound states.

Of these, the $\Upsilon(1S)$ ground state and its radial excitations ($\Upsilon(2S), \Upsilon(3S), \ldots$) are readily produced in e^+e^- collisions. The measured hadronic cross sections as a function of energy for the first four Υ states are shown in Figure 3.1 The $\Upsilon(4S)$ is the first of these states that is heavier than the mass of a $B\bar{B}$ pair. As such, it decays almost entirely ($>$ 96%) via $\Upsilon(4S) \to B\bar{B}$. Of these decays to B meson pairs, they are split almost evenly into $\Upsilon(4S) \to B^+B^-$ and $\Upsilon(4S) \to B^0\bar{B}^0$ (approximately 52% and 48%, respectively) [1]. Accelerators that exploit these favored decays through the process $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ are known as B factories. Two B factory experiments have been underway over the last decade: the BaBar experiment at the PEP-II collider at SLAC in California, and the Belle experiment at the KEKB collider in Tsukuba, Japan.

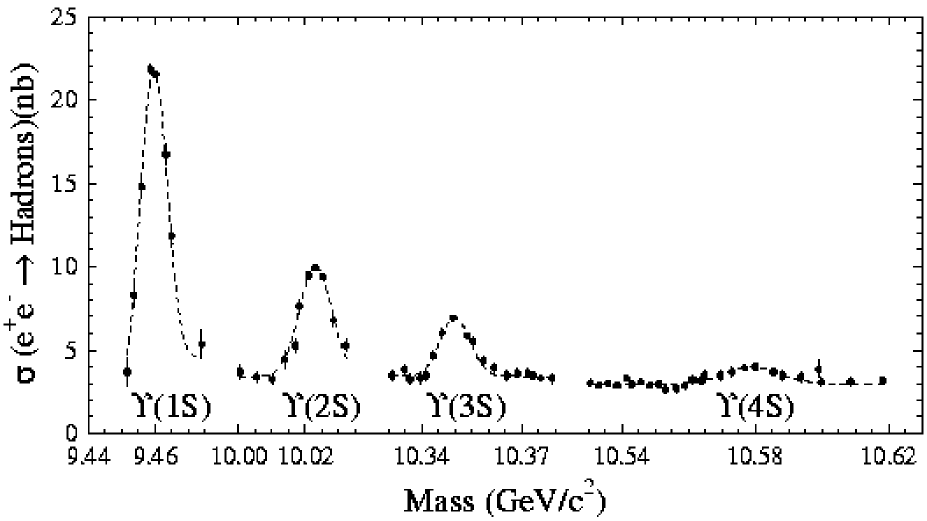

Figure 3.1: The Υ states, as observed via the cross section for $e^+e^- \rightarrow \mathrm{hadrons}$, as measured by the CUSB collaboration [26, 27] at CESR.

3.2 The KEKB B Factory

KEKB is an electron-positron collider located at the High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan [28]. The collider, shown schematically in Figure 3.2, consists of an 8 GeV electron storage ring, known as the high energy ring (HER), and a 3.5 GeV positron storage ring, known as the low energy ring (LER), each approximately 3 km in circumference.[1] Electrons and positrons from the two rings collide with a 22 mrad crossing angle at the interaction point (IP) and a nominal CM energy tuned to the $\Upsilon(4S)$ resonance at 10.58 GeV.

At the time of its design, the luminosity of KEKB was targeted at $1 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. It began operation in 1998 and achieved design luminosity in May 2003. Operational improvements and equipment upgrades led to significant increases in luminosity, culminating in a world record of $2.11 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in June 2009. During its operating period from 1998 through 2010, KEKB delivered an integrated luminosity of more than $1\ \mathrm{ab}^{-1}$. Most of this luminosity was delivered at the energy of $\Upsilon(4S)$, giving a total of over 771 million $B\bar{B}$ pairs. The subsequent B decays were recorded by the Belle detector, located at the IP.

[1]The asymmetry between the two beams was chosen to create a net boost in the lab frame, allowing measurement of decay-time differences between B^0 and $\bar{B}^0$, known as time-dependent CP violation studies. However, this type of analysis is not the subject of this work.

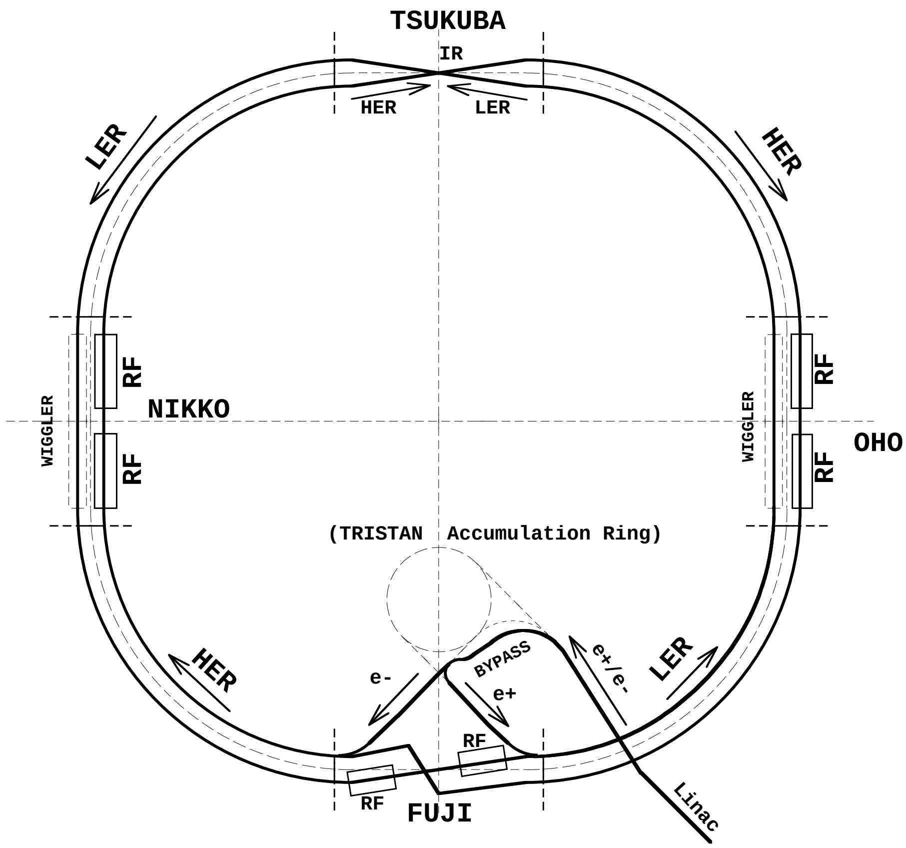

Figure 3.2: A schematic of the KEKB collider complex. Belle is located in the Tsukuba interaction region.

3.3 The Belle Detector

The Belle detector [29], shown in Figure 3.3, is a collection of sub-detectors built around the IP of the KEKB collider. The coordinate system for Belle is defined with the z-axis antiparallel to the positron beam and the x-axis pointing inward, toward the center of the KEKB storage rings. For convenience, we often use polar coordinates (θ, ϕ, and r) with polar angle θ defined as the angle away from the z-axis. The detector subsystems cover a full 2π in ϕ and three ranges in θ: the barrel region ($34^\circ < \theta < 127^\circ$), the forward endcap ($17^\circ < \theta < 34^\circ$), and the backward endcap ($127^\circ < \theta < 150^\circ$). Information from the Belle subdetectors is combined and used to search for specific B meson decays via a technique known as reconstruction.

3.3.1 Decay Reconstruction

Our primary tool to identify specific decays is known as reconstruction. The basic principle is to utilize kinematics and known particle properties to search for a specific process. For example, to identify a decay $X \to Y + Z$, we can search for the decay products Y and Z (preferrably coming from a common decay vertex). However, there could be other processes, such as $X \to Y + Z + \alpha$ or $F \to Y + Z$. These processes also have Y and Z in the final state, so it is not enough to simply identify the final state particles.

We can increase our certainty that we have observed the process of interest if we can verify the conservation of four-momentum:

$$(p_X)_\mu = (p_Y)_\mu + (p_Z)_\mu \tag{3.3.1}$$

Squaring the above relation and expanding, using natural units,

$$\begin{aligned}
&(p_X)_\mu (p_X)^\mu = (p_Y + p_Z)_\mu (p_Y + p_Z)^\mu \\
&m_X = \sqrt{(p_Y)_\mu (p_Y)^\mu + (p_Z)_\mu (p_Z)^\mu + (p_Y)_\mu (p_Z)^\mu + (p_Z)_\mu (p_Y)^\mu} \\
&m_X = \sqrt{m_Y^2 + m_Z^2 + 2(E_Y E_Z - \mathbf{p}_Y \cdot \mathbf{p}_Z)}
\end{aligned} \tag{3.3.2}$$

where E, $\mathbf{p}$, and m are the relevant energy, three-momentum, and rest mass, respectively. The right hand side of the above relation is known as the invariant mass of the Y, Z system. If Y and Z are truly products of the decay $X \to Y + Z$, then their invariant mass should be consistent with the mass of X.

As can be seen from above, the invariant mass is completely specified by measurements of m and $\mathbf{p}$ for each decay product, since the energy can then be deduced from

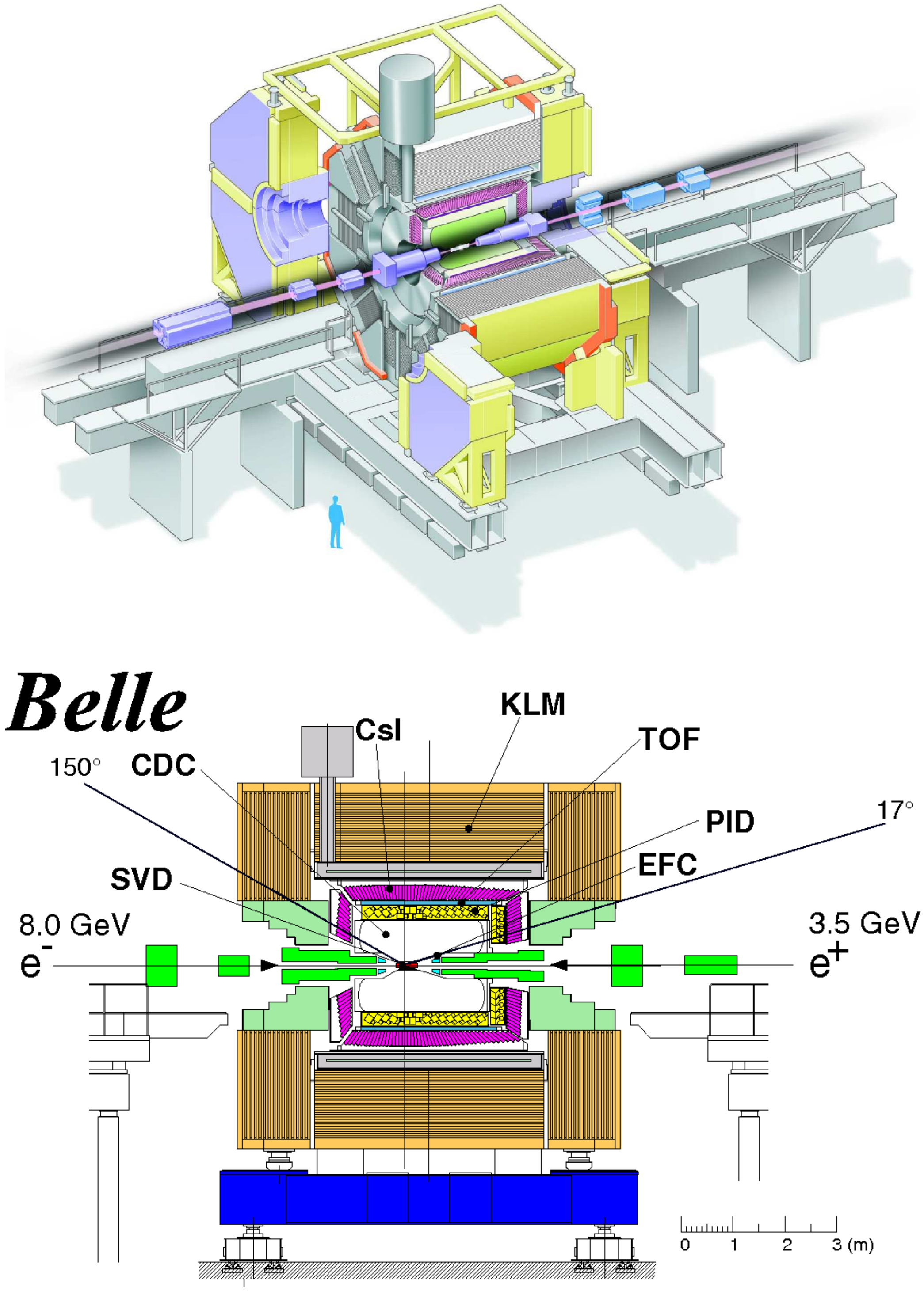

Figure 3.3: Isometric (top) and side (bottom) views of the Belle detector. In the isometric view, a human figure is added for scale.

relativistic relations. Thus, each Belle subsystem is designed to determine one or more of these values. Momentum of charged particles can be measured through magnetic spectroscopy. The momentum of neutral particles is determined by an energy measurement, performed using electromagnetic calorimetry. Finally, the mass of each decay product can be determined by systems that perform particle identification. All systems must be triggered to collect the desired data with a high efficiency while keeping the background trigger rate to a manageable level. Finally, the information from all these systems must be recorded and processed to allow the end user to reconstruct B meson decays in a given analysis.

3.3.2 Magnetic Spectroscopy

A charged particle traveling in a magnetic field follows a helical path with a radius of curvature given by

$$R = \frac{|\mathbf{p}_T|}{0.3B} \tag{3.3.3}$$

where $\mathbf{p}_T$ is the component of the particle momentum transverse to the magnetic field, given in GeV/c, and B is the magnitude of the magnetic field, given in Tesla. In Belle, a 1.5 T magnetic field is supplied along the z-axis by a superconducting solenoid with length 4.4 m and an inner radius of 1.7 m. Two detectors are used to track charged particles through the magnetic field: the silicon vertex detector (SVD) and the central drift chamber (CDC).

Silicon Vertex Detector (SVD)

The SVD is placed just outside the beam pipe, and as such is the closest detector to the IP. As its name suggests, its primary purpose is to supply spatial information on charged particles to determine the B-meson decay vertices. It also contributes to the overall tracking system. It consists of a series of reversed bias diode strips with their central region depleted of charge carriers. When a charged track passes through this region, it creates electron-hole pairs that drift to the electrodes, thus creating a current and indicating a hit. Data from two SVD configurations is used in this decay study. The first configuration, SVD1, was used through 2003. SVD1 covers the angular range $23° < \theta < 139°$ with three concentric cylindrical layers, each with a number of double-sided silicon strip detectors (DSSDs). Strips on each side of a DSSD are arranged perpendicular to those on the other side. This allows measurement of the z-position on one side and the $r - \phi$ position on

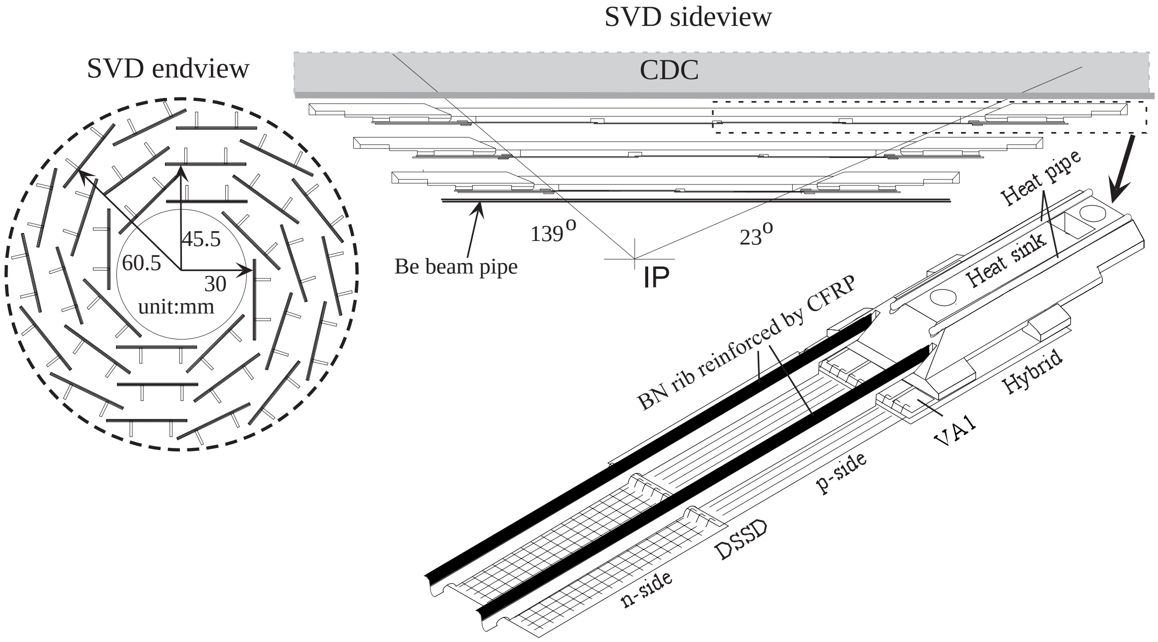

Figure 3.4: The first SVD configuration, SVD1, used through summer 2003.

the other. This configuration is shown in Figure 3.4. In the summer of 2003, SVD1 was removed and the SVD2 configuration was installed. SVD2 has improved angular acceptance, from $17^\circ < \theta < 150^\circ$ and a total of four concentric cylinders of DSSD detectors, with the innermost layer 0.5 cm closer to the IP than SVD1.

The performance of each configuration is measured by the resolution on the point of closest approach to the IP, known as the impact parameters. This is divided into two pieces: the $r - \phi$ and z resolutions. The $r - \phi$ resolution, $\sigma_{r\phi}$, of each configuration is given by

$$\begin{aligned}
[\mathrm{SVD1}]\sigma_{r\phi}(\mu m) &= 19.2 \oplus \tfrac{54.0}{p\beta \sin^{3/2}\theta} \\
[\mathrm{SVD2}]\sigma_{r\phi}(\mu m) &= 21.9 \oplus \tfrac{35.5}{p\beta \sin^{3/2}\theta}
\end{aligned} \tag{3.3.4}$$

and the z resolution is given by

$$\begin{aligned}
[\mathrm{SVD1}]\sigma_{z}(\mu\mathrm{m}) &= 42.2 \oplus \tfrac{44.3}{p\beta \sin^{5/2}\theta} \\
[\mathrm{SVD2}]\sigma_{z}(\mu\mathrm{m}) &= 27.8 \oplus \tfrac{31.9}{p\beta \sin^{5/2}\theta}
\end{aligned} \tag{3.3.5}$$

where in both cases the $\oplus$ operator indicates the addition of the two components of the resolution in quadrature.

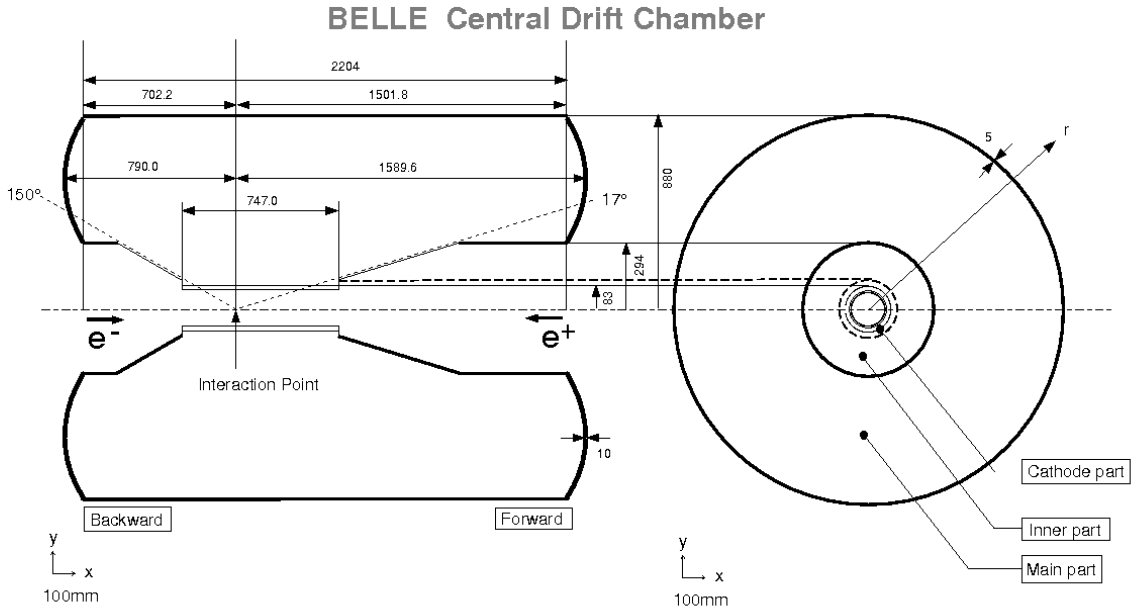

Figure 3.5: Overview of the CDC structure. Lengths are in mm.

Central Drift Chamber (CDC)

The next detector outward from the SVD is the CDC, shown in Figure 3.5. The CDC covers the angular range from $17^\circ < \theta < 150^\circ$. Each of its 8400 drift cells consists of a sense wire, held at a fixed positive voltage, surrounded by field wires, held at a fixed negative voltage. The CDC is filled with a 50% helium, 50% ethane (C_2H_6) mixture. Charged particles passing through the gas ionize electrons, which are accelerated by the electric field. As they accelerate, they produce secondary ionizations in the gas, and the resulting avalanche is collected by the sense wires. The time of the hits in the CDC relative to the event start time is combined with the known drift-velocity of electrons in the gas, and the distance of closest approach to the sense wire is derived. The series of hits left by a particle as it passes through the CDC and, optionally, hits from the SVD, allow the helix parameters of the charged track to be reconstructed. To maximize the z-resolution of the track fits, some of the drift cells are rotated by a small angle away from the z-axis.

Ultimately, the CDC is able to measure the transverse momentum, p_T, of most tracks to better than 1%. The p_T resolution of the CDC alone is given by

$$\frac{\sigma_{p_T}}{p_T}(\%) = (0.28 p_T) \oplus \left(\frac{0.35}{\beta}\right) \tag{3.3.6}$$

with p_T given in GeV/c. If information from the SVD is added, the resolution improves to

$$\frac{\sigma_{p_T}}{p_T}(\%) = (0.19 p_T) \oplus \left(\frac{0.30}{\beta}\right) \tag{3.3.7}$$

The CDC also provides information to the particle identification (PID) system. We will discuss this more in Section 3.3.4.

3.3.3 Electromagnetic Calorimetry

Electrons interacting in a material lose energy via ionization loss and bremsstrahlung radiation. If the energy of the electron is sufficient, the bremsstrahlung photons can cause e^+e^- pair-production. The resulting e^+ and e^- can create further ionizations and bremsstrahlung photons, continuing the process. Thus, a high energy incident electron or photon causes an electromagnetic shower. If the material is doped with a fluor, the ionization energy losses from the shower are converted into visible light, which can in turn be measured by a photodetector.

Most of the photons detected at Belle are the final products of cascade decays. As a result, they have relatively low energies of $\lesssim 500\ \mathrm{MeV}$. However, some decay modes have high energy photons (up to $\sim 4\ \mathrm{GeV}$) that are direct products of a two-body B decay, such as $B \to K^*\gamma$. Belle's calorimeters must therefore perform with excellent resolution over a wide energy range.

Electromagnetic Calorimeter (ECL)

The ECL, shown in Figure 3.6, is Belle's primary electromagnetic calorimeter. It consists of 8,736 cesium iodide crystals, doped with thallium as a fluor (CsI(Tl)), divided over three regions: the barrel ($32.2^\circ < \theta < 128.7^\circ$, 6642 crystals), the forward endcap ($12.4^\circ < \theta < 31.4^\circ$, 1152 crystals), and the backward endcap ($130.7^\circ < \theta < 155.1^\circ$, 960 crystals). The shape of each crystal varies by position in the ECL, but a typical crystal has a tower-like shape (see Figure 3.7), with a front face of $55\ \mathrm{mm} \times 55\ \mathrm{mm}$, a rear face of $65\ \mathrm{mm} \times 65\ \mathrm{mm}$, and a length of 30 cm. This length corresponds to approximately

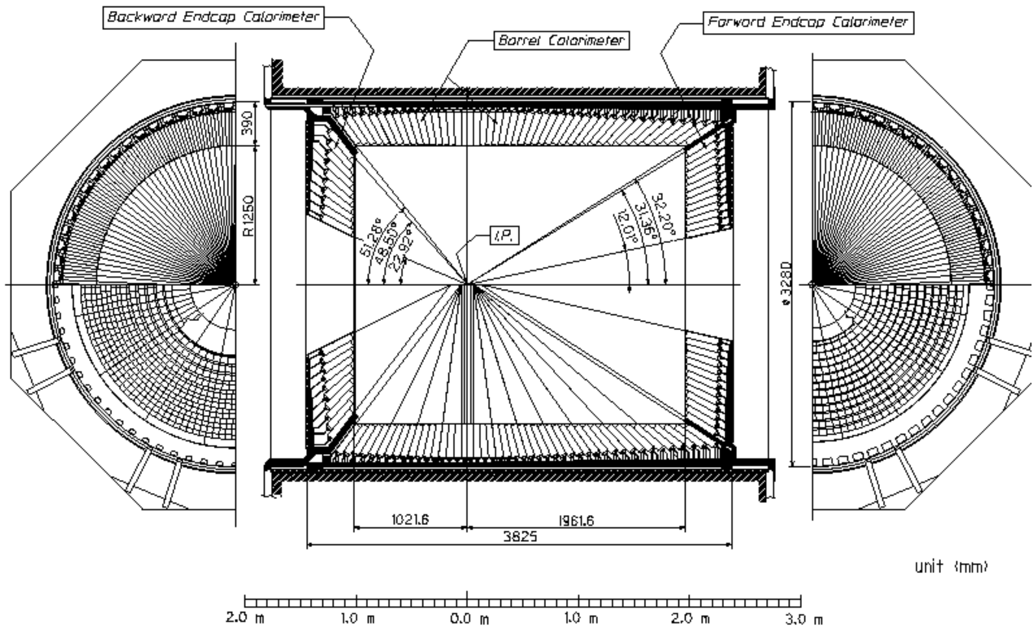

Figure 3.6: Overall configuration of the Belle ECL.

16.2 radiation lengths, or enough to reduce the energy of an incident particle by about $1/e^{16.2}$. This length ensures that showers from high energy particles are contained within the crystals, thus maintaining the energy resolution.

The angular coverage is the same as that of the CDC, allowing charged tracks to be matched with clusters of hits in the ECL. Photons can thus be identified by an ECL shower with no matching charged track from the CDC and a shower profile consistent with an electromagnetic cascade.[2] The position resolution of the ECL is

$$\sigma(\mathrm{mm}) = 0.27 \oplus \frac{3.4}{\sqrt{E}} \oplus \frac{1.8}{\sqrt[4]{E}} \tag{3.3.8}$$

and the energy resolution is

$$\frac{\sigma_E}{E}(\%) = \left(\frac{0.066}{E}\right) \oplus \left(\frac{0.81}{\sqrt[4]{E}}\right) \oplus 1.34 \tag{3.3.9}$$

with E given in GeV.

[2]The characteristic profile of an electromagnetic shower is notably narrower than that of a shower produced by an incident hadron.

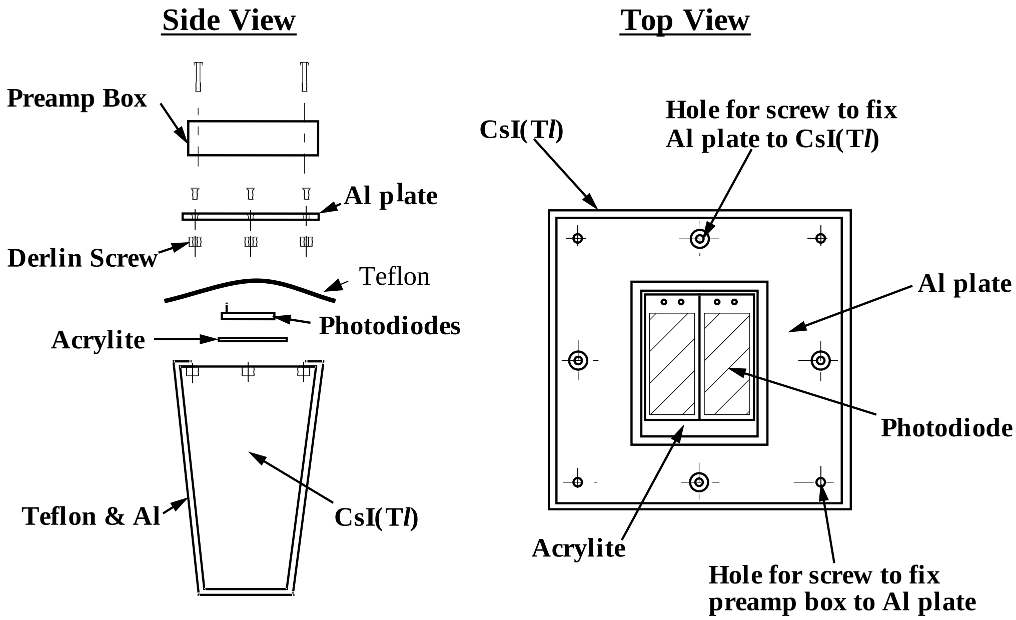

Figure 3.7: A typical ECL crystal assembly.

Extreme Forward Calorimeter (EFC)

Figure 3.8 shows the extreme forward calorimeter, which detects electromagnetic showers in the far forward and backward regions: $6.4^\circ < \theta < 11.5^\circ$ and $163.3^\circ < \theta < 171.2^\circ$, respectively.

The EFC's proximity to the beam pipe and the IP results in higher backgrounds than the ECL. To compensate, the EFC uses crystals of bismuth germanate ($Bi_4Ge_3O_{12}$, or BGO) due to their higher radiation tolerance. The EFC is not used in decay reconstruction. However, its geometric location allows it to act as a beam mask to reduce radiation backgrounds to the CDC, as well perform luminosity and background monitoring.

3.3.4 Particle Identification (PID)

Fits to the helix of a charged track can determine the particle momentum with high precision, but provide no information on the mass, and thus the identity of the particle. However, there are a number of measurements that can distinguish between particle types.

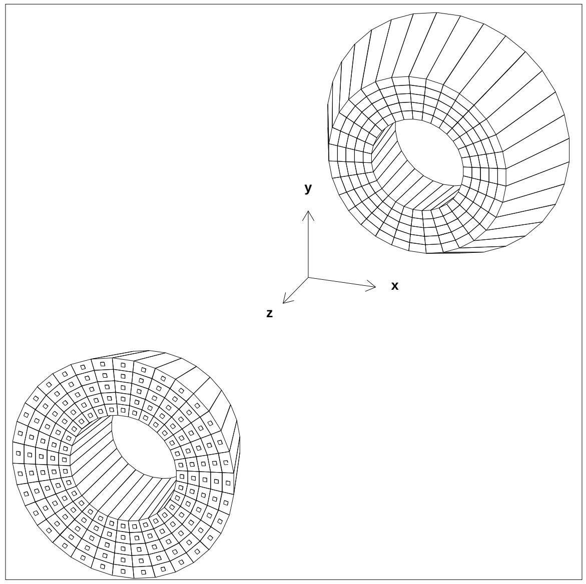

Figure 3.8: An isometric view of the forward and backward EFC.

One such measurement is the ionization energy loss of the particle per unit length traversed in a medium, dE/dx. This energy loss is given by the Bethe-Bloch formula,

$$\frac{dE}{dx} = \frac{4\pi N_0 z^2 \alpha^2}{mv^2} \frac{Z}{A} \left\{ \ln \left[\frac{2mv^2}{I(1-\beta^2)} \right] - \beta^2 \right\} \tag{3.3.10}$$

where m is the electron mass, v and z are the charge (in units of the electron charge) and velocity of the particle, N_0 is Avagadro's number, Z and A are the atomic number and mass number of the material, and I is an ionization potential of the medium. Equation 3.3.10 depends only on the velocity of the charged particle (v and β), so a measurement of dE/dx can in principle determine the velocity of the particle. Since we already measureme the momentum from the track helix, we can calculate the mass.

$$m = |\mathbf{p}| \sqrt{\frac{1}{\beta^2} - 1} \tag{3.3.11}$$

In practice, the mass is not directly calculated. Rather, the dE/dx distributions themselves are used to distinguish the particles.

Hits in the CDC provide measurements of dE/dx at Belle. Distributions of hits corresponding to various particle species are shown in Figure 3.9. The resulting PID capabilities are excellent below momenta of 1 GeV/c, and show some limited discrimination

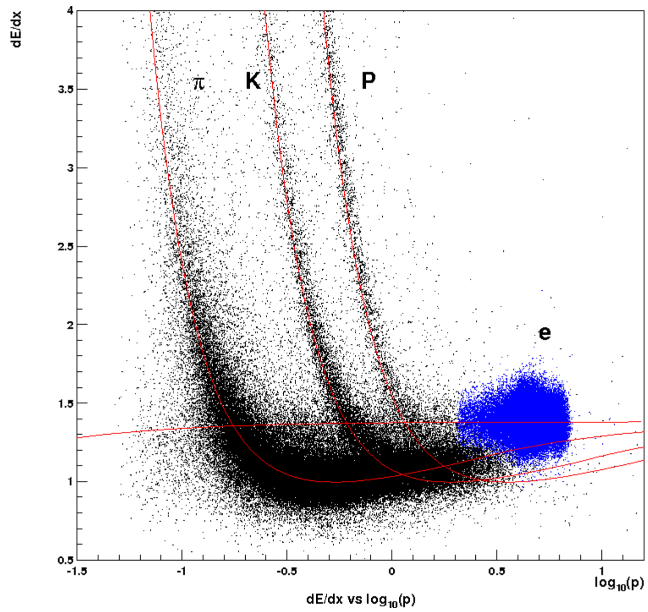

Figure 3.9: Measured distributions of dE/dx versus the log of particle momentum for pions, kaons, protons, and electrons, taken with the CDC. The scattered points are measured data and the lines are expectations from the Bethe-Bloch formula, Eq. 3.3.10.

power at high momentum. In the intermediate region, PID is supplemented by other subsystems: the Aerogel Cherenkov Counter and the Time-of-Flight Counter (TOF). Further, since the muon mass ($\sim 106\ \mathrm{GeV}/c^2$) is so close to the charged pion mass ($\sim 140\ \mathrm{GeV}/c^2$), muons must be identified by another subsystem, the K_L and μ Detector.

Aerogel Cherenkov Counter (ACC)

If a charged particle passes through a dielectric medium with index of refraction, n, at a velocity faster than c/n, then it emits coherent radiation known as Cherenkov radiation. If the index of refraction is suitably chosen, then pions at a given momentum may exceed c/n and emit Cherenkov radiation while kaons at the same momentum would not. This is the operating principle of the Belle Aerogel Cherenkov Counter.

The ACC layout is shown in Figure 3.10. It consists of 960 modules laid out throughout the barrel and endcap regions, each one a thin aluminum box containing two principal components: a stack of ultralight aerogel with index of refraction ranging from 1.01 to 1.03, and one or two fine mesh photomultipler tubes (PMTs) to detect Cherenkov

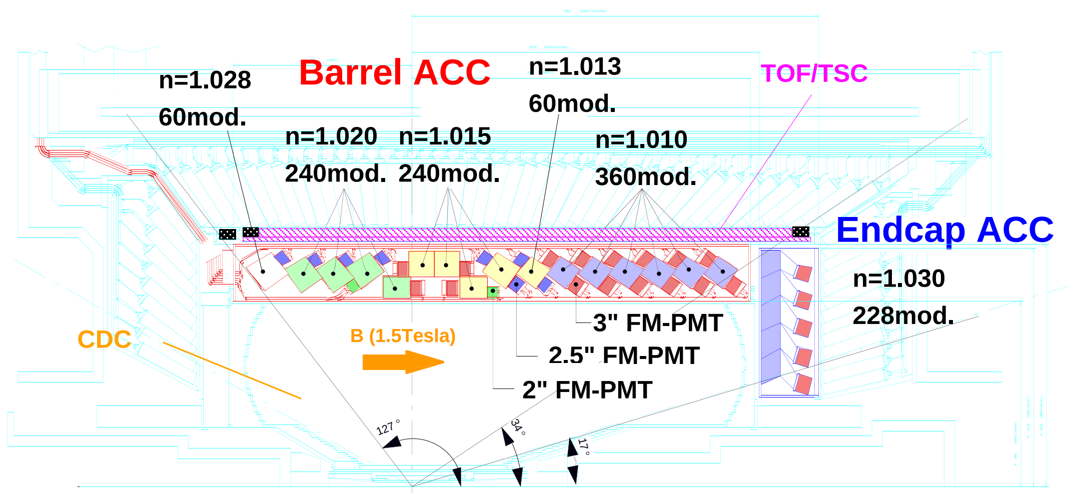

Figure 3.10: The arrangement of dedicated PID modules (barrel and endcap ACC and TOF) within Belle. Aerogel blocks are read out by one or two fine mesh PMTs.

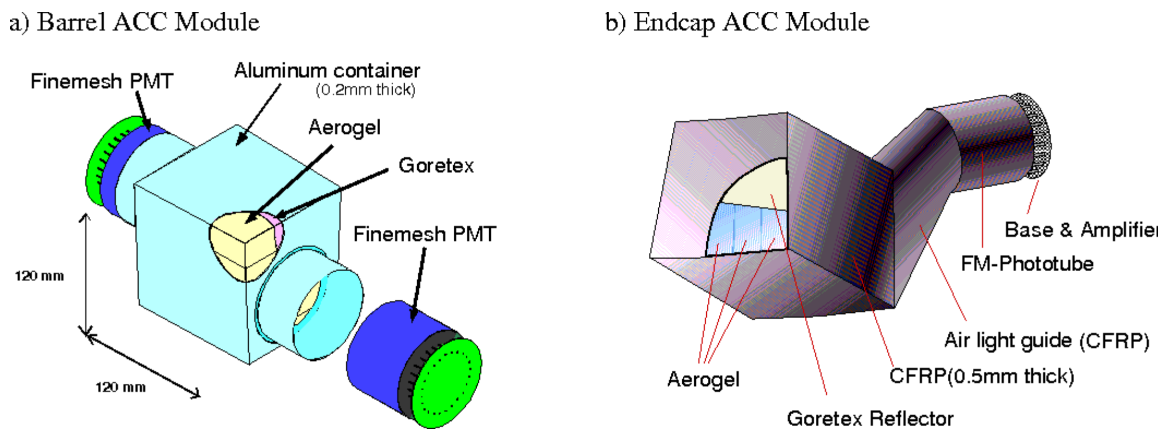

Figure 3.11: ACC modules for the barrel (left) and endcap (right).

light. Example modules are shown in Figure 3.11. The ACC's index of refraction allows it to differentiate pions and kaons in the approximate momentum range of 1 to 4 GeV$/c$.

Time-of-Flight Counter (TOF)

As seen previously in Eq. 3.3.11, a measurement of particle velocity can specify the particle mass, given that the momentum is already measured by the tracking systems. The TOF counter measures this velocity directly by determining the amount of time it takes for the particle to travel from the IP to the TOF system.

The TOF is constructed from 128 bars of plastic scintillators, laid out in a barrel configuration just outside of the ACC (see Figure 3.10). Two PMTs read out the scintillation light from the scintillators at either end of each bar. A thin trigger scintillator (TSC) is placed just inside each pair of TOF modules to provide triggering for the TOF, and is read out via a single PMT, which is coupled to the TSC via a plastic light guide. The layout of these counters is shown in Figure 3.12.

The timing resolution of each end of the TOF depends on the hit position within the TOF scintillator. As expected, the timing is significantly better for the PMT that is closer to the impact position of the charged particle. When timing from both ends of the bar is combined, the overall time resolution is of order 100 ps regardless of impact position, as shown in Figure 3.13. Figure 3.13 also shows the separation performance of the TOF, which provides significant discrimination between π and K for $p < 1.5$ GeV$/c$.

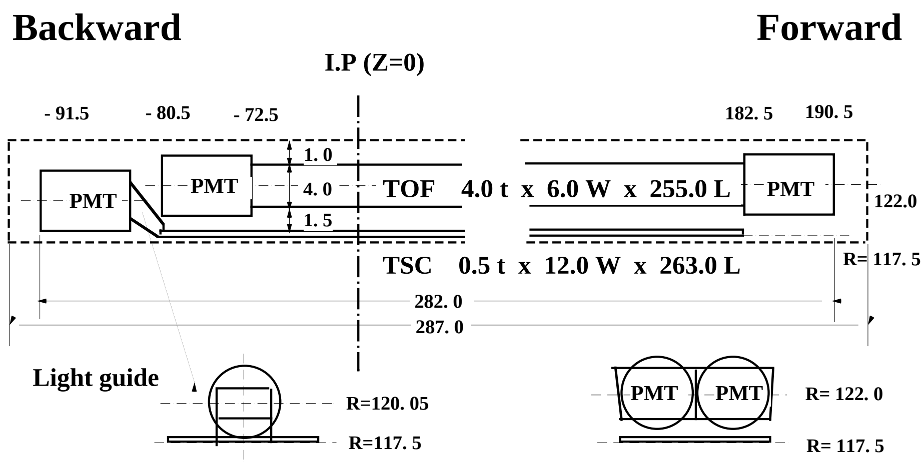

Figure 3.12: Dimensions and layout of a TOF/TSC module.

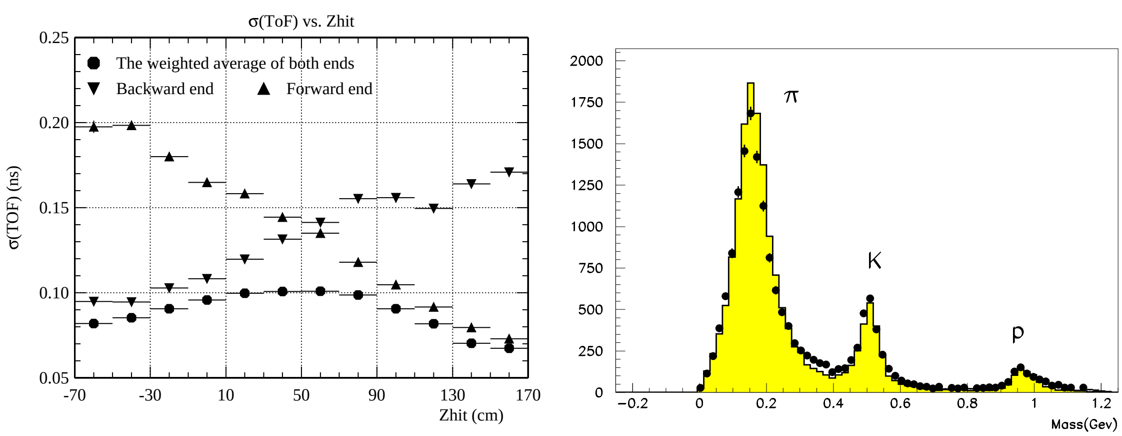

Figure 3.13: (Left) Timing resolution verus particle z impact position on the TOF as measured using $e^+e^- \rightarrow \mu^+\mu^-$ events. (Right) Mass distributions for π, K, and p as measured by the TOF for low momentum tracks ($p < 1.25$ GeV$/c$).

