A MEASUREMENT OF INCLUSIVE RADIATIVE PENGUIN DECAYS OF B MESONS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the

Graduate School of The Ohio State University

By

Jana B. Lorenc, B.S., M.S.

* * * * *

The Ohio State University

2002

Dissertation Committee:

Klaus Honscheid, Adviser

Richard D. Kass

Stephen Pinsky

Linn Van Woerkem

Rama K. Yedavalli

Approved by

_____

Adviser
Department of Physics

ABSTRACT

We have measured the branching fraction and photon energy spectrum for the inclusive radiative pengiun decay $b \rightarrow s\gamma$ for the full CLEO II and CLEO II.V datasets. We find $\mathcal{B}(b \rightarrow s\gamma) = (3.21 \pm 0.43 \pm 0.27^{+0.18}_{-0.10}) \times 10^{-4}$, where the first error is statistical, the second is systematic, and the third is from theory corrections. We also obtain first and second moments of the photon energy spectrum above 2.0 GeV, $\langle E_\gamma \rangle = 2.346 \pm 0.032 \pm 0.011$ GeV, and $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2 = 0.0226 \pm 0.0066 \pm 0.0020$ GeV^2, where the errors are statistical and systematic. From the first moment we obtain the HQET parameter $\overline{\Lambda} = 0.35 \pm 0.08 \pm 0.10$ GeV to order $1/M_B^3$ and $\beta_0 \alpha_s^2$.

This is dedicated to my parents, Miroslav and Bozena, and to my husband, Gregg.

Dear Friends and Family,

Without you and your tireless encouragement and support, none of this would have been possible. It's a shame that the acknowledgments always seem to be left for last, because it allows me, once again, to take for granted all of the people that have had so much influence on my life. Hopefully, you understand how much your friendship has meant to me.

I'd like to thank Klaus Honscheid for everything that he has taught me. Under his guidance, I learned the finer points of hardware design, and I'm grateful for the opportunities he has given me. Working with Ed Thorndike has been a great experience. His enthusiasm for physics, his insight, and his ability to explain any aspect of an analysis clearly, concisely and repeatedly is what made this project possible. His stories and advice made even the most difficult times more pleasant. It's been a great pleasure learning to do physics under his wing.

Designing and installing the CLEO III detector was some of the best fun I had in graduate school, and no one was more integral to this than Chris Bebek. No matter how busy he was, he always seemed to have time to explain some piece of hardware design or help me debug a board. His crate monitor became my crate monitor, and I am proud that I could take over that project.

Mats Selen is the reason that I'm in physics today. His enthusiasm for discovery and his sense of fun in doing research are contagious. I have so many wonderful U of I memories to thank him for, that there isn't enough room here to write them all down, but I specifically remember building the Cosmic Ray Test Stand in Physics Lab 1, the joy of Physics Van (and especially Van R&D, which usually involved massive sparks or violent explosions), and many a pleasant barbecue. I will also never forget that he was the one who first introduced me to the soft serve ice cream at the Red Caboose in Groton.

Gregg's contribution actually falls into two categories. First, as a co-worker, he's made just about every bit of CLEO work seem like fun, even if it's something as mundane as dressing cables or replacing power supplies. However, he's also the nicest, most unselfish person I've every met, and I feel lucky that I get to spend every day with him. He's taught me how to see the humor in any situation, how to take things as they come and how to enjoy life. Not only does he have a sixth sense about when to give me chocolate or take me for ice cream, but he's always ready with a kind word or a hug when I need it. I'd like to thank him for all of his love and support, not just for this project, but always. I wish I could write more here, but I don't really have to words or the space to express how I feel. Gregg, I hope that you understand and know how I feel without my having to commit the words to paper. You and your family are awesome. When your turn comes, I'll be right behind you...with a case of Skyline Chili to perk you right up.

Véronique Boisvert has helped me by sharing her experiences and teaching me what it means to be a model citizen at the lab. She was always willing to answer

any question, no matter how trivial, and she always had good advice to give. I'm going to miss having her around to tell me what to do, but I guess we all have to learn to find our own way at some point. Going to lunch with Véronique and Silvia was always a blast and our lunches helped to give me a sense of perspective that I wouldn't otherwise have had. Thank you!

Now that Véronique and Silvia have moved on to bigger and better things, my new partners in crime are Hanna Mahlke-Kruger, Lauren Hsu, Anto Romano, Nadia Adam, and Selina Li. Whether having fun at EYH, taking a Step class or salsa lessons, going for ice cream, or going to lunch, these ladies have been tremendously supportive of all of my endeavors. They've made me laugh when I needed to and given me encouragement and positive feedback when it was warranted. An extra special thank you goes to Hanna for reading over my thesis and giving me lots of wonderful feedback.

Tom Meyer and Janis Chang have been fantastic friends and neighbors. I have them to thank for the one race I ran in Ithaca...without them, I would never have believed I could actually run 3K! Tom is a superb physicist, and his questions and insight are always valuable (especially at practice talks!). He was one of the people that made CLEO III installation special. Janis has provided chocolate and therapy on many crucial occasions, and has been a good friend ever since I moved out to Ithaca. Cayce and Gavin and Katie and Steve, among others, have also made life in Ithaca more pleasant. Their support, whether it was reading my thesis, encouraging me to exercise, or just listening to me ramble, was invaluable. They put the happy in Happy Hour.

There are numerous people at OSU who made the first three years in Columbus lively and entertaining. Chul Gwon and Dallas Trinkle both contributed in their own way to make problem sets bearable and long days at the lab seem like fun. Chul is a very caring and entertaining officemate, and I'd like to thank him for helping me get through the first three years of grad school with a smile and teaching me how not to make an utter fool of myself on the basketball court. Dallas, well, what can I say about Dallas? He always made me laugh and introduced me to the Breakfast of Champions: stale Peeps and Coca-Cola. I hope you'll always save a dance for me, guys. Eric Eckhart, Dirk Hufnagel, and Terry Hart have also been good officemates who were always willing to engage in long conversations on just about any topic. Todd Pedlar and Tim Wilksen have taken over this role in Ithaca, and they never seem to get too upset when I come into the office, breathless with excitement, babbling about some new bird I just saw. Tolerant office mates are hard to come by and I've been very fortunate in that respect.

Jesse Ernst, Dan Hennessy, and Adam Lyon round out the $b \rightarrow s\gamma$ team. I learned so much from these three, that half of the information in this thesis should have a reference to them in the Bibliography. I read Jesse's thesis when I first came to graduate school, and I remember thinking how cool it was at the time. His was

the first analysis to catch my imagination, so I feel like I've come full circle by working on this project. I've always admired his creativity, his physics knowledge, and his patience for explaining any aspect of the analysis several times over. On top of that, he's an all-around good guy. Dan is an inspiration to me. His work ethic is unbeatable, and he seems to know something about everything. How do you do that, Dan? Adam was always full of good advice and stories. Furthermore, he often dropped what he was doing to help me sort out one issue or another. I just can't thank them enough. If I turn out half as good as any one of these three, I'll be thrilled.

I'd like to thank Dan Peterson for useful discussions about the inner workings of the detector and for cheering me up during the writing of my thesis. Andreas Warburton has helped me understand all sorts of nasty stuff like statistics and submitting jobs to the queues. I always remember the $-V now. Jean Duboscq stands out as someone who always knew the right thing to say, no matter what the problem. Chris Jones has gotten me over my fear of C++ by answering all of my questions with a smile. The Minnesotans have also improved the quality of life in Ithaca. Whether hanging out at the Plantation Inn or roasting marshmallows at a bonfire, somehow they were able to keep me calm in the weeks leading up to my thesis defense...no small feat, either.

Margee Carrier and Gloria LeFave were undisputedly my heroes during the days of CLEO III installation. Together, they made even the most mundane tasks seem like fun, and they taught me that where there is a will and a crimping tool, there is a way to connect any kind of cable to any kind of connector.

I'm truly lucky to be surrounded by so many kind and gifted people. The people I regularly interact with at Wilson and OSU, even though I don't explicitly mention them all by name, have answered questions, made suggestions, and their influence is apparent throughout the thesis.

Last, but certainly not least, my parents are the best anyone could ever ask for. They have been supportive of all of my endeavors, whether it was music lessons, the theater, or graduate school in physics. Thank you for letting me find my own way, and giving me so much to be happy about. If I accomplish anything in this world, it's thanks to the lessons they taught me when I was growing up. Their influence is immeasurable since it permeates every aspect of my life. They've been such an important part of my life that I can't possibly list all of the things I'm thankful for. I hope that by dedicating my work to them, they realize how much they mean to me.