Overall K/π Separation

For each charged track, information from dE/dx, the ACC, and the TOF are combined to give a likelihood ratio for the K hypothesis, given the K and π hypotheses:[3]

$$PID(K) = \frac{P_K^{dE/dx} P_K^{\mathrm{ACC}} P_K^{\mathrm{TOF}}}{P_K^{dE/dx} P_K^{\mathrm{ACC}} P_K^{\mathrm{TOF}} + P_\pi^{dE/dx} P_\pi^{\mathrm{ACC}} P_\pi^{\mathrm{TOF}}} \tag{3.3.12}$$

where P_K and P_π represent probability distributions for kaons and pions in the appropriate subsystem. The likelihood ratio for the π hypothesis can also be calculated as

$$PID(\pi) = \frac{P_\pi^{dE/dx} P_\pi^{\mathrm{ACC}} P_\pi^{\mathrm{TOF}}}{P_K^{dE/dx} P_K^{\mathrm{ACC}} P_K^{\mathrm{TOF}} + P_\pi^{dE/dx} P_\pi^{\mathrm{ACC}} P_\pi^{\mathrm{TOF}}} = 1 - PID(K) \tag{3.3.13}$$

These likelihood ratios can be used to select out those particles more likely to be K or π. Exact efficiencies and fake rates depend on the value of $PID(K)$ or $PID(\pi)$ that is used. A typical selection for kaons accepts those tracks with $PID(K) > 0.6$. Figure 3.14 shows a measured distribution of $PID(K)$ for varying momenta, as well as kaon efficiencies and the rates for pions to be misidentified as kaons using the typical selection criteria.

K_L and μ Detector (KLM)

The KLM identifies particles that would otherwise not be stopped within the Belle detector: K_L and μ. It consists of alternating layers of detectors and the iron of the flux return of the Belle solenoid. The detectors are pairs of glass-electrode-resistive plate counters (RPCs), each consisting of two parallel plates separated by a gas filled gap. Ionizing particles passing through the gap create a discharge in the gap that is read out by external pickup strips. An example RPC "superlayer" is shown in Figure 3.15 (left). Typical spatial resolutions of the superlayers are a few centimeters.

Hadrons can cause interactions within the KLM layers, and can thus be detected. Hadrons consistent with a K_L interaction are identified by the lack of a matching track in the CDC. In contrast, muons should have a matching track in the CDC. Further, μ interactions can be distinguished from hadronic interactions by their increased penetration depth and characteristically narrower clusters. Typical efficiencies for identifying muons are shown in Figure 3.15 (right).

[3] Other, similarly constructed likelihood ratios are available for proton identification, electron identification, etc. We focus on K/π separation since it is the most relevant to this analysis.

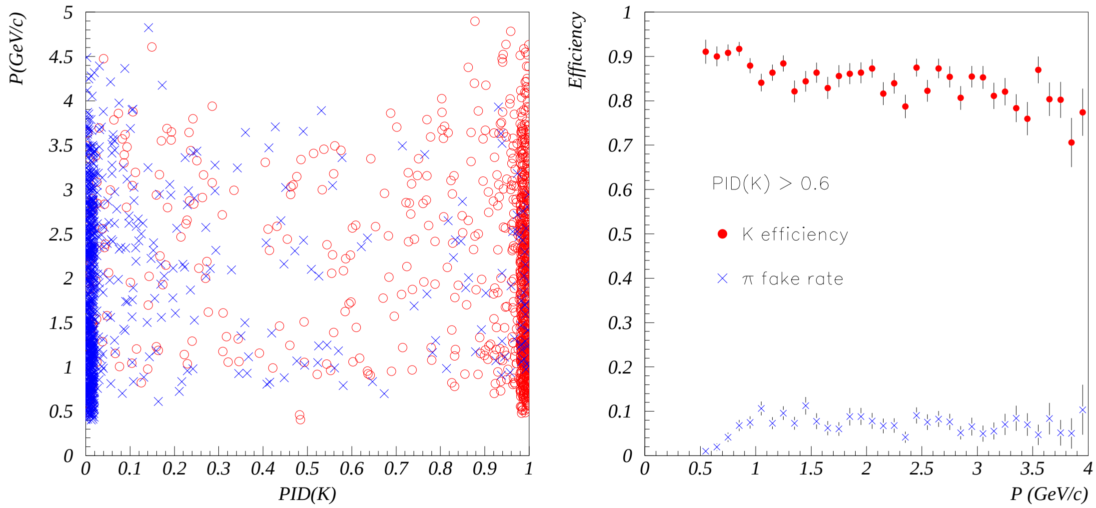

Figure 3.14: (Left) Distributions of $PID(K)$ as a function of momentum for kaons (open circles) and pions (cross-points). (Right) Kaon efficiencies and fake rates to misidentify pions as kaons. Both plots use $D^{*+} \to D^0\pi^+$ decays, which can be selected with excellent signal to noise ratio without using PID measurements.

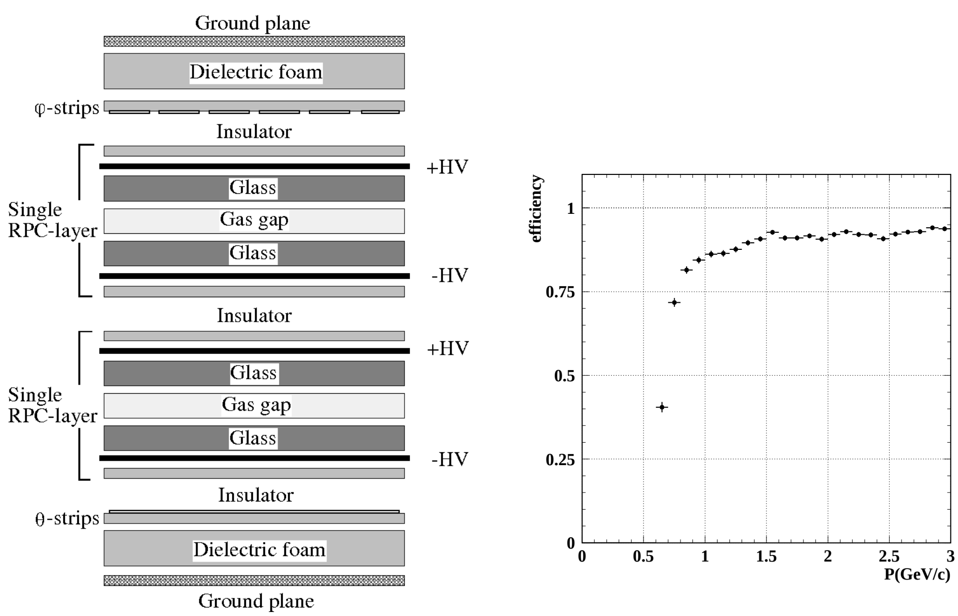

Figure 3.15: (Left) A KLM "superlayer," consisting of two RPC layers with pickup strips above and below. (Right) Typical muon identification efficiencies in the KLM.

3.3.5 Triggering and Data Collection

The high luminosity environment of KEKB creates up to MHz trigger rates in various subsystems, mostly from accelerator induced background events. In order to collect data from desired events, such as $e^+e^- \rightarrow B\bar{B}$, the trigger system must first make some immediate selection to identify an event as potentially interesting. This trigger decision is then fanned out to the various subsystems and the data from each detector is collected by the data acquisition system for offline event reconstruction and analysis.

Trigger System

The trigger system is responsible for identifying the events for recording and permanent storage. For physics studies and analyses, the primary events of interest are $e^+e^- \rightarrow B\bar{B}$, $e^+e^- \rightarrow q\bar{q}(q = u, d, s, c)$, $e^+e^- \rightarrow \tau^+\tau^-$, and $e^+e^- \rightarrow \gamma\gamma$. The trigger system also must record a fraction of the total number of Bhabha scattering ($e^+e^- \rightarrow e^+e^-$) and mu-pair events ($e^+e^- \rightarrow \mu^+\mu^-$), as they are useful for luminosity monitoring and detector calibration. The overall trigger rate should not exceed ~500 Hz due to data acquisition limits, so the trigger requirements must suppress triggers due to KEKB beam backgrounds and other sources.

An initial Level-1 (L1) trigger is obtained by combining various information from each subdetector into the global decision logic (GDL), as shown in Figure 3.16. The GDL for the L1 trigger is implemented on programmable logic devices, and provides a trigger decision 2.2 μs after the crossing time of the e^- and e^+ bunches. The L1 trigger is over 99% efficient for hadronic events, such as $B\bar{B}$.

Another trigger, known as the Level-3 (L3) trigger, can be implemented in the online computer software. It implements a fast track-finding algorithm and rejects events that are inconsistent with an interaction at the IP. It preserves hadronic events with high effiency ($> 99\%$) while rejecting $50\% - 60\%$ of L1 events. Events passing the L1 and, optionally, the L3 trigger are recorded to raw data files. A final, L4, trigger is applied offline on the raw data files, further reducing the background by a factor of roughly five and maintaining the hadronic event efficiency. Events passing the L4 criteria are processed into data files suitable for physics analysis by collaborators.

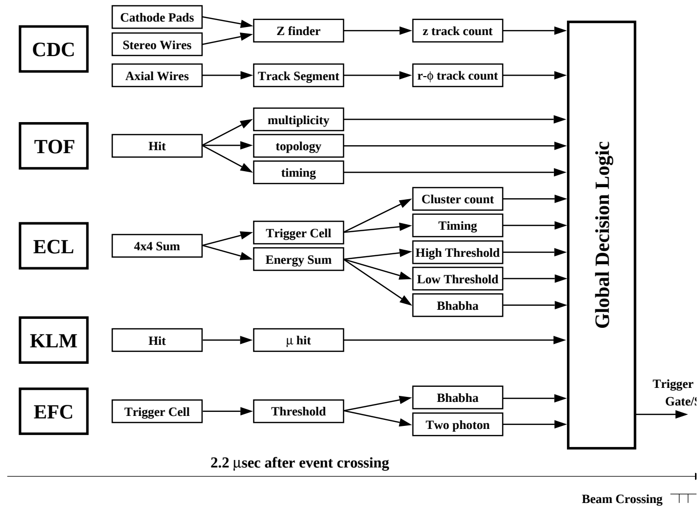

Figure 3.16: Schematic of the Belle Level-1 trigger system.

Data Acquisition (DAQ)

The Belle DAQ system, shown schematically in Figure 3.17, collects data from the seven Belle subsystems (SVD, CDC, ACC, TOF, ECL, KLM, EFC) after a L1 trigger. This data collection runs in parallel to increase the overal DAQ rate, allowing events to be recorded at up to $\sim$500 Hz. Most of the Belle systems utilize charge-to-time (Q-to-T) schemes, with a final readout performed by a time-to-digital (TDC) converter. The exceptions are the SVD and KLM, which use flash analog-to-digital converters (Flash ADC) and direct TDC readout, respectively.

Data from the subsystems is collected and assembled into events by an event-builder. Data sizes vary with the particle multiplicity of the event, but a typical size for a $B\bar{B}$ or $q\bar{q}$ event is 30 kB. After events are built, they are passed to the online computing farm, where the L3 trigger can be applied. Events to be recorded are then written out to offline storage.

Software, Simulation, and Analysis

A raw event consists of the TDC and ADC values of the various subsystems, described previously. A suite of software developed by Belle collaborators takes this data, determines and applies calibration corrections for each subsystem (e.g., applies alignment

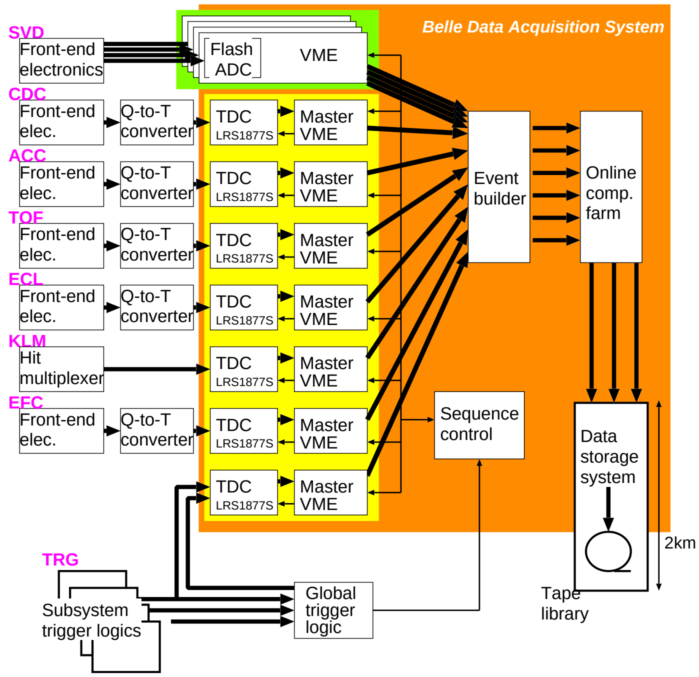

Figure 3.17: Schematic of the Belle DAQ system.

constants, calibrates for gain variations, etc.), and performs momentum or energy measurements for the detected particles. For charged tracks, a helix fit is performed to derive a momentum, and the track is extrapolated to the PID systems to search for associated information. For neutrals, the ECL or KLM can provide information on their energy and/or their direction of flight. Four-momenta and associated information related to particle identification is stored in data files suitable for final analysis by collaborators.[4]

Monte Carlo (MC) simulations play a vital role in optimizing and characterizing physics analyses. Expected distributions of particles and four-momenta for selected events or classes of events can be generated using the EvtGen package [30].[5] These distributions are passed to a full detector simulation, called GSIM, implemented in the GEANT [31] framework, and files are generated for analysis that match the actual data formats. End users then write reconstruction codes, often using components from the existing Belle

[4]The raw data files are stored permanently for potential reprocessing, in the event that the processing procedures change or improve. For example, recent improvements in the charged track identification and reconstruction have led to significant increases in efficiency for low momentum charged tracks. All data taken after the installation of SVD2 was reprocessed to benefit from these changes.

[5]Other event generators can be used, but EvtGen is typical for $B\bar{B}$ and $q\bar{q}$ events.

software libraries, to identify their decay channels of interest. These analysis codes are then tested on MC to evaluate their efficiency and optimize against expected backgrounds. Once finalized, the analysis can be performed on the Belle data sample.

Chapter 4

Analysis

4.1 Analysis Overview

We measure the branching fraction of $B \to X_s\eta$ through a semi-inclusive technique where the X_s is reconstructed as a system of a kaon and up to four pions. As with any hadronic B decay, a large background is expected from continuum $q\bar{q}$ production. In fact, the cross-section for $e^+e^- \to q\bar{q}$ is roughly three times that of $e^+e^- \to \Upsilon(4S)$. There are also so called "generic" backgrounds from $b \to c$ processes with similar or identical final states. To reduce both of these backgrounds we employ a technique used in similar analyses [15, 20, 21] and suppress these backgrounds with a minimum requirement on the η momentum in the CM frame (> 2.0 GeV$/c$). We also apply specific rejection criteria to further reduce these backgrounds. A small background is expected from some rare B decay processes. All backgrounds are studied using a set of MC samples before examining the data sample.

Selection criteria are chosen to effectively reconstruct the signal decays while reducing expected background contributions. Charged kaons and pions are identified using the particle identification systems described in the last chapter. K^0_S, π^0, and η mesons are selected primarily based on the invariant mass of their respective decay products. Typical selections for Belle analyses require the invariant mass to be within $\pm 2.5\sigma$ of the nominal mass, where σ is the measured invariant mass resolution. There are some exceptions to this rule, where selections are tightened to provide better agreement between control samples in data and Monte Carlo, or loosened based on information from previous Belle studies. Though these basic criteria provide an effective selection of signal events, some backgrounds must be targeted directly using specific suppression techniques. Because

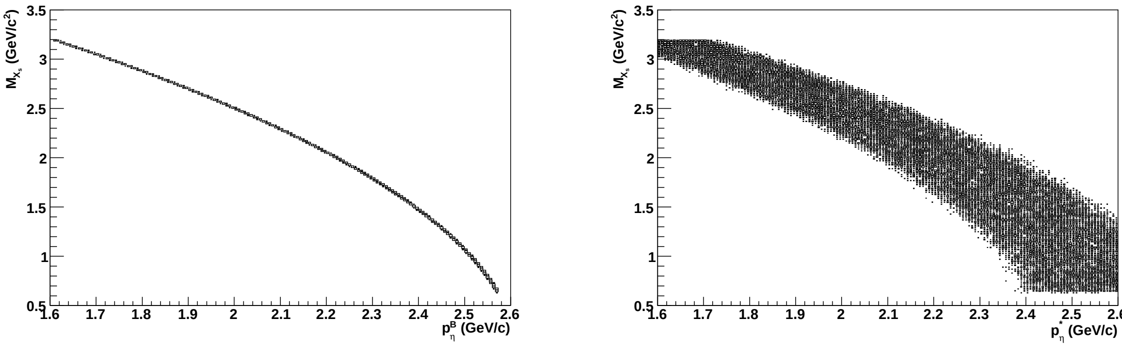

Figure 4.1: (Left) The kinematic relation between the X_s mass and the η momentum in the B rest frame. (Right) The same relation in the CM frame. In both cases, X_s mass was generated using EvtGen up to $3.2\ \mathrm{GeV}/c^2$, and no detector or reconstruction effects are included.

such techniques also may significantly reduce the signal efficiency, an optimization is performed to maximize the expected statistical significance of the observation, as estimated from the quantity $N_S/\sqrt{N_S + N_B}$, where N_S is the expected number of signal events, as determined from MC simulations, and N_B is the expected number of background events, as determined from either MC simulation or a data control region.

Our final analysis is limited by the cut on the η momentum. Due to the kinematics of the $B \to X_s\eta$ decay, in the B rest frame there is a direct relation between the momentum of the η and the mass of the X_s system recoiling against it. This relation is smeared somewhat by the finite B momentum in the CM frame. The B rest frame and CM frame distributions are shown in Figure 4.1. Ultimately, the requirement on the η momentum creates an approximate upper limit on the observable X_s mass spectrum of $2.6\ \mathrm{GeV}/c^2$. We determine the signal yields below this cutoff in $200\ \mathrm{MeV}/c^2$ wide bins of X_s mass. Despite all background suppression techniques, some contamination is expected during the final fitting procedure. The remaining charm backgrounds are modeled for the fit, and the expected contributions from rare backgrounds are subtracted from the fitted yields. The fitted yields are then combined with an efficiency determined from the signal MC to determine a branching fraction. The yields can also be measured separately for $B^-, \bar{B}^0$ and B^+, B^0 to calculate a direct CP asymmetry.

Code	Decay mode	Code	Decay mode
1000	$B^+ \to K^+\eta$	1010	$B^0 \to K^+\pi^-\eta$
1001	$B^+ \to K^+\pi^0\eta$	1011	$B^0 \to K^+\pi^-\pi^0\eta$
1020	$B^+ \to K^+\pi^+\pi^-\eta$	1030	$B^0 \to K^+\pi^-\pi^+\pi^+\eta$
1021	$B^+ \to K^+\pi^+\pi^-\pi^0\eta$	1031	$B^0 \to K^+\pi^-\pi^+\pi^+\pi^0\eta$
1040	$B^+ \to K^+\pi^+\pi^-\pi^+\pi^-\eta$	0100	$B^0 \to K_s^0\eta$
0110	$B^+ \to K_s^0\pi^+\eta$	0101	$B^0 \to K_s^0\pi^0\eta$
0111	$B^+ \to K_s^0\pi^+\pi^0\eta$	0120	$B^0 \to K_s^0\pi^+\pi^-\eta$
0130	$B^+ \to K_s^0\pi^+\pi^-\pi^+\eta$	0121	$B^0 \to K_s^0\pi^+\pi^-\pi^0\eta$
0131	$B^+ \to K_s^0\pi^+\pi^-\pi^+\pi^0\eta$	0140	$B^0 \to K_s^0\pi^+\pi^-\pi^+\pi^-\eta$

Table 4.1: A list of the decay modes reconstructed in this analysis, along with their associated numerical codes.

4.1.1 Semi-Inclusive Reconstruction

As mentioned in Section 1.1.5, a completely inclusive analysis involves a measurement over all possible final states. In practice, this can be a difficult feat, as there may be final states that are undetectable, or have a detection efficiency so low that they cannot be measured reliably. As an alternative, we use a semi-inclusive reconstruction technique[1] introduced by CLEO for analysis of $b \to s\gamma$ [32]. We search for the η meson of $B \to X_s\eta$ by reconstructing the decay $\eta \to \gamma\gamma$. The X_s is reconstructed as a kaon and up to four pions, of which at most one may be neutral. A total of 18 decay modes and their charge conjugates are reconstructed. For some plots and figures, we use a four digit code to represent each mode. The code corresponds to:

$$N_{\text{mode}} = N_{K^\pm} * 1000 + N_{K_s} * 100 + N_{\pi^\pm} * 10 + N_{\pi^0} * 1 \tag{4.1.1}$$

The reconstructed modes are listed explicitly in Table 4.1 with their associated codes.

Of course, there are other modes that could be classified as $B \to X_s\eta$. The fraction of modes that are valid $B \to X_s\eta$ decays but we do not reconstruct are referred to as "missing" modes, and the effect of these missing modes must be accounted for in the final reconstruction efficiency. For example, half of the K^0 or $\bar{K}^0$ mesons from the X_s system will be detected as K_S via the decay $K_S \to \pi^+\pi^-$. The other half manifest as K_L, which we do not attempt to detect in this analysis. Other missing modes occur when there are more than four total pions, or more than one π^0. The final reconstruction efficiency, ϵ, can

[1] This technique is also known as the "sum of exclusive modes" method.

be expressed as

$$\begin{aligned}\epsilon &= \epsilon_{\text{rec}}\mathcal{B}(X_s \to X_s^{\text{rec}}) \\ &= \epsilon_{\text{rec}}[1 - \mathcal{B}(X_s \to X_s^{\text{miss}})]\end{aligned} \tag{4.1.2}$$

where ϵ_{rec} is the efficiency to detect the reconstructed modes, and $\mathcal{B}(X_s \to X_s^{(\text{rec,miss})})$ are the branching fractions for X_s to decay into reconstructed or missing final states, respectively. Both values are estimated from the MC. The estimate of missing modes is discussed further in Section 5.2.2.

4.2 Analysis Samples - Data and MC

Signal and background distributions are first studied using MC samples. All samples are produced using the EvtGen generator to simulate underlying physics events. The generated events are then passed to GSIM to model the detector response and produce data files with formatting identical to the experimental data. The actual data is divided into ranges that reflect an operating period of KEKB and Belle, known as "experiments." Each experiment has associated run conditions that are recorded and imposed into the GSIM framework to produce MC that accurately reflects the KEKB and Belle conditions at the time. Samples of MC that are produced for collaboration-wide use are usually divided into "streams." Each stream corresponds to a set of data equal in size to the Belle data sample.

4.2.1 Experimental Data

The data sample used for this study consists of 604.6 fb^{-1} collected at the $\Upsilon(4S)$ resonance. This corresponds to Belle experiments 7-55. It includes approximately 657 million $B\bar{B}$ pairs, or 85% of the final Belle $\Upsilon(4S)$ data. This sample is nearly 200 times larger than the data sample used for the previous CLEO search for $B \to X_s\eta$.

4.2.2 Signal MC

The primary signal Monte Carlo sample consists of 6 million events (generated in proportion to the luminosity of each experiment number) in the mode $B \to X_s\eta$, where the X_s system is composed of a strange quark and a light anti-quark. The signal MC includes an extra package, known as PYTHIA [33], which fragments the X_s into hadrons. The X_s

is generated with a flat mass distribution in the range $0.6373\ \mathrm{GeV}/^2 < M_{X_s} < 3.2\ \mathrm{GeV}/c^2$. The lower bound corresponds to the approximate threshold for $K\pi$ production and is the lowest allowed mass by PYTHIA.

Two other PYTHIA-based samples with varying X_s mass spectra have also been generated to test the effects of the assumed mass spectrum on the reconstruction and signal extraction procedures. We leave discussion of these models to the next chapter, as they are motivated by studies of the systematic errors.

Since these PYTHIA-based models will never produce exclusive $B \to K\eta$, we supplement them by generating a separate sample of 250,000 events each in the modes $B^- \to K^-\eta$ and $\bar{B}^0 \to \bar{K}^0\eta$. Furthermore, we generate a total of 500,000 separate Monte Carlo events for the decay $B \to K^*(892)\eta$, since this mode is large, well measured [34], and is known to contribute to our signal.

4.2.3 Continuum ($q\bar{q}$) MC

One stream of continuum MC for experiments 7-55 is used to study expected $q\bar{q}$ backgrounds. This sample corresponds to 604.55 fb^{-1}, approximately the same size as expected continuum contributions for the data sample used in this study.

4.2.4 Generic ($B\bar{B}$) MC

Three streams of $b \to c$ MC, generated for experiments 7-55, are used to study the charm backgrounds. This is equivalent to 1.81 ab^{-1}, or 3 times the data sample used in this analysis.

4.2.5 Rare B Decay MC

The entire Belle rare MC sample of $24.6\ \mathrm{ab}^{-1}$ equivalent (40.7 times the data sample) is used to study backgrounds from $b \to s, u, d$ processes. For further study of radiative backgrounds, we use a MC sample of $B \to X_s\gamma$ and $B \to K^*(892)\gamma$ corresponding to 27.8 ab^{-1}, or approximately 46 times the data sample. To study possible backgrounds from $B \to X_d\eta$, we prepare a sample of 500,000 events with a flat M_{X_d} spectrum. The X_d system is fragmented into hadrons by PYTHIA.

4.2.6 Control Sample

A Monte Carlo sample of $B \to D\pi^{\pm}$ modes was generated in order to study possible discrepancies between data and MC. The sample consists of 1.76 million decays of $B^0 \to D^-\pi^+$ and 3.18 million decays of $B^+ \to \bar{D}^0\pi^+$, in accordance with their ratios of branching fractions (2.68×10^{-3} and 4.84×10^{-3}, respectively).

4.3 Event Selection

This section details the event selections that are applied. In all cases, they are applied to both the MC samples and the data in the same manner.

4.3.1 Pre-Selections (Skims)

Before basic selection criteria are applied, one or more pre-selections can be performed. These are known as skims. By limiting our analysis to events that pass these skims, we can focus on a smaller set of data, reducing the processing time and complications that can occur from handling the full Belle data.

Hadronic Event Skim

A skim known as HadronB isolates hadronic events from other processes, such as $e^+e^- \to \tau^+\tau^-$, $e^+e^- \to \mu^+\mu^-$, etc. This both reduces the overall data size and removes significant non-hadronic backgrounds for hadronic analyses. The skim makes requirements on the charged track multiplicity, the calorimeter cluster multiplicity, the total visible energy, the energy deposited in the calorimeter, the overall momentum balance in the z-direction, and the average ECL cluster energy. It also requires that the event vertex is consistent with the known IP to avoid events caused by beam particles striking the beam pipe or residual gas molecules. The optimization and performance of specific criteria are described in References [35] and [36]. This skim retains about 18% of the raw data, while retaining more than 99% of $B\bar{B}$ events, as determined from the MC.

High Momentum η Skim

Since we are searching for high momentum η mesons, we can significantly reduce the size of the data for analysis by removing events without such an η candidate. This skim reconstructs η mesons in the $\eta \rightarrow \gamma\gamma$ mode only, using only γ's with $E > 50$ MeV to avoid combinatorial backgrounds from low energy photons. A loose requirement is imposed on the two-photon invariant mass, $480\ \mathrm{MeV}/c^2 < M_{\gamma\gamma} < 620\ \mathrm{MeV}/c^2$, and the $\gamma\gamma$ candidate must have $p_\eta^{cm} > 1.6\ \mathrm{GeV}/c$. Roughly $\sim 10\%$ of data events that pass the hadronic skim also pass this skim.

Full Reconstruction Skim

This skim is only used on the control sample MC and data. This skim accepts events in which a B meson can be completely reconstructed in one of a number of common decay modes. Although this skim was developed primarily to detect decays with missing energy by accounting for all the other B decay products, it is of use for this analysis since the reconstructed modes include our control decays: $\bar{B}^0 \rightarrow D^+\pi^-$ and $B^- \rightarrow D^0\pi^-$.

4.3.2 Particle Selection and Reconstruction

η Mesons

Each γ used to reconstruct the η is required to have an energy greater than 200 MeV. In order to veto background photons from $\pi^0 \rightarrow \gamma\gamma$, each γ must satisfy a likelihood ratio requirement of $\mathcal{L}_{\pi^0} < 0.8$, where $\mathcal{L}_{\pi^0}$ is calculated from an existing Belle study [37]. Pairs of γ candidates passing these criteria are combined and their invariant mass is required to be in the range $480\ \mathrm{MeV}/c^2 < M_{\gamma\gamma} < 620\ \mathrm{MeV}/c^2$. Because the mass of the η is known more precisely than Belle's η mass resolution, the momenta of the photons are recalculated with the mass of the η fixed to the nominal [1] value. This procedure is known as a mass-constrained fit.

Some radiative decays contain a high energy photon that can be combined with low energy photons to create a fake η. In such a case, the asymmetry in the energy between the two photons, defined as

$$A = \frac{|E_{\gamma_1} - E_{\gamma_2}|}{|E_{\gamma_1} + E_{\gamma_2}|} \tag{4.3.1}$$

will be large. To suppress such fake η candidates, we require this asymmetry to be less than 0.6.[2]

[2] We note that our η mass selection is consistent with $\pm 2\sigma$, where σ is the detector resolution. This and the energy asymmetry cut of $A < 0.6$ are both more restrictive than the standard Belle values of $\pm 2.5\sigma$ and $A < 0.9$. Our study of radiative B backgrounds indicated some possible contributions that were not well

Charged Tracks

Charged tracks used in the X_s reconstruction are required to satisfy the impact parameter requirements $dr < 1$ cm and $dz < 10$ cm. This requirement removes poorly measured tracks. Identification of charged tracks is made using the $PID(K)$ variable derived from the PID systems. $K^{\pm}$'s are required to satisfy $PID(K) > 0.6$, and $\pi^{\pm}$'s must satisfy $PID(K) < 0.9$.

Neutral Kaons

K_s^0 candidates are reconstructed from $\pi^+\pi^-$ pairs. The vertex of each pair must be well reconstructed and displaced from the interaction point, and the reconstructed K_S^0 momentum vector must be aligned with the $\pi^+\pi^-$ vertex vector [38]. They are required to have an invariant mass in the range $482\ \mathrm{MeV}/c^2 < M_{\pi^+\pi^-} < 514\ \mathrm{MeV}/c^2$.[3]

Neutral Pions

Neutral pions are reconstructed from pairs of photons. Each γ daughter from the π^0 is required to satisfy the standard γ energy requirement ($E_\gamma > 50$ MeV) if it was detected in the barrel region, or a more restrictive $E_\gamma > 100$ MeV requirement if it was detected in the endcap. To reduce misreconstructed photons, we require that 90% of the energy of each γ candidate be contained in the central 3×3 array of calorimeter crystals of the 5×5 array of crystals used to reconstruct the photon. This eliminates misreconstructions due to hadronic showers, which have a larger characteristic lateral size than their electromagnetic counterparts.

When the photons are combined, we require that their invariant mass lie in the range $120\ \mathrm{MeV}/c^2 < M_{\gamma\gamma} < 150\ \mathrm{MeV}/c^2$, and their lab frame momentum satisfies $p_{\pi^0} > 300\ \mathrm{MeV}/c$. As in the η reconstruction, a mass constrained fit is performed using the known π^0 mass.

---

modeled by MC. These requirements were made more restrictive to avoid potential systematic errors from these backgrounds.

[3] The mass resolution for K_S^0 candidates is roughly $\sigma \sim 4\ \mathrm{MeV}/c^2$, so this mass range corresponds to $\pm 4\sigma$. This is larger than typical cut values of 2.5σ to 3σ, but the other K_S^0 selection criteria are known to provide a very pure sample of K_S^0's [39].

X_s System

The X_s is reconstructed from a $K^{\pm}$ or K_s^0 and up to four π's, of which up to one may be a π^0. If more than one charged track is used to form the X_s, a vertex fit is performed and must succeed, or the candidate is rejected.

B Mesons

The invariant mass of the B meson is connected to its energy and momentum in the CM frame (E^* and $\mathbf{p}^*$) by the usual relativistic relation $m^2 = E^{*2} + |\mathbf{p}^*|^2$, so it might seem prudent to calculate the total momentum and energy for the proposed B daughters. However, calculating the invariant mass in this fashion does not fully utilize our knowledge of the e^+e^- collision parameters. Since the only outgoing particles in the collision are two B mesons, we know they must each carry half of the CM energy. By using this energy in our calculation, we can improve our resolution on $M_{\rm bc}$ by a factor of 5-10 relative to the conventional invariant B mass. Therefore, it is more convenient to study two related variables: the beam-constrained mass, $M_{\rm bc}$ and the energy difference, ΔE. These are defined as:

$$M_{\rm bc} = \sqrt{(E^*_{\rm beam})^2 - |\mathbf{p}^*_B|^2} \tag{4.3.2}$$

$$\Delta E = E^*_B - E^*_{\rm beam} \tag{4.3.3}$$

where $E_{\rm beam}$ is half the total CM energy of the e^+e^- system, and is calibrated independently for each experimental run. For details of this calibration see Appendix B. True B candidates will have a beam-constrained mass near the B mass, 5.28 GeV/c^2, and an energy difference near zero.

We form B candidates from combinations of X_s and η. They are required to have a beam constrained mass $M_{\rm bc} > 5.23$ GeV$/c^2$ and an energy difference $|\Delta E| < 0.5$ GeV.

4.3.3 Signal Region & Fit Region

We define a signal region, where our expected $B \to X_s\eta$ events are concentrated, in the following variables: η mass before the mass constrained fit ($M_{\gamma\gamma}$), the η momentum in the CM frame ($p_\eta^{\rm cm}$), the energy difference (ΔE), and the beam constrained mass ($M_{\rm bc}$). The requirements are:

- $p_\eta^{cm} > 2.0$ GeV$/c$

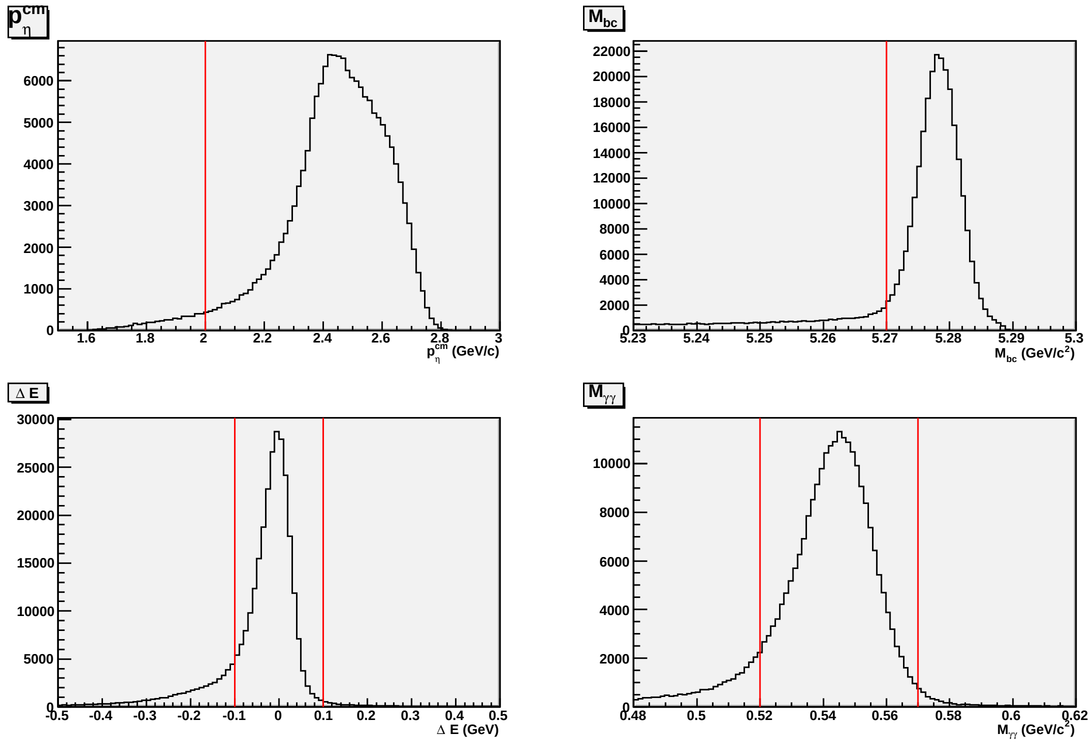

Figure 4.2: Signal regions in various observables, shown for signal MC assuming a flat X_s mass distribution. (Upper left) A center of mass η momentum distribution in GeV/c. The region to the right of the line is the signal region. (Upper right) $M_{\rm bc}$ distribution in GeV/c^2, with the signal region to the right of the red line. (Lower left) ΔE distribution in GeV, with the signal region defined between the two lines. (Lower right) An $M_{\gamma\gamma}$ distribution for η candidates in GeV/c^2, with the signal region defined between the lines.

- $M_{\rm bc} > 5.27\ {\rm GeV}/c^2$
- $-0.1\ {\rm GeV} < \Delta E < 0.1\ {\rm GeV}$
- $520\ {\rm MeV}/c^2 < M_\eta < 570\ {\rm MeV}/c^2$

The signal region cuts are shown graphically in Figure 4.2 using the signal Monte Carlo sample.

We also define a fit region, which includes all the above requirements except the $M_{\rm bc}$ requirement. The signal is ultimately extracted through a fitting procedure, described in Section 4.5, that is performed over events in the fit region.

4.3.4 Control sample selection criteria

The reconstructed modes and selection criteria for the $B \to D\pi^{\pm}$ control modes are nearly identical to those of $B \to X_s\eta$. The D meson is reconstructed as X_s, and the requirement for η momentum is instead applied to the charged π. To select D decays, we also impose a cut on the mass of the X_s of $|M_{X_s} - M_D| < 20\,\mathrm{MeV}/c^2$.

4.3.5 Multiple Candidates & Best Candidate Selection

Due to the large number of reconstructed modes, there may be many B candidates per event. However, to avoid counting one event multiple times, we must select only one candidate per event.

After the selection criteria listed above, there are an average of nine B candidates per event. In events where more than one B candidate can be formed, the candidate with the lowest $\chi^2 = \chi^2_{\mathrm{vtx}} + \chi^2_{\Delta E}$ is chosen, where χ^2_{vtx} is obtained from the X_s vertex fit, if performed, and $\chi^2_{\Delta E} = \frac{\Delta E^2}{\sigma^2_{\Delta E}}$, where $\sigma_{\Delta E}$ is based on asymmetric gaussian fits to each mode separately in signal MC. The ΔE resolution is affected most strongly by the addition of a π^0 to the X_s. For modes without a π^0, typical ΔE resolutions are 62 MeV for $\Delta E < 0$ and 29 MeV for $\Delta E > 0$. For modes with a π^0, the respective resolutions are 76 MeV and 31 MeV.

With no further cuts applied, this candidate selection chooses a completely correctly reconstructed B meson in approximately 38% of signal MC events where at least one B candidate was identified.[4] If we perform the candidate selection, then apply signal region cuts as well as continuum suppression cuts and vetoes (described in the next section), we choose a correctly reconstructed B meson in 56% of the remaining events. Of events with at least one candidate in the signal region, 79.5% are correctly reconstructed. The multiplicities without and with signal region cuts applied are shown in Figure 4.3.

Performing the candidate selection procedure before the fit region requirements or any background suppression techniques are applied represents a potential loss of efficiency. However, it is not necessarily inappropriate, as it may also reduce the contribution of backgrounds.[5] The ordering of this selection relative to other criteria was studied using

[4]Percentages of correctly chosen B candidates can be compared to those expected from randomly selecting a candidate. In this case, since there an average of nine B candidates per event, we expect that a random choice will be correct about 11% of the time.

[5]For example, consider a background event with nine fake B candidates, eight of which would be eliminated by background suppression cuts. If these cuts are applied first, then the one remaining candidate will

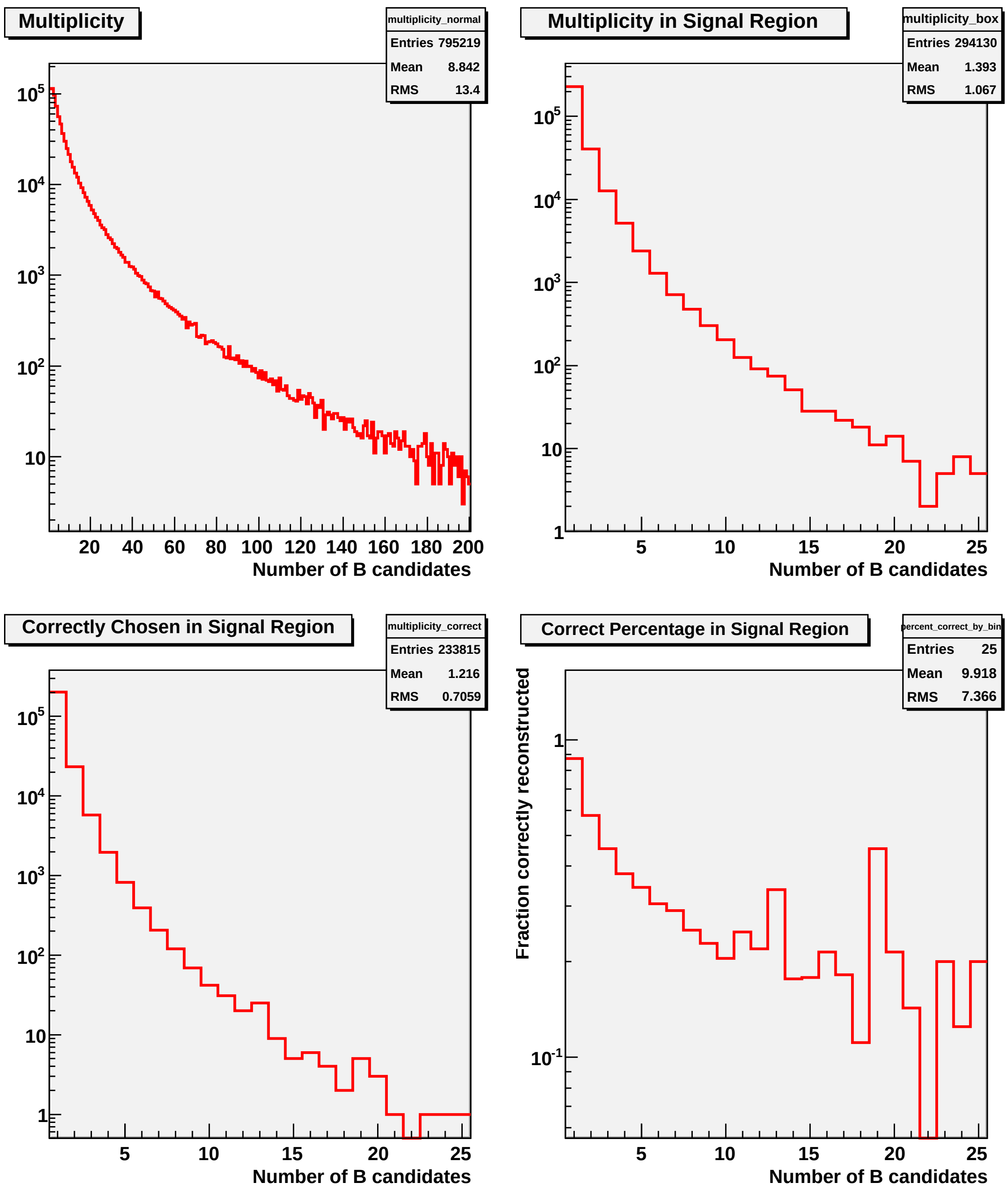

Figure 4.3: (Above) B candidate multiplicities for events with at least one candidate passing the standard cuts (left) and the standard cuts plus signal region cuts (right). (Below) Number of signal region events with a B candidate chosen correctly, as a function of candidate multiplicity (left), and fraction of signal region events with a correctly reconstructed B (right).

the expected statistical significance of the measurement. By performing candidate selection first, the expected statistical significance is 4.1σ. If instead the fit region criteria are applied first, as well as all background suppression criteria, the expected statistical significance drops to 3.5σ. More details on this study are available in Reference [40]. Ultimately, this ordering was chosen as it did not notably affect the expected significance of the measurement, and it was more expedient to perform computationally.

4.4 Background Suppression

Though some of the selection criteria described above help to reduce backgrounds from other processes and decays, we also apply specific suppression techniques that target each classification of background directly. We describe these below.

4.4.1 Continuum ($q\bar{q}$) Background

Continuum backgrounds are kinematically quite different than $B\bar{B}$ events. Since B mesons are spinless, their decay products are distributed nearly isotropically in the rest frame of the B meson. In the CM frame, the B mesons are produced nearly at rest, so this isotropic distribution is mostly preserved. In contrast, light quark-antiquark pairs produced in the e^+e^- continuum are generated back-to-back in the CM frame with high momentum. The quarks then fragment into hadrons along the original quark momentum vectors, creating back-to-back jets of hadrons.

To distinguish between the more spherical $B\bar{B}$ events and the jet-like $q\bar{q}$ events, we use a combination of probability density functions (PDFs) of variables that describe the event topology. One set of such variables is the Fox-Wolfram moments [41]. The moments are defined as

$$H_l = \sum_{i,j} \frac{|\mathbf{p}_i||\mathbf{p}_j|}{s} P_l(\cos\phi_{ij}) \tag{4.4.1}$$

where the indices i and j run over all particles produced in the event, $\mathbf{p}$ are the three-momenta of each particle, $\sqrt{s}$ is the total CM-frame energy, P_l are the Legendre polynomials, and ϕ_{ij} is the angle between particles i and j. Belle utilizes a modified form of these variables for extra discrimination between signal and continuum events [42]. Rather than

always remain and ultimately be chosen as the best candidate. If candidate selection is performed first, this event will be removed if any of the eight candidates are chosen.

including all particles in a single moment, separate moments are defined for those particles associated with the signal B and the other B in the event. For those particles that are associated with the other B, they are further subdivided into charged particles, neutral particles, and a pseudo-particle with a momentum that represents the total missing momentum from the event. For the moments including the signal B candidates, they are defined as

$$R_l^{\text{sig}} = \left\{ \begin{array}{ll} \frac{1}{E_{\text{beam}}-\Delta E} \sum_{i,j} |\mathbf{p}_i| P_l(\cos\phi_{ij}) & l = 0, 2, 4 \\ \frac{1}{E_{\text{beam}}-\Delta E} \sum_{i,j} Q_i Q_j |\mathbf{p}_j| P_l(\cos\phi_{ij}) & l = 1, 3 \end{array} \right\} \tag{4.4.2}$$

where E_{beam} is the half of the total CM energy, ΔE is the energy difference for the signal B, and the Q's are the charge of the corresponding particle. The index i runs over the signal B daughters. The index j runs over various sets of particles. For one set of five moments, j runs over the other charged particles in the event. For the next set, j runs over the other neutral particles in the event. For the final set, j corresponds to the missing momentum pseudo-particle. This results in a total of 15 moments, but only 11 non-zero moments, since those involving Q are only relevant for charged particles. Another set of five moments is defined as

$$R_l^{\text{other}} = \left\{ \begin{array}{ll} \frac{1}{(E_{\text{beam}}-\Delta E)^2} \sum_{i,j} |\mathbf{p}_i| P_l(\cos\phi_{ij}) & l = 0, 2, 4 \\ \frac{1}{(E_{\text{beam}}-\Delta E)^2} \sum_{i,j} Q_i Q_j |\mathbf{p}_j| P_l(\cos\phi_{ij}) & l = 1, 3 \end{array} \right\} \tag{4.4.3}$$

where in this case the indices i and j run over all the particles that are not associated with the signal B, including the pseudo-particle for the missing momentum. In addition to these 16 modified moments, another variable is included, the sum of the transverse momentum of all visible particles.

The distributions of these 17 variables are correlated to the quality of reconstruction of the overall event. This reconstruction quality can be quantified by the amount of apparent mass that is missing from the event. To account for these correlations, the distributions of the 17 variables are binned into seven regions of missing mass. Within each missing mass region, they are linearly combined into a Fisher discriminant [43], $\mathcal{F}$, with weights that maximize the separation between the signal and $q\bar{q}$ background. Distributions of the missing mass and $\mathcal{F}$ for each associated bin are shown in Figure 4.4 for both signal MC and continuum MC.

The polar angle of the candidate B meson flight direction, $\cos(\theta_B)$, can be used for further discrimination. Since they decay from the vector meson $\Upsilon(4S)$, true B mesons follow a $1 - \cos^2(\theta_B)$ distribution, while the combinatorial backgound from continuum is

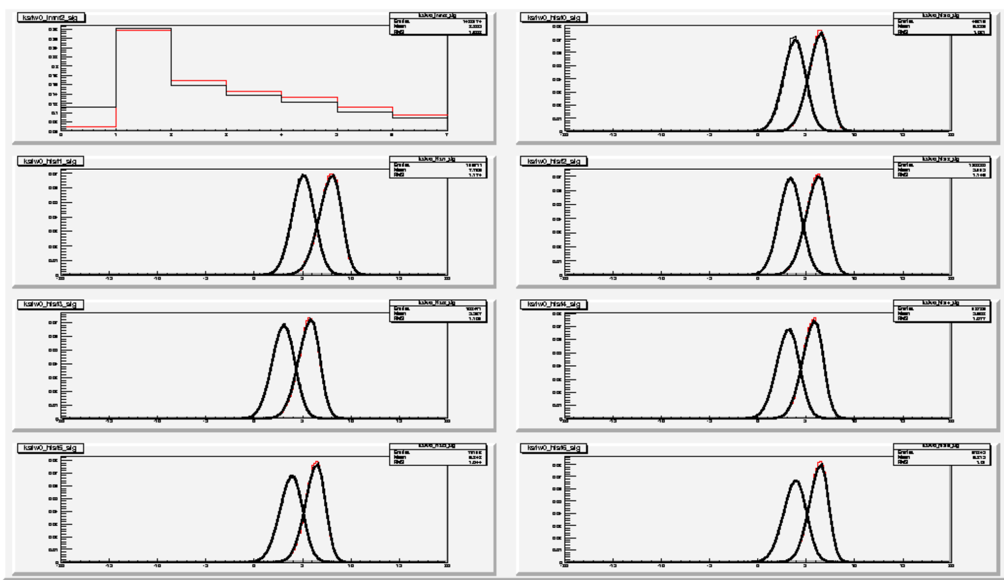

Figure 4.4: (Upper left) Distribution of events into missing mass bins for continuum MC (blue) and signal MC (red). (Others) Distributions of Fisher discriminants corresponding to each bin of missing mass, based on the modified Fox-Wolfram moments, for continuum Monte Carlo (blue, left gaussian in fits), and signal MC (red, right gaussian in fits).

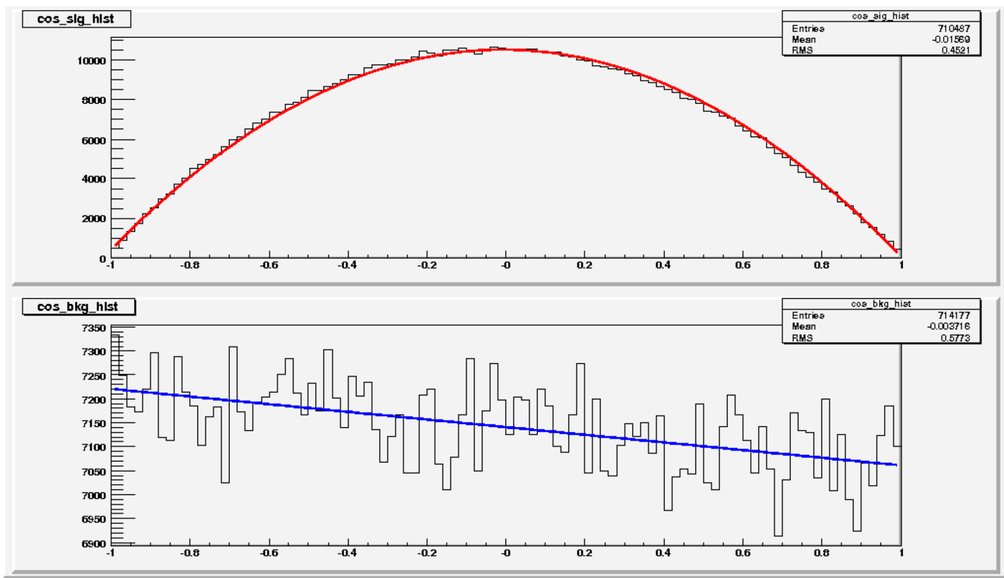

Figure 4.5: Distributions of $\cos(\theta_B)$ for signal MC (top) and continuum MC (bottom).

distributed nearly isotropically. The signal and continuum MC distributions of $\cos(\theta_B)$ are fitted to second and first order polynomials, respectively, and are shown in Figure 4.5.

Lastly, Δz, the difference in logitudinal displacements of the decay vertices for the signal B candidate and the remaining B is much narrow for continuum events, which are expected to have a common vertex, than for true B mesons, which have a finite flight length. Δz is fitted as a sum of three Gaussians for both signal and continuum. The mean of the three Gaussians is fixed to the same value. This value, along with the width of each of the three Gaussians, is fitted from the MC. Distributions and the fitted PDFs can be found in Figure 4.4.1. Separate PDFs are used for SVD1 data and SVD2 data.

These three sets of PDFs are combined into a likelihood ratio (LR) defined as:

$$\mathrm{LR} = \frac{\mathcal{P}_{sig}(\mathcal{F})\mathcal{P}_{sig}(cos(\theta_B))\mathcal{P}_{sig}(\Delta z)}{\mathcal{P}_{sig}(\mathcal{F})\mathcal{P}_{sig}(cos(\theta_B))\mathcal{P}_{sig}(\Delta z) + \mathcal{P}_{q\bar{q}}(\mathcal{F})\mathcal{P}_{q\bar{q}}(cos(\theta_B))\mathcal{P}_{q\bar{q}}(\Delta z)} \tag{4.4.4}$$

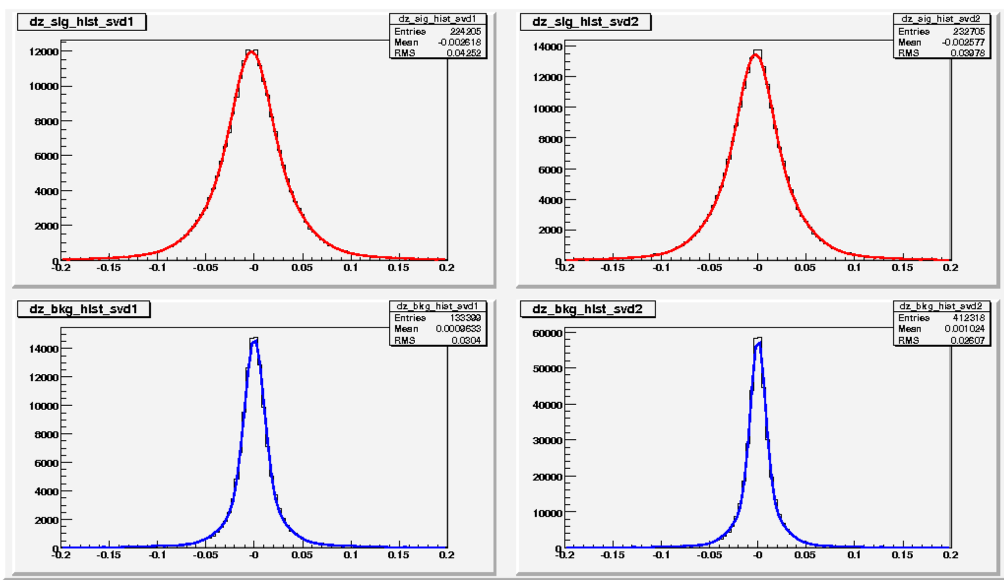

Figure 4.6: PDFs of Δz for signal MC (top) and continuum MC (bottom), shown separately for SVD1 data (left) and SVD2 data (right).

where $\mathcal{P}_{sig}$ is a PDF based on signal MC, and $\mathcal{P}_{q\overline{q}}$ is a PDF based on continuum MC. Overall likelihood ratio distributions from MC are shown in Figure 4.7.

The LR criteria are studied with the aid of the Belle flavor tagging algorithm [44], which analyzes the remaining particles not included in the signal reconstruction and determines the flavor (b or $\bar{b}$) of the other B meson in the event. The tagging algorithm also determines a quality factor for the tagging, r, which ranges from zero for no discrimination and unity for unambiguous flavor identification. For events where the signal B candidate arises from combinatorial backgrounds or misreconstructions, the other B in the event is unlikely to be well tagged, and will thus have a low quality factor. In some cases, the tagging algorithm will even disagree with our signal reconstruction (e.g., our candidate is a B^+ and the charge of the tag B is identified as $+1$). In these cases, we can apply more stringent continuum suppression requirements. For correctly reconstructed signal events, the remaining particles unused in the signal B are more likely to be tagged consistently with the signal B and with a high quality factor. For these events, we can make less severe requirements on the continuum suppression likelihood ratio. The selection criteria on the likelihood ratio vary over 6 bins of tagging quality, with their final values determined by maximizing the overall figure of merit, defined as:

$$FOM = \frac{N_{sig}}{\sqrt{N_{sig} + N_{q\overline{q}} + N_{B\overline{B}}}} \tag{4.4.5}$$

where N_{sig}, $N_{q\overline{q}}$, and $N_{B\overline{B}}$ are the number of expected signal, continuum, and generic events in the signal region passing the likelihood ratio selection. The number of signal events is calculated using the signal MC and an assumed branching fraction of $\mathcal{B}(B \to X_s\eta) = 4.2\times10^{-5}$. This is based on expectations from the QCD anomaly scenario, in which the branching ratio should be suppressed relative to $B \to X_s\eta'$ by a factor of $\tan^2\theta \approx 0.1$, where θ is the $\eta-\eta'$ mixing angle in the octet-singlet basis. The likelihood ratio criteria and overall figure of merit are listed in Table 4.2. The selections are 34.3% efficient on signal MC and suppress 99.5% of contributions from continuum MC.

4.4.2 Generic $B\bar{B}$ Backgrounds

Because of the inclusive nature of the X_s reconstruction, there are many modes that do not arise from $B \to X_s\eta$, but that do share the same final state. The $b \to c$ transitions of this type can be categorized into two subgroups, direct $B \to X_c\eta$ decays, such as $B^0 \to \bar{D}^0\eta$, and cascade decays $B \to X_c \to X_s\eta$. The cascade decays are heavily suppressed

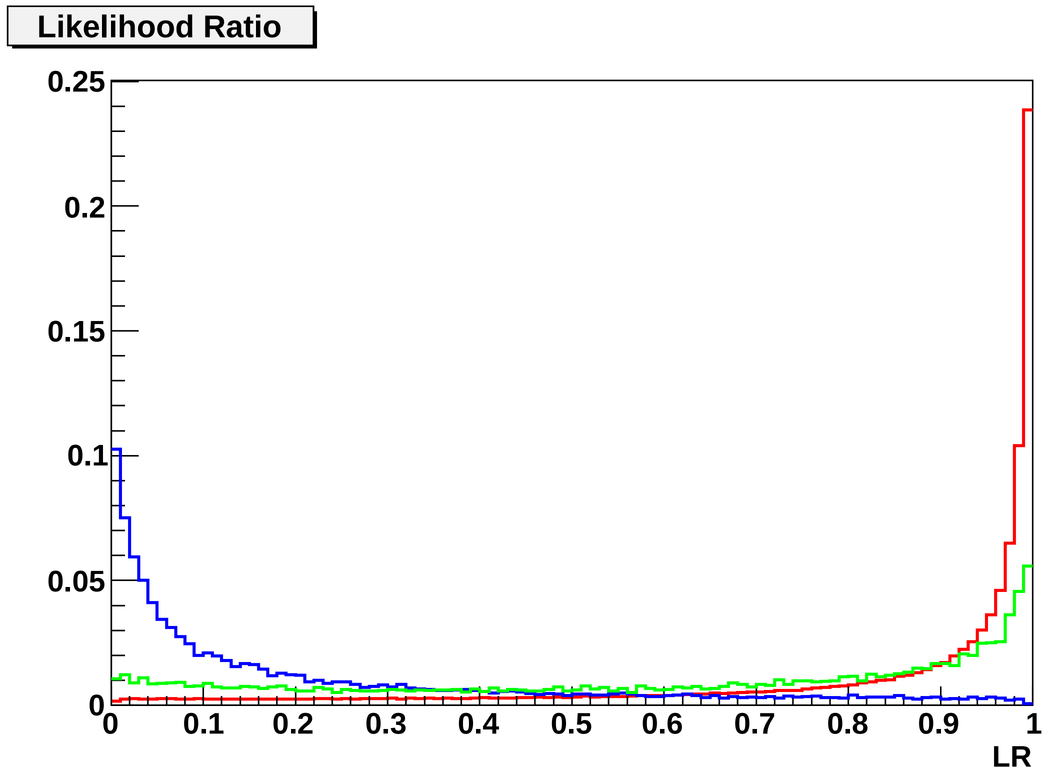

Figure 4.7: Likelihood ratio distributions, integrated over all bins of tag quality, for Monte Carlo samples of signal (red), continuum (blue), and generic $b \to c$ events (green).

B Flavor	Tag quality $\times$ tag charge $\times$ tag charge	LR	B flavor	Tag quality $\times$ tag charge $\times$ tag charge	LR
B^0/B^+	$[-1.00, 0.30]$	0.99	$\bar{B}^0/B^-$	$[-1.00, -0.95)$	0.67
	$(0.30, 0.50]$	0.97		$[-0.95, -0.85)$	0.89
	$(0.50, 0.70]$	0.96		$[-0.85, -0.70)$	0.94
	$(0.70, 0.85]$	0.94		$[-0.70, -0.50)$	0.96
	$(0.85, 0.95]$	0.89		$[-0.50, -0.30)$	0.97
	$(0.95, 1.00]$	0.67		$[-0.30, 1.00]$	0.99

Overall FOM	4.99

Table 4.2: Miniminum likelihood ratio requirements in bins of tag quality. If the flavor of the signal B is unknown, as is the case for neutral B candidates with a K_s^0, then the absolute value of tag quality times tag charge is used, with the likelihood cuts from the B^0/B^+.

Pseudo reconstruction	σ (MeV/c^2)
$D^0 \to Kn\pi^{\pm}$	5.4
$D^0 \to Kn\pi^{\pm}\pi^0$	17.8
$D^+ \to Kn\pi^{\pm}$	5.0
$D^+ \to Kn\pi^{\pm}\pi^0$	12.5
$D^0 \to K_s^0\eta$	12.5
$D_s^+ \to \eta\pi^+$	11.7
$\eta_c(1S) \to \eta\pi^+\pi^-$	34.0

Table 4.3: Pseudo-reconstructions and mass resolutions for charm background vetoes.

by the requirement of p^*_η, but nonetheless some remain, particularly in the high X_s mass region.

To suppress the direct $B \to X_c\eta$ backgrounds, as well as any tails of the cascade decays, we apply a series of rejections, commonly called vetoes, based on pseudo-reconstructions from within the B decay products. As an example, to veto candidates from $B^0 \to \bar{D}^0\eta$, we reconstruct $\bar{D}^0$ from combinations of the kaon and pions used in the X_s reconstruction. If any of these $\bar{D}^0$ candidates fall within a specified invariant mass range, then the B candidate is vetoed. A list of pseudo-reconstructions and mass windows used for vetoes is given in Table 4.3. The width of each pseudo-mass is determined from gaussian fits to MC simulation and veto windows are typically set at $\pm 2.5\sigma$ from the nominal mass. Table 4.4 shows approximate expected numbers of events in the signal region from selected modes before and after the vetoes. We describe the treatment of the remaining $b \to c$ events in Section 4.5.

4.4.3 Rare B Backgrounds

Three distinct backgrounds are identified from rare $b \to u, d, s$ transitions. The first comes from the very $B \to X_s\eta'$ decays that motivated this analysis. In these decays, a subsequent $\eta' \to \eta\pi^+\pi^-$ decay can lead to a $B \to X_s\eta$ final state. As with the charm cascades, these events are suppressed due to their characteristically lower η momentum. Nevertheless, we veto possible remaining backgrounds using a method similar to the charm vetoes. We reconstruct $\eta' \to \eta\pi^+\pi^-$, where the η from $B \to X_s\eta$ is combined with any two charged pions in the event. If the mass is within $100\ \mathrm{MeV}/c^2$ of the known η' mass, we veto the event.

Mode	$N_{\text{expected}}^{\text{pre-veto}}$	$N_{\text{expected}}^{\text{post-veto}}$	PDG $\mathcal{B}$	Belle MC $\mathcal{B}$
$B^0 \to \overline{D^0}\eta$	523	151	$(2.02 \pm 0.35) \times 10^{-4}$	1.65×10^{-4}
$B^+ \to \overline{D^0}\pi^+\eta$	485	130		
$B^0 \to \overline{D^*(2007)^0}\eta$	176	93	$(1.8 \pm 0.6) \times 10^{-4}$	1.78×10^{-4}
$B^+ \to \overline{D^*(2007)^0}\pi^+\eta$	308	146		
$B^0 \to D^-\pi^+\eta$	380	115		
$B^0 \to D^*(2010)^-\pi^+\eta$	430	155		
$B^+ \to \overline{D^0}\rho^+$	155	80	$(1.34 \pm 0.18) \times 10^{-2}$	1.34×10^{-2}
$B^+ \to \overline{D^*(2007)^0}\rho^+$	58	24	$(9.8 \pm 1.7) \times 10^{-3}$	15.5×10^{-3}
$B^+ \to \overline{D^0}\pi^+$	56	20	$(4.84 \pm 0.15) \times 10^{-3}$	4.96×10^{-3}
$B^+ \to \overline{D^0}D_s^+$	62	14	$(10.0 \pm 1.7) \times 10^{-3}$	9.2×10^{-3}
$B^0 \to D^-D_s^+$	41	11	$(6.5 \pm 1.3) \times 10^{-3}$	8.66×10^{-3}
$B^+ \to D_s^-K^+\pi^+$	82	24	$< 7 \times 10^{-4}$	25.6×10^{-4}
$B^+ \to \eta_c(1S)K^+$	46	14	$(9.1 \pm 1.3) \times 10^{-4}$	9.41×10^{-4}
$B^0 \to \eta_c(1S)K_s^0$	19	7	$(9.9 \pm 1.9) \times 10^{-4}$	11.87×10^{-4}
$B^0 \to \eta_c(1S)K^*(892)^0$	27	8	$(1.6 \pm 0.7) \times 10^{-3}$	1.60×10^{-3}
$B^0 \to D^-\rho^+$	60	28	$(7.5 \pm 1.2) \times 10^{-3}$	7.91×10^{-3}
$B^0 \to D^*(2010)^-\rho^+$	53	23	$(6.8 \pm 0.9) \times 10^{-3}$	7.14×10^{-3}

Table 4.4: Expected background contributions for selected $b \to c$ modes, based on generic MC events found in the signal region, before and after vetoes are applied. N_{expected} is calculated by counting events in the signal region. Branching fractions are listed as tabulated by the Particle Data Group (PDG) [1], and as used by Belle for MC generation.

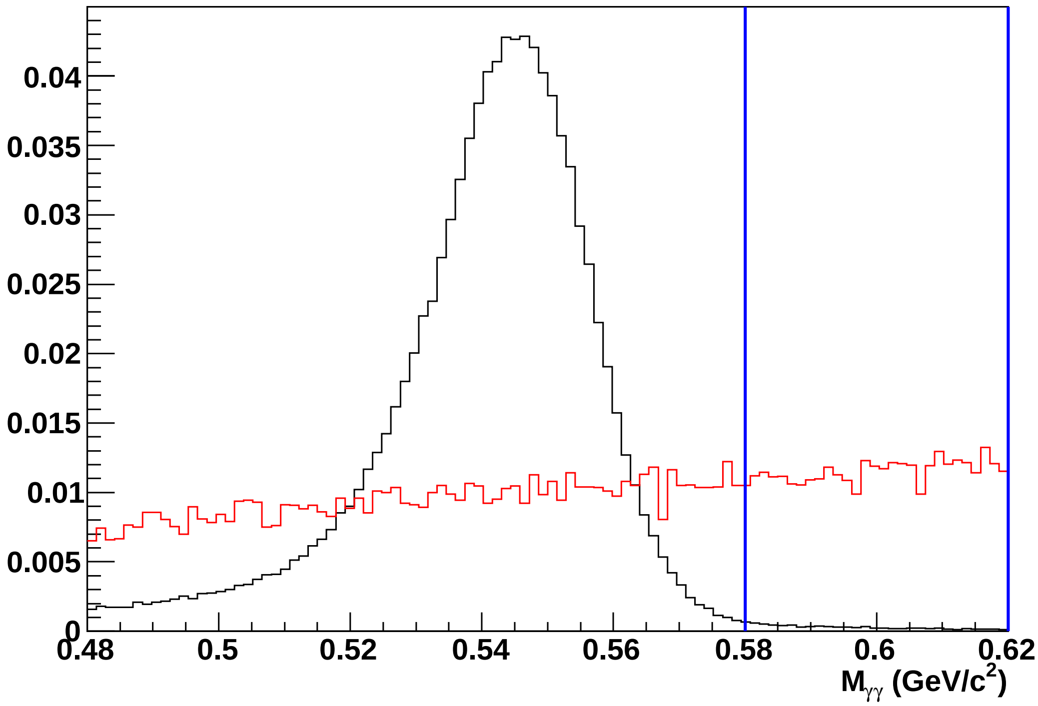

Figure 4.8: Distributions of $M_{\gamma\gamma}$, in GeV/c^2 for signal MC (black) and $b \to s\gamma$ MC (red). The area between the vertical lines is defined as the $M_{\gamma\gamma}$ sideband region, which is used to study the potential radiative backgrounds.

The second class of rare background is from $B \to X_s\gamma$, where we may misreconstruct an η from the high energy γ and a low energy background photon. Both the energy asymmetry and $M_{\gamma\gamma}$ requirements already described are the only criteria that are specifically used to reduce this background. However, we note that the distribution of $M_{\gamma\gamma}$ for such misreconstructions is quite different than for those with a true η. $M_{\gamma\gamma}$ distributions for $B \to X_s\gamma$ and $B \to X_s\eta$ are shown in Figure 4.8. It is clear that the signal region requirement for $M_{\gamma\gamma}$ significantly reduces these backgrounds. We can also define a sideband region, $0.58\ \mathrm{MeV}/c^2 < M_{\gamma\gamma} < 0.62\ \mathrm{MeV}/c^2$, where we can compare measured backgrounds from $B \to X_s\gamma$ to our expectations from the MC. We search for these backgrounds using a fitting procedure that will be described in more detail in the next section.

The last rare background comes from $B \to X_d\eta$, where one of the pions in the X_d system is misidentified as a kaon. The mis-assignment of the kaon mass to a pion will shift the ΔE distribution, so our ΔE signal region requirements help to reduce these backgrounds.

As with the $b \to c$ backgrounds, we still must characterize the remaining rare backgrounds for the final fitting procedure. We describe this procedure, including the background treatments, in the next section.

4.5 Maximum Likelihood Fit

We extract our signal yields in 200 MeV/c^2 wide bins of M_{X_s}. For each bin, we perform a one-dimensional extended unbinned maximum likelihood fit to $M_{\rm bc}$. The likelihood function is defined as

$$\mathcal{L} = \frac{1}{N!} e^{-(N_{\rm sig}+N_{B\bar{B}}+N_{q\bar{q}})} \times \prod_{i=1}^{N} N_{\rm sig}\mathcal{P}^i_{\rm sig} + N_{B\bar{B}}\mathcal{P}^i_{B\bar{B}} + N_{q\bar{q}}\mathcal{P}^i_{q\bar{q}} \tag{4.5.1}$$

where i runs over all the B candidates, N represents the yield, either total (no subscript) or from a given category (with subscript), and $\mathcal{P}$ represents a PDF for the given component (signal, $B\bar{B}$, or $q\bar{q}$). Although we do not include a component in the fit to model rare backgrounds, we expect them to contribute to the $B \to X_s\eta$ signal. These contributions are subtracted from the fitted signal yields. In this section, we describe the fitting functions used for each background and the methods used to determine how much rare background to subtract.

4.5.1 Signal Component

The signal component is modeled as a Gaussian. The yield of this Gaussian is floated to determine the amount of signal, and both the mean and width are fixed based on values from the signal MC. The MC sample varies by the X_s region. For the lowest M_{X_s} bin, from 0.4-0.6 GeV/c^2, only $K^+\eta$ and $K^0_S\eta$ can contribute, so the signal is modeled using dedicated MC samples of these modes. For the bin from 0.8-1.0 GeV/c^2, the signal is known to be dominated by the exclusive decay $B \to K^*\eta$, so again a dedicated exclusive MC sample is used. For all other bins, the PYTHIA MC with a flat X_s distribution is used. For all MC samples, separate fits are performed for the SVD1 and SVD2 data samples, since the resolutions and means may have slight differences due to the changes in hardware.

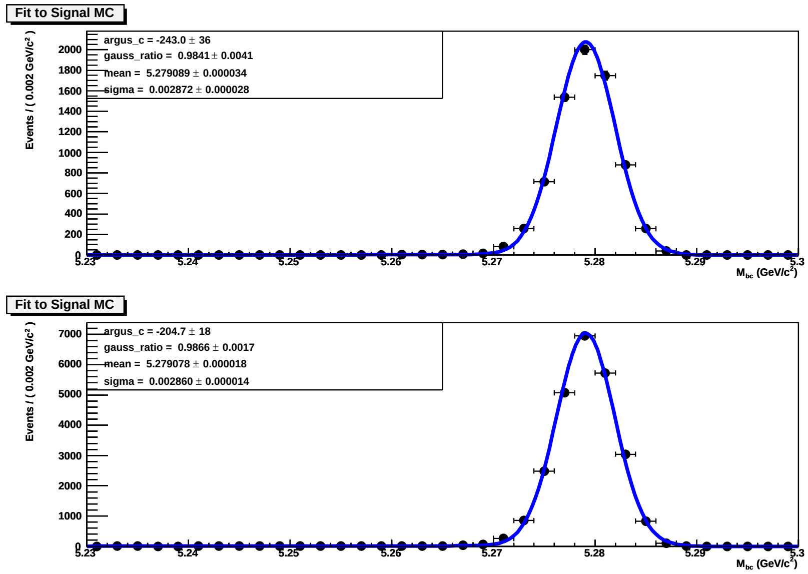

Figure 4.9: $M_{\rm bc}$ fits to the signal MC generated with exclusive $B \to K\eta$ under conditions for SVD1 (above) and SVD2 (below). These fits are used to set the mean and sigma for the signal Gaussian in the bin from 0.4 - 0.6 GeV$/c^2$.

Fits to these MC samples are shown in Figures 4.9, 4.10, and 4.11. After determining the means and widths from MC, we adjust them slightly for the final fit to account for differences between the MC and data. This is described in more detail in Section 4.5.5

4.5.2 Continuum / Combinatorial Component

Combinatorial backgrounds, where a B candidate arises from random combinations of particles, have a distinctly different shape in $M_{\rm bc}$ than correctly reconstructed B decays, allowing them to be distinguished relatively easily during the fitting procedure. These backgrounds arise largely from $q\bar{q}$ events, but there is some contribution from other events as well. These contributions are modeled by an empirically determined PDF known as an ARGUS function [45], defined as

$$\mathcal{P}(M_{\rm bc}) = t\sqrt{1-t^2}e^{\alpha(1-t^2)} \tag{4.5.2}$$

where $t = M_{\rm M_{bc}}/E^*_{\rm beam}$, and α is known as the ARGUS shape parameter. Both the shape and the yield of this PDF are allowed to float during the fits to data.

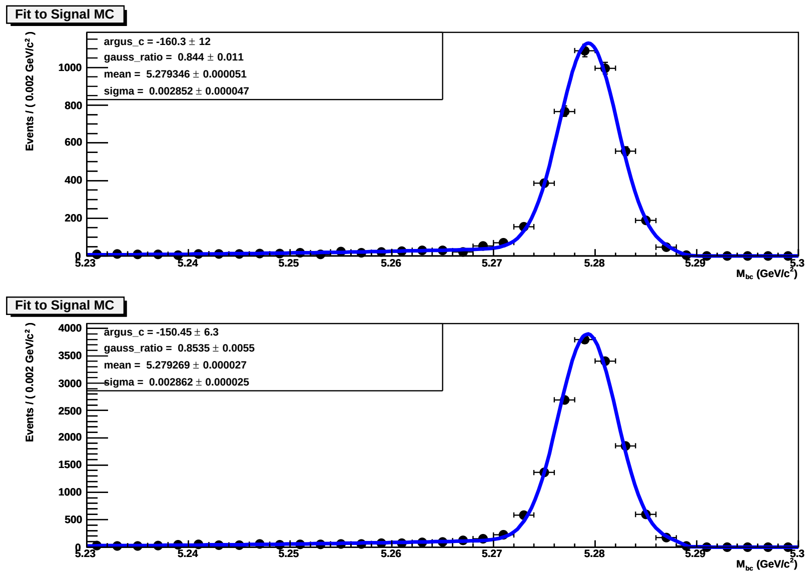

Figure 4.10: $M_{\rm bc}$ fits to the signal MC generated with exclusive $B \to K^*\eta$ under conditions for SVD1 (above) and SVD2 (below). These fits are used to set the mean and sigma for the signal Gaussian in the bin from 0.8 - 1.0 GeV/c^2.

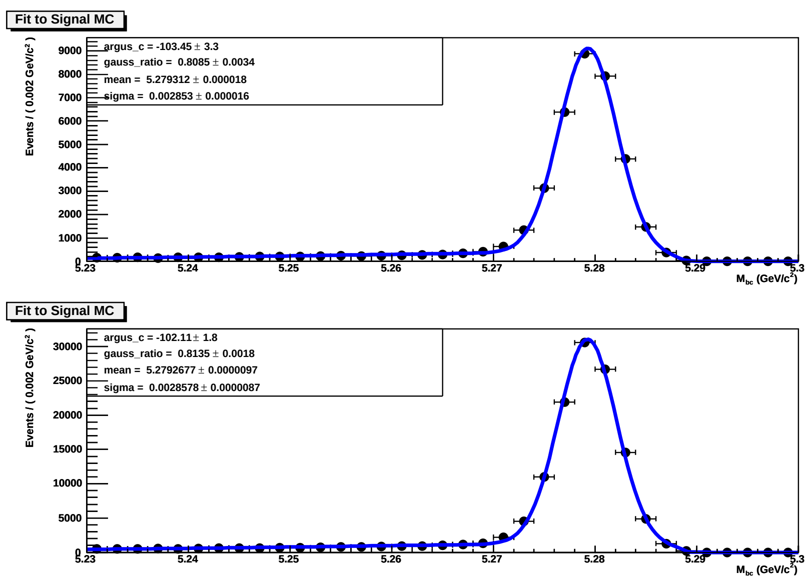

Figure 4.11: $M_{\rm bc}$ fits to the PYTHIA based signal MC with a flat X_s mass distribution under conditions for SVD1 (above) and SVD2 (below). The fit is over the entire X_s mass range (0.4 - 2.6 GeV/c^2), and is used to set the mean and sigma for the signal Gaussian in the bins from 0.6 - 0.8 GeV/c^2 and from 1.0 - 2.6 GeV/c^2.

4.5.3 $B\bar{B}$ Components

Based on classifications of the largest contributing backgrounds in the generic MC, the $B\bar{B}$ PDF is divided into five components: $B^0 \rightarrow \bar{D}^0\eta$, $B^0 \rightarrow \bar{D}^{*0}\eta$, $B^+ \rightarrow D^{(*)-}\pi^+\eta$, $B^0 \rightarrow \bar{D}^{(*)0}\pi^+\eta$, and all other $b \rightarrow c$ backgrounds.

$B\bar{B}$ PDF Shapes

Each of these is modeled by an ARGUS function plus a Gaussian, with the relative fraction of Gaussian to ARGUS determined by fits to the generic MC. Because the shape of the combinatorial component of these backgrounds varies by X_s mass bin, two sets of fits are performed to determine the PDF components. One set of fits is performed over the full range of M_{X_s}, 0.4-2.6 GeV/c^2. These fits, shown in Figure 4.12, are used to fix the Gaussian parameters of each $B\bar{B}$ PDF. Once these parameters are determined, they are fixed to these values and fits are performed in each bin of M_{X_s}. These bin-dependent fits are used to determine the ARGUS shape and the relative amount of Gaussian to ARGUS background for each component in each bin. These fits are shown in Figures 4.13-4.23.

$B\bar{B}$ PDF Normalizations

The shapes of the five $B\bar{B}$ fit PDFs are not distinct enough from one another to allow their normalizations to be floated for the fit, so they must be fixed. This is relatively straightforward in the case when the decays have known branching fractions. However, a glance at Table 4.4 reveals that for some decays, the known branching fractions are either poorly measured or completely unknown.

The $B^0 \rightarrow \bar{D}^0\eta$ and $B^0 \rightarrow \bar{D}^{*0}\eta$ modes have been previously measured by both the Belle and BaBar collaborations, though with relatively small data samples ($140\ \mathrm{fb}^{-1}$ and $89\ \mathrm{fb}^{-1}$, respectively), and with results that are somewhat inconsistent. For these modes, we fix their normalizations to those expected from the Belle branching fractions [46].