Thank you,

Jana

The author was born in Proviso East Township, Illinois, on July 28, 1974 as the only child of Miroslav and Bozena Lorenc. She grew up in Berwyn, a suburb of Chicago, where she spent many a happy summer playing baseball and ghost-in-the-graveyard in the alley behind her house. After graduating from Morton West High School, the author went on to study physics at the University of Illinois at Urbana-Champaign where she completed a Bachelor of Science in Engineering Physics. Hoping to work on the CLEO experiment, she entered the Ph.D. program at The Ohio State University in the summer of 1996 and began her research with Klaus Honscheid. She earned a Master of Arts degree in March 1999. Soon after, she moved to Ithaca, NY to work on the CLEO III detector installation. She got married on August 7, 1999 to Gregg Thayer and her life has been that much richer and happier ever since. She will now continue her research under the expert guidance of Ed Thorndike at the University of Rochester.

TABLE OF CONTENTS

Page

LIST OF TABLES

xiii

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Motivation

Particle physicists seek to understand the fundamental structure of observable matter and to develop a comprehensive theory that describes the interactions between elementary particles. The idea that all matter can be reduced to elementary components has been around since the days of the early Greeks who first introduced the idea of fundamental and indivisible particles called atoms. Over the course of many centuries this idea was explored much further. Mendeleev's periodic table organized atoms according to their properties and revealed a pattern. The pattern evident in the Periodic Table led to the discovery of nuclei which was followed by the discovery of the protons and neutrons glued together inside the nucleus of the atom. With the advent of particle accelerators, physicists could create their own subatomic particles. In the mid-20th century, many elementary particles were discovered and classified according to their properties. Likewise, these properties revealed patterns that seemed to indicate that these particles were composed of a simple set of elementary point-like particles.

By observing such patterns, physicists obtain clues about the rules by which these pieces were put together. Applying the scientific method, physicists examine these patterns and make hypotheses, create experiments to test these hypotheses, and then interpret the results to determine what the patterns indicate about the world around us. Physicists have collected these results and combined them in a theory that describes the behavior of these fundamental building blocks of the material world. This chapter describes the basics of this theory, which is known as the "Standard Model" (SM).

1.2 The Standard Model

The Standard Model is the most successful attempt to date at integrating all of the basic laws that govern the structure of matter and the interactions between

particles. It describes the behavior of all known subatomic particles using a single theoretical framework.

1.2.1   The Particles

The world around us consists of three kinds of elementary particles: quark, leptons, and force carrier particles called gauge bosons. All of our understanding of the properties of matter derives from the interactions of these particles. Each particle has an antimatter partner with identical mass but opposite electric charge. To the resolution achievable by current accelerators ($r \approx 10^{-19}$ m), these particles show no internal structure. There are several key properties that need to be determined in order to uniquely identify a particle:

- *Mass*

  A particle is usually first identified by measuring its mass which in principle can be determined using Newton's law by observing the acceleration of the particle under the influence of a force. This simple scenario breaks down when quark masses are considered because they are not accessible to direct experimental measurement. The precise value of the mass of a quark depends on the context, whether it is confined to a baryon or a meson.

- *Spin*

  A particle's spin is its intrinsic angular momentum, a purely quantum mechanical property similar to orbital angular momentum. Quarks and leptons have half-integer spin, while gauge bosons have integer spin. There is a connection between spin and symmetry, first noted by Pauli: the wave function of a system of n identical particles with half-integer spin, called fermions, changes sign if any two particles are interchanged. The wave function of a system of n identical particles with integer spin, called bosons, remains unchanged under the interchange of two particles. This connection between spin and symmetry leads to the Pauli Exclusion Principle which states that one quantum mechanical state can be occupied by only one fermion. This principle is profoundly important in all of particle physics.

- *Electric Charge*

  Many types of particles possess an electric charge. The total charge of a subatomic particle determines the strength of its interaction with the electromagnetic force.

- *Other Quantum Numbers*

  The most familiar quantum numbers are those associated with electrons in an atom like energy, angular momentum, and spin. When studying subatomic

particles, it is necessary to come up with some new quantum numbers that describe the states of quarks and leptons such as parity (space inversion) and particle-antiparticle conjugation.

Table 1.1 and Table 1.2 list what are believed to be the elementary particles that make up all matter, the quarks and the leptons. All of these particles are fermions and carry a spin of 1/2. Furthermore, although they have mass, they are point-like particles that have no size.

Quark	Name	Charge	Mass (GeV/c^2)	Weak Isospin
u	up	$+2/3$	$0.001 - 0.005$	$+1/2$
d	down	$-1/3$	$0.003 - 0.009$	$-1/2$
c	charm	$+2/3$	$1.15 - 1.35$	$+1/2$
s	strange	$-1/3$	$0.075 - 0.170$	$-1/2$
t	top	$+2/3$	$169 - 179$	$+1/2$
b	bottom	$-1/3$	$4.0 - 4.4$	$-1/2$

Table 1.1: The properties of quarks [1].

Our everyday world is made up of just three of these building blocks: the up quark, the down quark, and the electron. This set of particles is all that is necessary to form protons and neutrons. Together with electrons, protons and neutrons combine to form atoms which group together to form molecules and so on.

There are a total of six "flavors" of quarks and six leptons which are grouped into three "generations" of pairs. The three pairs form doublets of weak charge. Although individual quarks have fractional charge, they combine such that hadrons always have a net integer charge. Free quarks have never been observed because quarks are confined to hadrons, which are either mesons (a bound state containing a quark and an anti-quark) or baryons (containing three quarks or three antiquarks). This confinement is due to the fact that quarks, unlike leptons, carry an additional kind of charge known as color and follow the rules of quantum chromodynamics (QCD). The proton (uud) and the neutron (udd) are examples of baryons. There are three color charges and corresponding anti-color charges called red, green, and blue. The existence of a color quantum number is required to explain the existence of observed hadrons that combine three identical fermions in a completely symmetric ground state. Were it not for color, this situation would violate the Pauli Exclusion

Lepton	Charge	Mass (MeV/c^2)	Weak Isospin
ν_e	0	< 0.003	$+1/2$
e	-1	0.511	$-1/2$
ν_μ	0	< 0.19	$+1/2$
μ	-1	106	$-1/2$
ν_τ	0	< 18.2	$+1/2$
τ	-1	1777	$-1/2$

Table 1.2: The properties of leptons [1].

Principle. Mesons consist of a color-anticolor pair, and baryons consist of three quarks, each with a different color quantum number, so that observed baryons and mesons are always "colorless". Of specific interest to this analysis is the a b quark with a lighter anti-quark: B^- ($b\overline{u}$), B^+ ($\overline{b}u$), B^0 ($\overline{b}d$), and $\overline{B}^0$ ($b\overline{d}$). Similarly, K mesons, or kaons, contain a strange quark and a light anti-quark. K mesons whose quark spins are in the triplet configuration with total angular momentum $J = 1$ are called K^* mesons. Pions (π) are composed of the two lightest quarks: π^+ ($u\overline{d}$), π^- ($d\overline{u}$), and π^0 ($u\overline{u}$, $d\overline{d}$).

There are six leptons (and six anti-leptons), three of which have a non-zero electric charge. They appear to be point-like particles without any internal structure. The electron, the muon (μ) and the tau (τ) have electric charge and no color charge. The other three leptons are neutrinos (ν), which have no electric charge and carry only weak charge. They only participate in the weak interaction, so they rarely interact with matter and are difficult to detect.

1.2.2 The Forces

Now that we know the fundamental building blocks of nature, we can turn our attention to how these fundamental particles interact. Interactions are mediated by integer-spin particles called gauge bosons, which are listed in Table 1.3

It is an impressive demonstration of the unifying power of physics to realize that all the phenomena observed in the natural world can be attributed to the effects of just four fundamental forces: gravity, electromagnetism, weak force and strong force. Gravity and electrogmagnetism are the only forces that have significant effects at macroscopic distances. The effects of the weak and strong forces are confined to within 10^{-15} m of their sources. When two particles interact, they are said to exchange bosons. A brief description of the fundamental forces is given below.

Interaction	Gravity	Weak	Electromagnetic	Strong
Mediator	graviton	$W^{\pm}$, Z^0	γ (photon)	g (gluon)
Range	∞	$< 10^{-18}$	∞	$\sim 10^{-15}$
Mass (GeV/c^2)	0	80.4 ($W^{\pm}$), 91.2 (Z^0)	0	0
Charge	0	0 (Z^0), 1 ($W^{\pm}$)	0	0
Spin	2	1	1	1
Acts on	All particles	Leptons, Hadrons	Charged particles	Hadrons

Table 1.3: The properties of gauge bosons [1].

- **Gravity**
  Although its interaction is of infinite range, gravity is very weak. The gravitational force is described classically by Newton's law of universal gravitation, and its relativistic generalization is Einstein's theory of relativity. However, both of these are still classical theories and no satisfactory quantum theory of gravity has yet been formulated. A description of gravity is not included in the SM. In addition, the graviton, the gauge boson for gravity, has not been found [8]. Fortunately, the effects of gravity are negligible in most particle physics processes when compared to the other three interactions.