The three-body decays $B \rightarrow \bar{D}^{(*)}\pi\eta$ are examples of unmeasured modes, and appear in the Belle MC decay table only through a PYTHIA fragmentation entry that is meant to produce generic $B \rightarrow DX$ decays. For these decays, we have no measurements available to guide our normalizations, so we instead use the data itself. Since we have implemented vetoes on the D and D^*, we can look inside these vetoed selections to find an enhanced sample of our generic decays. By examining binned M_{bc} distributions in the

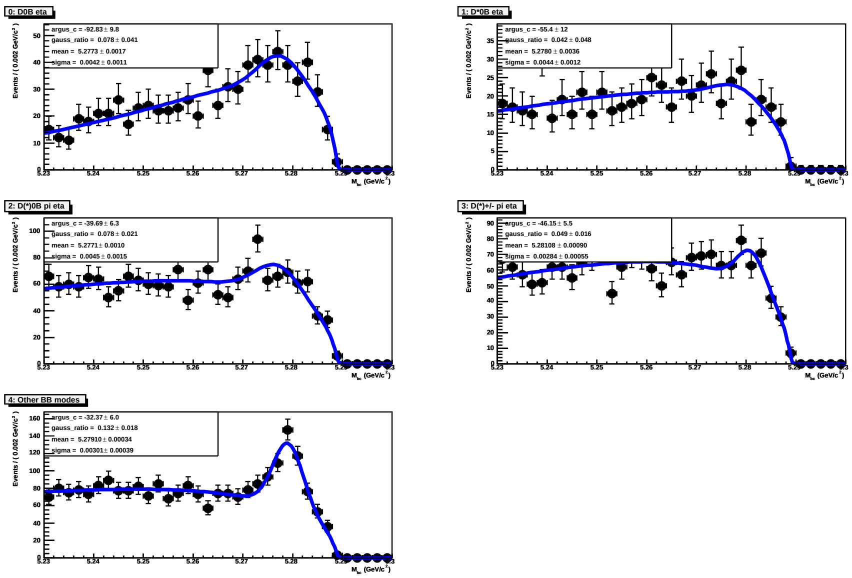

Figure 4.12: $M_{\rm bc}$ fits to the $b \to c$ MC samples over the entire X_s mass range (0.4 - 2.6 GeV/c^2). These fits are used only to determine the Gaussian shape parameters for each component of the $b \to c$ backgrounds: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

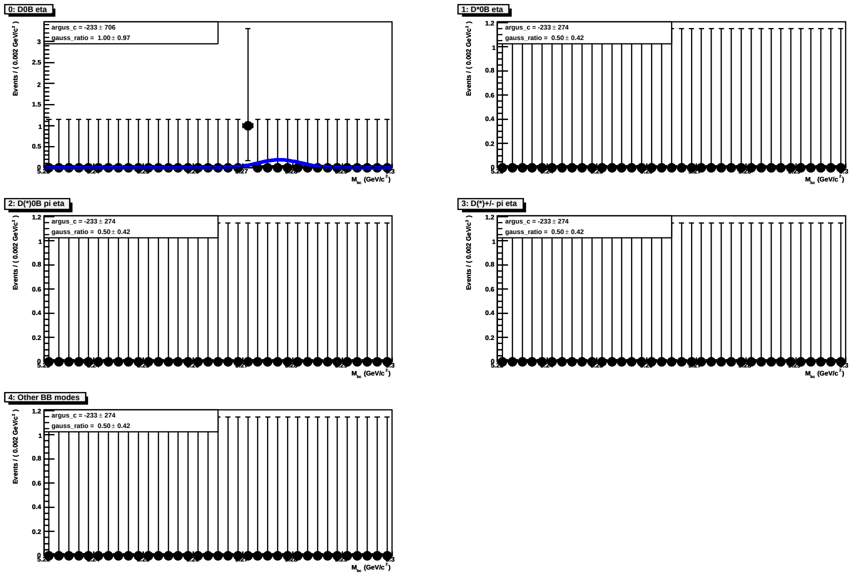

Figure 4.13: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 0.4 GeV/c^2 - 0.6 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

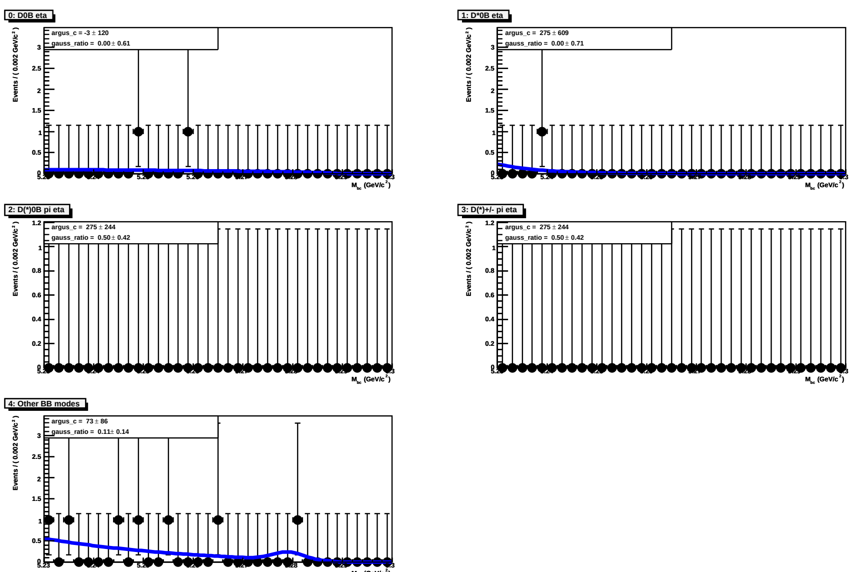

Figure 4.14: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 0.6 GeV/c^2 - 0.8 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

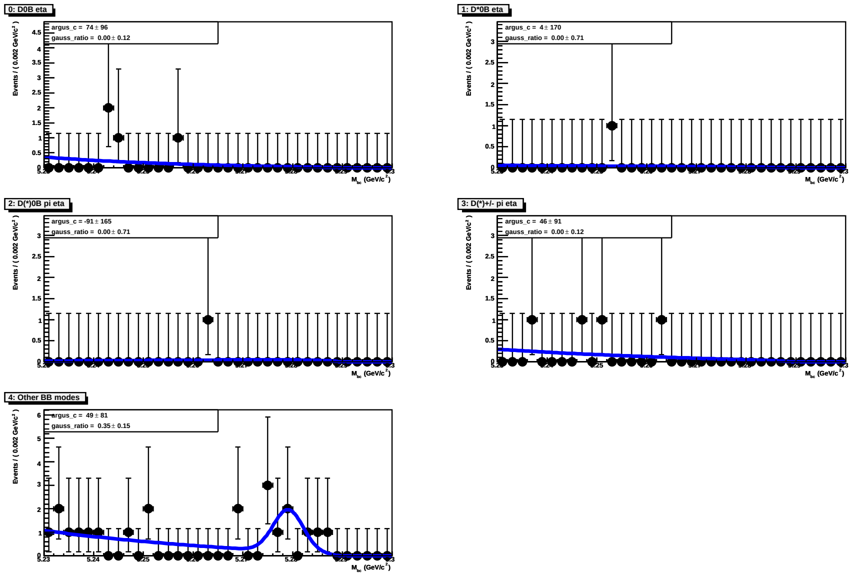

Figure 4.15: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 0.8 GeV/c^2 - 1.0 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

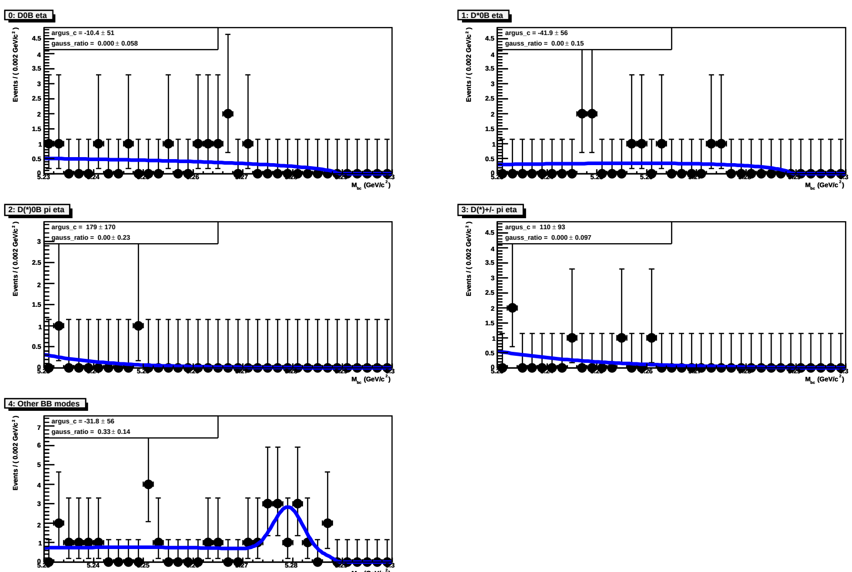

Figure 4.16: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 1.0 GeV/c^2 - 1.2 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

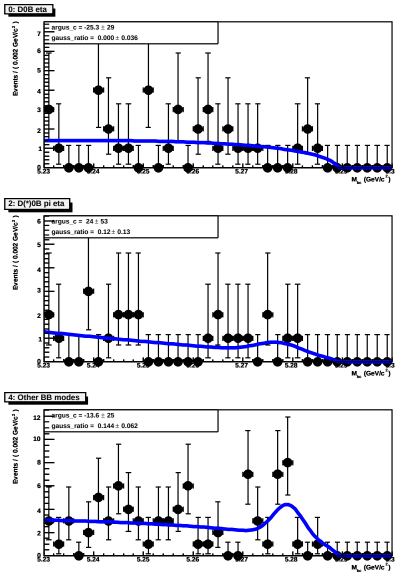

Figure 4.17: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 1.2 GeV/c^2 - 1.4 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

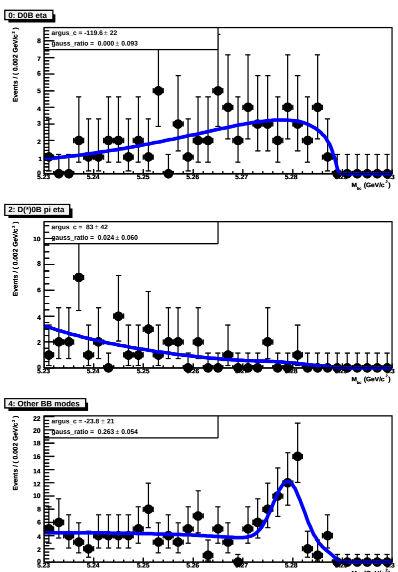

Figure 4.18: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 1.4 GeV/c^2 - 1.6 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

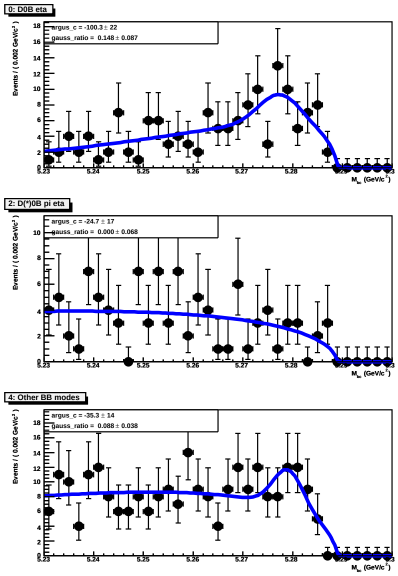

Figure 4.19: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 1.6 GeV/c^2 - 1.8 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

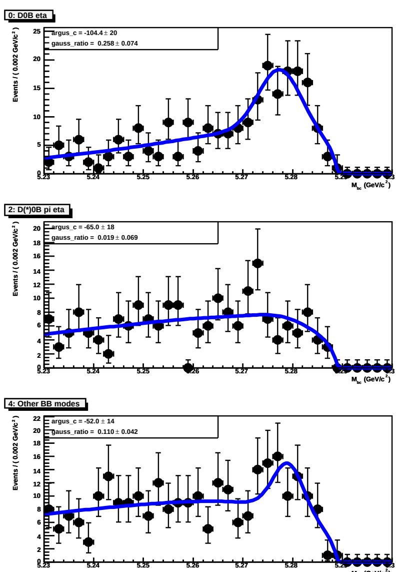

Figure 4.20: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 1.8 GeV/c^2 - 2.0 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

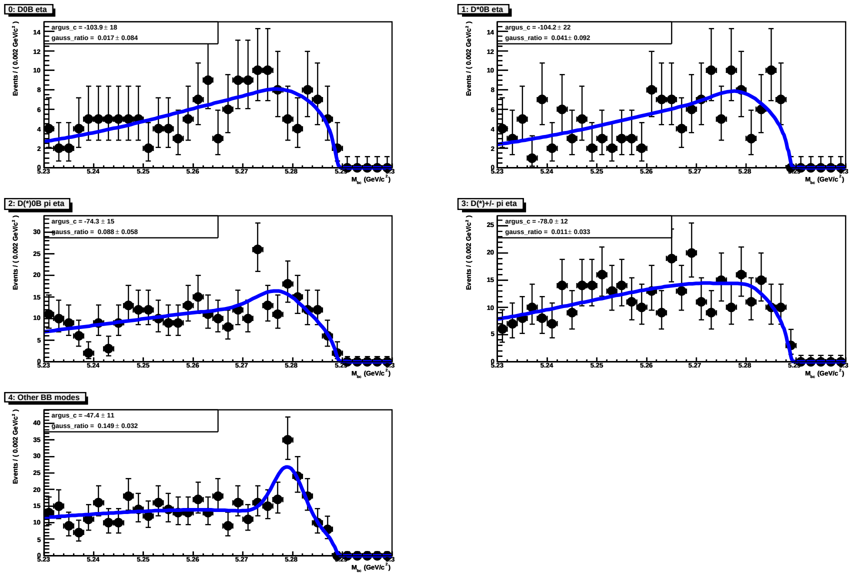

Figure 4.21: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 2.0 GeV/c^2 - 2.2 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

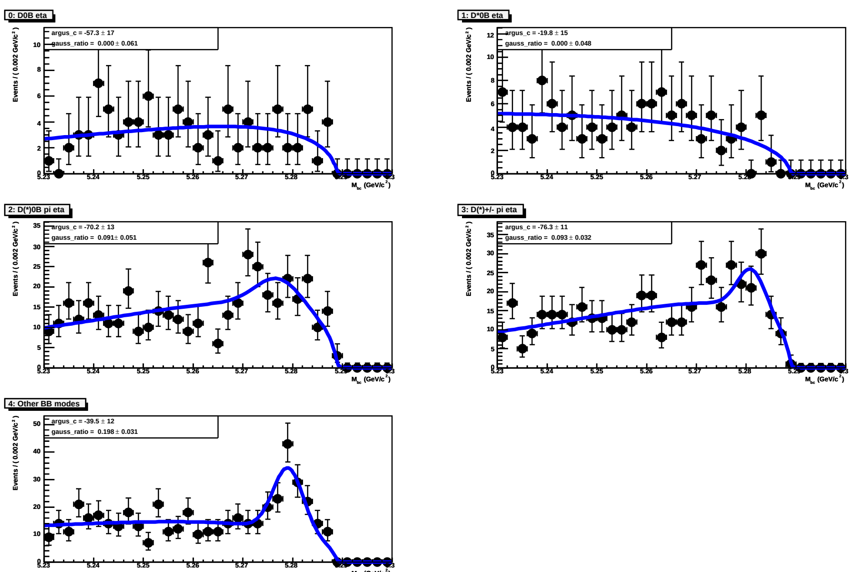

Figure 4.22: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 2.2 GeV/c^2 - 2.4 GeV/c^2. The fits correspond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

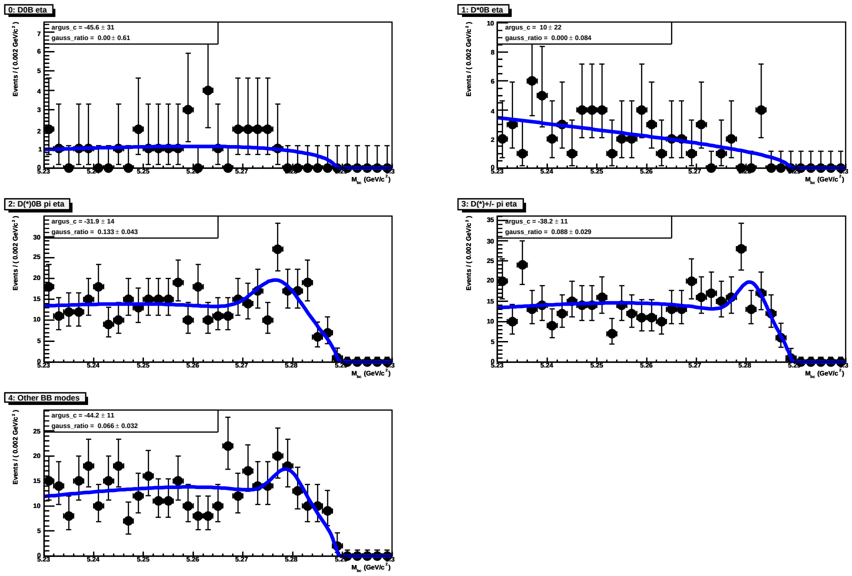

Figure 4.23: $M_{\rm bc}$ fits to the $b \to c$ MC samples for the X_s mass range 2.4 GeV/c^2 - 2.6 GeV/c^2. The fits correpond to: $B^0 \to \bar{D}^0\eta$ (upper left), $B^0 \to \bar{D}^{*0}\eta$ (upper right), $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (middle left), $B^0 \to \bar{D}^{(*)-}\pi^+\eta$ (middle right), and all other $b \to c$ modes (lower left).

veto windows, we can adjust the amount of these generic decays in MC, and compare to the data distributions until a best-fit is found. For each veto window, we define a χ^2_n as

$$\chi^2_n = \sum_i \left(\frac{N^i_{\rm MC} - N^i_{\rm data}}{\sqrt{N_{\rm data}}} \right)^2 \tag{4.5.3}$$

where $N_{\rm MC}$ and $N_{\rm data}$ are the number of events in each bin i of the $M_{\rm bc}$ histograms, for the data or the combined MC samples, respectively. Since various modes may contribute to more than one veto window, we optimize by minimizing the sum over all veto windows,

$$\chi^2 = \sum_n \chi^2_n \tag{4.5.4}$$

This χ^2 can be minimized with respect to a scaling factor that the MC normalization for each three-body mode must be multiplied by to give the best agreement with data. There are four scaling factors in this fitting procedure, but we further simplify by making the assumption that these scaling factors should be identical for D and D^* modes. Thus, the scaling factor for $B^+ \to \bar{D}^0\pi^+\eta$ and $B^+ \to \bar{D}^{*0}\pi^+\eta$ are assumed to be identical. The same assumption is made for $B^0 \to D^-\pi^+\eta$ and $B^0 \to D^{*-}\pi^+\eta$.

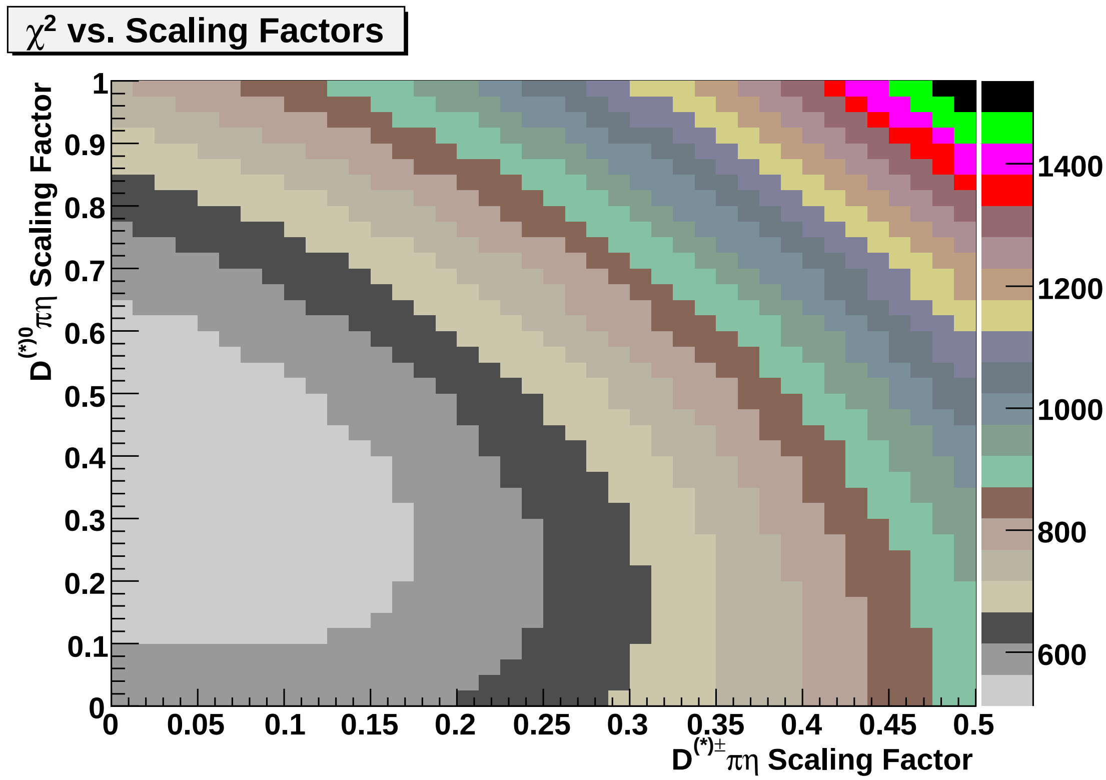

Figure 4.24: Distributions of χ^2, as defined in the text, as a function of the fractional amounts (relative to the generic MC) of $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (y-axis) and $B^0 \to D^{(*)-}\pi^+\eta$ included (x-axis). The minimum of the plot corresponds to the point (0.0,0.4).

The optimization is performed in two steps. We first perform a rough scan of the two dimensional parameter space to locate an approximate minimum. This scan is shown in Figure 4.24. Once the approximate minimum is located, each of the two parameters is varied independently and a second order polynomial fit is performed to the χ^2 profile to determine a more precise minimum. These fits are shown in Figure 4.25.

The remainder of generic contributions, included in a single PDF component, are fixed to their MC expectations. The scaling factors relative to the Belle MC are shown numerically for each PDF component in Table 4.5. Each of these scaling factors, whether fixed from an existing measurement or from the χ^2 procedure, is subject to some uncertainty, leading to potential systematic errors. We reserve discussion of these errors until the next chapter. However, we note here that if we scale the MC using these factors, we obtain reasonable agreement with the data in the veto windows, as shown in Figure 4.26.

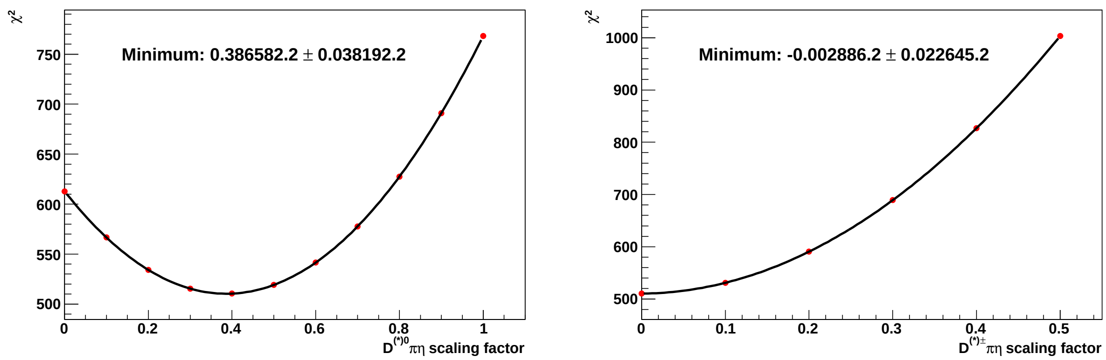

Figure 4.25: Distributions of χ^2, as defined in the text, as a function of the fractional amount (relative to the generic MC) of $B^+ \to \bar{D}^{(*)0}\pi^+\eta$ (left) and $B^0 \to D^{(*)-}\pi^+\eta$ included (right). The points represent an explicit χ^2 calculation, while the curve is a second order polynomial fit to the data to determine a minimum.

Background mode	Scaling factor	Source of value
$B^0 \to \overline{D^0}\eta$	1.07	previous Belle [[46]]
$B^0 \to \overline{D^{*0}}\eta$	0.79	previous Belle [[46]]
$B^0 \to D^-\pi^+\eta$	0.00	veto window χ^2
$B^0 \to D^{*-}\pi^+\eta$	0.00	veto window χ^2
$B^+ \to \overline{D^0}\pi^+\eta$	0.39	veto window χ^2
$B^+ \to \overline{D^{*0}}\pi^+\eta$	0.39	veto window χ^2
Other $b \to c$ decays	1.00	generic MC default

Table 4.5: List of $b \to c$ modes and normalization scaling factors relative to the generic MC.

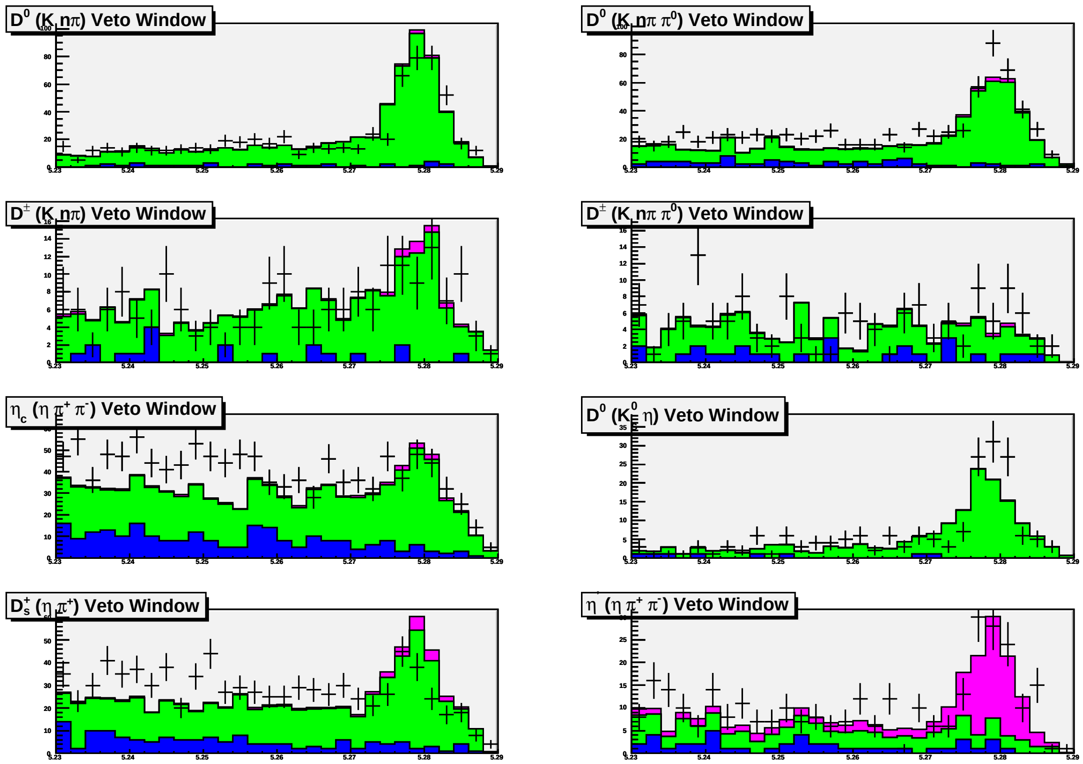

Figure 4.26: $M_{\rm bc}$ distributions in all veto windows defined in this analysis. The text indicates the pseudo-reconstruction used for the veto. Points with errors represent the data, and histograms represent the expectations from continuum (blue), generic (green), and rare (magenta) MC samples. Generic MC expectations are obtained using the scaling factors described in the text.

X_s Mass Range (GeV/c^2)	$B \to X_s\eta'$ Expected Events
0.4 – 0.6	0.00 ± 0.01
0.6 – 0.8	0.00 ± 0.01
0.8 – 1.0	0.02 ± 0.02
1.0 – 1.2	0.00 ± 0.01
1.2 – 1.4	0.01 ± 0.04
1.4 – 1.6	0.00 ± 0.07
1.6 – 1.8	0.06 ± 0.13
1.8 – 2.0	0.03 ± 0.13
2.0 – 2.2	0.09 ± 0.20
2.2 – 2.4	0.46 ± 0.20
2.4 – 2.6	0.00 ± 0.21
Total	0.68 ± 0.41

Table 4.6: Expected contributions from $B \to X_s\eta'$ to the signal yield for the full data sample.

4.5.4 Rare Background Subtraction

The expected contributions from rare backgrounds are generally smaller than their $b \to c$ counterparts. In most cases, one or fewer events per X_s mass bin are expected from these sources. Thus, rather than modeling them with a separate PDF, we subtract these small yields from the yield obtained in the final fit. The method of estimating the amount of the three identified rare backgrounds is unique to each.

$B \to X_s\eta'$

Of the rare backgrounds, the inclusive $X_s\eta'$ decay is the easiest to handle. First, the production of the η through a secondary decay of the η' causes most of these events to fail the p^*_η requirement. Second, we have also vetoed the remaining backgrounds explicitly through the $\eta' \to \eta\pi^+\pi^-$ decay. We can see from the $M_{\rm bc}$ distributions in this veto window that there is no significant disagreement between the data and MC, so we are confident that we can use the MC to accurately assess the contribution of these backgrounds. We perform fits of a Gaussian PDF plus an ARGUS background function to the $B \to X_s\eta'$ component of the rare MC in bins of M_{X_s}, and find the yields given in Table 4.6. The total contribution over all bins of X_s mass is less than one event.

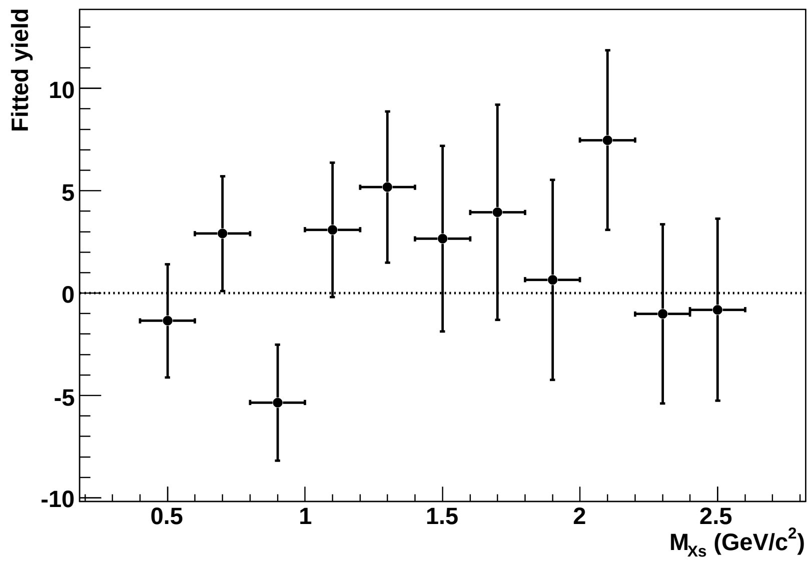

Figure 4.27: Yields of $M_{\rm bc}$ fits performed in the $M_{\gamma\gamma}$ sideband region for the full data sample. The errors are statistical only. This control sample is sensitive to $B \to X_s\gamma$ backgrounds.

$B \to X_s\gamma$

Backgrounds from radiative B decays should not peak in the $M_{\gamma\gamma}$ distribution of the candidate η mesons. Thus, we use the previously defined sideband region (see Figure 4.8) in $M_{\gamma\gamma}$ to verify our expectations from MC. We perform $M_{\rm bc}$ fits in bins of X_s mass using the same procedure as for the final signal fit, but in the sideband region of $M_{\gamma\gamma}$. In the rare MC, this procedure gives a yield of very nearly zero in all bins. We obtain yields in the data shown in Figure 4.27, based on individual fits that are shown in Figures 4.28, 4.29, and 4.30. The data yields are consistent with zero within errors, and therefore also consistent with the MC expectations. Thus, we assume that the MC also correctly models the behavior in the signal region. Because the background PDFs were not easily modeled, we do not use fits to the MC in the signal region to estimate the contributions. Instead, we simply count the number of events expected from the MC in the final signal sample. The results, shown in Table 4.7, indicate contributions of less than half an event in each M_{X_s} bin.

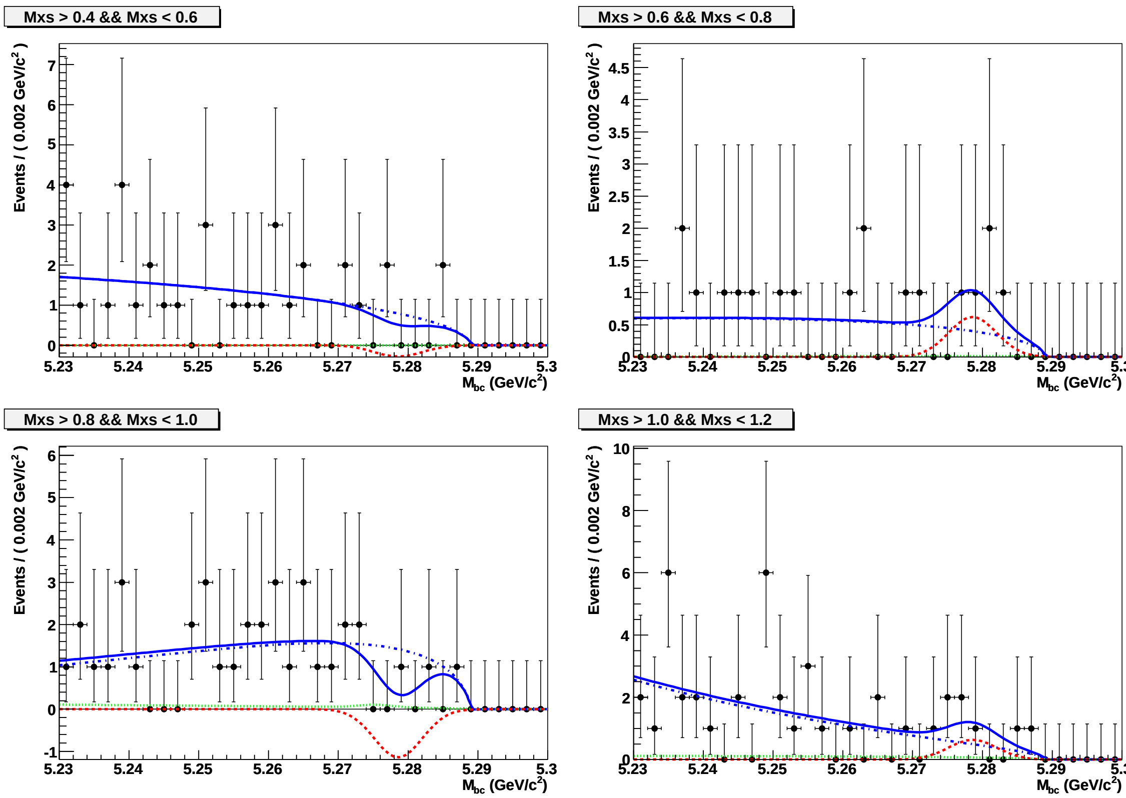

Figure 4.28: $M_{\rm bc}$ fits to the $M_{\gamma\gamma}$ sideband data, for M_{X_s} ranges from 0.4-1.2 GeV/c^2. The points with errors represent the data, and the lines represent the combined PDF (solid blue), signal PDF (dashed red), combinatorial PDF (dash-dotted blue), and the generic PDF (dotted green).

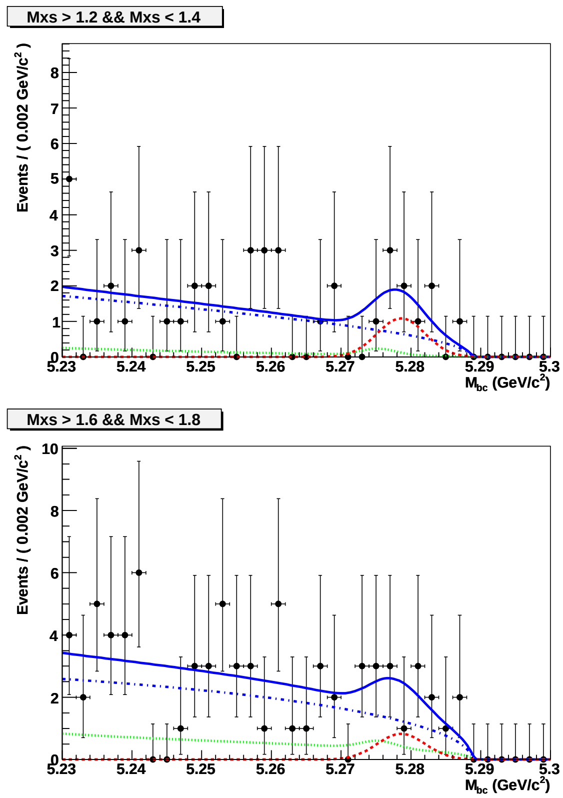

Figure 4.29: $M_{\rm bc}$ fits to the $M_{\gamma\gamma}$ sideband data, for M_{X_s} ranges from 1.2-2.0 GeV/c^2. The points with errors represent the data, and the lines represent the combined PDF (solid blue), signal PDF (dashed red), combinatorial PDF (dash-dotted blue), and the generic PDF (dotted green).

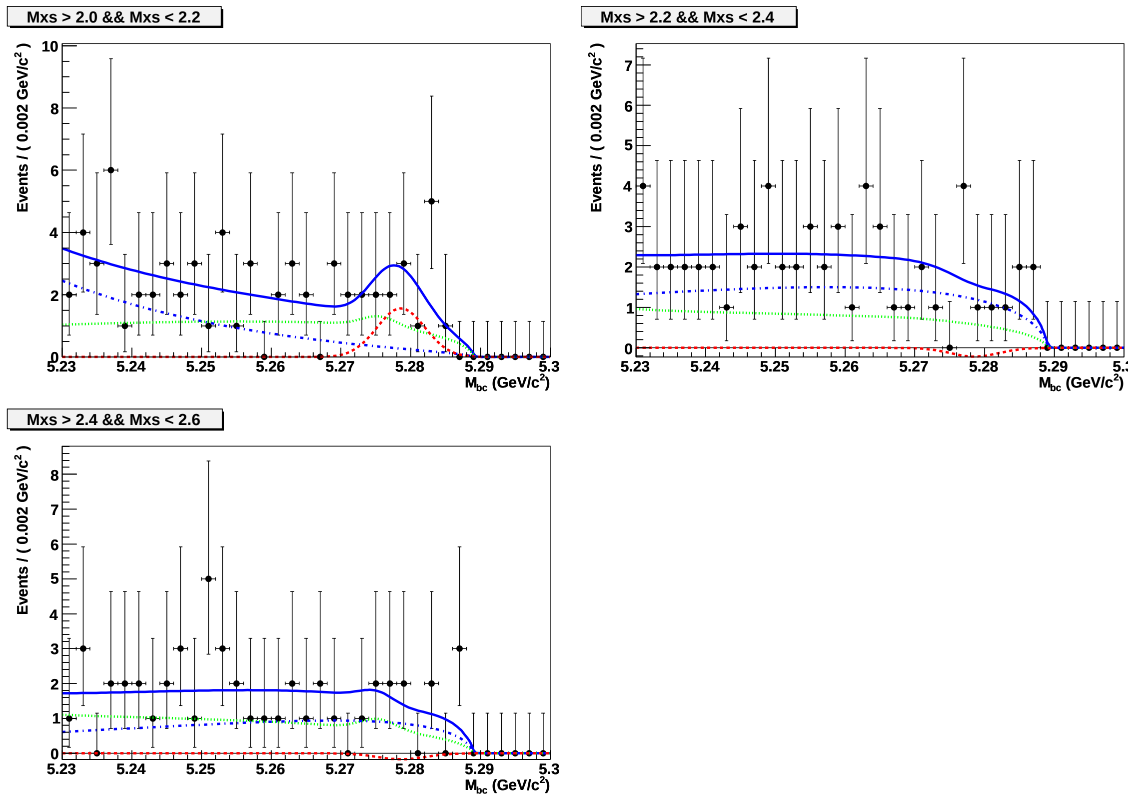

Figure 4.30: $M_{\rm bc}$ fits to the $M_{\gamma\gamma}$ sideband data, for M_{X_s} ranges from 2.0-2.6 GeV/c^2. The points with errors represent the data, and the lines represent the combined PDF (solid blue), signal PDF (dashed red), combinatorial PDF (dash-dotted blue), and the generic PDF (dotted green).

X_s Mass range (GeV/c^2)	$b \to s\gamma$ Expected events
0.4 – 0.6	0.09 $\pm$ 0.04
0.6 – 0.8	0.07 $\pm$ 0.04
0.8 – 1.0	0.20 $\pm$ 0.07
1.0 – 1.2	0.07 $\pm$ 0.04
1.2 – 1.4	0.22 $\pm$ 0.07
1.4 – 1.6	0.17 $\pm$ 0.06
1.6 – 1.8	0.24 $\pm$ 0.07
1.8 – 2.0	0.17 $\pm$ 0.06
2.0 – 2.2	0.20 $\pm$ 0.07
2.2 – 2.4	0.28 $\pm$ 0.08
2.4 – 2.6	0.15 $\pm$ 0.06
Total	**1.85 $\pm$ 0.2**

Table 4.7: Expected contributions from $b \to s\gamma$ in the signal region.

$B \to X_d\eta$

Though a few exclusive modes of the $B \to X_d\eta$ background are measured (e.g., $B^+ \to \pi^+\eta$), many are completely unknown. In some Belle analyses, backgrounds that result from particle misidentification can be distinguished from the signal through their shifted ΔE distribution. In these cases, the signal is extracted through a two-dimensional fit to $M_{\rm bc}$ and ΔE. Unfortunately, our use of ΔE in our best candidate selection causes the ΔE distributions to be biased and unsuitable for fitting.

To obtain an estimate of how much of these backgrounds to subtract, we can repeat our analysis for $B \to X_s\eta$, but replacing the charged kaon in X_s with a pion to create the final state $X_d\eta$. We can then repeat our fitting procedure on data to determine the yield of $X_d\eta$ events. Once we know the fitted yields, we can use our dedicated MC sample of $B \to X_d\eta$ events and study how these modes are reconstructed as $X_s\eta$ with our normal analysis. This study results in a ratio that translates our $X_d\eta$ yields into an $X_s\eta$ yield.

The results of the $X_d\eta$ reconstruction and fits to the data are shown in Figure 4.31, with individual fit results for each M_{X_s} bin in Figures 4.32-4.35. We note that for these fits we have forced the normalizations of the generic PDFs to zero, and we do not subtract out any possible backgrounds to the $X_d\eta$ signal. Further, we use the MC study of X_s/X_d ratios to divide the total backgrounds evenly among all bins (except for the first bin, which

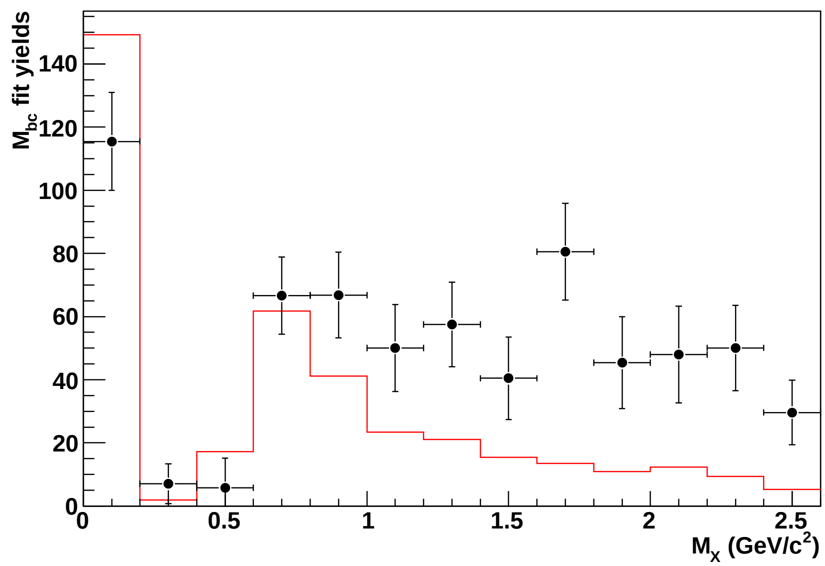

Figure 4.31: Fitted yields for $X_d\eta$ in the data (points with errors), as well as the expectations from the $X_d\eta$ components of the rare MC (red histogram).

we take from the rare MC since the $B^+ \to \pi^+\eta$ branching fraction is measured). Both of these procedures represent potential sources of systematic error that we will treat in the next chapter. Results of the MC study of X_s/X_d reconstructions are shown in Figure 4.36 and Table 4.8. These results are combined with the $X_d\eta$ fit yields to give the amount of backgrounds to ultimately be subtracted, shown in Table 4.9.

4.5.5 Control Fits

Before moving on to fitting the data, we must account for possible differences in the fit functions between the MC data and the experimental data. These may arise from small miscalibrations of the MC parameters relative to the experimental parameters, or from other MC modeling or assumptions that do not describe the data exactly. We study these by using a decay of a similar topology but a much larger branching fraction, so that we can compare the expectations from MC with a data sample of high signal-to-noise ratio. In our case, $B \to D\pi^+$ decays are appropriate, with the D reconstructed in all the modes that we use for our X_s.

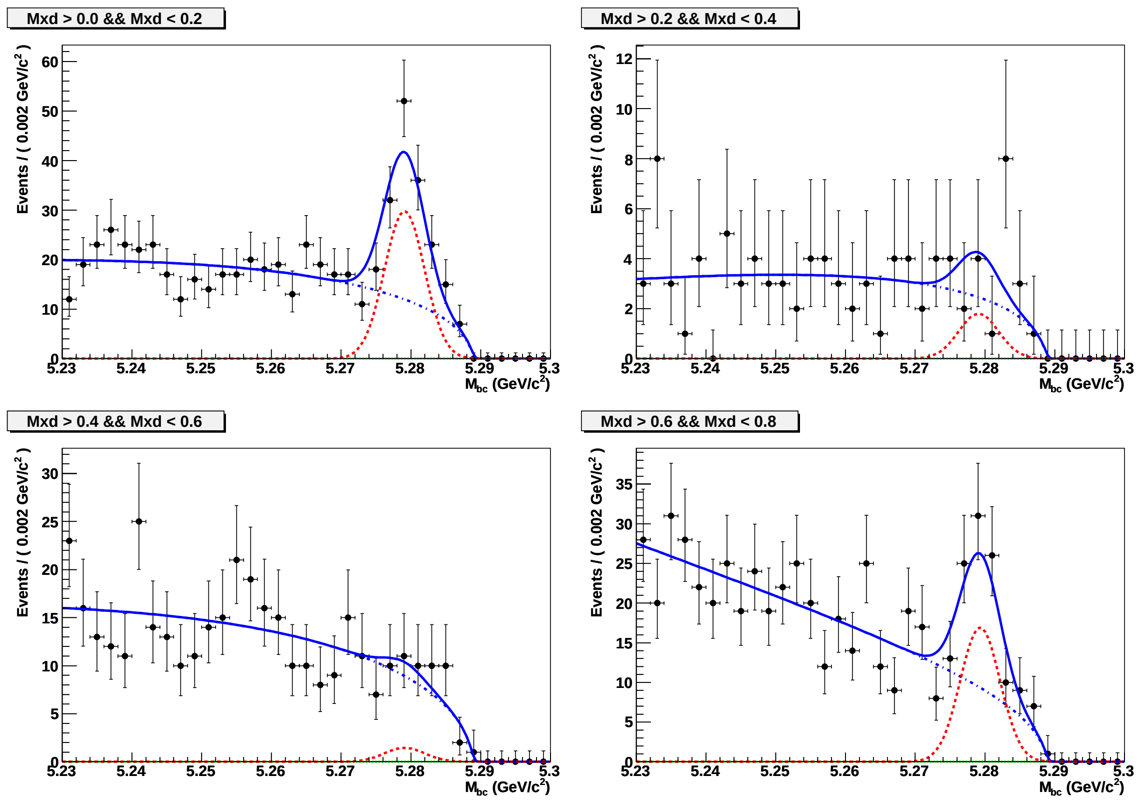

Figure 4.32: $M_{\rm bc}$ fits for the $X_d\eta$ reconstruction in data, for X_d mass from 0.0-0.8 GeV$/c^2$. The points with errors are the data, and the lines represent the total PDF (solid blue), the signal Gaussian (dashed red), and the combinatorial ARGUS (dash-dotted blue).

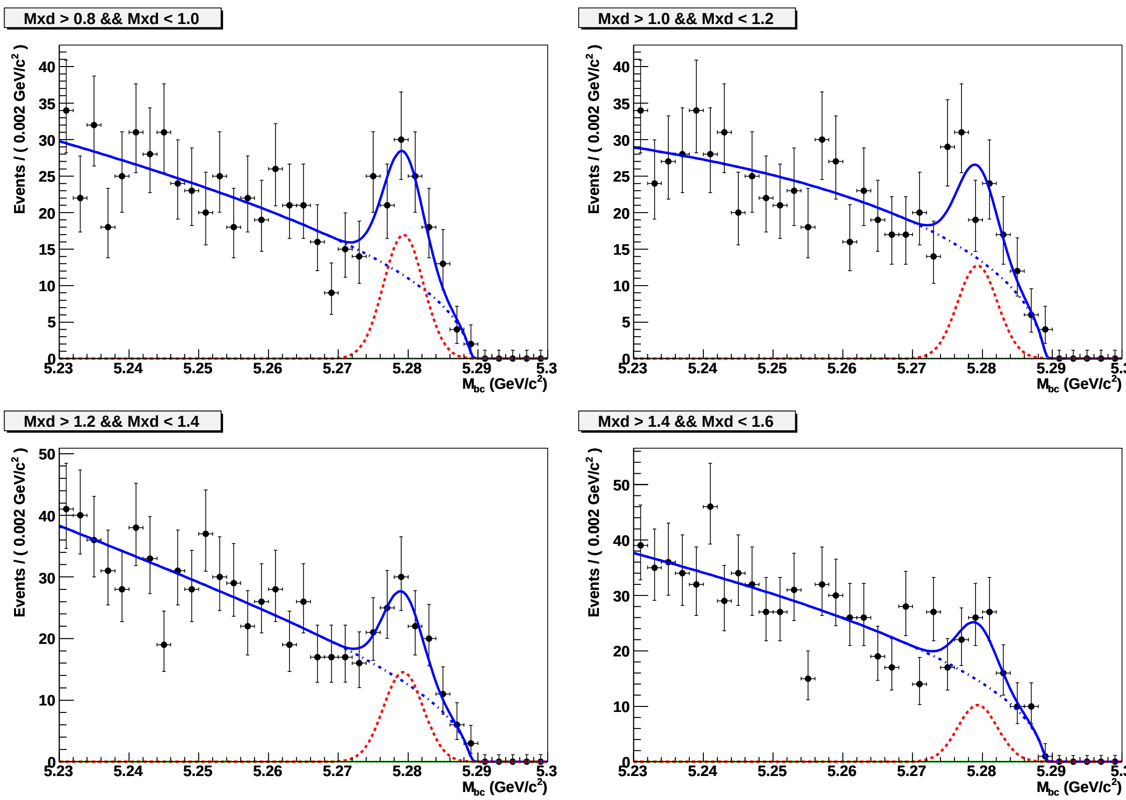

Figure 4.33: $M_{\rm bc}$ fits for the $X_d\eta$ reconstruction in data, for X_d mass from 0.8-1.6 GeV$/c^2$. The points with errors are the data, and the lines represent the total PDF (solid blue), the signal Gaussian (dashed red), and the combinatorial ARGUS (dash-dotted blue).

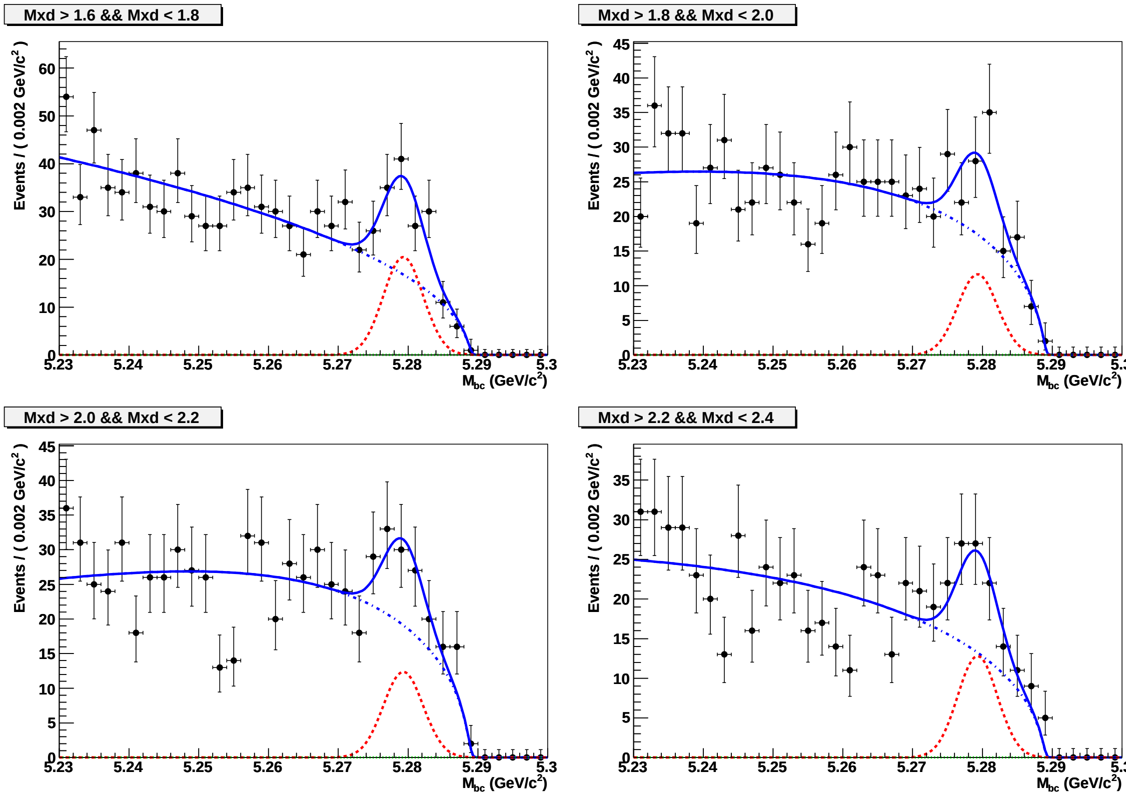

Figure 4.34: $M_{\rm bc}$ fits for the $X_d\eta$ reconstruction in data, for X_d mass from 1.6-2.4 GeV$/c^2$. The points with errors are the data, and the lines represent the total PDF (solid blue), the signal Gaussian (dashed red), and the combinatorial ARGUS (dash-dotted blue).

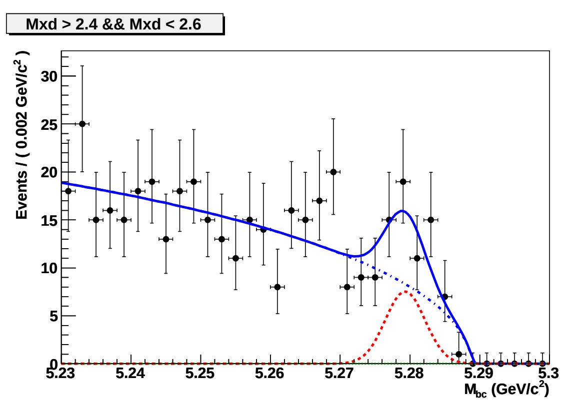

Figure 4.35: $M_{\rm bc}$ fits for the $X_d\eta$ reconstruction in data, for X_d mass from 2.4-2.6 GeV$/c^2$. The points with errors are the data, and the lines represent the total PDF (solid blue), the signal Gaussian (dashed red), and the combinatorial ARGUS (dash-dotted blue).

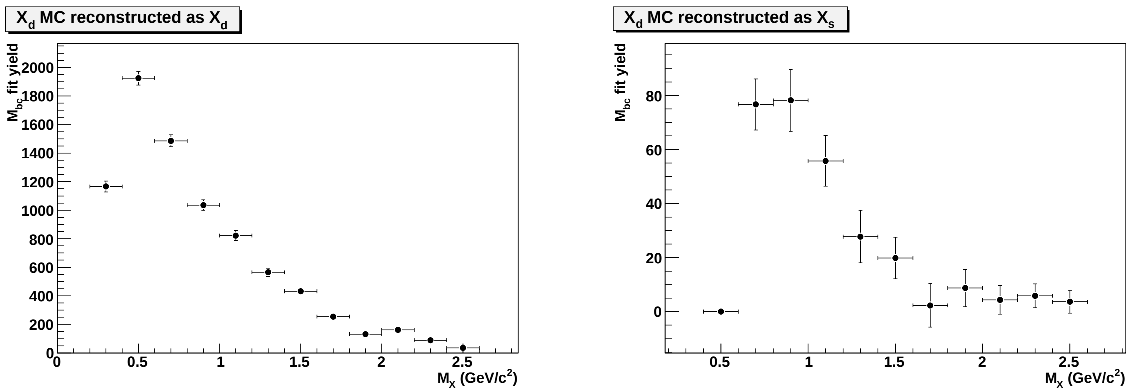

Figure 4.36: Fitted yields for the $X_d\eta$ MC sample reconstructed as $X_d\eta$ (left), as described in the text, and the same sample reconstructed as $X_s\eta$ using the standard signal reconstruction (right).

X_d, X_s Mass range (GeV/c^2)	$X_d\eta$ Reconstruction yield	$X_s\eta$ Reconstruction yield
0.2 – 0.4	1166 $\pm$ 38	
0.4 – 0.6	1925 $\pm$ 47	0.0 $\pm$ 0.7
0.6 – 0.8	1486 $\pm$ 42	77 $\pm$ 9
0.8 – 1.0	1036 $\pm$ 37	78 $\pm$ 11
1.0 – 1.2	822 $\pm$ 34	56 $\pm$ 9
1.2 – 1.4	565 $\pm$ 29	28 $\pm$ 10
1.4 – 1.6	432 $\pm$ 26	20 $\pm$ 8
1.6 – 1.8	255 $\pm$ 21	2.3 $\pm$ 8.0
1.8 – 2.0	131 $\pm$ 14	8.7 $\pm$ 6.9
2.0 – 2.2	162 $\pm$ 15	4.3 $\pm$ 5.3
2.2 – 2.4	89 $\pm$ 11	5.9 $\pm$ 4.4
2.4 – 2.6	35 $\pm$ 7	3.7 $\pm$ 4.2
Total	8105$\pm$ 103	283 $\pm$ 25.3

Ratio of X_s to X_d	0.035 $\pm$ 0.003

Table 4.8: $M_{\rm bc}$ fit yields to the $B \to X_d\eta$ MC sample.

X_s Mass range (GeV/c^2)	$B \to X_d\eta$ Subtracted events
0.4 – 0.6	5.18 $\pm$ 0.39
0.6 – 0.8	1.91 $\pm$ 0.23
0.8 – 1.0	1.91 $\pm$ 0.23
1.0 – 1.2	1.91 $\pm$ 0.23
1.2 – 1.4	1.91 $\pm$ 0.23
1.4 – 1.6	1.91 $\pm$ 0.23
1.6 – 1.8	1.91 $\pm$ 0.23
1.8 – 2.0	1.91 $\pm$ 0.23
2.0 – 2.2	1.91 $\pm$ 0.23
2.2 – 2.4	1.91 $\pm$ 0.23
2.4 – 2.6	1.91 $\pm$ 0.23
Total	**24.28 $\pm$ 2.33**

Table 4.9: Estimated numbers of events from $B \to X_d\eta$ to be subtracted from the final signal yields. The lowest mass bin is the rare MC expectation from $B \to \pi\eta$, and the other mass bins are the scaled expectation from the data-driven estimate of $B \to X_d\eta$.

	$\delta\overline{x}$ (MeV/c^2)	r_σ
SVD1	0.0	1.00
SVD2	0.2	1.03

Table 4.10: Correction factors applied to the signal MC PDF parameters to obtain the final signal PDF, as determined from control sample comparisons.

To account for differences in detector configurations, we perform $M_{\rm bc}$ fits independently for SVD1 and SVD2. Since the signal is quite strong, we can allow the mean and width of the signal Gaussian to float in the fit to the data as well as the MC. We then compare the two fits to determine if there is any notable shift in the mean of $M_{\rm bc}$, which we denote as $\delta\overline{x}$, or any fractional change in the width, denoted as r_σ. The fit results are shown in Figure 4.37 and the resulting correction factors are shown in Table 4.10. The SVD1 samples are found to be consistent within errors for both the mean and width of the Gaussian. The SVD2 samples show a small shift in the mean of the Gaussian, around 0.2 MeV/c^2, as well as a small underestimation of width, by about 3%. These corrections are applied to the fitted $X_s\eta$ signal PDF parameters when fitting to the experimental data.

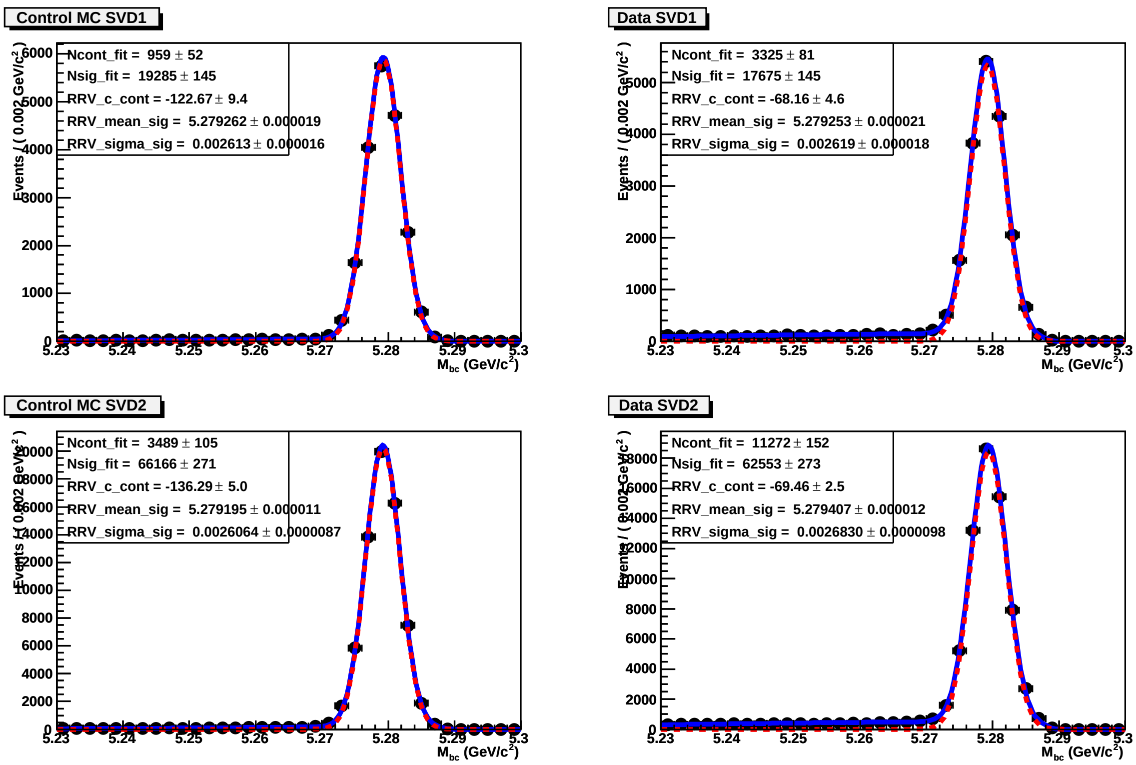

Figure 4.37: A comparison of $M_{\rm bc}$ fit results for SVD1 (above) and SVD2 (below) between the control sample MC (left) and the data from the full reconstruction skim (right).

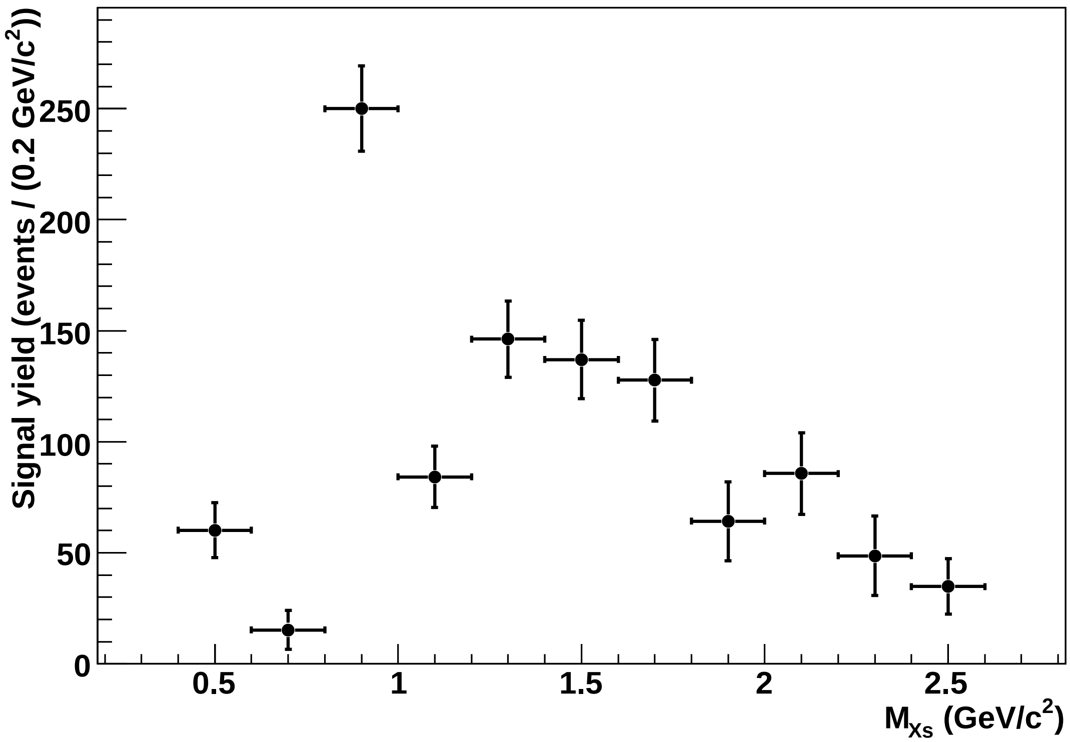

Figure 4.38: Background subtracted $M_{\rm bc}$ data fit yields for each X_s mass bin. Errors are statistical only.

4.6 Results

4.6.1 Fits to Data

The fitting procedure is performed over the data sample of 657×10^6 $B\bar{B}$ pairs, and the expected rare backgrounds are subtracted. The results of the fits are summarized graphically in Figure 4.38 and numerically in Table 4.11. Individual fits corresponding to each bin of M_{X_s} are shown in Figures 4.39, 4.40, and 4.41. Even before we perform further analysis on this spectrum to determine an M_{X_s} dependent branching fraction, we note a few features. First, the known $K\eta$ and $K^*\eta$ signals are evident, as expected. We also see a clear signal at higher X_s mass. Though some of this can be accounted for by the known decays $B \to K^*_{0,2}(1430)\eta$, the signal above $1.8\ \mathrm{GeV}/c^2$ cannot be explained by these decays, and thus does not correspond to any known decay.

X_s Mass range (GeV/c^2)	Fit yield		
0.4 - 0.6	60.2	$\pm$	12.4
0.6 - 0.8	15.3	$\pm$	8.8
0.8 - 1.0	250.0	$\pm$	19.2
1.0 - 1.2	84.2	$\pm$	13.8
1.2 - 1.4	146.3	$\pm$	17.2
1.4 - 1.6	137.0	$\pm$	17.6
1.6 - 1.8	127.7	$\pm$	18.4
1.8 - 2.0	64.2	$\pm$	17.8
2.0 - 2.2	85.7	$\pm$	18.4
2.2 - 2.4	48.6	$\pm$	17.9
2.4 - 2.6	34.8	$\pm$	12.5

Table 4.11: Summary for the fit yields (with expected rare contributions subtracted), along with their statistical errors for each bin.

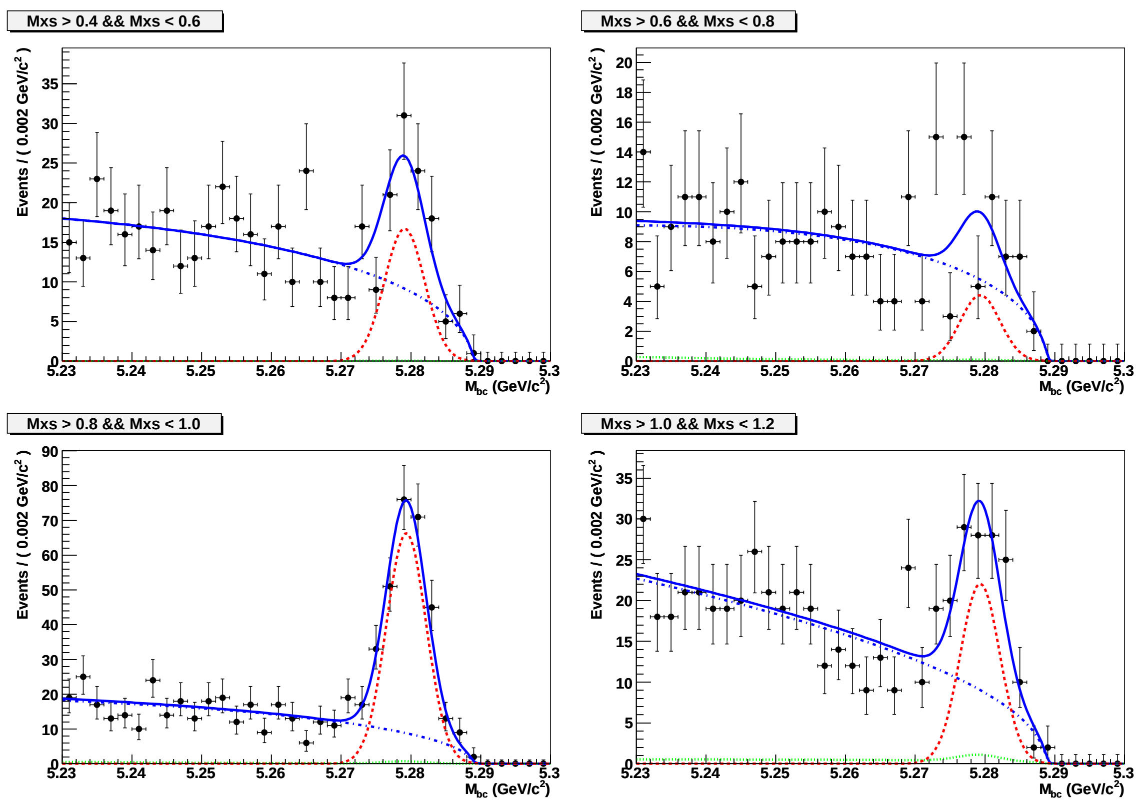

Figure 4.39: $M_{\rm bc}$ fits for the X_s mass bins from 0.4-1.2 GeV/c^2. The points are data and the lines are the total fit (solid blue), the signal component (dashed red), combinatorial component (dash-dotted blue), and the $B\bar{B}$ components (dotted green).

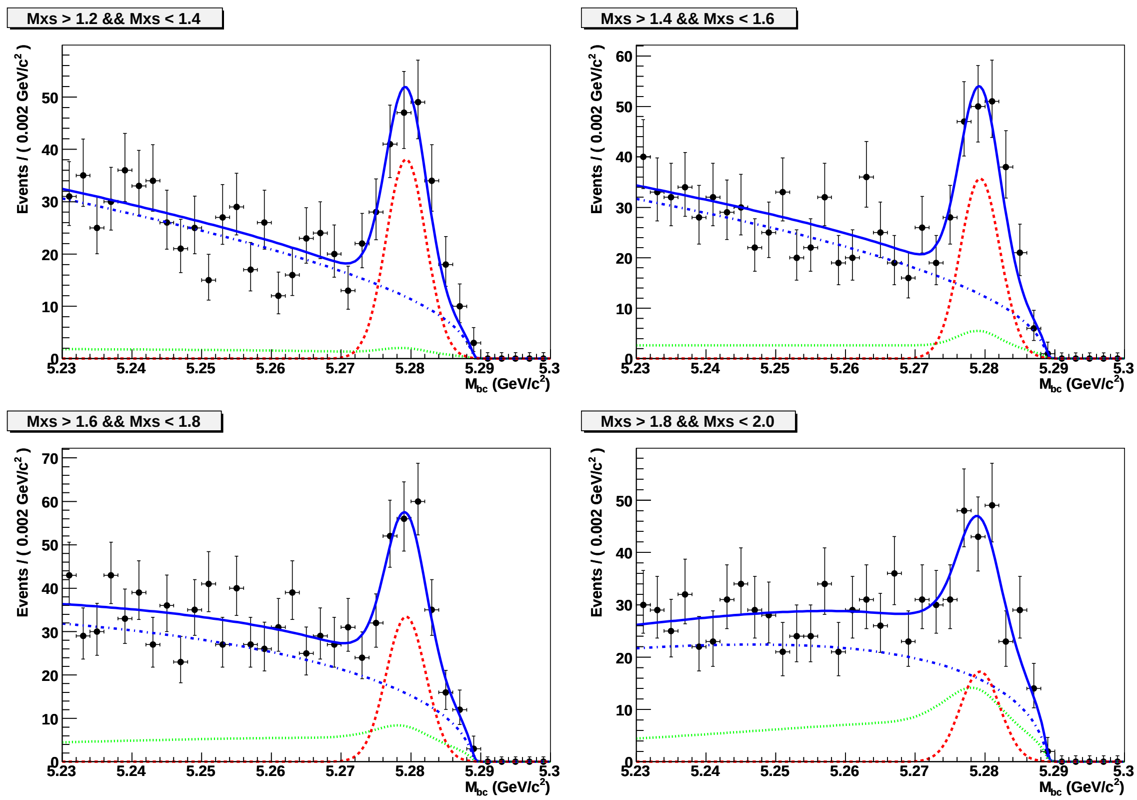

Figure 4.40: M_{bc} fits for the X_s mass bins from 1.2-2.0 GeV/c^2. The points are data and the lines are the total fit (solid blue), the signal component (dashed red), combinatorial component (dash-dotted blue), and the $B\bar{B}$ components (dotted green).

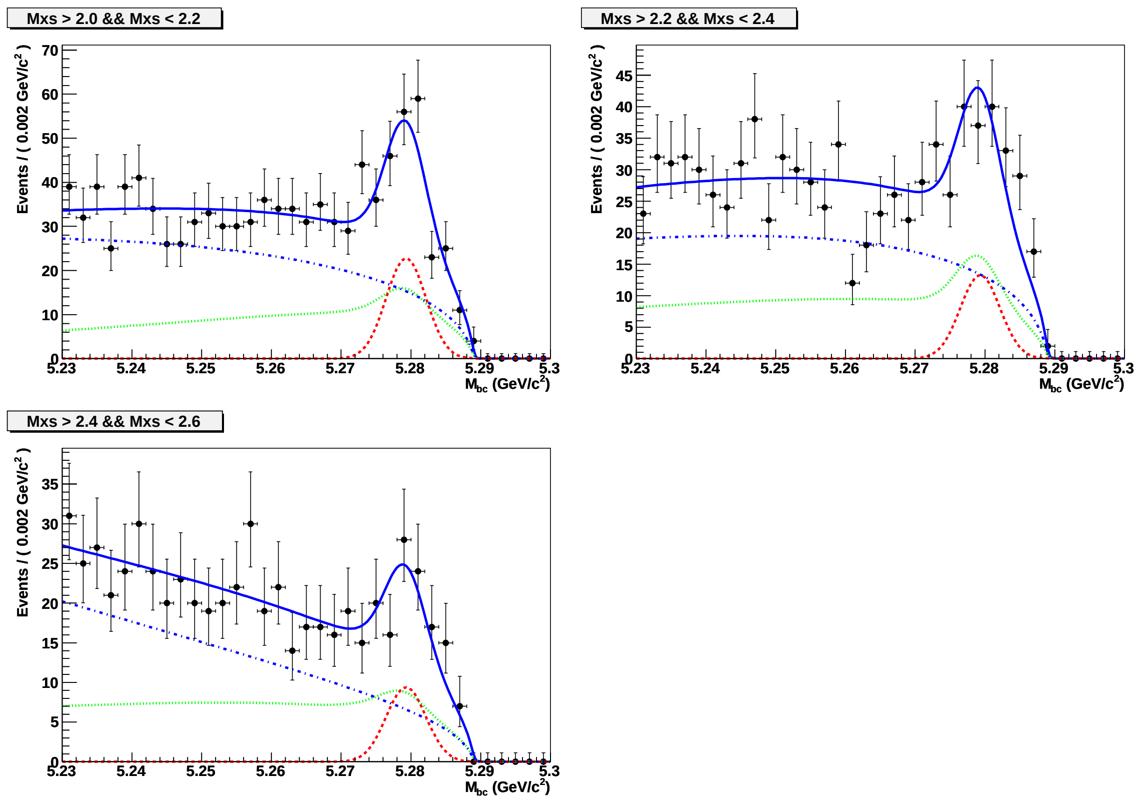

Figure 4.41: $M_{\rm bc}$ fits for the X_s mass bins from 2.0-2.6 GeV$/c^2$. The points are data and the lines are the total fit (solid blue), the signal component (dashed red), combinatorial component (dash-dotted blue), and the $B\bar{B}$ components (dotted green).

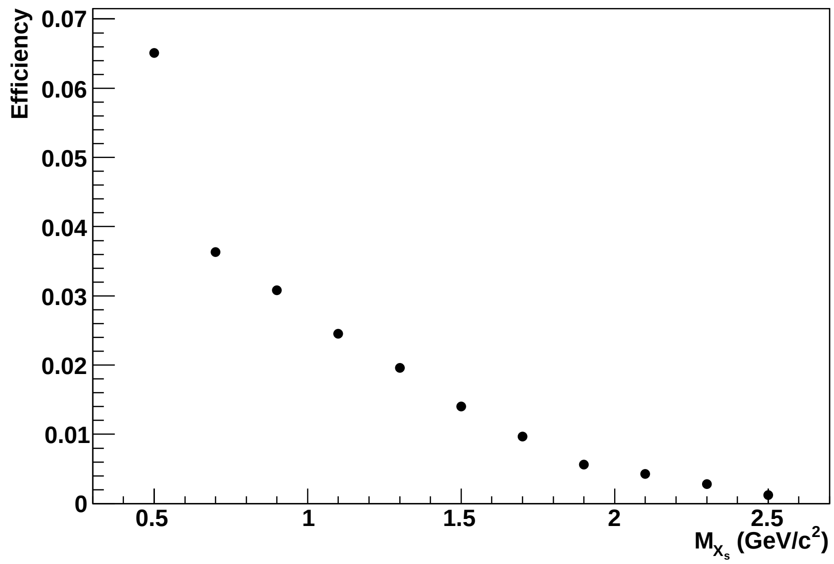

Figure 4.42: Reconstruction efficiency for $B \to X_s\eta$, not including $\mathcal{B}(\eta \to \gamma\gamma)$.

4.6.2 Efficiency & Branching Fraction

To determine a branching fraction for a given X_s mass bin, we must combine our fitted signal yield (with expected backgrounds subtracted) with the efficiency as determined from the flat X_s signal model (or $K\eta$ and $K^*\eta$ in the appropriate mass bins). The branching fraction is calculated as

$$\mathcal{B}(B \to X_s\eta)_i = \frac{N^i_{\mathrm{sig}} - N^i_{X_s\gamma} - N^i_{X_d\eta} - N^i_{X_s\eta'}}{2N_{B\overline{B}}\epsilon_i r_i \mathcal{B}(\eta \to \gamma\gamma)} \tag{4.6.1}$$

where the index i corresponds to an X_s mass bin, $N_{B\bar{B}}$ is the total number of $B\bar{B}$ pairs produced in the data sample (see Appendix C), ϵ is the efficiency as determined from the signal MC, r_i is a bin-dependent correction factor based on differences in efficiency between data and MC (discussed in more detail below), and the $\eta \to \gamma\gamma$ branching ratio accounts for the fact that we force $\eta \to \gamma\gamma$ in our signal MC, but in reality this decay only occurs for ~39% of η mesons [1].

The efficiency is calculated from fits to the signal MC. The fitted yield is determined, and the efficiency is calculated as this yield divided by the number of generated $B \to X_s\eta$ decays. The bin-by-bin efficiency is shown in Figure 4.42.

	$N_{\text{standard}}/N_{\text{loose\_LR}}$	$N_{\text{standard}}/N_{\text{no\_BCS}}$
data	0.528 ± 0.003	0.901 ± 0.005
MC	0.548 ± 0.002	0.902 ± 0.005
r (Data/MC)	0.963 ± 0.006	0.998 ± 0.007

Table 4.12: Ratios of yields for the standard and loose cuts for the requirements on continuum suppression LR and best candidate selection (BCS), calculated with the $B \to D\pi^+$ control sample in data and MC. The corresponding correction factor is calculated as the ratio of the data/MC values.

The r_i correction factors are similar to the adjustments made to the signal PDF using the control sample in that they account for possible systematic differences between the MC and the experimental data. Each r_i is the product of eight individual correction factors due to: π^0 reconstruction, η reconstruction, K_S^0 reconstruction, tracking, particle identification for charged π and K, the continuum suppression LR, and the best candidate selection. All but the last two are taken from previous Belle studies.[6] The correction factors for the LR and candidate selection are unique to this analysis, and must be studied using the $B \to D\pi^+$ control sample. For the LR correction, we define a set of loose LR criteria that are relaxed by 10% across all bins of flavor tag quality. We then calculate the ratio of yields in the control sample using the nominal LR cut and the loose LR cut. This calculation is done for both MC and data, and the data ratio divided by the MC ratio is our LR correction factor. We follow a similar procedure for the candidate selection, but with our loose cut defined as removing the candidate selection entirely. The results of these studies are shown in Table 4.12. No correction factor is applied for the candidate selection, since the calculated value is consistent within errors with one. All correction factors and their sources are shown in Table 4.13.

Combining the fitted yields with efficiencies and associated corrections as per Equation 4.6.1, we calculate a differential branching fraction as a function of X_s mass, $d\mathcal{B}/dM_{X_S}$, as shown in Figure 4.43 and tabulated numerically in Table 4.14.

[6] π^0 and η reconstruction efficiencies are studied using an inclusive η sample [47]; K_S^0 reconstruction is studied by comparing data for $D^+ \to K_S^0\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ [48]; tracking efficiency is studied using inclusive D^* decays [49, 50]; particle identification for both K and π are studied using inclusive D^* decays in which the decay kinematics allow high efficiency identification of $K^\pm$ and $\pi^\pm$ without the use of the PID systems [51].

M_{X_s} range (GeV/c^2)	π^0	$\pi^\pm$ ID	$K^\pm$ ID	Total
0.4 – 0.6	1.000	1.000	0.975	0.923
0.6 – 0.8	0.994	0.984	1.004	0.929
0.8 – 1.0	0.995	0.979	1.001	0.924
1.0 – 1.2	0.992	0.977	1.004	0.920
1.2 – 1.4	0.991	0.973	1.003	0.915
1.4 – 1.6	0.989	0.971	1.003	0.911
1.6 – 1.8	0.987	0.970	1.002	0.908
1.8 – 2.0	0.989	0.967	1.002	0.907
2.0 – 2.2	0.987	0.967	1.003	0.905
2.2 – 2.4	0.988	0.966	1.001	0.904
2.4 – 2.6	0.987	0.966	1.000	0.902

Table 4.13: Summary of efficiency correction factors from various sources. Those that are identical in all bins are not shown, but are included in the total. These are: η reconstruction (0.979), K_S^0 reconstruction (1.003), tracking (1.000), best candidate selection (1.000), and continuum suppression likelihood ratio (0.963).

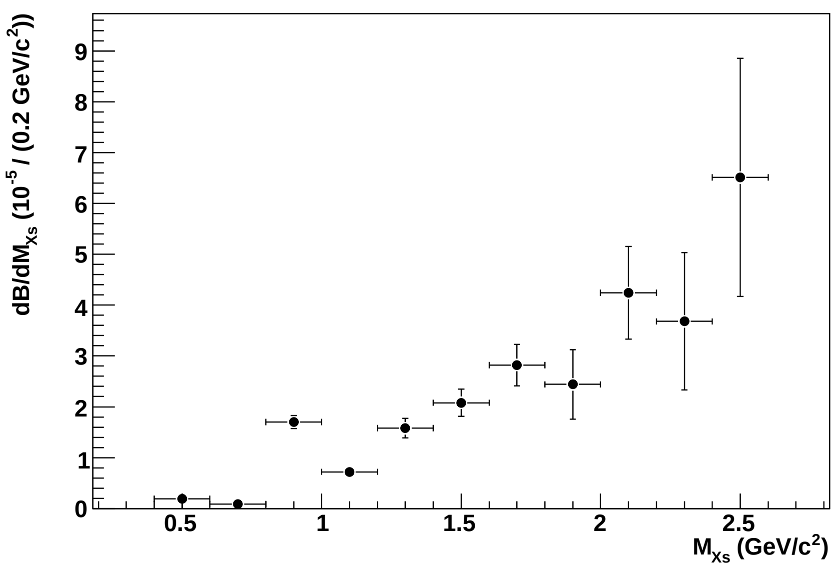

Figure 4.43: Calculated differential branching fraction for $B \to X_s\eta$ as a function of X_s mass. Errors are statistical only.

$M_{X_s}(\mathrm{GeV}/c^2)$	$\mathcal{B}(B \to X_s\eta)(10^{-6})$
0.4–0.6	1.9 ± 0.4
0.6–0.8	0.9 ± 0.5
0.8–1.0	17.0 ± 1.3
1.0–1.2	7.2 ± 1.2
1.2–1.4	15.8 ± 1.9
1.4–1.6	20.8 ± 2.7
1.6–1.8	28.2 ± 4.1
1.8–2.0	24.4 ± 6.8
2.0–2.2	42.4 ± 9.1
2.2–2.4	36.8 ± 13.5
2.4–2.6	65.1 ± 23.4

Table 4.14: Measured M_{X_s} dependent branching fractions for $B \to X_s\eta$. Errors are statistical only.

Mode	Previous Belle $\mathcal{B}(10^{-6})$	This study $\mathcal{B}(10^{-6})$
$B^+ \to K^{*+}(K^0\pi^+)\eta$	$22.6^{+3.1}_{-2.9}$	28.2 ± 3.9
$B^+ \to K^{*+}(K^+\pi^0)\eta$	$20.1^{+4.1}_{-3.9}$	22.7 ± 5.2
$B^0 \to K^{*0}(K^+\pi^-)\eta$	16.9 ± 1.5	19.7 ± 2.1
$B^0 \to K^{*0}(K^0\pi^0)\eta$	$16.7^{+6.3}_{-5.6}$	16.5 ± 6.4

Table 4.15: Mode-by-mode branching fraction comparisons for $B \to K^*\eta$. For the previous Belle result, the errors are combined statistical and systematic uncertainties. For the results of this study, the errors are statistical only.

4.6.3 Branching Fraction Cross Check with $B \to K^*\eta$

A natural question that arises after the calculation of the branching fraction is whether it is consistent with previous measurements. Although the inclusive $B \to X_s\eta$ mode has not been previously measured, one of its exclusive contributions, $B \to K^*\eta$, is a known reference to which we can compare. Unlike our inclusive study, where all final states are combined for the fit, we divide the various $K^*\eta$ modes into the following four final states to calculate branching fractions for each independently: $B^+ \to K^{*+}(K^0_S\pi^+)\eta$, $B^+ \to K^{*+}(K^+\pi^0)\eta$, $B^0 \to K^{*0}(K^+\pi^-)\eta$, and $B^0 \to K^{*0}(K^0_S\pi^0)\eta$. We perform the fits within 1.5 times the measured width of the K^* resonance (± 75 MeV/c^2), and obtain the results shown in Figure 4.44. The calculated branching fractions are compared to those previously measured by Belle [34] in Table 4.15. All results agree within errors.

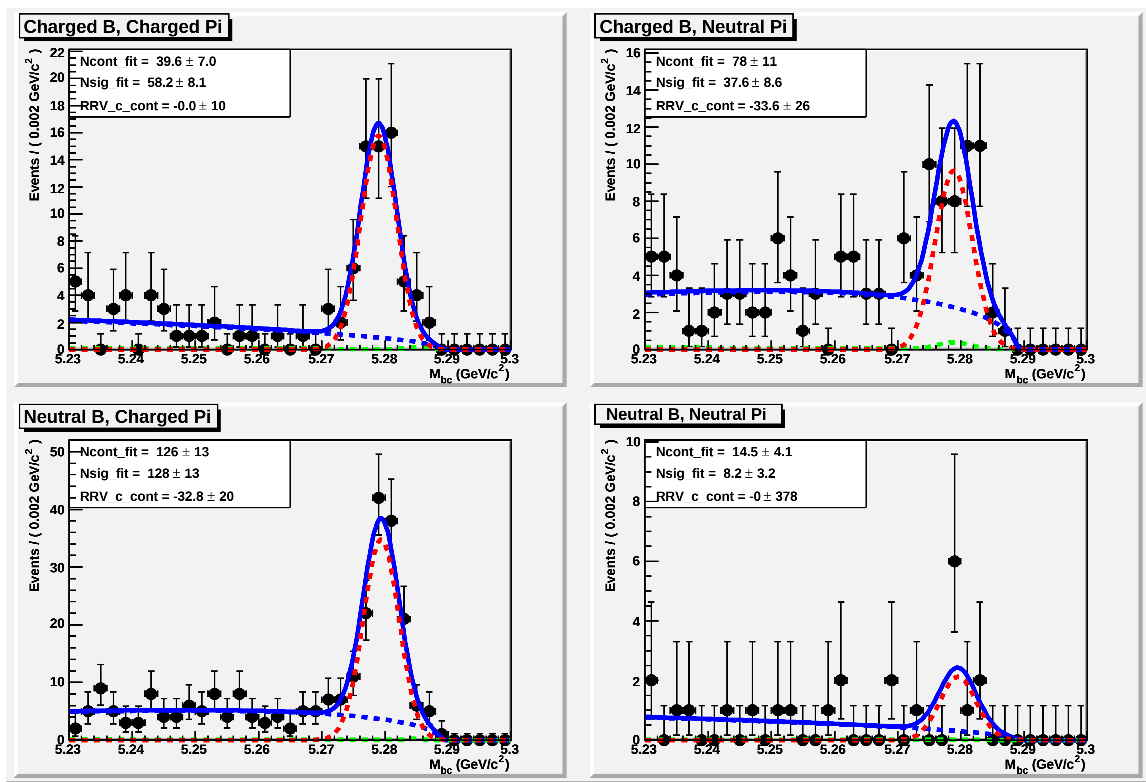

Figure 4.44: $M_{\rm bc}$ fit results to the full data sample for the following modes: $B^+ \to K^{*+}(K^0_S\pi^+)\eta$ (top left), $B^+ \to K^{*+}(K^+\pi^0)\eta$ (top right), $B^0 \to K^{*0}(K^+\pi^-)\eta$ (bottom left), and $B^0 \to K^{*0}(K^0_S\pi^0)\eta$ (bottom right). The red dashed lines, blue dash-dotted lines, and green dotted lines represent the signal, combinatorial background, and generic background contributions. The solid blue line is the total fit result.

4.6.4 Direct CP Asymmetry

We can perform similar fits to measure the direct CP asymmetry of $B \to X_s\eta$. Because the efficiency for $B \to X_s\eta$ should be the same whether it originated from a b or $\bar{b}$ decay, we can rewrite the original definition, Equation 1.3.5, as the following:

$$A_{\rm CP} = \frac{N_b - N_{\overline{b}}}{N_b + N_{\overline{b}}} \tag{4.6.2}$$

Thus, we need only determine the asymmetry in the signal yields between the b and $\bar{b}$ modes. However, the flavor of the b quark cannot be determined for all X_s modes. Thus, we perform this measurement using only modes that have a self-tagged state (i.e., the b quark flavor of the B meson can be completely determined from the final state). This excludes 5 out of 18 modes that have zero net charge and include a K^0_S.