- **Electromagnetic force**
  The electromagnetic interaction is governed by the theory of quantum electrodynamics (QED). It is mediated by the photon (γ) which couples to all electrically charged particels. It is this force that allows atoms to bind together to form larger structures.

- **Strong force**
  The strong force is responsible for holding quarks together to form hadrons. It is governed by the theory of quantum chromodynamics (QCD) and interacts only with particles that carry color charge. Quarks have color charge while leptons do not, so only quarks can interact via the strong force. This force is responsible for binding the protons and neutrons in the nucleus together and for confining quarks inside hadrons. Quark color is the source of fields called color fields, just as the electric charge is the source of electromagnetic fields. The quantum excitations of these color fields are called "gluons", and the strong force is mediated by the massless gluon. A unique feature of the strong force is that the gluon itself carries a color charge, implying that gluons can interact with themselves. Because of this direct coupling, the color field between quarks and antiquarks does not spread out as the distance between

them is increased. Free quarks and color fields have not been observed because they are eternally confined to reside in "colorless" states. While the strong force is very strong for quarks that are far apart, it is much weaker for quarks that are very close together, so even though the quarks inside a hadron are tightly bound together, they are still able to wiggle around inside. As a quark is pulled out of a hadron, the force between the quark and other quarks in the nucleon is so strong that there is enough energy to form a new $q\overline{q}$ pair, thus maintaining the colorless state.

- **Weak force**
  Weak interactions are responsible for the decay of massive quarks and leptons into lighter quarks and leptons. When a quark or lepton changes type, it is said to change flavor. All flavor changes are due to the weak interaction. The weak force is responsible for β-decay, the $d \rightarrow u$ transition which causes the neutron (composed of three quarks, udd) to disintegrate into a proton (uud), an electron, and a neutrino. The weak interaction is mediated by the massive vector bosons, the $W^{\pm}$ and Z^0, and is of very short range.

Efforts to unify the last three of these interactions into one global description have resulted in the Standard Model. Experiments have verified its predictions to good precision and as yet, there is no piece of experimental evidence to contradict it. However, even though the Standard Model has resisted all attacks by experimental data, there are many unanswered questions about the physics territory beyond its frontiers. There are many input parameters to the SM such as the fermion masses and the coupling constants that cannot be derived from first principles and can only be determined experimentally. In addition, there is no way as yet to combine a theory of gravity with quantum mechanics and integrate it into the Standard Model. The focus of particle physics is the experimental verification or the refutation of predictions that are based on the Standard Model and possible extensions to it.

1.2.3   Theory of Weak Interactions

The weak interaction couples to particles that carry "weak charge", called weak isospin. All quarks and leptons couple to the weak force through allowed weak vertices pictured in Figure 1.1. The charge conjugate processes are also allowed.

A b ($\overline{b}$) quark with charge $-1/3$ ($+1/3$) can convert into a c ($\overline{c}$) quark with charge $+2/3$ ($-2/3$) through the emission of a W^- (W^+). Although the outgoing quark carries the same color as the incoming quark, the quark flavor has changed. Quark flavor is not conserved in weak interactions!

Furthermore, only the weak interaction violates "mirror symmetry" or parity invariance. The parity operation is equivalent to an inversion of space coordinates. Prior to 1956, physicists believed that the laws of physics were the same for a process
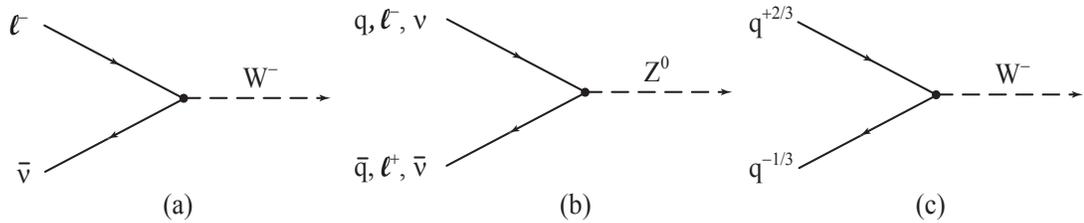
Figure 1.1: Diagrams for allowed weak vertices. Diagram (a) is an example of a charged vertex for leptons; (b) is the neutral vertex for leptons, and (c) is the charged vertex for quarks. There is no flavor changing neutral vertex for quarks! The charge conjugates of these reactions are also allowed, although they are not explicitly pictured.

and its mirror image. In 1956, Lee and Yang pointed out that no experiment had actually been done to test this principle in *weak* decays (although the strong and electromagnetic processes had provided plenty of evidence that parity was conserved in those interactions) [9]. Later that year, C.S. Wu and her collaborators performed an experiment designed to observe a spatial asymmetry in the β-decay of the Cobalt 60 atom [10]. In this experiment, a collection of Co^{60} atoms were aligned so that their spins pointed in one direction. The directions of the emitted electrons from the β-decay were recorded. The experiment found that the electrons were almost always emitted in the direction of the nuclear spin, violating parity. In the case of the weak interaction, the process and its mirror image are distinguishable, and the weak force responsible for β-decay can tell its right hand from its left. The "handedness", or helicity, of a particle refers to a particle's spin component in the direction of motion. A particle of spin $\frac{1}{2}$ can have a helicity of $+1$, "right-handed", or -1, "left-handed".

Weak decays are inherently left-handed and can only couple to particles that have spins pointing opposite to their direction of motion. Over the course of several experiments, it was determined that all neutrinos are left-handed and all antineutrinos are right-handed. In general, leptons can interact via the electromagnetic and the weak force. Neutrinos, having no charge, can only interact through the weak interaction. In fact, any massless fermion that engages in weak interactions must be left-handed due to the structure of the weak interaction current. This has implications for the way we group leptons and quarks.

The weak interaction prefers its leptons to be left-handed, such that if a lepton, such as an electron, were actually massless it would act only on left-handed electrons. (For a massive particle, we can always move to a frame of reference in which the

direction of motion changes, thus changing the helicity of the particle.) Therefore, it makes sense to split the lepton wave function into separate right- and left-handed components and group them separately. See Table 1.4 for the new grouping of leptons.

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$$

$$e_R \qquad \mu_R \qquad \tau_R$$

Table 1.4: Three Generations of Leptons. The absence of right-handed neutrinos assumes that the neutrinos are massless.

Returning to the discussion of quarks, we now have some basis for understanding why the weak interaction does not conserve flavor. The six quarks are eigenstates of the strong force:

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

However, the eigenstates of the weak interaction are not the same as those of the strong interaction. For the purposes of the weak force, the quark generations are skewed:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L$$

$$u_R, \ d_R \qquad c_R, \ s_R \qquad t_R, \ b_R$$

The weak interaction eigenstates d', s', and b' are linear combinations of the strong interaction eigenstates (or flavor eigenstates) d, s, and b. By convention, the u, c, and t quarks are unmixed and the unprimed letters represent the strong rather than the weak eigenstates. The mixing between the weak and strong eigenstates is expressed with the Cabibbo-Kobayashi-Maskawa matrix [11],[12]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L .$$

The CKM matrix can be thought of as a rotation from the quark mass (strong) eigenstates, d, s, and b, to a set of new states, d', s', and b', with diagonal couplings to u, c, and t. If the matrix were diagonal, then there would be no cross-generational decay.

The CKM matrix is one of the foundations of the Standard Model. It is a unitary matrix that can be described completely by three Euler-type angles and a complex phase [13]. The unitarity of the CKM matrix implies several relations among its elements that are important to the decay $b \to s\gamma$ as we will show in Section 2.1. The constraints of unitarity connect the individual elements of the matrix, so picking a particular value for one element restricts the range of others. The values of the CKM matrix elements, like fermion masses, cannot be derived from any other quantities but are fundamental input parameters of the Standard Model and cannot be predicted. The elements of this matrix can only be determined experimentally, and the magnitudes have been measured to various degrees of accuracy as shown [1]:

$$
\begin{pmatrix}
0.9742 \text{ to } 0.9757 & 0.219 \text{ to } 0.226 & 0.002 \text{ to } 0.005 \\
0.219 \text{ to } 0.225 & 0.9734 \text{ to } 0.9749 & 0.037 \text{ to } 0.043 \\
0.004 \text{ to } 0.014 & 0.035 \text{ to } 0.043 & 0.9990 \text{ to } 0.9993
\end{pmatrix}.
$$

The magnitude of a CKM element is determined from experimental information on the corresponding quark flavor transition. A matrix element like V_{tb} indicates the strength of the coupling between the two quarks in question, in this case the t quark and the b quark. The diagonal elements of the CKM matrix are clearly dominant (near 1). The large value of V_{tb} reflects the experimental fact that top quark preferentially decays into the bottom quark. Likewise, the value of V_{cs} indicates that charm quarks tend to decay to strange quarks. As you go farther off diagonal, the values get smaller. The elements furthest off the diagonal are the smallest. A perusal of the CKM matrix shows that the strength of the decay amplitude is strongest within each generation. If kinematically allowed to decay, a quark will decay preferentially within its own generation. It is possible, however, for a quark to jump to the next generation, as in the case of the $b \to c$ decay. Sometimes, even less frequently than in the previous case, a quark can jump across two generations as in the case of the $b \to u$ decay. The b quark and particles that contain it, like the B meson, cannot decay to the quark within their generation because it is not kinematically allowed. (The t quark is heavier than the b.) They can only decay to lighter quarks that are not in their generation. This makes the decay less likely to happen, so these particles tend to live a long time. More details on electroweak interactions can be found in many books and papers [14],[11],[12],[8],[15].

1.2.4 FCNC in Rare B Decays

Notice that in the CKM matrix, there is no element called V_{bs} that indicates a coupling between the b quark and the s quark, yet this thesis is about the decay $b \to s\gamma$, a $b \to s$ transition! "Flavor-changing neutral current" (FCNC) processes, are processes in which a quark changes flavor without changing its charge, for example $b \to s$, $s \to d$. Naïvely, one would expect such processes to proceed at tree level,

mediated by the Z^0 boson. For example, the decay $b \rightarrow se^+e^-$ should proceed via the tree diagrams shown in Figure 1.2.
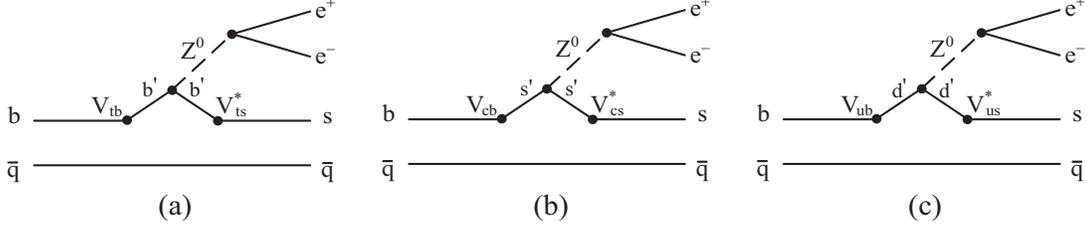
Figure 1.2: Diagrams for hypothetical $b \rightarrow se^+e^-$ decay.

Except for the CKM matrix elements, these three diagrams are identical. We can write the amplitude for (a) as $V_{tb}V_{ts}^*\mathcal{T}$. Likewise, (b) can be written as $V_{cb}V_{cs}^*\mathcal{T}$ and (c) as $V_{ub}V_{us}^*\mathcal{T}$, where $\mathcal{T}$ is the common piece and is the same for all three. Amplitudes add, and the sum of the three pieces is $(V_{tb}V_{ts}^* + V_{cb}V_{cs}^* + V_{ub}V_{us}^*)\mathcal{T}$. But because the CKM matrix is unitary, the term in parentheses vanishes. This demonstrates one aspect of the GIM mechanism — flavor changing neutral current processes *vanish at tree level* [16].

Even though $b \rightarrow s$ decays are not allowed at tree level, it is possible for this transition to occur through other means: a b quark can emit and then reabsorb a W thus changing its flavor twice in a loop process called a penguin decay (see Figure 1.3). In these loop processes, effective FCNC are possible, but they happen much more rarely than the tree level $b \rightarrow c$ decays. In $b \rightarrow s\gamma$, a photon is emitted from one of charged particles, either the quark lines or the W.

1.3   A Preview of Things to Come

The analysis we have done is a measurement of the rate and spectrum of the $b \rightarrow s\gamma$ decay. Since $b \rightarrow s\gamma$ is a rare process forbidden at tree level in the Standard Model and mediated by a loop diagram, it is a sensitive probe of new physics beyond the Standard Model. Experimentally, it has no competition from a tree level diagram, which makes it possible to measure the rate for this decay. With $b \rightarrow s\gamma$, we have the potential to measure a rare decay whose rate can place constraints on New Physics and whose spectrum can shed light on the fundamental QCD interactions that bind b quarks into B mesons. The rest of this thesis is organized in the following way:

- **Chapter 2:** Describes the theory of $b \rightarrow s\gamma$.
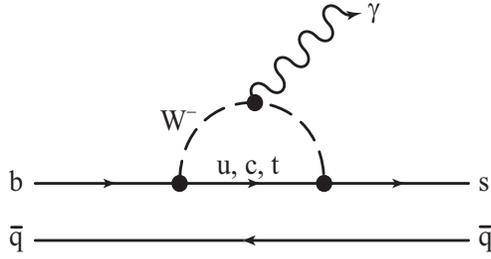
Figure 1.3: Diagram for the penguin decay $b \to s\gamma$.

- **Chapter 3:** Describes particle detection and identification using the CLEO II and CLEO II.V detectors.

- **Chapter 4:** Describes the major backgrounds and presents an overview of the strategy for eliminating them.

- **Chapter 5:** Describes how we distinguish between continuum and signal-like events to obtain a yield that includes signal events and $B\overline{B}$ backgrounds. Furthermore, it describes the B backgrounds and how we eliminate them to obtain the final $b \to s\gamma$ yield.

- **Chapter 6:** Describes how we calculate the efficiency for finding $b \to s\gamma$ events and how we evaluate the systematic errors.

- **Chapter 7:** We present the photon energy spectrum for $b \to s\gamma$, the branching fraction, and the first and second moments of the spectrum.

- **Chapter 8:** We interpret the results obtained in Chapter 7 and use the moments to extract HQET parameters.

CHAPTER 2

Theory of $b \to s\gamma$

2.1 Anatomy of Penguin Decays

The $b \to s\gamma$ decay is a flavor changing neutral current (FCNC) process that was demonstrated to vanish at tree level in Chapter 1. Here, we go beyond the tree level and consider one-loop diagrams, known as penguin diagrams.
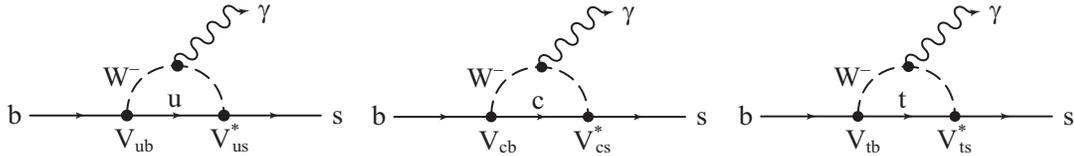
Figure 2.1: Diagrams contributing to the penguin decay $b \to s\gamma$.

The three diagrams in Figure 2.1 all contribute to $b \to s\gamma$. They differ among themselves in two ways: the CKM matrix elements and the quark inside the loop. The quarks inside the loop differ only in their mass, so we can write the amplitude of each diagram as $V_{ub}V_{us}^* P(m_u)$, $V_{cb}V_{cs}^* P(m_c)$, and $V_{tb}V_{ts}^* P(m_t)$, where $P(m)$ describes the loop common to the three, a function of the mass of the quark inside the loop. If the masses m_t, m_c, and m_u were all the same, then the factors $P(m)$ would all be the same and the sum of the three amplitudes would vanish by the unitarity of the CKM matrix, just as it did for the tree level diagrams in Figure 1.2. However, because the masses are *not* the same, the factors $P(m)$ differ, and the sum of the three amplitudes does not vanish.

Nevertheless, there is some cancelation, and the actual sum is indeed smaller than the sum of the magnitudes of the three terms. From the values of the CKM

matrix elements in Section 1.2.3 we see that $|V_{ub}V_{us}^*| \ll |V_{cb}V_{cs}^*| \approx |V_{tb}V_{ts}^*|$. Since the CKM matrix is unitary, $V_{cb}V_{cs}^* \approx -V_{tb}V_{ts}^*$, and $V_{ub}V_{us}^* \approx 0$, to a good approximation. The sum of the three amplitudes becomes $V_{tb}V_{ts}^*[P(m_t) - P(m_c)]$. The degree of suppression shrinks as the difference in masses increases. This demonstrates the other aspect of the GIM mechanism — flavor changing neutral current processes are suppressed for diagrams beyond tree level [16]. Because the top quark is so much heavier than the charm and up quarks, the GIM suppression for $b \to s\gamma$ is relatively weak.