To account for possible mis-reconstruction of a B in the opposite flavor, we define the dilution factor, D,

$$D = \frac{1}{1-2w} \tag{4.6.3}$$

where w is the wrong-tag fraction, the fraction of events where a B was reconstructed in the wrong flavor. This wrong-tag fraction must be calculated using the MC. An event is incorrectly tagged if the best candidate is reconstructed in a flavor with a b quark, but actually originated from a meson with a $\overline{b}$ quark, and vice versa. The wrong-tag fractions and corresponding dilution factors are calculated bin-by-bin in X_s mass in the signal region for the signal MC sample for each M_{X_s} bin. For the bins that use the PYTHIA MC samples as signal MC, we use the flat M_{X_s} MC to determine the wrong-tag fraction. These values can be found in Table 4.16.

Using these dilution factors, we calculate the final measured asymmetry as

$$A_{\rm CP} = DA_{\rm CP}^{\rm raw} = \frac{1}{1-2w}\frac{N_b - N_{\overline{b}}}{N_b + N_{\overline{b}}} \tag{4.6.4}$$

where the raw asymmetry, $A_{\rm CP}^{\rm raw}$, is determined directly from the fit. As with the branching fraction analysis, the expected backgrounds complicate our extraction of $A_{\rm CP}$. In the presence of backgrounds with a nonzero value of $A_{\rm CP}$, we can rewrite the expression as

$$A_{\rm CP}^{\rm true} = \frac{A_{\rm CP}^{\rm measured}N_{\rm total} - \sum_i A^i_{\rm CP}N^i_{\rm bg}}{N_{\rm total} - \sum_i N^i_{\rm bg}} \tag{4.6.5}$$

In the lowest mass bin, effects of $B \to \pi^+\eta$ backgrounds are subtracted. Since the world average measurement of $A_{\rm CP}$ for this process is -0.16 ± 0.07, we can subtract this contribution from the $A_{\rm CP}$ value that we determine. In the other bins, we subtract $B \to X_d\eta$

X_s Mass range (GeV/c^2)	Wrong-tag fraction, w	Dilution factor, D
0.4 – 0.6	0.0000 $\pm$ 0.0000	1.0000 $\pm$ 0.0001
0.6 – 0.8	0.0022 $\pm$ 0.0022	1.0045 $\pm$ 0.0007
0.8 – 1.0	0.0047 $\pm$ 0.0006	1.0095 $\pm$ 0.0011
1.0 – 1.2	0.0095 $\pm$ 0.0006	1.0194 $\pm$ 0.0060
1.2 – 1.4	0.0128 $\pm$ 0.0008	1.0263 $\pm$ 0.0085
1.4 – 1.6	0.0148 $\pm$ 0.0010	1.0305 $\pm$ 0.0092
1.6 – 1.8	0.0174 $\pm$ 0.0013	1.0362 $\pm$ 0.0054
1.8 – 2.0	0.0216 $\pm$ 0.0019	1.0451 $\pm$ 0.0065
2.0 – 2.2	0.0244 $\pm$ 0.0023	1.0514 $\pm$ 0.0166
2.2 – 2.4	0.0185 $\pm$ 0.0026	1.0384 $\pm$ 0.0076
2.4 – 2.6	0.0234 $\pm$ 0.0045	1.0492 $\pm$ 0.0328

Table 4.16: Wrong tag fractions and their corresponding dilution factors as calculated from signal MC.

events, but these do not have a known A_{CP}. We thus assume that for these M_{X_S} bins, the A_{CP} of the subtracted backgrounds is zero. The final results for A_{CP} as a function of M_{X_s} are shown in Figure 4.45 and summarized in Table 4.17. We also determine the value of A_{CP} for the full X_s mass range, and the mass range above the narrow kaonic resonances. As with the branching fractions, we must incorporate systematic effects to obtain a true estimate of our uncertainties.

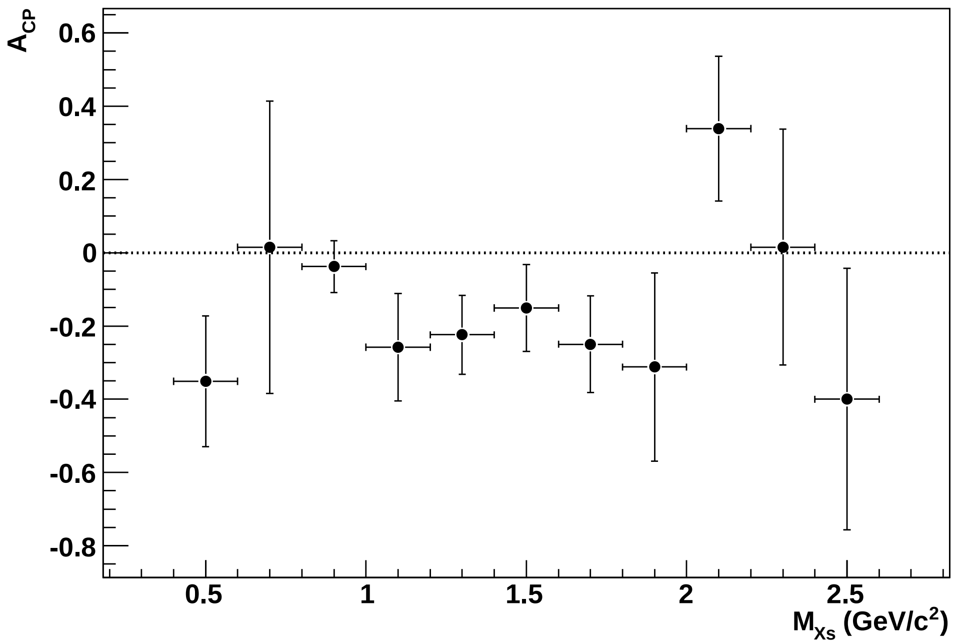

Figure 4.45: A_{CP} values measured for the data, in bins of X_s mass. Errors are statistical only.

X_s Mass range (GeV/c^2)	$A_{\mathrm{CP}}(B \to X_s\eta)$
0.4 – 0.6	-0.35 $\pm$ 0.18
0.6 – 0.8	0.02 $\pm$ 0.40
0.8 – 1.0	-0.04 $\pm$ 0.07
1.0 – 1.2	-0.26 $\pm$ 0.15
1.2 – 1.4	-0.22 $\pm$ 0.11
1.4 – 1.6	-0.15 $\pm$ 0.12
1.6 – 1.8	-0.25 $\pm$ 0.13
1.8 – 2.0	-0.31 $\pm$ 0.26
2.0 – 2.2	0.34 $\pm$ 0.20
2.2 – 2.4	0.02 $\pm$ 0.32
2.4 – 2.6	-0.40 $\pm$ 0.36
0.4 – 2.6	-0.13 $\pm$ 0.04
1.0 – 2.6	-0.15 $\pm$ 0.06

Table 4.17: Direct CP asymmetries measured in the data, bin-by-bin in X_s mass. Errors are statistical only.

Chapter 5

Systematic Errors

Our results from the last chapter give the nominal values for the measurements of this work. The uncertainties reflect only the statistical uncertainties on the fits used to calculate the branching fraction and direct CP asymmetry. Until now, we have not discussed uncertainties that could systematically affect our results. In general, these come from uncertainties in our analysis procedures, potential discrepancies between the MC simulation and the data, and uncertainties due to our modeling of the X_s system. We break up our systematic uncertainties in the branching fraction into two parts: those that are related to our modeling of the X_s system and those that are not. Since modeling effects result in much smaller uncertainties in the direct CP asymmetry, we report all $\mathcal{A}_{\mathrm{CP}}$ systematic uncertainties together.

In addition, we also define a significance for each measurement. The significance is defined as

$$\mathcal{S} = \sqrt{-2\ln\left(\frac{\mathcal{L}_0}{\mathcal{L}_{\max}}\right)} \tag{5.0.6}$$

where $\mathcal{L}_{\max}$ is the maximum likelihood obtained from the fit, and $\mathcal{L}_0$ is the maximum likelihood obtained when the parameter of interest is fixed to zero. For the branching fraction, $\mathcal{L}_0$ is evaluated by fixing the signal yield to zero. For the direct CP asymmetry, it is evaluated by fixing the fitted $\mathcal{A}_{\mathrm{CP}}^{\mathrm{raw}}$. Systematic effects are included in significance estimates by convolving the likelihood function with a Gaussian of width determined by the additive systematic errors.

5.1 Branching Fraction - Non-model Systematic Uncertainties

Within the non-modeling systematic uncertainties, we further distinguish between additive and multiplicative effects. Additive effects are those that can affect the signal yield, and arise from the fitting procedures and the rare background subtraction. We assign systematic errors due to the PDF shapes and normalizations and each rare background subtraction.

5.1.1 PDF Shapes and Normalizations

Uncertainties due to PDF parameters that are fixed in the fitting procedure are estimated by determining an uncertainty on these parameters, then repeating the fitting procedure and examining how the final fit yields are affected. The details are different depending on the PDF of interest.

Signal PDFs

In Section 4.5.5 we used a $B \to D\pi^+$ control sample to estimate possible differences between signal PDF parameters in data and MC. The presence of such differences indicates a systematic error in the MC relative to the data. To account for these differences, we repeat the fitting procedure on the data, but we allow the signal PDF parameters to vary by plus or minus the observed differences between the control sample data and MC. Because the SVD1 data sample was consistent with MC, we use the SVD2 differences.

$B\bar{B}$ PDFs

We separate errors from the $B\bar{B}$ backgrounds into two components: those due to the fixed PDF shapes, and those due to our choices of PDF normalizations. Errors due to the PDF shapes are obtained by performing a series of fits where each shape parameter that was fixed in the final fit (Gaussian mean, Gaussian sigma, ARGUS shape, ratio of Gaussian to ARGUS) is allowed to vary by the symmetric statistical errors obtained from the fit to the MC.

Our choices of normalizations for the $B\bar{B}$ PDFs are related to the χ^2 procedures discussed earlier, so they play a vital role in estimating the associated systematic uncertainties. In the case of the modes $B^0 \to \bar{D}^0\eta$ and $B^0 \to \bar{D}^{*0}\eta$, we chose normalizations based

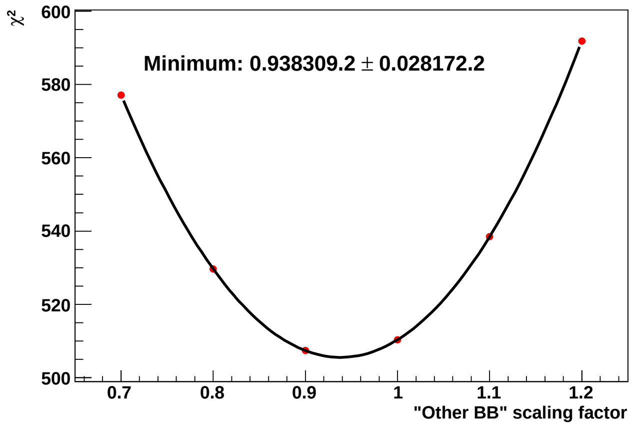

Figure 5.1: Distribution of χ^2 as a function of the fraction (relative to the nominal generic MC) of "other" BB modes.

on the χ^2 procedure. To estimate a systematic uncertainty related to this choice, we repeat the χ^2 procedure, but allow the normalizations for these modes to vary.[1] We compare the difference between the normalization values obtained from this minimization and those based on the previous Belle measurement. The difference between these is the amount that we vary the final normalizations of these modes.

We estimate effects related to the normalizations for the $B \to D^{(*)}\pi\eta$ modes using the one dimensional fits to the χ^2 profile (see Figure 4.25) to estimate an uncertainty on the obtained normalizations. This uncertainty is the amount by which these normalizations are allowed to vary.

For the PDF consisting of all other $b \to c$ backgrounds, we perform another χ^2 minimization where the normalization for these modes is allowed to vary.[2] The results of this minimization are shown in Figure 5.1. We vary the final normalization for this PDF by the difference between the value obtained from the χ^2 study, 0.94, and unity, which is the expected value from the MC.

We summarize the various minimization results and the variations used for the systematics in Table 5.1.

[1]During this procedure, we fix the the normalizations for the three body $D^{(*)}\pi\eta$ modes to their nominal values.

[2]During this minimization, all other $B\bar{B}$ modes are fixed to their nominal values.

Background mode	χ^2 Result	Nominal	Variation	Source
$B^0 \to \bar{D}^0\eta$	0.87±0.12	1.07	±0.20	χ^2 / Belle difference
$B^0 \to \bar{D}^{*0}\eta$	1.31±0.37	0.79	±0.52	χ^2 / Belle difference
$B^0 \to D^{(*)-}\pi^+\eta$	0.00+0.02	0.00	+0.02	χ^2 minimum error
$B^+ \to \bar{D}^{(*)0}\pi^+\eta$	0.39±0.04	0.39	±0.04	χ^2 minimum error
Other $b \to c$ decays	0.94±0.03	1.00	±0.06	χ^2 / MC difference

Table 5.1: Summary of scaling factors for $b \to c$ modes, as obtained from the χ^2 procedure and the nominal values used in the fit, as well as the variations used for systematic studies and their sources.

M_{X_s} (GeV/c^2)	Sig. shape	BB shape	BB norm.	Sub total	Sub total (%)
0.4–0.6	0.97	0.18	0.04	0.98	1.63%
0.6–0.8	0.71	0.31	0.04	0.77	5.06%
0.8–1.0	3.46	1.01	0.15	3.61	1.44%
1.0–1.2	1.82	1.33	0.18	2.26	2.69%
1.2–1.4	2.94	1.64	0.23	3.37	2.31%
1.4–1.6	2.49	2.76	0.81	3.80	2.77%
1.6–1.8	1.89	4.77	1.75	5.42	4.24%
1.8–2.0	1.53	7.74	3.33	8.56	13.33%
2.0–2.2	1.44	5.24	1.91	5.76	6.72%
2.2–2.4	3.27	6.05	1.82	7.11	14.64%
2.4–2.6	1.29	4.33	0.73	4.57	13.14%

Table 5.2: Positive systematic errors related to fitting. All entries are in number of events, except for the last column, which are percentages relative to the nominal yields.

Summary

Using the methods described above, the fits are repeated and the differences from the nominal values are used to assign systematic uncertainties on the fitted yields. These are summarized independently for positive and negative errors in Tables 5.2 and 5.3. The subtotals for fitting uncertainties are determined by adding the individual errors (in numbers of events) in quadrature.

5.1.2 Rare Backgrounds

The subtracted rare backgrounds are also subject to uncerainties, and should be reflected in the final result. For both the $X_s\eta'$ and $X_s\gamma$ modes, we conservatively allow the

M_{X_s} (GeV/c^2)	Sig. Shape	BB Shape	BB Norm.	Sub Total	Sub Total (%)
0.4-0.6	1.10	0.07	0.04	1.10	1.83%
0.6-0.8	0.71	0.43	0.04	0.83	5.46%
0.8-1.0	3.77	1.06	0.15	3.92	1.57%
1.0-1.2	1.96	1.37	0.18	2.40	2.85%
1.2-1.4	3.19	2.25	0.23	3.92	2.68%
1.4-1.6	2.78	3.04	0.81	4.20	3.07%
1.6-1.8	2.09	4.55	1.75	5.31	4.16%
1.8-2.0	1.60	6.91	3.37	7.85	12.22%
2.0-2.2	1.60	6.22	1.97	6.71	7.84%
2.2-2.4	3.22	6.16	1.79	7.18	14.78%
2.4-2.6	1.43	6.02	0.78	6.24	17.92%

Table 5.3: Negative systematic errors related to fitting. All entries are in number of events, except for the last column, which are percentages relative to the nominal yields.

amount of each background to vary by $\pm 100\%$. For the former, this variation reflects the large uncertainty in the known branching fraction.

The sources of uncertainties on the subtracted $X_d\eta$ backgrounds vary by M_{X_s} bin. For the lowest bin, which is expected to contain only contributions from $B^+ \to \pi^+\eta$, we vary the amount of the subtraction by $\pm 30\%$, the relative disagreement between the data fits and the MC expectation (as seen in Figure 4.31). For the other bins, where the MC expectation is not based on a measured branching fraction, we must use our data-driven procedures to estimate the uncertainties. As mentioned previously, these estimates did not include potential backgrounds to $B \to X_d\eta$ modes, so our fitted yields can be considered upper limits. As such, we allow the subtracted $X_d\eta$ contributions to vary by -100% in all bins.

For the positive variation, we quantify the uncertainty arising from the assumption of a uniform distribution of events throughout all X_s mass bins. To do this, we perform a study of the dedicated $X_d\eta$ MC sample, in which we determine a relationship between the reconstructed X_s mass and generated X_d mass, shown in Figure 5.2. Using the fractional distributions shown in the lower part of the figure, we can recalculate how the X_d events migrate into X_s mass bins. We determine the difference in number of events in this estimate versus the estimate in which the events are uniformly distributed. For the systematic variation, we use the larger of this difference and the uncertainty in the uniformly distributed values. These results are summarized in Table 5.4.

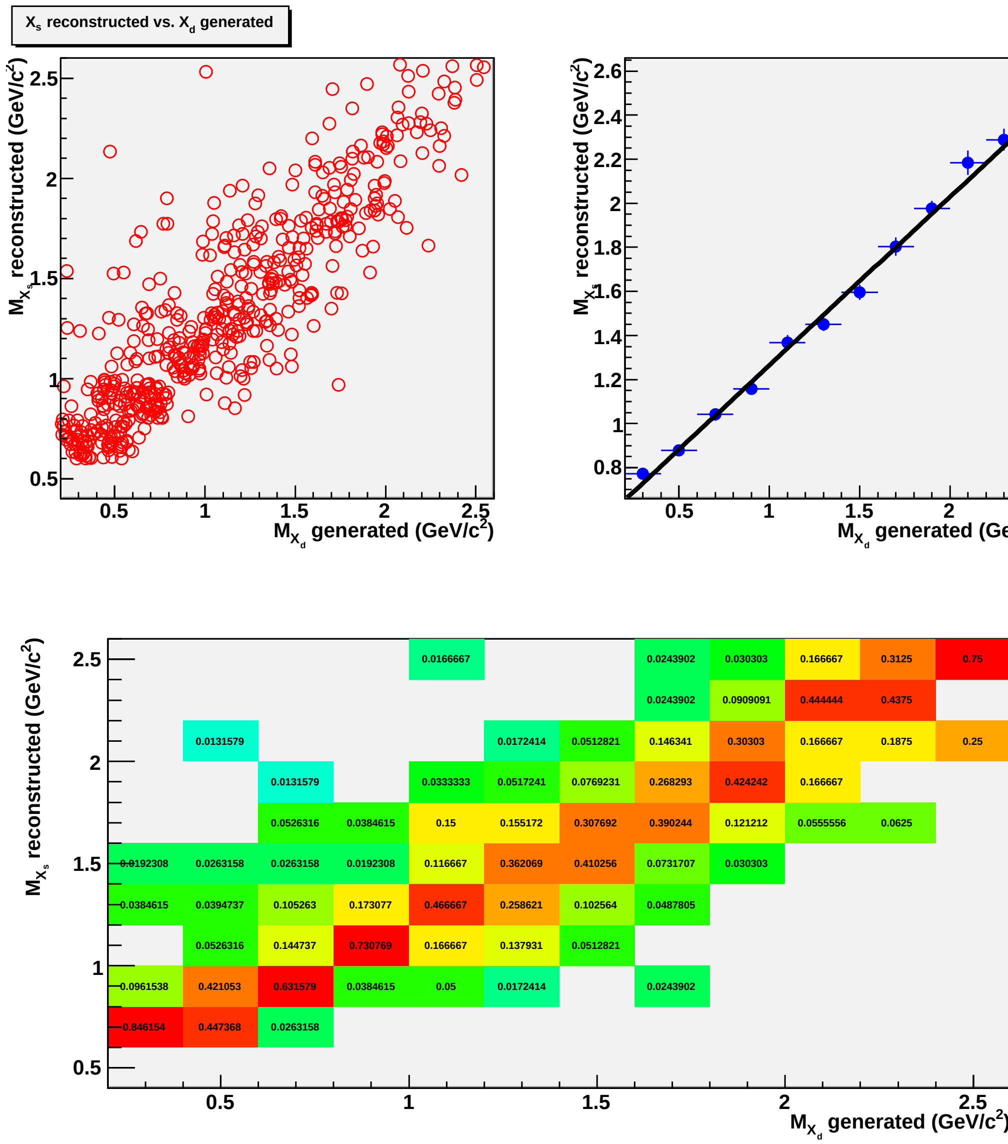

Figure 5.2: Reconstructed X_s mass versus generated X_d mass in the signal region for the $B \to X_d \eta$ PYTHIA Monte Carlo sample, represented as a scatter plot (upper left), a profile histogram with a linear fit (upper right), and numerically (bottom).

M_{X_s} (GeV/c^2)	N_{X_d} (standard)	N_{X_d} (weighted)	$\delta N_{X_d}+$	$\delta N_{X_d}-$
0.4–0.6	5.18	5.18	1.55	1.55
0.6–0.8	1.91	0.36	0.23	1.91
0.8–1.0	1.91	1.86	0.23	1.91
1.0–1.2	1.91	2.69	0.78	1.91
1.2–1.4	1.91	2.28	0.37	1.91
1.4–1.6	1.91	1.88	0.23	1.91
1.6–1.8	1.91	2.71	0.80	1.91
1.8–2.0	1.91	2.01	0.23	1.91
2.0–2.2	1.91	1.87	0.23	1.91
2.2–2.4	1.91	1.72	0.23	1.91
2.4–2.6	1.91	1.75	0.23	1.91

Table 5.4: Expected $X_d\eta$ backgrounds using the standard method, the weighting based on the $X_d \to X_s$ study, and the final variations used for systematics.

The final systematic uncertainties due to rare $B\bar{B}$ backgrounds are the combination of each of the three backgrounds in quadrature. These are summarized in Table 5.5.

5.1.3 Multiplicative Uncertainties

The final class of non-model related uncertainties are due to factors that multiply the efficiency in the denominator of Equation 4.6.1: the efficiency correction factors, the branching fraction of $\eta \to \gamma\gamma$, and the total number of $B\bar{B}$ pairs.

The uncertainties from the correction factors for K_S^0 reconstruction, π^0 reconstruction, η reconstruction, charged K and π identification, and tracking all come from the same Belle studies used to estimate these correction factors [47, 48, 49, 50, 51]. For those that we calculated using our control sample (LR and candidate selection), we assign an uncertainty based on the difference between our correction factors and one. The uncertainty due to the branching fraction of $\eta \to \gamma\gamma$ is based on the world average measurement uncertainty. Finally, the uncertainty on $N_{B\bar{B}}$ is based on dedicated measurements for each experiment number, explained in more detail in Appendix C. Since all these terms are multiplicative, we combine the uncertainties by adding the relative errors in quadrature. The values for each term, and the combined total, are shown in Table 5.6.

M_{X_s} (GeV/c^2)	$N_{X_s\gamma}$ ($\pm$)	$N_{X_s\eta'}$ ($\pm$)	$N_{X_d\eta}$ ($+$)	$N_{X_d\eta}$ ($-$)	Subtotal ($+$)	Subtotal ($-$)	Sub Total ($+\%$)	Subtotal ($-\%$)
0.4–0.6	0.09	0.00	1.55	1.55	1.56	1.56	2.58%	2.58%
0.6–0.8	0.07	0.00	0.23	1.91	0.24	1.91	1.57%	12.51%
0.8–1.0	0.20	0.02	0.23	1.91	0.30	1.92	0.12%	0.77%
1.0–1.2	0.07	0.00	0.78	1.91	0.78	1.91	0.93%	2.27%
1.2–1.4	0.22	0.01	0.37	1.91	0.43	1.92	0.29%	1.31%
1.4–1.6	0.17	0.00	0.23	1.91	0.29	1.92	0.21%	1.40%
1.6–1.8	0.24	0.06	0.80	1.91	0.84	1.93	0.66%	1.51%
1.8–2.0	0.17	0.03	0.23	1.91	0.29	1.92	0.45%	2.99%
2.0–2.2	0.20	0.09	0.23	1.91	0.32	1.92	0.37%	2.24%
2.2–2.4	0.28	0.46	0.23	1.91	0.59	1.98	1.21%	4.09%
2.4–2.6	0.15	0.00	0.23	1.91	0.28	1.92	0.79%	5.50%

Table 5.5: Systematic errors related to subtraction of rare backgrounds.

M_{X_s} (GeV/c^2)	K_S^0 ($\pm$%)	π^0 ($\pm$%)	$\pi^\pm$ ($\pm$%)	$K^\pm$ ($\pm$%)	Tracking ($\pm$%)	Total ($\pm$%)
0.4–0.6	1.08	0.00	0.00	0.74	0.77	5.07
0.6–0.8	1.07	0.40	0.37	0.70	1.60	5.29
0.8–1.0	1.08	0.38	0.38	0.72	1.81	5.36
1.0–1.2	1.06	0.57	0.49	0.76	2.04	5.47
1.2–1.4	1.03	0.65	0.59	0.80	2.35	5.61
1.4–1.6	1.02	0.70	0.68	0.84	2.62	5.74
1.6–1.8	1.01	0.74	0.78	0.87	2.92	5.91
1.8–2.0	0.92	0.62	0.92	0.88	3.25	6.07
2.0–2.2	0.96	0.72	0.93	0.92	3.20	6.07
2.2–2.4	0.95	0.75	0.97	0.95	3.31	6.14
2.4–2.6	0.89	0.76	0.98	0.97	3.40	6.19

Table 5.6: Systematic errors related to factors that multiply the efficiency. Errors that are identical in all bins are not shown, but are included in the total. These are: η reconstruction (2.69%), continuum suppresion likelihood ratio (3.68%), candidate selection (0.75%), $\mathcal{B}(\eta \to \gamma\gamma)$ (0.51%), and $N_{B\bar{B}}$ (1.36%).

5.2 Branching Fraction - Modeling Systematic Uncertainties

Until now, we have not investigated the effects of our assumption of signal model. Our assumption of a flat X_s mass spectrum may affect our final results, and these effects should be included in our uncertainty. Further, our use of the PYTHIA generator to perform the hadronization of the X_s system may also introduce some bias.

5.2.1 Spectrum Model

We investigate the effects of our flat M_{X_s} MC by generating two other MC models. The first uses an X_s system that has a mass distribution based on the QCD anomaly model [18]. The second assumes a three body decay of the type $B \to (u,d)\bar{s}\eta$ to calculate the η momentum, then uses conservation of momentum to assign an appropriate recoil mass. The mass spectra generated by these models, as well as the corresponding calculated efficiencies, are shown in Figure 5.3. The bin-by-bin efficiencies for each model are not significantly different, a result that is not surprising considering that the same PYTHIA fragmentation is used in each model, resulting in similar final state distributions for each M_{X_s} bin.

Despite the similarities in efficiency, the different mass distributions are connected to another problem: the migration of events between M_{X_s} bins. To quantitatively estimate an uncertainty due to this effect, in each reconstructed mass bin we utilize mass migration matrices for each model. An example for the model with flat X_s mass is shown in Figure 5.4 (the results for the other models can be found in Reference [40]). For each M_{X_s} bin, we take the difference between the total reconstructed efficiency and subtract the correctly reconstructed efficiency (the elements along the diagonals). We then use the average of this value over the three PYTHIA models. This value represents a bin-by-bin systematic error percentage.

5.2.2 PYTHIA Fragmentation

The use of PYTHIA to hadronize the X_s system into final state particles introduces a potential mismatch of final states observed in MC versus the data. If there is a significant disagreement between these final states, the efficiency will be calculated incorrectly and an error will be introduced into the branching fraction measurement. We

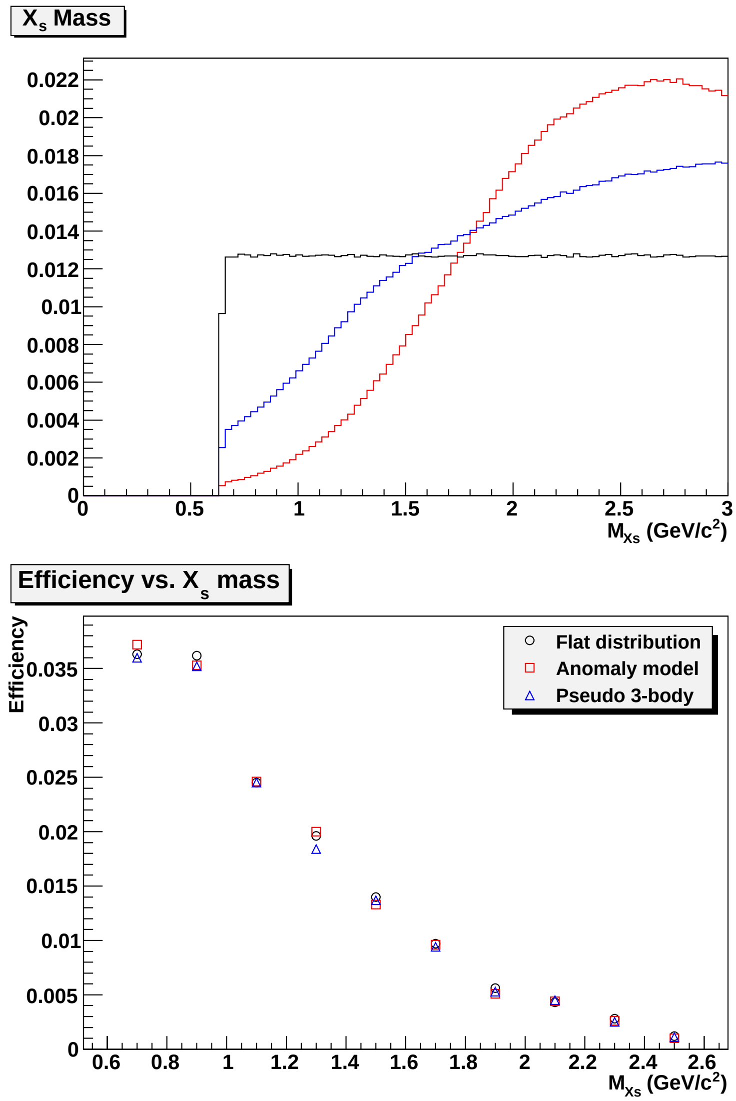

Figure 5.3: (Top) Generated distributions of M_{X_s} from PYTHIA MC using the flat (black), anomaly (red), and three-body distributions (blue). (Bottom) The associated efficiencies for each spectral model.

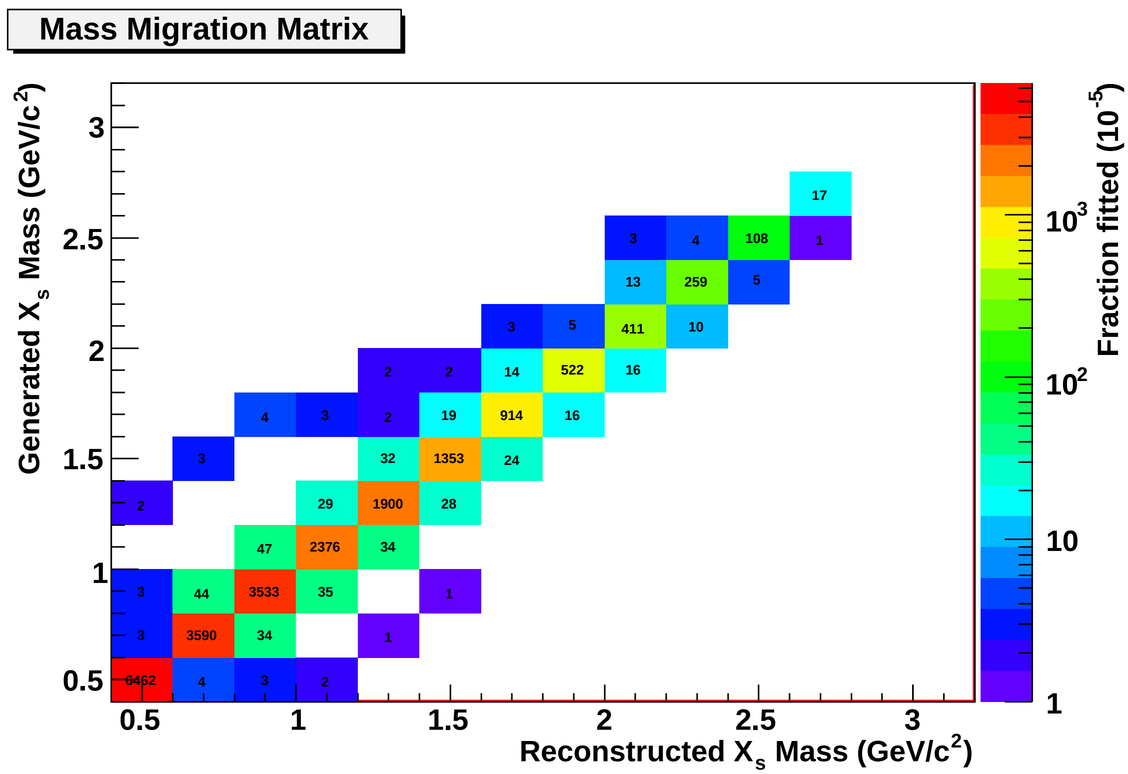

Figure 5.4: Migration matrix for the flat X_s mass model. Each element is calculated from the fitted gaussian yield divided by the number of generated events.

subdivide potential errors due to fragmentation effects into those introduced by the fraction of total X_s modes that are unreconstructed (also called "missing" modes) and those that are reconstructed.

We determine the fraction of missing modes for each M_{X_s} bin by using the signal MC. For modes that are missed because they have a K_L^0, but would have otherwise been reconstructed, we assign no systematic error, since these should be produced symmetrically with the corresponding K_S^0 modes. For all other missing modes, we allow the fraction to vary by $\pm 30\%$. This error is not applied to the $K\eta$ and $K^*\eta$ bins, since they are based on exclusive MC samples that do not use PYTHIA.

For the distributions of reconstructed modes, we can use the data directly to verify that PYTHIA is producing a proper distribution of final states. We divide the reconstructed states into the following categories: charged and neutral B modes, modes with and without a π^0, modes with a K_S^0 and those with a K^+, and modes with one or two total π's and those with three or four total π's. We compare the relative yields of these categories in data with the expectations from the PYTHIA-based MC. The results can be seen in Figure 5.5.[3] All mode categories agree within errors between the data and MC, except for the observed deficit of π^0 modes.

To account for this potential discrepancy in final state distributions, we assign a systematic uncertainty based on the following procedure. We perform fits of $M_{\rm bc}$ in bins of X_s mass, individually for modes with a π^0 and again for modes without a π^0, and calculate the following quantity:

$$f_{\pi^0} = \frac{N_{\pi^0}}{N_{\text{no }\pi^0} + N_{\pi^0}} = \frac{N_{\pi^0}}{N_{\rm total}} \tag{5.2.1}$$

where the N values represent the signal yield from the appropriate fit. We do the same for the PYTHIA-based signal MC with the flat X_s mass distribution. A comparison between data and MC is shown in Figure 5.6. We can rewrite the above expression for f_{π^0} in terms of the reconstruction efficiency, ϵ, as:

$$f_{\pi^0} = \frac{N_{\pi^0}}{\epsilon N_{\rm generated}} \tag{5.2.2}$$

where $N_{\rm generated}$ is the total number of $X_s\eta$ events in the signal MC sample. Rearranging this expression gives another formulation for the efficiency.

$$\epsilon = \frac{Y_{\pi^0}}{f_{\pi^0} N_{\rm generated}} \tag{5.2.3}$$

[3] Full $M_{\rm bc}$ fits for each category can be found in Reference [40].

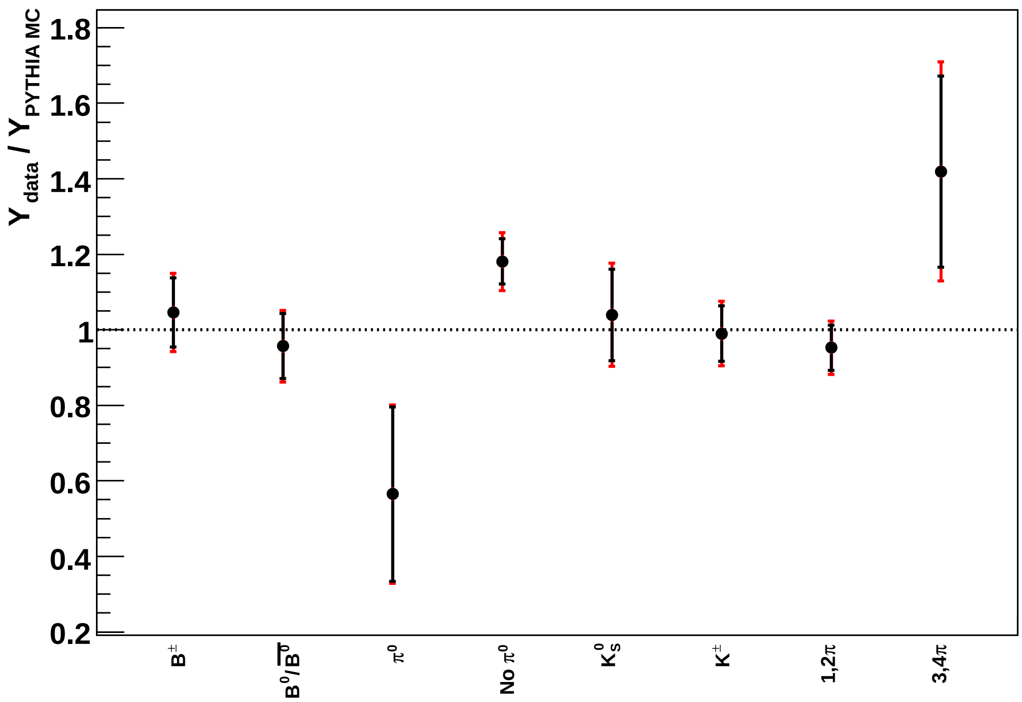

Figure 5.5: Ratio of expected $M_{\rm bc}$ fit yields between data and MC, divided by mode category. The first errors are statistical only, while the second include systematics due to fitting, K_s^0 reconstruction, π^0 reconstruction, charged K/π identification, and tracking.

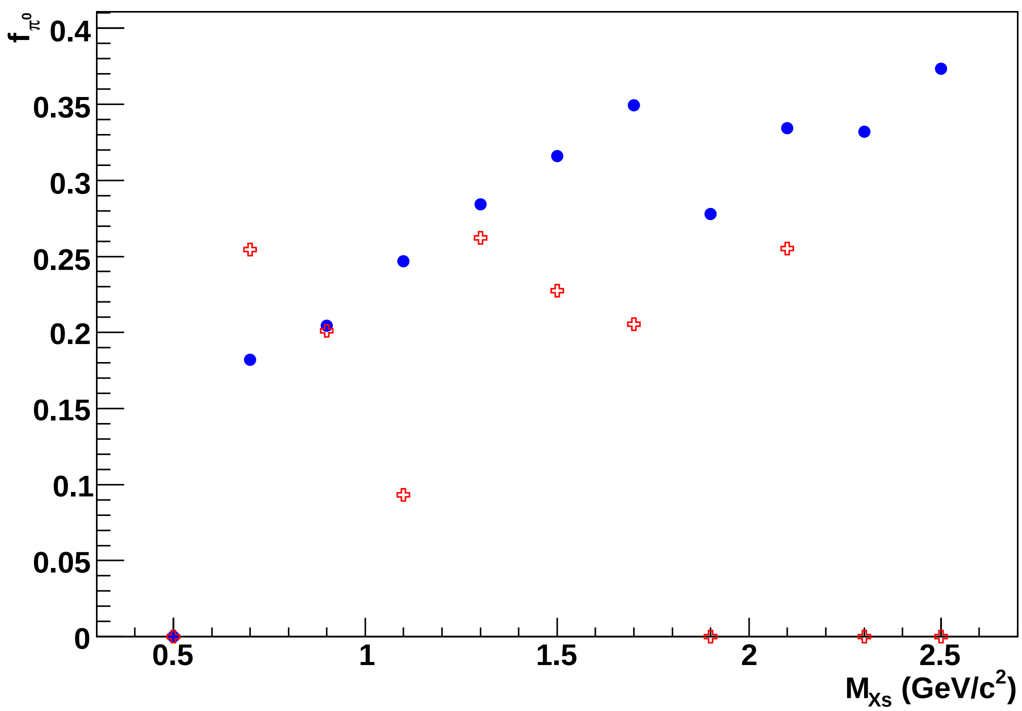

Figure 5.6: Fraction of yields in modes with a π^0 as a function of X_s mass, for data (red crosses) and PYTHIA MC (blue circles).

We can compute the efficiency for signal MC first directly. We then use the above expression with f_{π^0} determined from the data and both N_{π^0} and $N_{\mathrm{generated}}$ from the signal MC to compute an expected efficiency with an f_{π^0} determined from data. We do this for each bin of M_{X_s} and determine the difference in this recalculated efficiency and the original efficiency. The difference between the ratio of these two efficiencies and unity is used as the systematic error due to the PYTHIA fragmentation.

As with the missing mode systematic, this error is not applied to the $K\eta$ and $K^*\eta$ bins, since these signal models are not dependent on the fragmentation model.

5.2.3 Modeling Error Summary

The uncertainties introduced by the mass spectrum assumption and the PYTHIA fragmentation are summarized in Table 5.7. The relative errors are added in quadrature to determine a total error.

M_{X_s} (GeV/c^2)	Mass migration (±%)	Missing modes (±%)	PYTHIA (+%)	PYTHIA (-%)	Total (%)
0.4–0.6	0.70	0.00	0.00	0.00	± 0.70
0.6–0.8	2.20	0.07	9.62	0.00	$^{+9.87}_{-2.20}$
0.8–1.0	2.82	0.00	0.00	0.00	± 2.85
1.0–1.2	3.06	3.63	0.00	18.26	$^{+4.75}_{-18.87}$
1.2–1.4	4.54	4.47	0.00	2.64	$^{+6.37}_{-6.90}$
1.4–1.6	5.68	7.25	0.00	9.56	$^{+9.21}_{-13.28}$
1.6–1.8	5.30	10.24	0.00	18.39	$^{+11.53}_{-21.70}$
1.8–2.0	7.01	13.21	0.00	27.99	$^{+14.95}_{-31.74}$
2.0–2.2	5.76	16.33	0.00	10.85	$^{+17.32}_{-20.43}$
2.2–2.4	8.46	18.89	0.00	33.40	$^{+20.69}_{-39.29}$
2.4–2.6	7.35	21.06	0.00	37.37	$^{+22.31}_{-43.52}$

Table 5.7: Systematic errors related to the signal model, including mass migration, missing modes, and PYTHIA fragmentation.

5.3 Branching Fraction - Summary and Significance

We tabulate all uncertainties related to the fit yields in Table 5.8, as well as the significance of each yield including these systematic effects. We also calculate yields and significances for two larger regions: the full X_s mass range, which has a significance of 23, and a high mass region, $M_{X_s} > 1.8$ GeV/c^2, which has a significance of 7. It is important to note that although the branching fraction measurements suffer from large uncertainties, primarily due to modeling effects, the significance is only affected by those uncertainties related to the fitted yields. Thus, both our total result and the high mass excess can be regarded as statistically significant.

We also list the branching fractions and their final systematic uncertainties for each M_{X_s} bin in Table 5.9. In addition, we calculate a cumulative branching fraction for each X_s mass bin by adding together the results from various bins. For these cumulative branching fractions, we add uncorrelated errors (such as the statistical errors) in quadrature, and add correlated errors (such as the modeling error) linearly.