2.2 Making a Distinction Between $b \to s\gamma$ and $B \to X_s\gamma$

There is a distinction between the parton-level process $b \to s\gamma$ and the hadron-level process, $B \to X_s\gamma$. Since b quarks are bound by imperfectly understood strong interactions into hadrons, effects due to beyond-SM physics must be untangled from the effects of these strongly bound states. It is often difficult to tell the difference between beyond-SM physics and non-perturbative QCD, neither of which are perfectly understood.

The decay $b \to s\gamma$ is essentially very simple: the quark-level process is a two-body decay that produces a trivial photon energy spectrum consisting of a mono-energetic photon, a discrete line, with a mean of $\sim m_b/2$. The strange system carries away the rest of the energy. However, this simple picture is not the whole story.

In reality, we are not measuring the parton-level decay $b \to s\gamma$. Rather, we are observing $B \to X_s\gamma$ for which the picture is less simplistic. Exclusive decay amplitudes, like $B \to K^*\gamma$, depend not only on the underlying weak transition between the b and s quarks, but also on the wavefunctions which describe how the B and K^* mesons are put together in terms of their quark and gluon pieces. When a b quark decays to an s quark, the s quark combines with all of the other pieces that make up the B meson. These pieces consist of the $\overline{u}$ or $\overline{d}$ spectator quark and assorted gluons, collectively referred to as the "brown muck". These pieces combine to form strongly bound hadrons, X_s [17].

If the b quark mass is known, then it is possible to calculate a theoretical prediction for the rate of $b \to s\gamma$ (although it is a *very* difficult calculation) [18]. If one has calculated the quark-level process $b \to s\gamma$ and if one knows the momentum distribution of the b quark within the B meson, one can compute the rate for $B \to X_s\gamma$. The physical spectrum (i.e. $B \to X_s\gamma$), which is what we measure, is obtained by folding the $b \to s\gamma$ spectrum with the b quark momentum distribution, the effect of which is to broaden, by Doppler shifting, the width of the photon energy spectrum. It is important to note that *the total rate is unchanged by this folding*.

Thus, the shape of the photon energy spectrum gives information about the b quark mass and about the b quark momentum distribution within the B meson. The mean of the photon energy distribution is given by the first moment of the spectrum,

$\langle E_\gamma \rangle \sim m_b/2$. The width of the photon "line" comes from the Doppler broadening and from gluon radiation $b \to sg\gamma$. Both of these features of the spectrum are very unlikely to be influenced by new physics. Instead, new physics appears in the rate.

The parton-level, or quark-level, photon energy spectrum is determined using next-to-leading order QCD. The function that describes the b quark momentum distribution is known as the shape function. It is convoluted with the parton-level spectrum to obtain the physical photon energy spectrum in the rest frame of the B meson. The dominant sources of error on the prediction for the photon energy distribution are the poorly known value of the mass of the b quark and the effects of the soft gluon radiation. To leading order, the shape function, which cannot be determined from first principles, is a universal characteristic of the B meson governing the inclusive decay spectra in processes that have (practically) massless quarks in the final state, like $B \to X_s\gamma$ and $B \to X_u l\nu$ [19].

Experimentally, the motion of the B meson in the lab frame and our detector resolution present an additional wrinkle to the measurement of the photon energy spectrum. Not only is the b quark moving inside the B meson, but the B itself is moving in the frame of reference of the lab so the photon energy spectrum is Doppler broadened by the motion of the B, an effect that we can reverse if we know the momentum of the B. In addition, our detector resolution further smears the final measured spectrum. If we are to make a precise measurement of the underlying shape function, we need to disentangle all of these effects.

2.3 OPE and HQET

One of the main difficulties in analyzing the rate for the inclusive $B \to X_s\gamma$ decay is taking into account the gluon "muck" that surrounds the quarks the B meson. Physics at different distances (energy scales) must be analyzed using different theoretical approaches. At short distances much smaller than $1/\Lambda_{QCD}$, the strong interaction can be described perturbatively by the exchange of individual gluons. However, at distances of order $1/\Lambda_{QCD}$, quarks and gluons hadronize and QCD becomes non-perturbative. OPE and HQET are the tools, methods, and theories that are used to calculate the rate for $B \to X_s\gamma$.

The *Operator Product Expansion* (OPE) is a tool that can be used to identify the physics that is present at a given scale and then separate it out explicitly. The OPE enables the formulation of an effective field theory for a process. Effective field theories are based on the idea that, in a given decay, only certain degrees of freedom are important for understanding the physics. It is possible to use the kinematic properties of a system (like the mass/momentum of a B meson or a b quark) to better define the region in which the theory is valid [17],[20].

The strong force has the property of "asymptotic freedom", and this running coupling constant implies that the strong force is very strong for quarks that are

far apart and much weaker for quarks that are close together [14]. This means that while a quark is strongly confined to a meson, it is still free to rattle around inside. The borderline between these regions of interaction strength is called the QCD scale, where $\Lambda_{QCD} \approx 200$ MeV. When defining an effective field theory, one can define a scale for a particular process. At tree level, for example, the scale is set by the mass of the virtual W boson. However, because of the loop, the decay $b \to s\gamma$ has at least two different mass scales, the W mass and the mass of the b quark, m_b. At high energy scales (short distances), $\mu \sim M_W \sim 80$ GeV, quark decays are governed by Feynman diagrams such as the one depicted in Figure 2.2(a). To obtain an effective low energy theory relevant for the scale $\mu \sim m_b \sim 5$ GeV, heavy degrees of freedom must be integrated out to obtain an effective coupling for point-like interactions of initial and final state particles, as shown in Figure 2.2(b) [21].

Using OPE, the decay amplitude $\mathcal{M}$ for the decay of a B meson to some final state f ($B \to f$) can be expressed as [21]

$$\mathcal{M} = -\frac{G_F}{\sqrt{2}} V_{CKM} \sum_i \mathcal{C}_i \langle f|\mathcal{O}_i|B\rangle \left[1 + \mathcal{O}\left(\frac{m_b^2}{M_W^2}\right)\right].$$

where $\mathcal{C}_i$ are the Wilson coefficients, the effective coupling constants of the theory, that contain the information on the short-distance physics defined at some scale μ. All dependence on heavy masses $M >> \mu$ such as m_b, M_W or *the masses of new undiscovered heavy particles* is contained in the $\mathcal{C}_i$. This is what is meant by the statement that loop diagrams are sensitive to new physics beyond the SM. The choice of the scale determines the division between the hard QCD interactions included in the Wilson coefficients and the soft-QCD effects included in the matrix elements $\langle f|\mathcal{O}_i|B\rangle$. This effectively separates the physics at different energy scales.
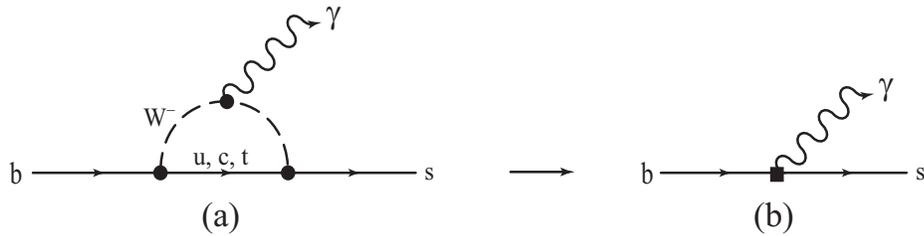
Figure 2.2: Diagrams for $b \to s\gamma$ representing (a) high energy quantum field theory (short-distance physics) and (b) low energy (long-distance) physics represented by point-like interactions.

Using OPE, we obtain a theoretical framework for the calculation of the branching fraction, set by the heavy quark expansion which predicts that, up to small bound-state corrections, the inclusive decay rate will agree with the parton-level rates for the decays of the b quark. Heavy Quark Effective Theory (HQET) is based on the principle that within a hadron, heavy quarks will not be affected by the light degrees of freedom. Within a B meson, the b quark is heavy when compared to Λ_{QCD}. The light degrees of freedom, consisting of quarks, antiquarks, and gluons, interact with the heavy quark through the exchange of gluons transferring momentum of order Λ_{QCD}. They are unable to resolve the features of the heavy quark. To these light quarks, the b quark is nothing more than a source of color and electric field. Perfect heavy quark symmetry would hold only for an infinitely heavy quark, however the b quark is not infinitely heavy. Therefore, we need to make corrections to this symmetry of order $\mathcal{O}(\Lambda_{QCD}/m_b)$. HQET provides a systematic way of calculating these corrections.

2.4 $b \to s\gamma$ Branching Fraction

The calculation of the rate for $b \to s\gamma$ is based on an operator product expansion with an effective Hamiltonian [19]

$$\mathcal{H}_{eff} = -\frac{G_F}{\sqrt{2}}V_{ts}^*V_{tb}\sum_{i=1}^{10}\mathcal{C}_i(\mu)\mathcal{O}_i.$$

where $\mathcal{O}_i$ are six-dimensional operators which govern the $b \to s\gamma$ transition, and V_{ts}, V_{tb} are CKM matrix elements. The $C_i(\mu)$ represent the corresponding Wilson coefficients and contain the relevant short-distance physics. The QCD corrections to the one-loop contribution to $\mathcal{B}(b \to s\gamma)$ increase the predicted rate by a factor of $2-3$ [18]. The dependence on the mass scale μ is a significant source of theoretical uncertainty.