5.4 Direct CP Asymmetry Systematic Uncertainties

In the branching fraction, we find that the largest components of the systematic errors are typically due to our modeling of the X_s system and the resulting uncertainties in

X_s Mass range (GeV/c^2)	Fit yield	$\mathcal{S}$
0.4 - 0.6	60.2 $\pm$ 12.4 $^{+1.8}_{-1.9}$	5.7
0.6 - 0.8	15.3 $\pm$ 8.8 $^{+0.8}_{-2.1}$	1.9
0.8 - 1.0	250.0 $\pm$ 19.2 $^{+3.6}_{-4.4}$	14.0
1.0 - 1.2	84.2 $\pm$ 13.8 $^{+2.4}_{-3.1}$	6.6
1.2 - 1.4	146.2 $\pm$ 17.2 $^{+3.4}_{-4.4}$	9.2
1.4 - 1.6	137.0 $\pm$ 17.6 $^{+3.8}_{-4.6}$	8.1
1.6 - 1.8	127.7 $\pm$ 18.4 $^{+5.5}_{-5.6}$	7.2
1.8 - 2.0	64.2 $\pm$ 17.8 $^{+8.6}_{-8.1}$	3.5
2.0 - 2.2	85.7 $\pm$ 18.4 $^{+5.8}_{-7.0}$	4.6
2.2 - 2.4	48.6 $\pm$ 17.9 $^{+7.1}_{-7.4}$	2.7
2.4 - 2.6	34.8 $\pm$ 12.5 $^{+4.6}_{-6.5}$	2.7
0.4 - 2.6	1054 $\pm$ 54 $^{+16}_{-18}$	23
1.8 - 2.6	233 $\pm$ 34 $^{+13}_{-15}$	7

Table 5.8: Summary of $M_{\rm bc}$ fit yields (with expected rare contributions subtracted), along with their statistical and systematic errors and significance (including systematics) for each bin.

M_{X_s} (GeV/c^2)	$\mathcal{B}(B \to X_s\eta)$ (10^{-5})	Cumulative $\mathcal{B}(B \to X_s\eta)$ (10^{-5})
0.4 – 0.6	0.19 ±0.04 ±0.01 $^{+0.00}_{-0.00}$	0.19 ±0.04 ±0.01 $^{+0.00}_{-0.00}$
0.6 – 0.8	0.09 ±0.05 ±0.01 $^{+0.01}_{-0.00}$	0.28 ±0.06 ±0.02 $^{+0.01}_{-0.00}$
0.8 – 1.0	1.70 ±0.13 $^{+0.09}_{-0.10}$ $^{+0.00}_{-0.00}$	1.99 ±0.15 ±0.11 $^{+0.01}_{-0.00}$
1.0 – 1.2	0.72 ±0.12 $^{+0.04}_{-0.05}$ $^{+0.03}_{-0.14}$	2.71 ±0.19 ±0.15 $^{+0.04}_{-0.14}$
1.2 – 1.4	1.58 ±0.19 ±0.10 $^{+0.10}_{-0.11}$	4.29 ±0.26 ±0.24 $^{+0.14}_{-0.25}$
1.4 – 1.6	2.08 ±0.27 $^{+0.13}_{-0.14}$ $^{+0.19}_{-0.28}$	6.36 ±0.38 $^{+0.36}_{-0.37}$ $^{+0.33}_{-0.52}$
1.6 – 1.8	2.82 ±0.41 ±0.21 $^{+0.33}_{-0.61}$	9.19 ±0.55 ±0.54 $^{+0.66}_{-1.14}$
1.8 – 2.0	2.44 ±0.68 $^{+0.36}_{-0.34}$ $^{+0.37}_{-0.78}$	11.63 ±0.88 $^{+0.76}_{-0.75}$ $^{+1.03}_{-1.91}$
2.0 – 2.2	4.24 ±0.91 $^{+0.38}_{-0.43}$ $^{+0.73}_{-0.87}$	15.87 ±1.26 $^{+1.03}_{-1.05}$ $^{+1.76}_{-2.78}$
2.2 – 2.4	3.68 ±1.35 $^{+0.59}_{-0.61}$ $^{+0.76}_{-1.45}$	19.55 ±1.85 $^{+1.35}_{-1.37}$ $^{+2.52}_{-4.22}$
2.4 – 2.6	6.51 ±2.34 $^{+0.95}_{-1.28}$ $^{+1.45}_{-2.83}$	26.06 ±2.99 $^{+1.91}_{-2.11}$ $^{+3.97}_{-7.05}$

Table 5.9: Summary of the branching fractions in each bin of X_s mass along with their statistical, systematic, and modeling errors. The last column shows the cumulative branching fraction, if all bins including and below the given mass are added.

efficiency. The situation is quite different in the direct CP asymmetry. Since the efficiencies for the b and $\bar{b}$ modes are expected to be almost identical, we can calculate $\mathcal{A}_{\rm CP}$ without making use of the efficiency. As a result, we define only a single systematic uncertainty.

Part of the uncertainty can be determined from the raw asymmetry and the dilution factor as,

$$\delta A_{\rm CP}(\text{syst.}) = \sqrt{(A_{\rm CP}^{\rm raw})^2(\delta D(\text{syst.}))^2 + D^2(\delta A_{\rm CP}^{\rm raw}(\text{syst.}))^2} \quad (5.4.1)$$

The uncertainties on the wrong-tag fractions dilution factors are somewhat model dependent. To account for this, we compare the nominal value calculated from the flat M_{X_s} MC and compare to that obtained from the QCD anomaly MC and the three-body MC. We use the maximum difference between these values and the flat M_{X_s} MC to assign an uncertainty on the wrong-tag fraction, w. We combine this systematic error in quadrature with the statistical error on w to give the systematic error on D. Uncertainties on the raw fitted asymmetry are estimated by repeating the fitting procedure and allowing all fixed PDF parameters to vary by amounts that are determined in the same way as those for the branching fraction fits. We also add another term based on the observed deviation of $A_{\rm CP}$ from zero in the signal MC sample, shown in Figure 5.7, which is guaranteed to have a null asymmetry.

To account for possible bias in the detector or the analysis procedure, which is not included in Equation 5.4.1, we measure the $\mathcal{A}_{\rm CP}$ in the $B \to D\pi^+$ control sample data, which is expected to exhibit only a small asymmetry. We find a value of -0.014 ± 0.004, consistent with the previous Belle measurement [52], but inconsistent with zero. The deviation from zero is added in quadrature to the errors described above.

Finally, we account for the subtracted backgrounds, which may have their own non-zero asymmetries. For the lowest mass bin, where we have a known background from $B^+ \to \pi^+\eta$, we vary the assumed value of $\mathcal{A}_{\rm CP}(B^+ \to \pi^+\eta)$ by the errors on the world average of the measured asymmetry. For all other X_s mass bins we allow the asymmtery of the subtracted backgrounds to vary from -1 to 1, to account for all possible asymmetry values.

5.4.1 Direct CP Asymmetry - Summary and Significance

We summarize the final values of $\mathcal{A}_{\rm CP}$, including systematic uncertainties, in Table 5.10. For the measurement over the full mass range, we calculate a significance, taking

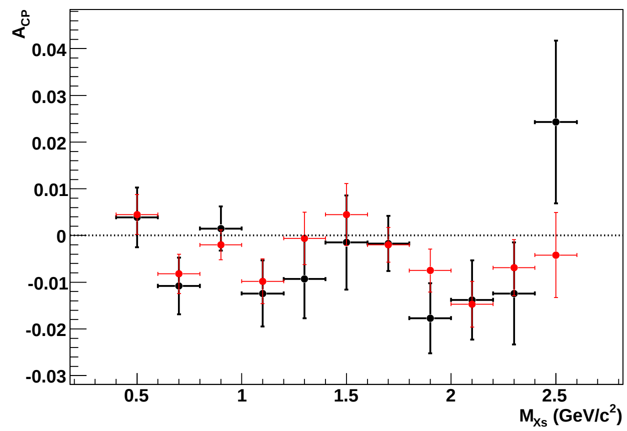

Figure 5.7: $A_{\rm CP}$ values of the signal MC sample as a function of M_{X_s}, using the standard cuts (black), and with the LR cut removed (red). The latter was studied to identify any possible bias due to the LR cuts.

into account all uncertainties, of 2.6. This corresponds to a non-zero direct CP asymmetry at over 99% confidence, and is a strong hint of direct CP violation in $B \to X_s\eta$.

M_{X_s} (GeV/c^2)	$A_{\rm CP}(B \to X_s\eta)$
0.4 – 0.6	-0.35 $\pm$ 0.18 $\pm$ 0.02
0.6 – 0.8	0.02 $\pm$ 0.40 $\pm$ 0.13
0.8 – 1.0	-0.04 $\pm$ 0.07 $\pm$ 0.02
1.0 – 1.2	-0.26 $\pm$ 0.15 $^{+0.03}_{-0.04}$
1.2 – 1.4	-0.22 $\pm$ 0.11 $^{+0.02}_{-0.03}$
1.4 – 1.6	-0.15 $\pm$ 0.12 $^{+0.02}_{-0.03}$
1.6 – 1.8	-0.25 $\pm$ 0.13 $^{+0.02}_{-0.03}$
1.8 – 2.0	-0.31 $\pm$ 0.26 $\pm$ 0.06
2.0 – 2.2	0.34 $\pm$ 0.20 $^{+0.04}_{-0.03}$
2.2 – 2.4	0.02 $\pm$ 0.32 $\pm$ 0.05
2.4 – 2.6	-0.40 $\pm$ 0.36 $^{+0.07}_{-0.12}$
0.4 – 2.6	-0.13 $\pm$ 0.04 $^{+0.02}_{-0.03}$
1.0 – 2.6	-0.15 $\pm$ 0.06 $\pm$ 0.03

Table 5.10: Direct CP asymmetries in bins of X_s mass. The first errors are statistical and the second are systematic.

Chapter 6

Summary and Conclusion

The final results, including all systematic uncertainties, for both the branching fraction and direct CP asymmetry are shown graphically in Figure 6.1, and can be found numerically in Tables 5.9 and 5.10 in the previous chapter. Over the full mass range, $0.4\ \mathrm{GeV}/c^2 < M_{X_s} < 2.6\ \mathrm{GeV}/c^2$, we find a branching fraction of $\mathcal{B}(B \to X_s\eta) = (26.1 \pm 3.0(\mathrm{stat})^{+1.9}_{-2.1}(\mathrm{syst})^{+4.0}_{-7.1}(\mathrm{model})) \times 10^{-5}$. The measured M_{X_s} dependent branching fractions are consistent with the known $B \to K\eta$ and $B \to K^*(892)\eta$ processes. In the high mass region, $1.8\ \mathrm{GeV}/c^2 < M_{X_s} < 2.6\ \mathrm{GeV}/c^2$, which is above any significant contributions from previously measured exclusive processes, we observe a signal with 7σ significance. The specific origin of this signal is unknown, and may be clarified by future exclusive measurements targeting this mass region. We also measure the CP asymmetry of $B \to X_s\eta$, both as a function of M_{X_s} and for the full mass range, where we find $\mathcal{A}_{CP} = -0.13 \pm 0.04^{+0.02}_{-0.03}$. The significance of the asymmetry measurement is 2.6σ, indicating that if the errors can be reduced, it may be a good candidate for future observation of direct CP asymmetry. The results of this analysis have been condensed into a manuscript that will be published in Physical Review Letters. The paper can be found in Appendix A.

The precision of our branching fraction measurements for the $X_s\eta$ process are limited primarily by systematic errors on the modeling of the X_s system. Regardless, we can still draw some conclusions from the observed rate and mass spectrum: the similarity in spectral shape to $B \to X_s\eta'$ and the lack of strong suppression of the $B \to X_s\eta$ branching fraction relative to the η' mode imply that the origin of the large contribution in the η' mode is also common to the η mode [24], and disfavors singlet-specific mechanisms, such as an intrinsic charm component of the η' or the QCD anomaly coupling [17, 18]. We hope that this measurement will help to guide future theoretical work, ultimately resulting in

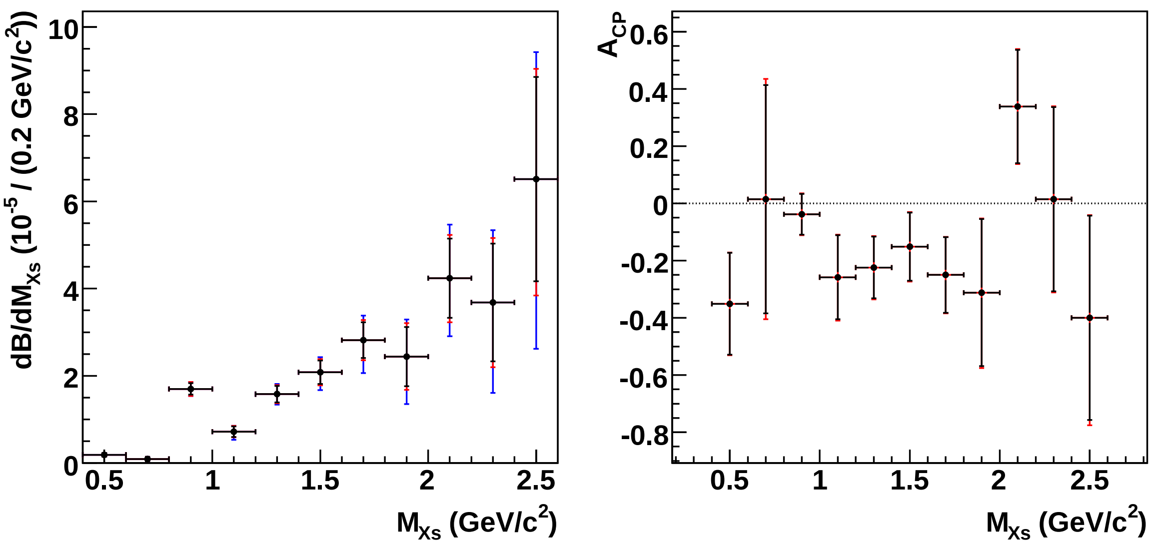

Figure 6.1: (Left) Final differential branching fraction of this study, as a function of M_{X_s}. From the smallest to largest error bars, the first include statistical effects only, then add systematic effects, then add modeling effects. (Right) Final direct CP asymmetry of this study, as a function of M_{X_s}. The first set of error bars includes statistical effects only, and the second set includes systematic effects.

more precise predictions for both the η and η' modes. Such predictions may also allow more accurate modeling of the X_s system in experimental studies, ultimately reducing the potential systematic errors of future analyses. If these decays continue to become better constrained both theoretically and experimentally, it will be clearer whether they can be explained within the SM, or whether they may be indicative of new physics.

Appendix A

Publication Preprint

Belle Preprint 2010-19
KEK Preprint 2010-30
arXiv:0910.4751
To appear in PRL

First Measurement of Inclusive $B \to X_s\eta$ Decays

K. Nishimura,[9] T. E. Browder,[9] I. Adachi,[10] H. Aihara,[46] K. Arinstein,[1,33] T. Aushev,[20,14] A. M. Bakich,[40] V. Balagura,[14] E. Barberio,[24] K. Belous,[13] V. Bhardwaj,[35] M. Bischofberger,[26] A. Bondar,[1,33] A. Bozek,[30] M. Bračko,[22,15] M.-C. Chang,[5] Y. Chao,[29] A. Chen,[27] K.-F. Chen,[29] P. Chen,[29] B. G. Cheon,[8] C.-C. Chiang,[29] I.-S. Cho,[50] Y. Choi,[39] J. Dalseno,[23,42] M. Danilov,[14] Z. Doležal,[2] A. Drutskoy,[4] S. Eidelman,[1,33] N. Gabyshev,[1,33] B. Golob,[21,15] H. Ha,[18] J. Haba,[10] T. Hara,[10] K. Hayasaka,[25] H. Hayashii,[26] Y. Horii,[45] Y. Hoshi,[44] W.-S. Hou,[29] H. J. Hyun,[19] T. Iijima,[25] K. Inami,[25] R. Itoh,[10] M. Iwabuchi,[50] Y. Iwasaki,[10] N. J. Joshi,[41] T. Julius,[24] J. H. Kang,[50] P. Kapusta,[30] H. Kawai,[3] T. Kawasaki,[32] H. Kichimi,[10] C. Kiesling,[23] H. J. Kim,[19] H. O. Kim,[19] M. J. Kim,[19] Y. J. Kim,[7] K. Kinoshita,[4] B. R. Ko,[18] P. Kodyš,[2] S. Korpar,[22,15] P. Križan,[21,15] T. Kumita,[47] A. Kuzmin,[1,33] Y.-J. Kwon,[50] S.-H. Kyeong,[50] J. S. Lange,[6] M. J. Lee,[38] S.-H. Lee,[18] J. Li,[9] C. Liu,[37] Y. Liu,[29] D. Liventsev,[14] R. Louvot,[20] A. Matyja,[30] S. McOnie,[40] K. Miyabayashi,[26] H. Miyata,[32] Y. Miyazaki,[25] G. B. Mohanty,[41] T. Mori,[25] E. Nakano,[34] M. Nakao,[10] H. Nakazawa,[27] Z. Natkaniec,[30] S. Nishida,[10] S. Ogawa,[43] T. Ohshima,[25] S. L. Olsen,[38,9] W. Ostrowicz,[30] G. Pakhlova,[14] C. W. Park,[39] H. Park,[19] H. K. Park,[19] K. S. Park,[39] R. Pestotnik,[15] M. Petrič,[15] L. E. Piilonen,[48] M. Röhrken,[17] S. Ryu,[38] H. Sahoo,[9] Y. Sakai,[10] O. Schneider,[20] C. Schwanda,[12] A. J. Schwartz,[4] K. Senyo,[25] O. Seon,[25] M. E. Sevior,[24] M. Shapkin,[13] C. P. Shen,[9] J.-G. Shiu,[29] F. Simon,[23,42] P. Smerkol,[15] A. Sokolov,[13] E. Solovieva,[14] M. Starič,[15] T. Sumiyoshi,[47] S. Suzuki,[36] Y. Teramoto,[34] K. Trabelsi,[10] S. Uehara,[10] T. Uglov,[14] Y. Unno,[8] S. Uno,[10] G. Varner,[9] K. E. Varvell,[40] K. Vervink,[20] C. H. Wang,[28] M.-Z. Wang,[29] P. Wang,[11] Y. Watanabe,[16] J. Wicht,[10] K. M. Williams,[48] E. Won,[18] Y. Yamashita,[31] M. Yamauchi,[10] C. C. Zhang,[11] Z. P. Zhang,[37] P. Zhou,[49] V. Zhulanov,[1,33] T. Zivko,[15] A. Zupanc,[17] and O. Zyukova[1,33]

(The Belle Collaboration)

[1]*Budker Institute of Nuclear Physics, Novosibirsk*
[2]*Faculty of Mathematics and Physics, Charles University, Prague*
[3]*Chiba University, Chiba*
[4]*University of Cincinnati, Cincinnati, Ohio 45221*
[5]*Department of Physics, Fu Jen Catholic University, Taipei*
[6]*Justus-Liebig-Universität Gießen, Gießen*
[7]*The Graduate University for Advanced Studies, Hayama*
[8]*Hanyang University, Seoul*
[9]*University of Hawaii, Honolulu, Hawaii 96822*
[10]*High Energy Accelerator Research Organization (KEK), Tsukuba*
[11]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing*
[12]*Institute of High Energy Physics, Vienna*
[13]*Institute of High Energy Physics, Protvino*
[14]*Institute for Theoretical and Experimental Physics, Moscow*
[15]*J. Stefan Institute, Ljubljana*
[16]*Kanagawa University, Yokohama*
[17]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe*
[18]*Korea University, Seoul*
[19]*Kyungpook National University, Taegu*
[20]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne*
[21]*Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana*
[22]*University of Maribor, Maribor*
[23]*Max-Planck-Institut für Physik, München*
[24]*University of Melbourne, School of Physics, Victoria 3010*
[25]*Nagoya University, Nagoya*
[26]*Nara Women's University, Nara*
[27]*National Central University, Chung-li*
[28]*National United University, Miao Li*
[29]*Department of Physics, National Taiwan University, Taipei*
[30]*H. Niewodniczanski Institute of Nuclear Physics, Krakow*
[31]*Nippon Dental University, Niigata*
[32]*Niigata University, Niigata*
[33]*Novosibirsk State University, Novosibirsk*
[34]*Osaka City University, Osaka*

[35]Panjab University, Chandigarh
[36]Saga University, Saga
[37]University of Science and Technology of China, Hefei
[38]Seoul National University, Seoul
[39]Sungkyunkwan University, Suwon
[40]School of Physics, University of Sydney, NSW 2006
[41]Tata Institute of Fundamental Research, Mumbai
[42]Excellence Cluster Universe, Technische Universität München, Garching
[43]Toho University, Funabashi
[44]Tohoku Gakuin University, Tagajo
[45]Tohoku University, Sendai
[46]Department of Physics, University of Tokyo, Tokyo
[47]Tokyo Metropolitan University, Tokyo
[48]IPNAS, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[49]Wayne State University, Detroit, Michigan 48202
[50]Yonsei University, Seoul

We report a first measurement of inclusive $B \to X_s\eta$ decays, where X_s is a charmless state with unit strangeness. The measurement is based on a pseudo-inclusive reconstruction technique and uses a sample of 657×10^6 $B\bar{B}$ pairs accumulated with the Belle detector at the KEKB e^+e^- collider. For $M_{X_s} < 2.6$ GeV/c^2, we measure a branching fraction of $(26.1\pm3.0(\text{stat})^{+1.9}_{-2.1}(\text{syst})^{+4.0}_{-7.1}(\text{model}))\times10^{-5}$ and a direct CP asymmetry of $\mathcal{A}_{CP} = -0.13\pm0.04^{+0.02}_{-0.03}$. Over half of the signal occurs in the range $M_{X_s} > 1.8$ GeV/c^2.

PACS numbers: 13.25.Hw, 13.30.Eg, 14.40.Nd

Decays of B mesons involving the $b \to s$ transition are an excellent tool for searches for physics beyond the Standard Model (SM). Theoretical treatments of these decays into exclusive hadronic final states, however, suffer from large uncertainties in the hadronization process. The uncertainties can be effectively reduced by leaving some of the final states in the calculation at the quark level, which corresponds to a measurement of an inclusive hadronic state X_s of unit strangeness.

Among such $b \to s$ decays, those involving the η and η' mesons exhibit unique properties due to interference between their underlying SU(3) octet and singlet components [1]. The CLEO collaboration reported the first measurement of inclusive $B \to X_s\eta'$ with an unexpectedly large branching fraction and an X_s spectrum that peaks at high X_s mass [2], a result confirmed in improved, higher-statistics measurements [3, 4]. Explanations included a large intrinsic $c\bar{c}$ component of the η' [5], the QCD anomaly mechanism [6] that couples two gluons to the flavor singlet component of the η', and also new physics sources [7]. The first is disfavored by the lack of an enhancement of $B \to \eta_c K$ relative to $B \to J/\psi K$ [8], while the second is disfavored by a measurement of $\Upsilon(1S) \to \eta' X$ [9], which indicates an $\eta' gg$ form factor that cannot explain the enhancement. A recent treatment [10] using soft collinear effective theory suggests that a measurement of the complementary process $B \to X_s\eta$ can elucidate the possible contribution from nonperturbative charm-penguin amplitudes or higher-order gluonic operators to both the η and η' processes. CLEO performed the only previous search with an upper limit of $\mathcal{B}(B \to X_s\eta) < 4.4 \times 10^{-4}$ [2].

In this Letter, we report a measurement of $B \to X_s\eta$ using a sample of 657×10^6 $B\bar{B}$ pairs accumulated with the Belle detector at the KEKB e^+e^- collider [11]. The Belle detector is a large solid-angle magnetic spectrometer and is described in detail elsewhere [12].

We reconstruct candidate B mesons using a pseudo-inclusive method, with the X_s composed of a K^+ or $K^0_S(\to \pi^+\pi^-)$ and up to four pions, of which at most one is a $\pi^0(\to \gamma\gamma)$. This gives a total of 18 reconstructed channels and their charge-conjugates [13]. Charged pions and kaons are selected based on information from the time-of-flight, aerogel Cherenkov, and drift chamber dE/dx systems. Typical efficiencies to correctly identify kaons (pions) are above 88% (98%), with misidentification rates for pions as kaons (kaons as pions) below 12% (4%). K^0_S candidates are required to have an invariant mass within 16 MeV/c^2 (4σ) of the K^0_S mass and a displaced vertex from the interaction point. For π^0 candidates, each daughter photon is required to have energy greater than 50 (100) MeV in the barrel (endcap) region and a shower shape consistent with a photon. The invariant mass of the photon pair must be within 15 MeV/c^2 (2.5σ) of the π^0 mass. The π^0 momentum is recalculated using the nominal π^0 mass. To suppress combinatorial backgrounds, we require π^0 candidates to have laboratory momenta greater than 300 MeV/c. Pions and kaons are combined to form an X_s.

Candidate η mesons are reconstructed in the $\eta \to \gamma\gamma$ mode from photons with $E_\gamma > 200$ MeV. The invariant mass of the γ-pair is required to lie between 520 MeV/c^2 and 570 MeV/c^2, or within 2σ of the nominal mass. We veto an η candidate if either of its photons can be combined with another photon in the event to form a candidate π^0. To suppress background from radiative B decays, we require the energy asymmetry of the two photons, defined as $|E_{\gamma_1} - E_{\gamma_2}|/|E_{\gamma_1} + E_{\gamma_2}|$, to be less than 0.6. The η mass constraint [14] is used to refit the momenta of the daughter photons. To suppress secondary η-mesons from $b \to c \to \eta$ chains, we retain only η candidates whose center-of-mass (CM) momentum satisfies $|\mathbf{p}_\eta^*| > 2.0$ GeV/c.

B meson candidates are formed from combinations of an X_s and an η. A beam-constrained mass, $M_{\rm bc} = \sqrt{E_{\rm beam}^2/c^4 - |\mathbf{p}_B^*|^2/c^2}$ and energy difference, $\Delta E = E_B - E_{\rm beam}$ are calculated, where $E_{\rm beam}$, $\mathbf{p}_B^*$, and E_B are the beam energy, B momentum, and B energy, all in the CM frame. The signal is obtained using fits to $M_{\rm bc}$ with $|\Delta E| < 0.1$ GeV.

We use a simulated signal Monte Carlo (MC) sample [15] consisting of $B \to K\eta$ for $M_{X_s} < 0.6$ GeV/c^2, $B \to K^*\eta$ for $M_{X_s} \in [0.8, 1.0]$ GeV/c^2, and $B \to X_s\eta$ in all other mass regions ($M_{X_s} \in [0.6, 0.8]$ GeV/c^2, and $M_{X_s} > 1.0$ GeV/c^2). For the $B \to X_s\eta$ component, fragmentation of the X_s system into hadrons is simulated by PYTHIA [16], assuming a model in which the X_s mass spectrum is flat from the $K\pi$ threshold up to 3.2 GeV/c^2. We find an average of approximately nine B candidates per event, with 10% of events having more than 20 candidates. We select the candidate with the lowest χ^2, with χ^2 defined as the sum of $\chi^2_{\Delta E} = (\Delta E/\sigma_{\Delta E})^2$, where the resolution σ_{Δ_E} is estimated separately for each reconstructed mode, and, if available, a reduced-χ^2 of a vertex fit that includes all X_s daughter charged tracks except those used as part of a K_S^0 candidate. The resolution in ΔE is asymmetric. It also varies by mode, with the most significant differences between modes with and without a π^0. For modes without (with) a π^0, typical ΔE resolutions are 62 (76) MeV for $\Delta E < 0$ and 29 (31) MeV for $\Delta E > 0$. After applying this procedure and applying continuum suppression cuts (described below), we select the correctly reconstructed B in 56% of simulated events.

The dominant background to $B \to X_s\eta$ comes from continuum production of quark pairs, $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$). These events have a jet-like topology, and are suppressed relative to the spherical $B\bar{B}$ events using a Fisher discriminant [17] formed from event shape variables [18, 19]. Further suppression is obtained by combining this Fisher discriminant with the cosine of the B flight direction in the CM frame and, when available, the displacement between the signal B and the other B in the event. This suppression is optimized as a function of b-flavor tag quality [20], and is approximately 34% efficient for the signal modes while suppressing over 99% of the continuum background.

Decays of the type $B \to X_c\eta$ and $B \to X_c \to X_s\eta$, where X_c is any state containing charm mesons, may have final states identical to the signal mode. We search among the candidate B decay products for combinations consistent with selected charm meson decays and veto the candidate if the mass of the reconstructed combination is within $\pm 2.5\sigma$ of the known mass. The modes and their veto widths are: $D^0 \to Kn\pi^\pm(\pi^0)$, 13.5 (44.5) MeV/c^2; $D^+ \to Kn\pi^\pm(\pi^0)$, 12.5 (31.3) MeV/c^2; $D^0 \to K_S^0\eta$, 31.3 MeV/c^2; $D_s^+ \to \eta\pi^+$, 29.3 MeV/c^2; and $\eta_c(1S) \to \eta\pi^+\pi^-$, 85.0 MeV/$c^2$. We also veto events with an $\eta' \to \eta\pi^+\pi^-$ candidate with an invariant mass $M_{\eta\pi\pi}$ within 100 MeV/c^2 of the nominal η' mass.

Signal yields are obtained using an extended unbinned maximum likelihood fit to $M_{\rm bc}$ in 200 MeV/c^2 bins of X_s mass up to 2.6 GeV/c^2. The probability density function (PDF) for the signal is taken as a Gaussian, with the mean and width determined from the appropriate signal MC sample ($K\eta$, $K^*\eta$, or $X_s\eta$) for the mass bin. The mean and width are calibrated for small differences between the data and MC using a $B \to D\pi^\pm$ control sample, with D reconstructed from a K and one to four π, with at most one π^0. All reconstructed modes are combined for the fit, and no attempt is made to separate correctly reconstructed B candidates and those with some missing or incorrectly attributed B daughters (self-cross-feed). Shapes for the charm contributions remaining after the vetoes are assigned based on a MC sample of generic $b \to c$ processes. Four separate PDFs are assigned for the largest charm backgrounds as identified in MC: $B^0 \to \bar{D}^0\eta$, $B^0 \to \bar{D}^{*0}\eta$, $B^0 \to D^{(*)-}\pi^+\eta$, and $B^+ \to \bar{D}^{(*)0}\pi^+\eta$. All other $b \to c$ backgrounds are combined into another PDF. Each charm PDF consists of a Gaussian component to describe the peaking in $M_{\rm bc}$, and an empirically determined parameterization (ARGUS function) [21] to describe non-peaking combinatorial contributions. The shape parameters are taken from the appropriate background MC sample. Normalizations of the modes $B^0 \to \bar{D}^{(*)0}\eta$ are based on the previous Belle measurement [22]. The branching fractions for the decays $B \to D^{(*)}\pi\eta$ are unknown, so their normalization is determined by a simultaneous χ^2 minimization based on the difference between the expected and observed $M_{\rm bc}$ distribution of the events in all eight veto windows. The normalization scaling of $D^-\pi^+\eta$ is assumed to be the same as of $D^{*-}\pi^+\eta$. A similar assumption is used for the $D^{(*)0}\pi\eta$ modes. The χ^2 technique is verified by repeating the optimization over the $B^0 \to \bar{D}^{(*)0}\eta$ modes, for which the results are consistent with the previous Belle measurement. This χ^2 is also used to study systematic errors on the normalizations of all charm PDFs. Normalization for the PDF that includes all other $b \to c$ modes is fixed to the MC expectation. The remaining combinatorial $q\bar{q}$ backgrounds are modeled with an ARGUS function. For the final fit, the signal yield and both the yield and shape parameter of the $q\bar{q}$ ARGUS PDF are allowed to vary.

Rare B decay backgrounds are studied with a dedicated MC sample, and include contributions from $B \to X_s\eta'$, $B \to X_s\gamma$, and $B \to X_d\eta$. These expected yields are subtracted from the fit yield to give a final yield. The expected yields for $B \to X_s\eta'$ and $B \to X_s\gamma$ are based on the known branching fractions, and are found to be less than 0.5 events in each X_s mass bin. The $B^+ \to \pi^+\eta$ branching fraction is also known, and the expectation is 5.2 events in the lowest bin of X_s mass. We estimate the contribution from other $B \to X_d\eta$ modes by repeating the reconstruction and the fitting procedure but replacing the K^+ candidate of X_s with a π^+ candidate. Performing these fits on data and using a dedicated $X_d\eta$ MC sample to estimate the rate to misreconstruct X_d as X_s, we estimate a total contamination of 19.1 ± 2.3 events from $X_d\eta$, distributed uniformly in the range $M_{X_s} \in [0.6, 2.6]$ GeV/c^2.

The fit to the full mass range, $M_{X_s} \in [0.4, 2.6]$ GeV/c^2, is shown in Fig. 1(a), and gives a background-subtracted yield of $1054 \pm 54^{+16}_{-18}$. We also define a high mass region, $M_{X_s} \in [1.8, 2.6]$ GeV/c^2, where the summed yield is $233 \pm 34^{+13}_{-15}$. Significances are determined in each mass bin by convolving the likelihood function with a Gaussian of width determined by the systematic errors on the yield. The maximum likelihood, $\mathcal{L}_{\max}$, and the likelihood at a signal yield of zero, $\mathcal{L}_0$, are used to determine the significance, which is defined as $\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\max})}$. The significance is 23 (7) for the full (high) X_s mass range.

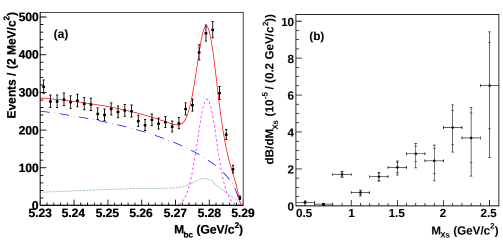

FIG. 1: (color online). (a) The $M_{\rm bc}$ distribution for the full mass range, $M_{X_s} \in [0.4, 2.6]$ GeV/c^2. The points with errors correspond to the data, while the curves correspond to the overall fit PDF (solid red), the signal PDF (dashed magenta), the sum of all $b \to c$ background PDFs (dotted green), and the combinatorial background PDF (dash-dotted blue). (b) Differential branching fraction, $d\mathcal{B}/dM_{X_s}$, for $B \to X_s\eta$. The error bars correspond to statistical error and total error. Error bars on the first two bins are smaller than the points.

Reconstruction efficiencies in bins of X_s mass range from 6.5% to 0.1%, not including the branching fraction for $\eta \to \gamma\gamma$; these results are based on the signal MC and assume equal production of B^+B^- and $B^0\bar{B}^0$ at $\Upsilon(4S)$. Efficiency losses are monotonic with an average of 30% efficiency loss with each increase in M_{X_s} bin. Figure 1(b) shows the differential branching fraction as a function of M_{X_s}. Table I gives the final results for each X_s mass bin. For the full M_{X_s} range,

TABLE I: Measured background-subtracted signal yields (N_S), branching fractions ($\mathcal{B}$), and CP asymmetry ($\mathcal{A}_{CP}$), for each M_{X_s} range. Uncertainties on N_S are statistical. Uncertainties on $\mathcal{B}$ are statistical, systematic, and modeling, respectively. The uncertainties for $\mathcal{A}_{CP}$ are statistical and systematic.

M_{X_s}(GeV/c^2)	N_S	$\mathcal{B}(10^{-6})$	$\mathcal{A}_{CP}(10^{-2})$
0.4–0.6	60 ± 12	$1.9 \pm 0.4 \pm 0.1 \pm 0.0$	$-35 \pm 18 \pm 2$
0.6–0.8	15 ± 9	$0.9 \pm 0.5 \pm 0.1^{+0.1}_{-0.0}$	$2 \pm 40 \pm 13$
0.8–1.0	250 ± 19	$17.0 \pm 1.3 \pm 1.0 \pm 0.0$	$-4 \pm 7 \pm 2$
1.0–1.2	84 ± 14	$7.2 \pm 1.2 \pm 0.5^{+0.3}_{-1.4}$	$-26 \pm 15^{+3}_{-4}$
1.2–1.4	146 ± 17	$15.8 \pm 1.9 \pm 1.0^{+1.0}_{-1.1}$	$-22 \pm 11^{+2}_{-3}$
1.4–1.6	137 ± 18	$20.8 \pm 2.7^{+1.3+1.9}_{-1.4-2.8}$	$-15 \pm 12^{+2}_{-3}$
1.6–1.8	128 ± 18	$28.2 \pm 4.1 \pm 2.1^{+3.3}_{-6.1}$	$-25 \pm 13^{+2}_{-3}$
1.8–2.0	64 ± 18	$24.4 \pm 6.8^{+3.6+3.7}_{-3.4-7.8}$	$-31 \pm 26 \pm 6$
2.0–2.2	86 ± 18	$42.4 \pm 9.1^{+3.9+7.3}_{-4.4-8.7}$	$34 \pm 20^{+4}_{-3}$
2.2–2.4	49 ± 18	$36.8 \pm 13.5^{+5.9+7.6}_{-6.1-14.5}$	$2 \pm 32 \pm 5$
2.4–2.6	35 ± 13	$65.1 \pm 23.4^{+9.5\ +14.5}_{-12.9-28.3}$	$-40 \pm 36^{+7}_{-12}$
0.4–2.6	1053 ± 54	$261 \pm 30^{+19+40}_{-21-71}$	$-13 \pm 4^{+2}_{-3}$
1.0–2.6	728 ± 48	$241 \pm 30^{+18+40}_{-20-71}$	$-15 \pm 6 \pm 3$
1.8–2.6	233 ± 34	$169 \pm 29^{+15+33}_{-18-59}$	$0 \pm 14 \pm 5$

we sum the individual contributions and find the following partial branching fraction $\mathcal{B}(B \to X_s\eta; M_{X_s} \in [0.4, 2.6]\ \text{GeV}/c^2) = (26.1 \pm 3.0^{+1.9+4.0}_{-2.1-7.1}) \times 10^{-5}$, where errors are statistical, (model-independent) systematic, and decay modeling. A large fraction of the inclusive signal occurs in the high mass region, where we find $\mathcal{B}(B \to X_s\eta; M_{X_s} \in [1.8, 2.6]\ \text{GeV}/c^2) = (16.9 \pm 2.9\ (\text{stat})^{+1.5}_{-1.8}\ (\text{syst})^{+3.3}_{-5.9}\ (\text{model})) \times 10^{-5}$.

The direct CP asymmetry is defined as $\mathcal{A}_{CP} = (\mathcal{B}^- - \mathcal{B}^+)/(\mathcal{B}^- + \mathcal{B}^+)$, where $\mathcal{B}^+(\mathcal{B}^-)$ is the partial branching fraction for B^+ or B^0 (B^- or $\bar{B}^0$). We measure this asymmetry in the subset of reconstructed modes in which the B flavor can be inferred from the final state (13 out of 18 modes). We adjust the fitted CP asymmetry to account for events that are reconstructed with the wrong B flavor by multiplying the raw fitted asymmetry by a correction factor. This factor is estimated from the signal MC, and ranges from unity to 1.05. The bin-by-bin results, as well as the results of separate fits for $\mathcal{A}_{CP}$ over the full X_s mass range and the range above the narrow kaonic resonances ($M_{X_s} \in [1.0, 2.6]$ GeV/c^2), are shown in Table I. For $M_{X_s} \in [0.4, 2.6]$ GeV/c^2, we find $\mathcal{A}_{CP} = -0.13 \pm 0.04^{+0.02}_{-0.03}$, with a significance of 2.6σ relative to a null asymmetry. All $\mathcal{A}_{CP}$ results that include the range $M_{X_s} \in [0.4, 0.6]$ GeV/c^2 are calculated with the assumption that the $B^+ \to \pi^+\eta$ backgrounds in this region contribute a CP asymmetry consistent with the existing measured world average [14].

Systematic errors on the fitted signal yield are dominated by PDF uncertainties. Uncertainties in the signal PDF parameters are studied using a $B \to D\pi$ control sample, while those due to normalizations and shapes for the $b \to c$ backgrounds are estimated by using comparisons between the veto window χ^2 procedure and one of the following: a repeated χ^2 procedure with relaxed assumptions, MC expectations, or, when available, previous measurements. Errors from the background subtractions are dominated by uncertainties in the estimate of backgrounds from $B \to X_d\eta$. Our estimates of these backgrounds may have included other small contributions, such as those from $B \to X_s\eta$, so we allow these estimates to vary by -100%. Positive uncertainties are estimated from the difference in expected yields assuming a flat distribution of X_d events in X_s mass versus those obtained from a MC study of cross-feed from X_d mass to X_s mass. In all cases the systematic uncertainties on the background-subtracted signal yields are at least a factor of two smaller than the statistical errors.

The model-independent systematic error includes contributions from the signal yield, the selection efficiency, the number of $B\bar{B}$ pairs, and the $\eta \to \gamma\gamma$ branching fraction [14]. For X_s mass bins above 1.8 GeV/c^2, the errors on the signal yields from uncertainties in the PDF shapes (primarily for the charm PDFs) dominate with a contributed relative uncertainty of 7-18%. For the lower X_s mass bins, the efficiency error is the largest contribution with a relative uncertainty of 5-6%. This error is the combination of individually determined contributions from control sample studies of the following: tracking, reconstruction of π^0, η, and K^0_S, particle identification, continuum suppression, and candidate selection.

We define an additional error due to modeling of the X_s system, which is studied in three parts. The first is due to the fraction of unreconstructed modes (e.g., modes with more than a total of four π's, more than one π^0, or more than one K). We vary these fractions by $\pm 30\%$ of the PYTHIA expectation and use the differences in efficiency to estimate an M_{X_s} bin-dependent uncertainty that rises with X_s mass from zero to $\pm 21.1\%$. The second is due to differences in the observed frequency of decay modes and those expected from PYTHIA. We find good agreement between data and MC in the relative amounts of charged and neutral B modes, modes with K^0_S and those with K^+, and modes with one or two total π's and those with three or four total π's. However, we find a significant excess of modes without a π^0 over those with a π^0, which we attribute to inaccuracies in the PYTHIA fragmentation. To quantify this uncertainty, we re-estimate the PYTHIA efficiencies with the fraction of π^0 modes adjusted to match data, and use the difference between this value and the nominal efficiency to assign an error. This error is usually only negative, due to the higher reconstruction efficiency for modes without a π^0, and is as large as -37% in the highest X_s mass bin. The final component of the modeling uncertainty is due to the assumed X_s mass spectrum. We study the efficiencies of other M_{X_s} signal MC samples where the spectrum rises toward high mass and assign errors based on the differences from the flat M_{X_s} MC. Using these samples, we also study the fractions of self-cross-feed candidates that are reconstructed with an incorrect X_s mass. These effects are small compared to the first two components of the modeling error.

The systematic error on $\mathcal{A}_{CP}$ includes contributions due to: uncertainties in the PDF parameters; possible detector and measurement biases are estimated from the measured $\mathcal{A}_{CP}$ of the $B \to D\pi$ control sample and the signal MC, respectively; uncertainty due to the signal model is studied by checking the fractions of events with incorrectly identified flavor using alternative M_{X_s} spectra; and possible contamination due to $B \to \pi\eta$ ($B \to X_d\eta$) decays is estimated by varying their A_{CP} by the measured uncertainty [14] ($\pm 100\%$).

In summary, we report the first measurement of the inclusive process $B \to X_s\eta$, and find a partial branching fraction of $\mathcal{B}(B \to X_s\eta; M_{X_s} \in [0.4, 2.6]\ \mathrm{GeV}/c^2) = (26.1 \pm 3.0(\mathrm{stat})^{+1.9}_{-2.1}(\mathrm{syst})^{+4.0}_{-7.1}(\mathrm{model})) \times 10^{-5}$. The measured M_{X_s} dependent branching fractions are consistent with the known $B \to K\eta$ and $B \to K^*(892)\eta$ processes [23]. In the high mass region, $M_{X_s} \in [1.8, 2.6]$ GeV/c^2, which is above any significant contributions from previously measured exclusive processes [24], we observe a signal with a 7σ significance. We also measure the CP asymmetry of $B \to X_s\eta$, both as a function of M_{X_s} and for the full mass range, where we find $\mathcal{A}_{CP} = -0.13 \pm 0.04^{+0.02}_{-0.03}$, consistent with rough theoretical expectations [10]. No theoretical prediction is currently available for the shape of the M_{X_s} spectrum. However, the similarity in spectral shape to $B \to X_s\eta'$ and the lack of strong suppression of the $B \to X_s\eta$ branching fraction relative to the η' mode imply that the origin of the large contribution in the η' mode is also common to the η mode [10], and disfavors η' specific mechanisms [5, 6].

We thank the KEKB group for excellent operation of the accelerator, the KEK cryogenics group for efficient solenoid operations, and the KEK computer group and the NII for valuable computing and SINET3 network support. We acknowledge support from MEXT, JSPS and Nagoya's TLPRC (Japan); ARC and DIISR (Australia); NSFC (China); MSMT (Czechia); DST (India); MEST, NRF, NSDC of KISTI (Korea); MNiSW (Poland); MES and RFAAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE (USA).

[1] H.J. Lipkin, Phys. Lett. B **254**, 247 (1991).
[2] T.E. Browder *et al.* (CLEO Collaboration), Phys. Rev. Lett. **81**, 1786 (1998).
[3] G. Bonvicini *et al.* (CLEO Collaboration), Phys. Rev. D **68**, 011101 (2003).
[4] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **93**, 061801 (2004).
[5] I. E. Halperin and A. Zhitnitsky, Phys. Rev. Lett. **80**,

438 (1998).
[6] D. Atwood and A. Soni, Phys. Lett. B **405**, 150 (1997).
[7] W. S. Hou and B. Tseng, Phys. Rev. Lett. **80**, 434 (1998).
[8] K. W. Edwards *et al.* (CLEO Collaboration), Phys. Rev. Lett. **86**, 30 (2001).
[9] M. Artuso *et al.* (CLEO Collaboration), Phys. Rev. D **67**, 052003 (2003).
[10] J. Chay, C. Kim, A. K. Leibovich, and J. Zupan, Phys. Rev. D **76**, 094031 (2007).
[11] S. Kurokawa and E. Kikutani, Nucl. Instr. Meth. A **499**, 1 (2003).
[12] A. Abashian *et al.* (Belle Collaboration), Nucl. Instr. Meth. A **479**, 117 (2002).
[13] Throughout this paper, the inclusion of the charge-conjugate mode decay is implied unless otherwise stated.
[14] C. Amsler *et al.* (Particle Data Group), Phys. Lett. B **667**, 1 (2008).
[15] We use the EvtGen B meson decay generator, D.J. Lange, Nucl. Instr. Meth. A **462**, 152 (2001). The detector response is simulated with GEANT, R. Brun *et al.*, GEANT 3.21, CERN Report DD/EE/84-1 (1984).
[16] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006).
[17] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, 2nd Edition (John Wiley and Sons, 2001).
[18] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
[19] S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).
[20] H. Kakuno *et al.*, Nucl. Instr. Meth. A **533**, 516 (2004).
[21] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).
[22] S. Blyth *et al.* (Belle Collaboration), Phys. Rev. D **74**, 092002 (2006).
[23] S.J. Richichi *et al.* (CLEO Collaboration), Phys. Rev. Lett. **85**, 520 (2000).
[24] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **97**, 201802 (2006).

Appendix B

Beam Energy Calculation

For studies of B mesons at Belle, the primary process of interest is

$$e^+e^- \to \Upsilon(4S) \to B\bar{B}$$

In the CM frame, the two B mesons are produced with equal energy and equal and opposite momentum. As a result, the nominal energy of each B meson is half of the total mass of the $\Upsilon(4S)$. Though the central value of this mass is known with high precision, 10.5794 ± 0.0012 GeV$/c^2$, the width of the $\Upsilon(4S)$, 20.5 ± 2.5 MeV$/c^2$ [1], coupled with drifts in the KEKB accelerator conditions, limit our a priori knowledge of the beam energy.

We can improve the determination of the beam energy by utilizing known B decays for calibration. Recalling Equation 4.3.2, the beam-constrained mass, $M_{\rm bc}$, can be calculated from the B three-momentum and beam energy in the CM frame: $\mathbf{p}_B^*$ and $E_{\rm beam}$, respectively. $M_{\rm bc}$ peaks at the B mass, assuming that the beam energy is correctly calculated. However, throughout a given set of experimental data, the operating conditions at KEKB may cause shifts in the beam energy away from the nominal value of ~ 5.29 GeV$/c^2$.

Despite changes in the beam energy, the boost to the CM frame remains the same. To see why, consider an arbitrary four-momentum and the boost necessary to move to the CM frame. To simplify, we choose a system of coordinates such that the three-momentum is aligned entirely in one-spatial direction. This gives

$$\begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} E \\ |\mathbf{p}| \end{pmatrix} = \begin{pmatrix} E^* \\ 0 \end{pmatrix} \tag{B.1}$$

where $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. This gives us a relation for β,

$$\beta = \frac{|\mathbf{p}|}{E} \tag{B.2}$$

that depends only on the magnitude of the three momentum and the energy, both in the original frame of references. In our electron-positron system, these quantities are given by

$$\begin{aligned} |\mathbf{p}| &= \sqrt{E_{\mathrm{HER}}^2 + E_{\mathrm{LER}}^2 - 2E_{\mathrm{HER}}E_{\mathrm{LER}}\cos\theta} \\ E &= E_{\mathrm{HER}} + E_{\mathrm{LER}} \end{aligned} \tag{B.3}$$

where we assume, to very good approximation, that the electron and positron are massless. By inspection, it is evident that as long as the energy of the high energy (electron) ring and low energy (positron) ring change by the same fraction, then the β (and therefore γ) for the boost will remain constant. We can also verify that the boost direction does not change. In the lab frame, the three-momentum of the e^+e^- system, $\mathbf{p}_{e\bar{e}}$, is given by

$$\mathbf{p}_{e\bar{e}} = \begin{pmatrix} E_{\mathrm{HER}}\sin\theta \\ 0 \\ E_{\mathrm{HER}}\cos\theta - E_{\mathrm{LER}} \end{pmatrix} \tag{B.4}$$

We can see that the x and z components will scale identically provided that, as before, E_{HER} and E_{LER} scale together. Under normal KEKB operation, we assume that these energies are always scaled together, so the boost vector remains constant.

Because the boost vector does not change over time, the calculation of $\mathbf{p}_B^*$ is unaffected by beam energy changes. This means that we can use the quantity $M_{\mathrm{bc}} - E_{\mathrm{beam}}^*$ to calibrate the beam energy. To do so, we use the following equality:

$$(E_{\mathrm{beam}}^*)_{\mathrm{nominal}}^2 - (M_{\mathrm{bc}})_{\mathrm{nominal}}^2 = (E_{\mathrm{beam}}^*)_{\mathrm{calibrated}}^2 - (M_{\mathrm{bc}})_{\mathrm{calibrated}}^2 \tag{B.5}$$

where "nominal" indicates the values using the nominal beam energy and "calibrated" indicates the values for the calibrated beam energy. Thus, $(E_{\mathrm{beam}}^*)_{\mathrm{nominal}}$ is $5.29\ \mathrm{GeV}/c^2$, $(M_{\mathrm{bc}})_{\mathrm{nominal}}$ is obtained from data using the nominal beam energy, $(E_{\mathrm{beam}}^*)_{\mathrm{calibrated}}$ is the value to be determined, and $(M_{\mathrm{bc}})_{\mathrm{calibrated}}$ is the known B meson mass ($\sim 5.279\ \mathrm{GeV}/c^2$).

To perform this calibration, we must choose B decays for which we can easily measure the means of their $(M_{\mathrm{bc}})_{\mathrm{nominal}}$ distributions. This is done by performing fits to each distribution, independently for each decay mode that is used. The fits use a PDF consisting of a Gaussian component for signal and an ARGUS component for combinatorial background. After obtaining the Gaussian mean from the fit procedure, the beam energy for each decay mode can be calculated from Equation B.5. The beam energy values from the different modes are combined using a standard weighted average:

$$\overline{(E_{\mathrm{beam}}^*)}_{\mathrm{calibrated}} = \frac{\sum_i((E_{\mathrm{beam}}^*)_{\mathrm{calibrated}}^i/\sigma_i^2)}{\sum_i(1/\sigma_i^2)} \tag{B.6}$$

where the index i runs over the various modes analyzed, and σ_i is the uncertainty in any $(E^*_{\mathrm{beam}})^i_{\mathrm{calibrated}}$. This uncertainty is related to the mean of the nominal beam constrained mass and its fitted uncertainty by

$$\sigma_i = \frac{(M_{\mathrm{bc}})^i_{\mathrm{nominal}}}{(E^*_{\mathrm{beam}})^i_{\mathrm{calibrated}}}\delta(M_{\mathrm{bc}})^i_{\mathrm{nominal}} \tag{B.7}$$

The overall statistical uncertainty on the weighted mean beam energy is also determined in the usual way,

$$\sigma = \frac{1}{\sqrt{\sum_i (1/\sigma_i^2)}} \tag{B.8}$$

For the fits to the nominal beam-constrained mass distributions, all PDF parameters are floated. To obtain good fits under these conditions, the signal modes must have a high signal-to-noise ratio. Originally, this calibration was performed using the decays $B^- \to D^0\pi^-$ and $B^- \to D^{*0}\pi^-(D^{*0} \to D^0\pi^0)$ decays, with D^0 reconstructed in the modes $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$, and $D^0 \to K^-\pi^+\pi^-\pi^+$ [53]. In 2009, an effort was undertaken to implement a new algorithm for charged track reconstruction at Belle. The new algorithm resulted in significant improvements in efficiency for charged track reconstruction, and as a result changed the signal-to-noise ratios in some of the listed modes. Ultimately, the entire sample of $\Upsilon(4S)$ data taken under SVD2 conditions (experiments 31-65), was reprocessed with this algorithm. At this time, the selection of modes used to calibrate the beam energy was revisited.

Investigations into the beam energy calculations revealed significant differences between the values calculated for different modes. An example of these discrepancies is shown in Figure B.1. It was discovered that modes including neutral pions had a systematic offset in beam energy relative to those that included only charged tracks. It is likely that this offset is due to fitting bias, as a Gaussian does not properly model the small energy leakage from the ECL for the $\pi^0 \to \gamma\gamma$ reconstruction. Though these offsets were small , $\mathcal{O}(0.5\ \mathrm{MeV})$, it was conservatively decided to remove these modes from the beam energy analysis to reduce systematic errors, leaving only two $B^- \to D^0\pi^-$ sub-decays in which to measure the beam energy: $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^-\pi^+$. A further complication arose in the $D^0 \to K^-\pi^+\pi^-\pi^+$ decay, in that the increased efficiency for charged tracks resulted in significantly higher combinatorial backgrounds, and therefore an overall lower signal-to-noise ratio (see Figure B.2). As the large background creates problems with fit stability, the selection criteria were also revisited. To further improve statistics, a previ-

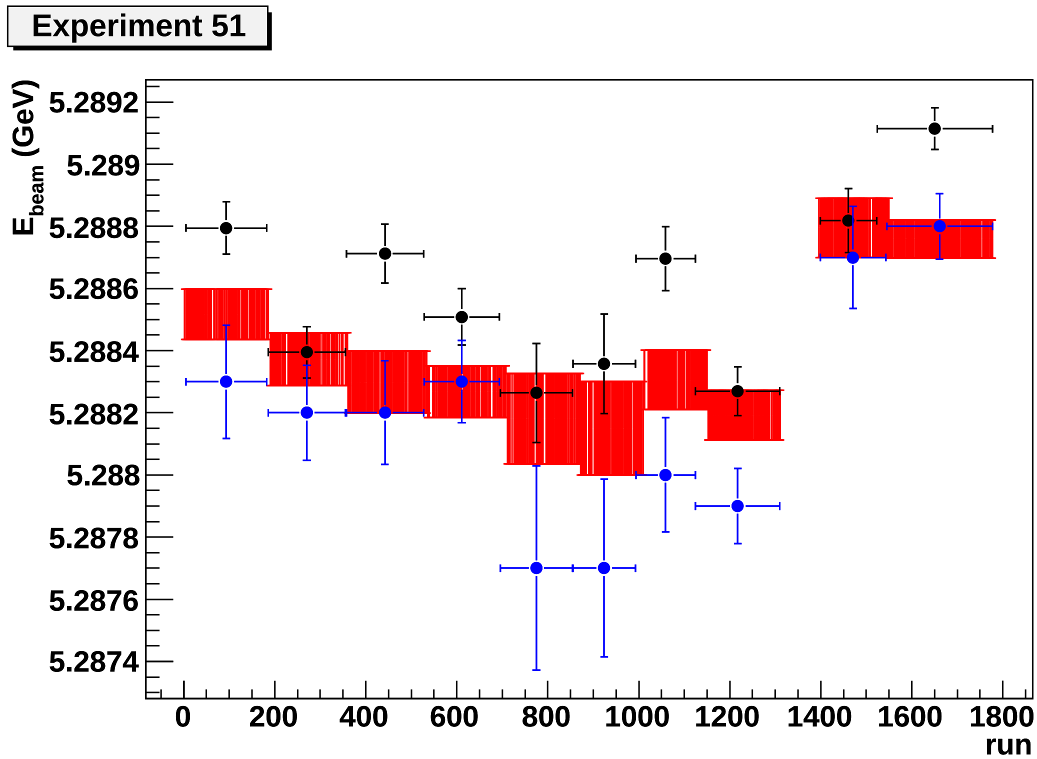

Figure B.1: Beam energy values by run number for experiment 51, using the existing beam energy procedures, for data processed with the old tracking (red bands), new tracking (black points with errors), and new tracking with only the $B^- \to D^0\pi^-(D^0 \to K^-\pi^+)$ decays (blue points with errors). All error bars are statistical only.

ously unused mode was also added: $B^0 \to D^-\pi^+(D^- \to K^+\pi^-\pi^-)$. The final selection criteria for the three modes used are shown in Table B.1. For all three reconstructed modes, in cases where multiple B candidates exist, the B candidate with its candidate D mass the closest to the nominal D mass is selected. The addition of loose impact parameter cuts for charged tracks resulted in the largest improvements to the signal-to-noise ratio. Sample distributions of the nominal beam-constrained mass with the final selection criteria are shown in Figure B.3.

The beam energy was re-calculated using these new procedures for experiments 31-65. In addition to the changes in reconstruction criteria, an effort was made to add systematic uncertainties to each beam energy measurement. These include a small bias observed in MC in reconstructed B momentum, but are dominated by the uncertainty in the world average of the B mass, $\sim 0.3\ \mathrm{MeV}/c^2$ [1]. Though the removal of the π^0 modes

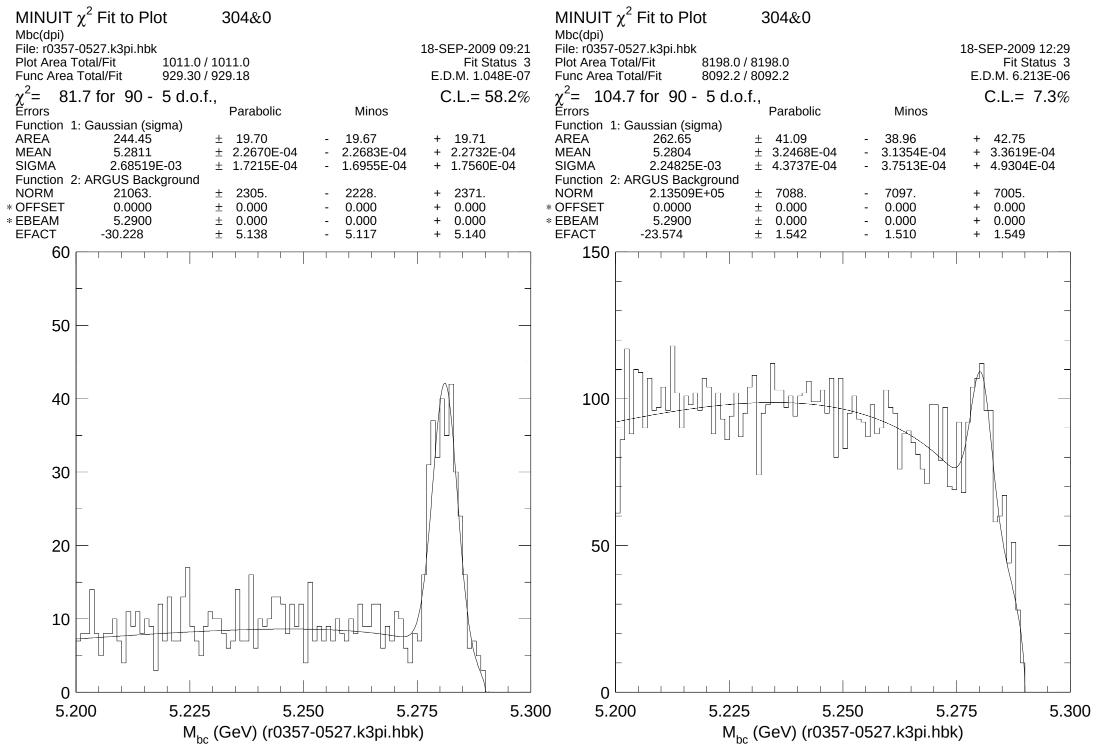

Figure B.2: M_{bc} distributions calculated with the nominal beam energy for the decay mode $B^- \to D^0\pi^-(D^0 \to K^-\pi^+\pi^-\pi^+)$, using data from runs 357-527 of experiment 51 with the old tracking (left) and new tracking (right). Both plots utilize the existing beam energy reconstruction selection criteria.

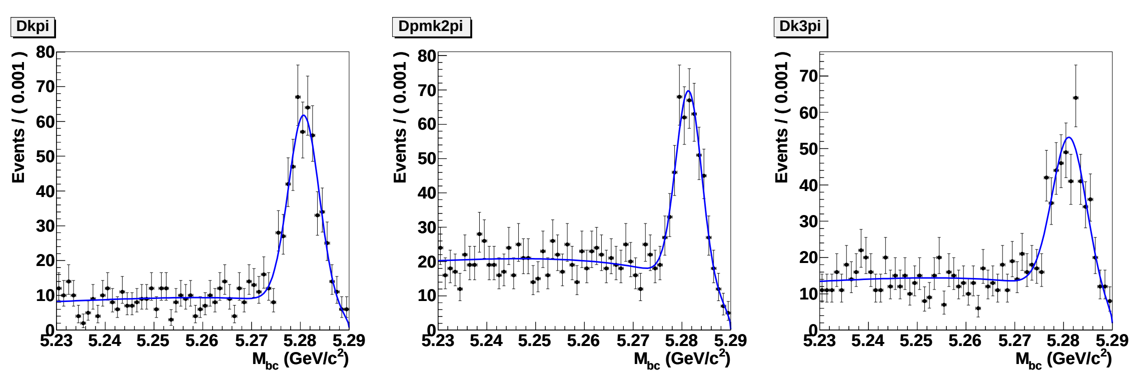

Figure B.3: Beam energy values by run number for experiment 51, using the new beam energy procedures. The run range is the same as that used for Figure B.2. The distributions correspond to the modes with a D reconstructed as $K^+\pi^-$ (left), $K^+\pi^-\pi^-$ (middle), and $K^+\pi^-\pi^+\pi^-$ (right).

Decay	Sub-decay	Criteria	Comment
		$R2 < 0.5$	$q\bar{q}$ suppression
		$PID(K)_K > 0.3$	kaon ID
	All decays	$PID(K)_\pi < 0.7$	pion ID
		$dr_{(K,\pi)} < 3$ cm	reject bad tracks
		$\lvert dz_{(K,\pi)}\rvert < 5$ cm	
$B^- \to D^0\pi^-$			
		$\lvert cos\theta_{\mathrm{sph}}\rvert < 0.9$	$q\bar{q}$ suppression
	$D^0 \to K^-\pi^+$	$\lvert\Delta E\rvert < 50$ MeV	
		$\lvert M_{K\pi} - M_D\rvert < 16.2$ MeV$/c^2$	
		$\lvert cos\theta_{\mathrm{sph}}\rvert < 0.6$	$q\bar{q}$ suppression
	$D^- \to K^-\pi^+\pi^-\pi^+$	$\lvert\Delta E\rvert < 40$ MeV	
		$\lvert M_{K\pi\pi\pi} - M_D\rvert < 13.3$ MeV$/c^2$	
$B^0 \to D^-\pi^+$			
		$\lvert cos\theta_{\mathrm{sph}}\rvert < 0.9$	$q\bar{q}$ suppression
	$D^- \to K^+\pi^-$	$\lvert\Delta E\rvert < 50$ MeV	
		$\lvert M_{K\pi\pi} - M_D\rvert < 16.2$ MeV$/c^2$	

Table B.1: Selection criteria for the reconstructions used to calculate the beam energy. $R2$ is the second Fox-Wolfram moment, normalized by the zeroth Fox-Wolfram moment, and θ_{sph} is the angle between the sphericity axis of the signal B daughters and the sphericity axis of the other B daughters.

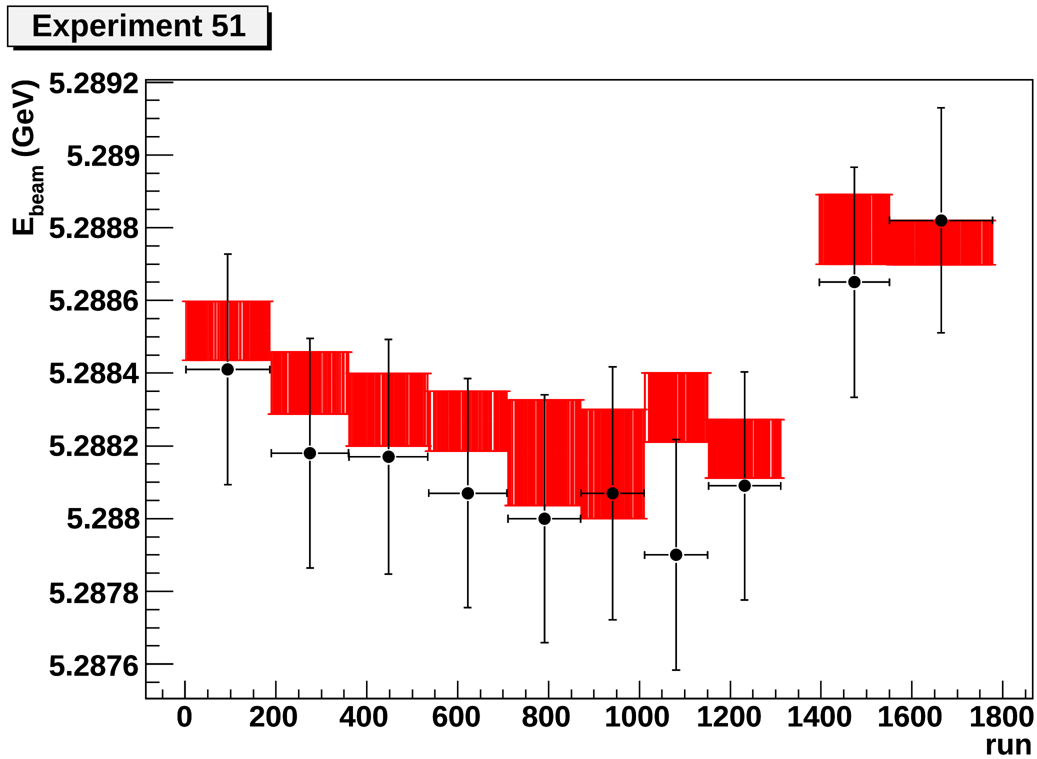

Figure B.4: Beam energy values by run number for experiment 51, for data processed with the old tracking and old procedures (red bands), and new tracking with new procedures (black points with errors). Error bars for the old procedures are statistical only, and those for the new procedures include systematic uncertainties.

caused an overall shift in the beam energy relative to the previously tabulated values, the new values remain consistent with the old values within their total uncertainty. The final results for experiment 51 are shown in Figure B.4. Results for experiments 31-65 have been updated in the Belle beam energy database.

Since the procedure for calculating beam energy already requires fits to the nominal beam-constrained mass distributions, we can also monitor the width of these distributions throughout a series of runs and experiments. These widths directly reflect the spreads in beam energy over the corresponding run ranges. Variations of up to 10% are visible, as shown in Figure B.5. The spread in beam energies is expected to be correlated with the bunch current of the LER. As such, we have examined these spreads as a function of the LER current. Because we combine many runs to gain adequate statistics for the $M_{\rm bc}$ fits, we cannot ensure that the LER current was stable during any particular run range.

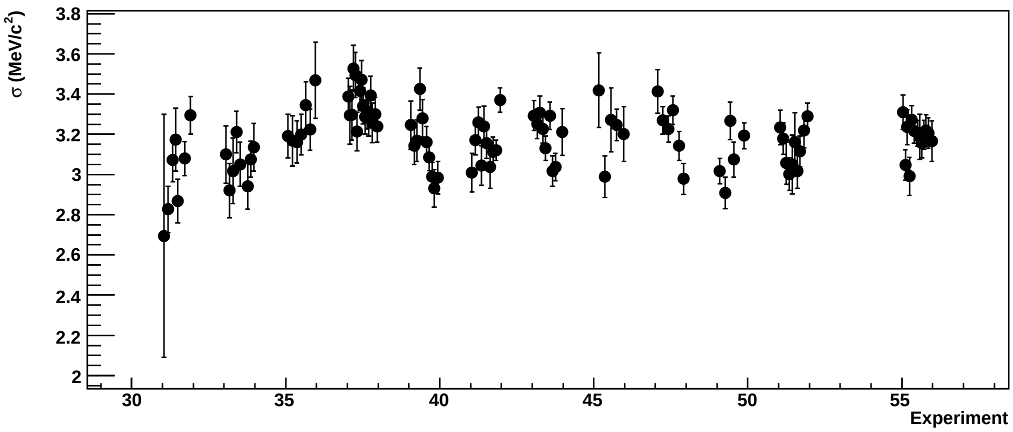

Figure B.5: Fitted widths of the nominal beam-constrained mass versus experiment number.

To avoid periods where the LER current may have been changing significantly, we restrict the study to points where the RMS of the LER current over the run range was less than 100 mA. The results are shown in Figure B.6. An approximate correlation is seen, though the uncertainties in both the fitted widths and the LER currents make it difficult to make a firm conclusion. A future study in which the run ranges are chosen specifically to isolate regions of stable beam current may be able to demonstrate this effect more dramatically.

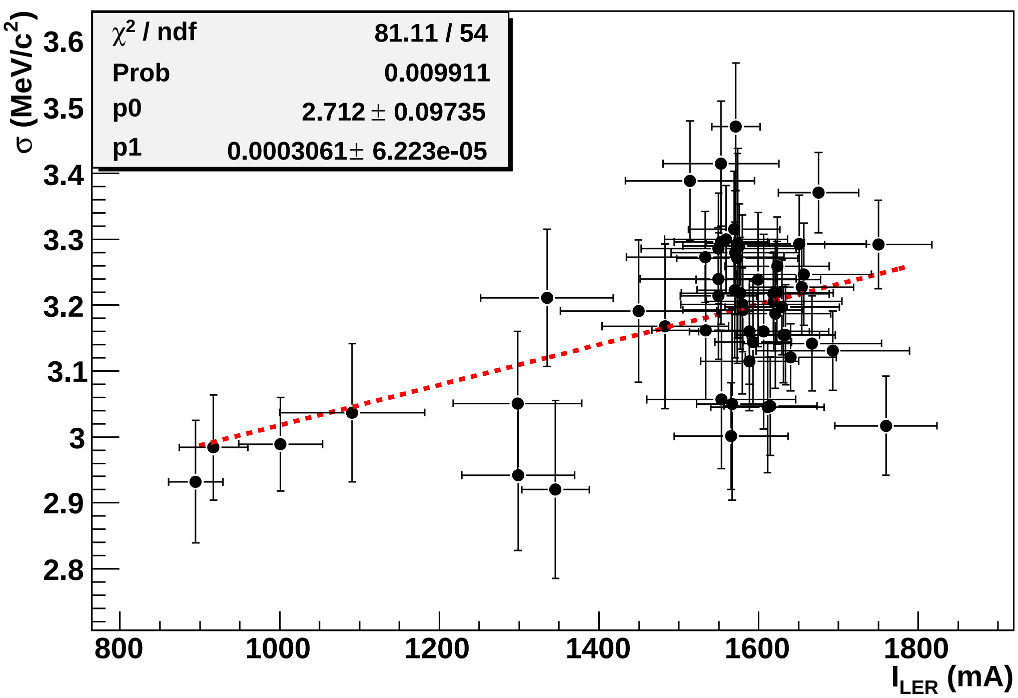

Figure B.6: Fitted widths of the nominal beam-constrained mass versus the average LER current for the corresponding run range.

Appendix C

Calculation of the Number of $B\bar{B}$ Pairs

The calculation of each branching fraction at Belle requires knowledge of the number of $B\bar{B}$ pairs in the data sample. The hadronic data sample includes contributions from both $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events and $e^+e^- \rightarrow q\bar{q}$ backgrounds. The total number of $B\bar{B}$ events produed on-resonance (at the $\Upsilon(4S)$ energy) is thus

$$N^{\mathrm{on}} = N^{\mathrm{on}}_{B\bar{B}} + N^{\mathrm{on}}_{q\bar{q}} \tag{C.1}$$

The number of $q\bar{q}$ events at a given CM energy $E^* = \sqrt{s}$ is given by

$$N_{q\bar{q}} = \frac{\sigma^0_{q\bar{q}}}{s}\mathcal{L} \tag{C.2}$$

where $\sigma^0_{q\bar{q}}$ is the cross section for $q\bar{q}$ production with the dependence on s removed and instead written explicitly, and $\mathcal{L}$ is the integrated luminosity. To get a numerical estimate of $N_{q\bar{q}}$ component in the on-resonance data, Belle takes approximately 10% of its data off-resonance, ~ 60 MeV below the $\Upsilon(4S)$. We can define the ratio, α, of the number of $q\bar{q}$ events in the on-resonance sample and the off-resonance sample.

$$\alpha = \frac{N^{\mathrm{on}}_{q\bar{q}}}{N^{\mathrm{off}}_{q\bar{q}}} = \frac{\mathcal{L}^{\mathrm{on}} s^{\mathrm{off}}}{\mathcal{L}^{\mathrm{off}} s^{\mathrm{on}}} \tag{C.3}$$

We can use this to rewrite Equation C.1 and solve for the number of $B\bar{B}$ pairs.

$$\begin{aligned} N_{B\bar{B}} &= N^{\mathrm{on}} - N^{\mathrm{on}}_{q\bar{q}} \tfrac{N^{\mathrm{off}}_{q\bar{q}}}{N^{\mathrm{off}}_{q\bar{q}}} \\ N_{B\bar{B}} &= N^{\mathrm{on}} - \alpha N^{\mathrm{off}}_{q\bar{q}} \end{aligned} \tag{C.4}$$

However, the rate of fermion production should be identical whether the fermions produced are quarks or leptons, so we can relate α to the number of on-resonance and off-resonance Bhabha or mu-pair events,

$$\alpha = \frac{N^{\text{on}}_{e\bar{e}}}{N^{\text{off}}_{e\bar{e}}} = \frac{N^{\text{on}}_{\mu\bar{\mu}}}{N^{\text{off}}_{\mu\bar{\mu}}} \tag{C.5}$$

In previous implementations of the $N_{B\bar{B}}$ calculations [35, 53], the value of α was determined independently for Bhabha events and mu-pair events. However, the number of mu-pair events for data taken after experiment 27 is not considered reliable, so we must use α only as obtained for Bhabha events.

We also have not yet included the overall efficiencies for detecting $B\bar{B}$ or $q\bar{q}$ events in the hadronic data samples. When these are explicitly written, we modify Equation C.4 into

$$N_{B\bar{B}} = \frac{N^{\text{on}} - r(\epsilon_{q\bar{q}})\alpha N^{\text{off}}_{q\bar{q}}}{\epsilon_{B\bar{B}}} \tag{C.6}$$

where $r(\epsilon_{q\bar{q}})$ is the ratio of the efficiency for $q\bar{q}$ events off-resonance to the efficiency for $q\bar{q}$ events on-resonance. Through experiment 55, the values of both $\epsilon_{B\bar{B}}$ and $r(\epsilon_{q\bar{q}})$ have been assumed constant, as estimated from early MC studies [35]. Since these studies were potentially outdated, especially in light of the new tracking algorithms introduced for data from experiment 61 and beyond, a new MC study was undertaken to assess the appropriate values of $\epsilon_{B\bar{B}}$ and $r(\epsilon_{q\bar{q}})$.

The efficiency of the hadronic skim was tested by applying the skim to MC samples of generic $B\bar{B}$ decays as well as on-resonance and off-resonance $q\bar{q}$ decays. Efficiencies for all experiments are shown in Figure C.1. Averages for experiments 61, 63, and 65, for which the $N_{B\bar{B}}$ calculation was to be performed, are shown in Table C.1, along with the previously used values. Over the entire experimental range of Belle, the $B\bar{B}$ efficiency varies by roughly 0.5%, and the $q\bar{q}$ efficiency by about 0.3%. Due to the significant variations, it was decided to use experiment-dependent efficiencies and efficiency ratios for the new calculations. It is also notable that the previous $q\bar{q}$ efficiencies are much lower (over 10%) than any of those obtained from MC. The exact source of this discrepancy is unknown, as much has changed in the Belle analysis tools and software libraries since the value was originally calculated in 2001. However, the important parameter for the $N_{B\bar{B}}$ calculation is the efficiency ratio $r(\epsilon_{q\bar{q}})$, shown for each experiment in Figure C.2. Despite the large changes in overall $q\bar{q}$ efficiency, the ratio shows variations of less than 0.5%.

Quantity	Previous value	New values		
		Exp. 61	Exp. 63	Exp. 65
$\epsilon_{B\bar{B}}$	0.9913	0.9897	0.9897	0.9893
$\epsilon_{q\bar{q}}^{\mathrm{on}}$	0.795	0.9092	0.9093	0.9088
$\epsilon_{q\bar{q}}^{\mathrm{off}}$	0.792	0.9092	0.9084	0.9086
$r(\epsilon_{q\bar{q}})$	0.9958	1.0001	0.9990	0.9998

Table C.1: Efficiencies and continuum efficiency ratio used for previous $N_{B\bar{B}}$ calculations. The previous value for the off-resonance continuum efficiency is inferred from the value of $r(\epsilon_{q\bar{q}})$ and $\epsilon_{q\bar{q}}^{\mathrm{on}}$.

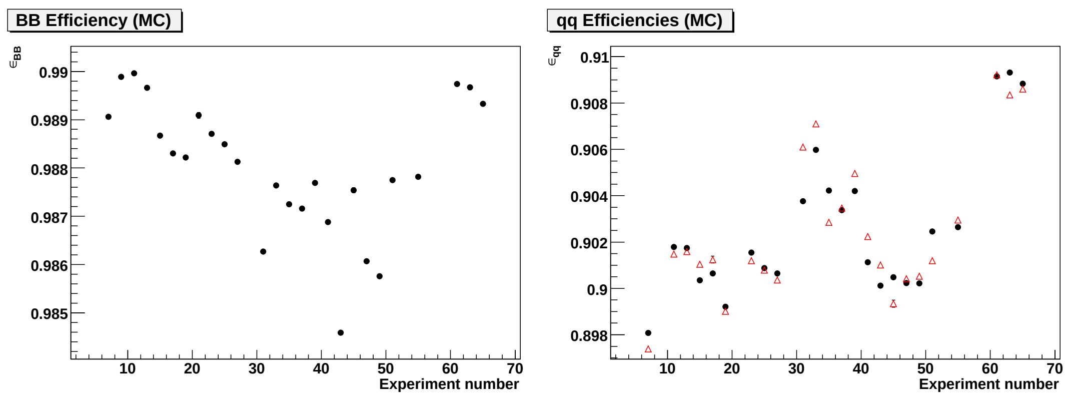

Figure C.1: Hadronic skim efficiencies versus experiment number for MC samples of $B\bar{B}$ events (left) and $q\bar{q}$ events (right). The $q\bar{q}$ efficiencies are shown for conditions on-resonance (black filled circles) and off-resonance (red open triangles).

The final tabulated values of $N_{B\bar{B}}$ for experiments 61, 63, and 65, are shown in Table C.2, using both the old efficiency values and the new experiment-dependent values. The relative differences in $N_{B\bar{B}}$ between the two calculations are in the range 1% to 1.4%, roughly the same size as the uncertainties on the calculated values.

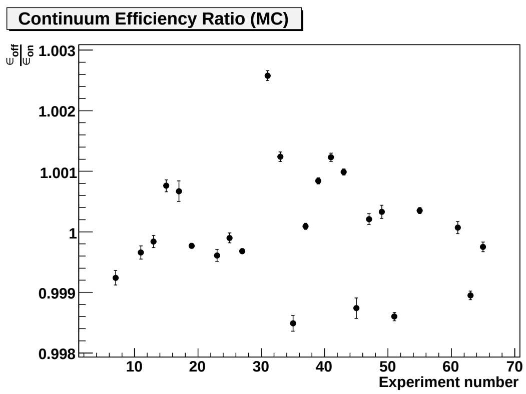

Figure C.2: Ratio of efficiencies of off-resonance $q\bar{q}$ events to on-resonance $q\bar{q}$ events versus experiment number.

	$N_{B\bar{B}}$ (10^6)		
	Exp. 61	Exp. 63	Exp. 65
Using old ϵ, r	(37.87 ± 0.59)	(35.90 ± 0.56)	(42.21 ± 0.67)
Experiment-dependent ϵ, r	(37.45 ± 0.56)	(35.62 ± 0.53)	(41.79 ± 0.63)

Table C.2: $N_{B\bar{B}}$ values for experiments 61, 63, and 65, using both the previous and new efficiencies and experiment-dependent efficiency ratios. Uncertainties include statistical and systematic effects.

Bibliography

[1] K. Nakamura *et al.*, J. Phys. **G37**, 075021 (2010).

[2] C. Burgess and G. Moore, *The Standard Model: A Primer* (Cambridge University Press, 2007).

[3] ALEPH, DELPHI, L3, OPAL, and SLD Collaborations, LEP Electroweak Working Group, and SLD Electroweak and Heavy Flavor Groups, Phys. Rep. **427**, 257 (2006).

[4] D. Perkins, *Introduction to High Energy Physics*, 4th ed. (Cambridge University Press, 2000).

[5] E. Klempt, hep-ph/0404270.

[6] S. Glashow, Nucl. Phys. **22**, 579 (1961).

[7] A. Salam and J. Ward, Phys. Lett. **13**, 168 (1964).

[8] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[9] CKM Fitter Group, J. Charles *et al.*, Eur. Phys. J. **C41**, 1 (2005), updated results and plots available at: http://ckmfitter.in2p3.fr.

[10] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **71**, 032001 (2005).

[11] V. Abazov *et al.*, Nucl. Instrum. Methods **A565**, 463 (2006).

[12] A. Breskin and R. Voss, *The CERN Large Hadron Collider: Accelerator and Experiments* (CERN, Geneva, 2009).

[13] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events* (Cambridge University Press, 2006).

[14] E. Kou, Phys. Rev. **D63**, 054027 (2001).

[15] CLEO Collaboration, T. E. Browder *et al.*, Phys. Rev. Lett. **81**, 1786 (1998), hep-ex/9804018.

[16] A. Datta, X. G. He, and S. Pakvasa, Phys. Lett. **B419**, 369 (1998), hep-ph/9707259.

[17] I. E. Halperin and A. Zhitnitsky, Phys. Rev. Lett. **80**, 438 (1998), hep-ph/9705251.

[18] D. Atwood and A. Soni, Phys. Lett. **B405**, 150 (1997), hep-ph/9704357.

[19] CLEO, M. Artuso *et al.*, Phys. Rev. **D67**, 052003 (2003), hep-ex/0211029.

[20] CLEO Collaboration, G. Bonvicini *et al.*, Phys. Rev. **D68**, 011101 (2003), hep-ex/0303009.

[21] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **93**, 061801 (2004), hep-ex/0401006.

[22] W.-S. Hou and B. Tseng, Phys. Rev. Lett. **80**, 434 (1998), hep-ph/9705304.

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

[24] J. Chay, C. Kim, A. K. Leibovich, and J. Zupan, Phys. Rev. **D76**, 094031 (2007), 0708.2466.

[25] S. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[26] CUSB Collaboration, T. Bohringer *et al.*, Phys. Rev. Lett. **44**, 1111 (1980).

[27] CUSB Collaboration, G. Finocchiaro *et al.*, Phys. Rev. Lett. **45**, 222 (1980).

[28] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods **A499**, 1 (2003).

[29] A. Abashian *et al.*, Nucl. Instrum. Methods **A479**, 117 (2002).

[30] D. J. Lange, Nucl. Instrum. Methods **A462**, 152 (2001).

[31] R. Brun, F. Carminati, and S. Giani, CERN-W5013.

[32] CLEO Collaboration, M. S. Alam *et al.*, Phys. Rev. Lett. **74**, 2885 (1995).

[33] T. Sjöstrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), hep-ph/0603175.

[34] Belle Collaboration, C. H. Wang *et al.*, Phys. Rev. **D75**, 092005 (2007), hep-ex/0701057.

[35] B. Casey, *Search for CP Violation in Charmless Hadronic B Meson Decays*, PhD thesis, University of Hawaii, 2001.

[36] B. Casey, Belle Internal Note 390: HadronB.

[37] P. Koppenburg, Belle Internal Note 665: An Improved π^0 and η Veto.

[38] F. Fang, Belle Internal Note 323: Study of $K_S^0 \to \pi^+\pi^-$ Selection.

[39] F. Fang, *Measurement of Branching Fractions and CP Violation in $B \to \eta_c K$ and Observation of $B^\pm - > p\bar{p}K^\pm$*, PhD thesis, University of Hawaii, 2003.

[40] K. Nishimura, Belle Internal Note 1066: Inclusive $B \to X_s\eta$ Branching Fraction and Mass Spectrum Measurement.

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

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

[43] P. H. R.O. Duda and D. Stork, *Pattern Classification*, 2nd ed. (Wiley and Sons, 2001).

[44] H. Kakuno *et al.*, Nucl. Instrum. Methods **A533**, 516 (2004).

[45] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. **B241**, 278 (1990).

[46] BELLE Collaboration, S. Blyth *et al.*, Phys. Rev. **D74**, 092002 (2006), hep-ex/0607029.

[47] S. Lin, Belle Internal Note 645: Update of π^0 systematics Using Inclusive η (78/fb).

[48] S. Lin, Belle Internal Note 901: Br. of B to hh with 447 Million BBbar pairs.

[49] P. Koppenburg, Belle Internal Note 621: A Measurement of the Track Finding Efficiency Using Partially Reconstructed D^* Decays.

[50] G. Majumder, Belle Internal Note 641: Uncertainty of track finding efficiency with embedded tracks.

[51] S. Nishida, Belle Internal Note 779: Study of kaon and pion identification using inclusive D^* sample.

[52] Belle Collaboration, K. Abe *et al.*, Phys. Rev. **D73**, 051106 (2006), hep-ex/0601032.

[53] S. Swain, *Search for Direct CP Violation in the Decay $B^- \to D_{CP}K^{*-}$*, PhD thesis, University of Hawaii, 2004.