For inclusive decays, the hadronic matrix elements of the local operators, $\mathcal{O}_i$ can be calculated using the heavy-quark expansion. If only leading-order expressions for the Wilson coefficients are used in the effective weak Hamiltonian, the branching ratio suffers from large perturbative uncertainties, therefore it is necessary to carry out the calculation to next-to-leading order. With this calculation, the theoretical uncertainty is reduced to about 10%. The theoretical prediction for the total branching fraction for $b \to s\gamma$ in the Standard Model is [18]

$$\mathcal{B}(B \to X_s\gamma) = (3.28 \pm 0.33) \times 10^{-4}$$

Recently, Gambino and Misiak [22] used a different choice for the charm quark mass. They change $m_c/m_b = 0.29 \pm 0.22$ to 0.22 ± 0.04 and obtain a different theoretical prediction for the branching fraction:

$$\mathcal{B}(B \to X_s \gamma) = (3.73 \pm 0.30) \times 10^{-4}$$

The interest in these decays comes from the possibility that, besides the W, other exotic particles could appear in the loop. There are a large number of references in the theoretical literature on the connections between beyond-SM physics and $b \to s\gamma$, including Higgs doublet models, SUSY models, Technicolor, extra dimensions, and models with composite W bosons. A measurement of $\mathcal{B}(b \to s\gamma)$ different from the SM value would indicate beyond-SM physics, and a measurement close to the SM value could constrain the parameters of these beyond-SM physics options [23].
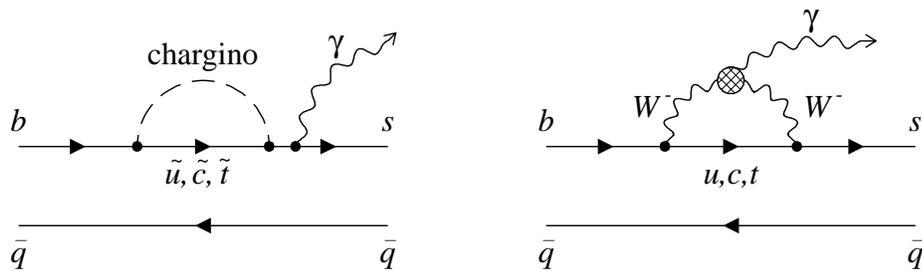
Figure 2.3: Possible diagrams contributing to beyond Standard Model $b \to s\gamma$ transitions.

2.5 $b \to s\gamma$ Photon Energy Spectrum

In contrast to the rate, the photon energy spectrum is insensitive to beyond-SM Physics. The mean energy of the photon, $\langle E_\gamma \rangle$, is to a good approximation, equal to half the b quark mass, m_b, while the mean square width of the energy distribution, $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2$, depends on the mean square momentum of the b quark within the B meson.

The beauty of $b \to s\gamma$ is that the branching fraction is sensitive to the new physics whereas the shape of the photon energy spectrum is determined by QCD dynamics and is insensitive to new physics beyond the SM. This shape function

represents the physics that *all B* decays to light quarks have in common, a feature that can be exploited to great effect in $b \rightarrow ul\nu$ decays.

The OPE and the parton model predictions are equivalent in the heavy quark limit, $m_b \rightarrow \infty$. Observables can be written as the parton model expectation plus non-perturbative corrections which cannot be calculated from first principles. In the HQET parameterization, it is possible to extract the parameters $\overline{\Lambda}$ and λ_1 from the moments of the spectrum. Physically, these parameters can be related to the mass of the b quark and the Fermi momentum of the b quark in the given scheme and can be used as an input to the theory to make higher order predictions. The shape function has a strong dependence on these input parameters, and studying the spectrum can give us insight into the actual shape of this function [19].

CHAPTER 3

CESR and CLEO

Our source of subatomic particles is the Cornell Electron-positron Storage Ring (CESR) which collides electrons (e^-) and positrons (e^+) together at a center of mass energy of about 10 GeV. When an electron and its antiparticle, the positron, collide, they annihilate to form a flash of energy which results in the creation of new matter. The CLEO detector identifies and measures the properties of the decay products of these collisions in an effort to better understand the fundamental constituents of matter and the laws governing their interactions. Below is a description of CESR and the two generations of the CLEO detector that were used to collect the data for this analysis.

3.1 Accelerator and Storage Ring

CESR is a symmetric e^+e^- collider 768 meters in circumference located about 12 meters below the the Cornell University campus in Ithaca, New York. This facility services two experiments, CLEO and the Cornell High Energy Synchrotron Source (CHESS), which uses the synchroton radiation emitted by the electron and positron beams to perform a variety of different experiments [24],[25],[26].

To create these collisions, electrons and positrons must first be created and accelerated to the required energy, and then forced to collide at an interaction point (IP) located inside the CLEO detector. The annihilation energy created when the e^- and e^+ collide, about 10.58 GeV, is just enough energy to produce a pair of B mesons.

Electrons are obtained by heating a filament. They are then accelerated to about 200 MeV by a microwave electromagnetic field in a 30 meter long vacuum pipe and injected into the synchrotron. Positrons are produced in the same linac by accelerating the electrons up to 150 MeV and allowing them to strike a tungsten target. The positrons are selected out of the resulting spray of particles, focused, accelerated in the remainder of the linac, and injected into the synchrotron [27].

The synchrotron takes the e^- (e^+) bunches from the linac and accelerates them to the energy at which they are to collide, about 5.29 GeV per beam. After only

19

4000 revolutions and 1/100 seconds in the synchrotron, the particles are up to energy and are transferred to the storage ring where the beams travel in counter-rotating bunches for almost two hours. As the particles circulate around the ring, they emit synchrotron radiation losing about 1.2 MeV per turn. This energy is replaced through several radio frequency cavities. Prior to March 1994, CESR was colliding the beams of seven equally spaced bunches head-on . After March 1994, CESR operated with trains of closely spaced bunches that collided with a small crossing angle. The ability to add more bunches to each train allowed an increase in luminosity, a measure of the rate with which e^+e^- collisions occur. Specifically, the luminosity is given by

$$\mathcal{L} = fn\frac{N_{e^+}N_{e^-}}{A},$$

where f is the revolution frequency, n is the number of bunches of each particle type, N_{e^+} and N_{e^-} are the number of electrons and positrons in each bunch, and A is the cross-sectional area of the beams [28]. The units for instantaneous luminosity are cm^{-2} s^{-1}, and the rate of producing events of a particular process is the product of the total cross section for the process and the instantaneous luminosity. The important quantity in particle physics is actually the integrated luminosity, often stated in pb^{-1}, where 1 barn $\equiv 10^{-28}$ m^2. Figure 3.1 shows a drawing of the storage ring.

3.2   The CLEO Detector

The purpose of this experiment is to see what comes out of the flash of energy resulting from particle-antiparticle annihilations. By measuring what kinds of particles are produced, how often certain particles are created, and how long they live, we can reconstruct some of the natural laws that govern the behavior of subatomic particles. To reconstruct all of the particles, charged and neutral, that come from a given e^+e^- interaction it is necessary to measure the properties of every particle that is produced with as much accuracy as possible. These properties are its energy and momentum at the vertex, its charge, its type, and its position as it moves through the detector. Since there are different particle types we are interested in and these types have different properties, we have to build a general purpose detector made up of several different components, each of which can measure some property of a particle with good accuracy.

The basic reactions which occur when subatomic particles encounter matter are the basis for all particle detectors. Depending on their charge, momentum, energy, and mass, particles lose different amounts of energy as they interact with the material in the detector. Penetrating subatomic particles see matter as an aggregate of electrons and nuclei, and reactions can occur with the atoms as a whole or with

CESR

SYNCHROTRON

WEST
TRANSFER
LINE

EAST
TRANSFER
LINE

LINAC   e⁻

e⁺

CHESS
WEST

CLEO II

CHESS
EAST

○ Positron Bunch - Clockwise

● Electron Bunch - Counter Clockwise
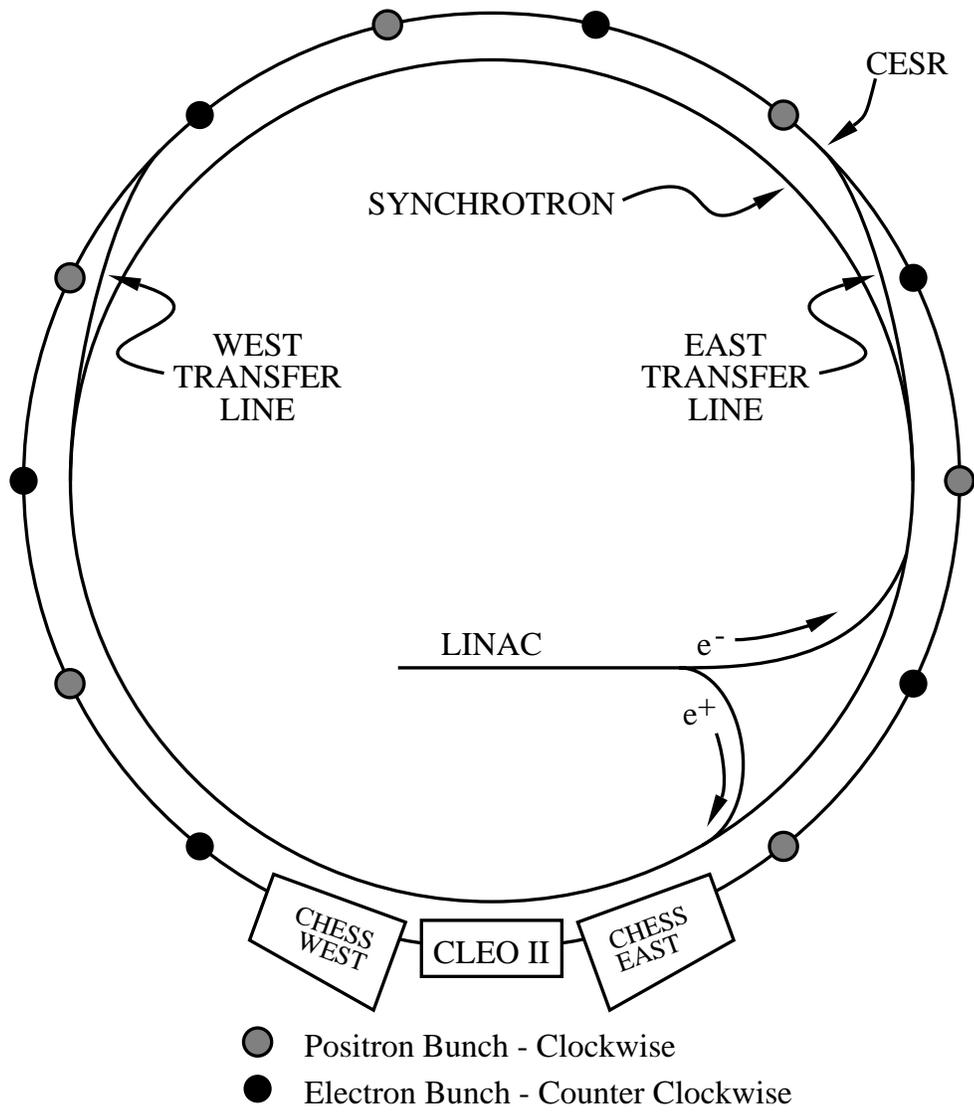
Figure 3.1: A schematic of the CESR storage ring.

their individual parts. An alpha particle entering a gold foil, for example, may scatter elastically from a nucleus via the strong force, collide electromagnetically with an atomic electron, or be absorbed in a nuclear reaction to produce other types of radiation. Similarly, we will be able to distinguish different particle species, by observing the traces they leave behind in the CLEO sub-detectors.

Between October of 1989 and February of 1998, the time period over which the data used in this analysis were collected, the CLEO detector existed in two different configurations, CLEO II and CLEO II.V. The difference between them is the removal of the central tracking chamber and the addition of a silicon vertex detector for better precision tracking in the inner detector. Particles created at the IP encounter the following detector components as they travel out through the CLEO detector from the e^+e^- collision point: the beam pipe, the tracking system, the time-of-flight counters, the electromagnetic crystal calorimeter, the 1.5 Tesla superconducting solenoid, and the muon identification system. A more complete description of both versions of the CLEO detector (CLEO II and CLEO II.V) can be found in references [29] and [30]. The detector components are described in the order in which the particles created by the e^+e^- interaction encounter them. Figure 3.2 shows the side view of the CLEO II.V detector.

3.2.1   Charged Particle Tracking - Inner Detector

About 2/3 of the particles produced in a collision are charged. We use the basic interactions of radiation with matter to "track" the points in space along the particle trajectory using a set of concentric drift chambers and a silicon detector.

A drift chamber consists of a gas filled volume strung with wires. These wires are arranged to form cells consisting of a central sense wire held at positive high voltage and surrounded by a cell "wall" comprised of wires at ground, forming a strong electric field within the drift cell. As a charged particle passes through the drift cell it ionizes some of the gas molecules in the cell volume. The electric field causes the released electrons to drift towards the sense wire. Near the sense wire, the electrons gain sufficient energy that they are able to ionize the chamber gas molecules that lie in their path, freeing more electrons in the process. These electrons can also ionize the chamber gas, causing an avalanche, thus amplifying the signal on the sense wire [31].

The readout electronics measure the height of the pulse and the time of its arrival. The timing information determines how far away from the anode wire the original ionization occurred. Using the wire positions and the drift time, obtained by measuring the time the electrons take to reach the sense wire, we map out a set of points through which the original particle passed. This set of positions is called a "track". The amount of charge collected on the sense wire is directly proportional to the energy loss of the particles. We obtain spatial information by measuring the

Helium Reservoir

Muon Chambers

Superconducting Coil
Barrel Shower Detector
Drift Chamber
SVX Detector
Micro-Beta Quadrupole
Vertex Detector
End Cap Time of Flight
Pole Tip Shower Detector
Time of Flight Scintillators

Magnet Yoke

raf

0   1   2   3   4

Meters
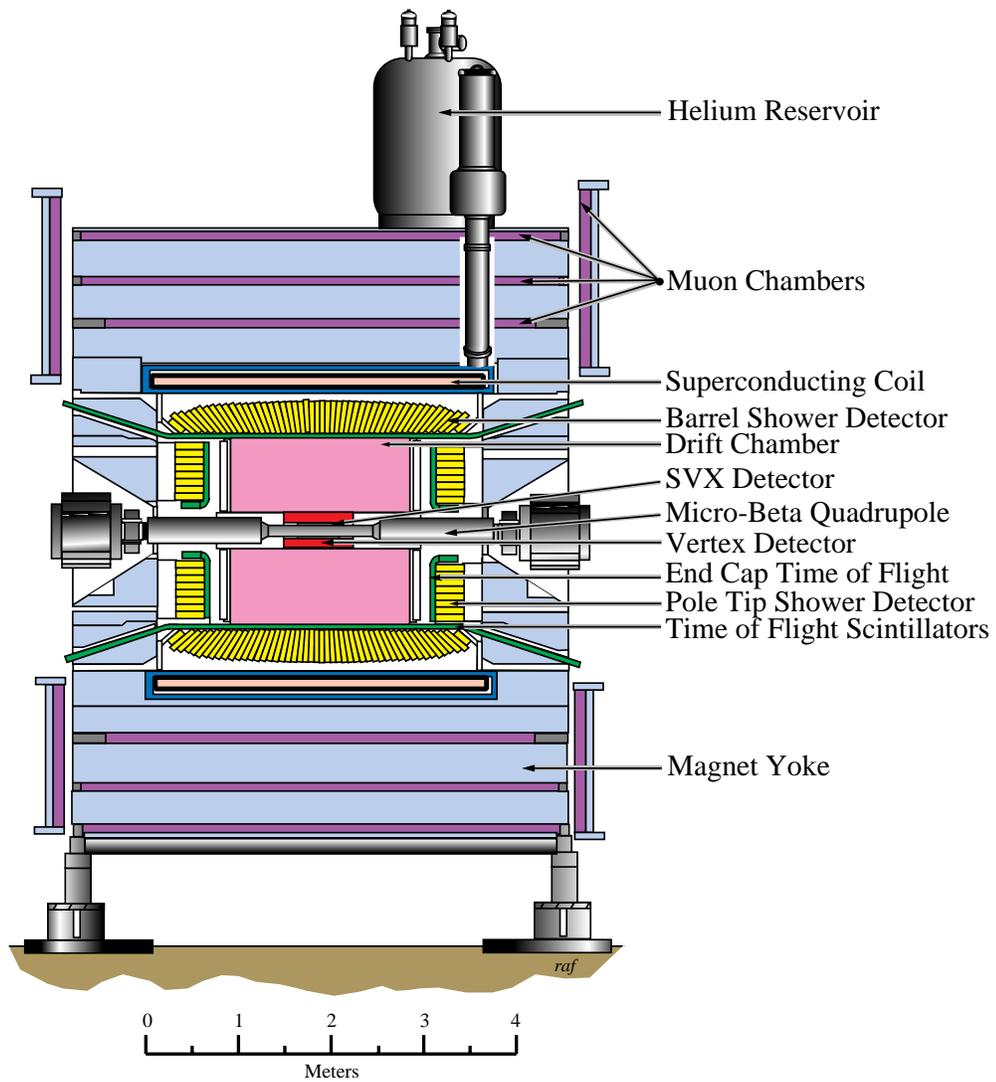
Figure 3.2: The side view of the CLEO II.V detector

drift time of the electrons coming from an ionizing event, and we also measure the specific energy loss, or energy loss per unit distance, (dE/dx) of the particle. Since the CLEO tracking chambers sit in a 1.5 Tesla magnetic field, a charged particle will move on a helical trajectory, and we can get full information about a particle's charge and momentum from the curvature of the particle's track.

During the first phase of CLEO II running, the charged particle tracking system was made up of three concentric, cylindrical tracking chambers aligned along the direction of the storage ring beams. Closest to the beam pipe was the Precision Tracking Layer (PTL) which extended from the beam pipe at 3.5 cm out to the Vertex Detector (VD). The Vertex Detector covers the radial region from 7.5 to 17.5 cm and is followed by the outer Drift Chamber (DR). In 1995, the PTL was exchanged for a Silicon Vertex Detector (SVX) after which the detector took on the name CLEO II.V. In the next section, we will describe the components of the CLEO detector.

PTL

The Precision Tracking Layers extend from the beam pipe at a radial distance of 3.5 cm out to the 7.5 cm. The PTL is a 6 layer straw tube drift chamber with 64 axial wires per layer, each layer offset from the adjacent layer by half a cell allowing us to determine whether the particle passed to the left of the central wire or to the right. The purpose of this detector is to measure the transverse particle direction near the interaction point. To define the boundaries of the cell and the electric field, this detector uses 384 tightly packed aluminized Mylar tubes held at ground. At the center of each tube is a sense wire made of gold plated tungsten and maintained at positive high voltage. A diagram of the PTL together with the VD is show in Figure 3.3. The single hit resolution of this detector is approximately 50 μm.

SVX

In the silicon vertex detector a particle traversing the active detector creates electron-hole pairs along its path. These charge carriers are then collected on fine sense strips in the silicon. The fine pitch strips allow the detector to be made smaller than traditional drift chambers and with an even finer resolution.

In April 1995, the Silicon Vertex Detector replaced the PTL as the innermost detector. The SVX is a device consisting of three layers of 300 μm double-sided silicon wafers, permitting the measurement of two coordinates of track position for each wafer with position resolution that is superior to that of the PTL.

The innermost layer of the SVX is located just outside the beam pipe at a radius of 2.35 cm. The second layer is at a radius of 3.25 cm, and the outermost layer is at 4.8 cm. The silicon vertex detector consists of 96 silicon wafers, read out on both sides. These wafers are arranged in eight octants of twelve wafers each. The

sensitive strips on the inside radius of the detector in each layer run parallel to the beam and measure the $r - \phi$ coordinates of particles passing through the detector. Strips on the outer radius of the detector run perpendicular to the direction of the beam and record $r - z$ information. Both electrons and holes are collected on sense strips. The detector consists of a total of 26,208 channels. The resolution for tracks at normal incidence is 12 μm on the $r - \phi$ side and about 30 μm for a z measurement.

Vertex Detector

The second of the three wire chambers was originally installed in CLEO in 1984. Extending radially from 8.4 cm to 16.0 cm, the vertex detector (VD) has 10 axial wire layers with a total of 800 sense wires and 2272 field wires arranged to form hexagonal cells providing radial particle track information. Figure 3.3 shows the cell structure of the vertex detector.

The primary purpose of this detector is to measure the transverse momentum for particles with a low p_T (transverse momentum). The drift distance sign ambiguity is resolved by staggering each layer by half a cell with respect to the previous layer. Longitudinal track information is obtained from cathode strips, segmented in ϕ and z, on both the inner and outer walls of the VD. In addition, the signal from the sense wires is detected and read out on both ends and turned into a z-measurement using the method of "charge division". The sense wires have enough resistance to create a measurable difference in current for different travel lengths along the wire. The difference in the charge is collected at the two ends, providing a z measurement. With a gas of 50% argon and 50% ethane, the single hit resolution is approximately 90μm in $r - \phi$ and about 750 μm in z.

Drift Chamber

The drift chamber (DR) is the outermost and the largest of the three CLEO wire chambers extending radially from 17.5 cm to 95 cm with a length of 2.15 m. There are a total of 12,240 sense wires and 36,240 field wires arranged in 51 layers of nearly square cells with dimensions approximately 14 mm $\times$ 14 mm. Mechanical support for the wires is provided by 3.175 cm thick aluminum endplates, material with 3% radiation length[1] that deteriorates neutral particle reconstruction in the endcap calorimeter.

There are 40 axial layers for precision measurements of dE/dx and the $x - y$ coordinates of charged particles. The DR also has 11 stereo layers at a small angle ($3° - 7°$) with respect to the z axis providing z measurements throughout the volume of the detector. The axial sense wires are staggered by a half cell from

---

[1]The mean distance over which a particles' energy is reduced by a factor e is called the radiation length. The radiation length is given either in g/cm^2 or in cm and is a useful measure of how much energy a given particle will deposit in another material as it passes through.
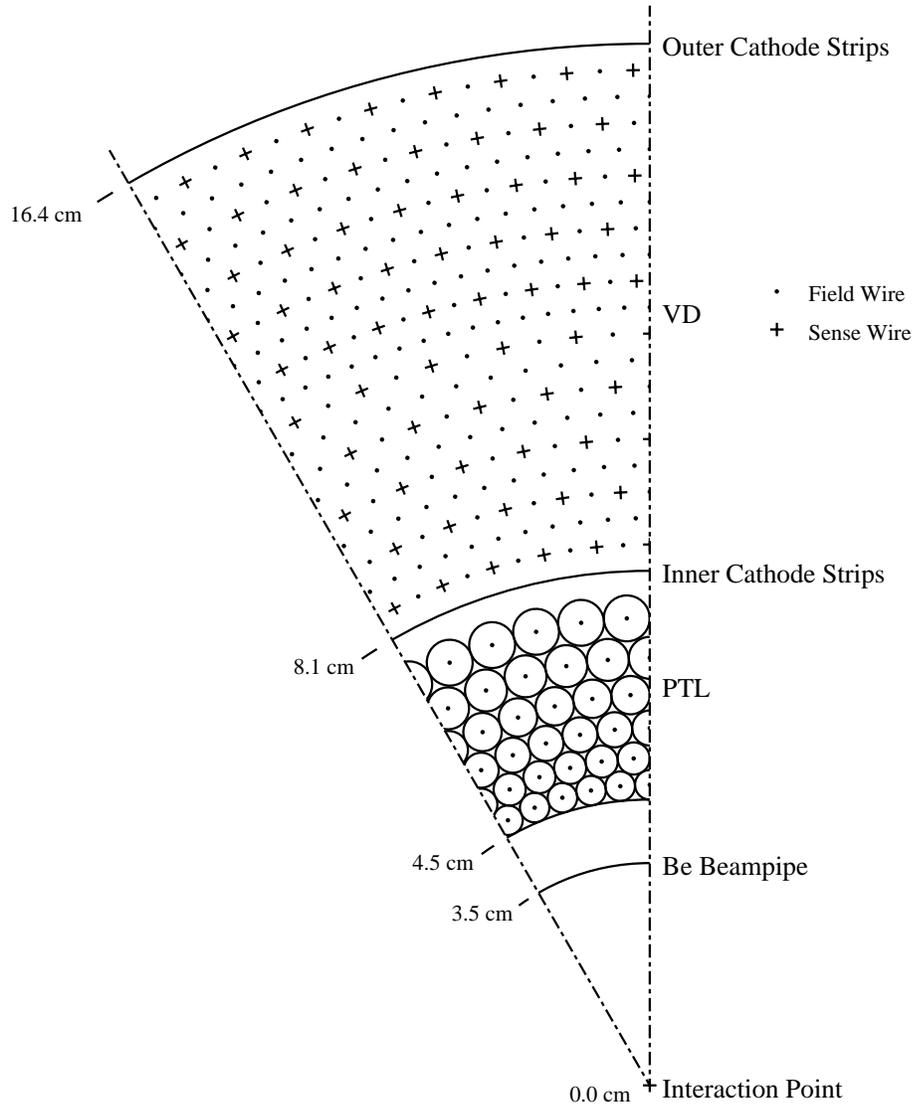
Figure 3.3: Cell structure of the CLEO II inner trackers.

16.4 cm

Outer Cathode Strips

· Field Wire
+ Sense Wire

VD

Inner Cathode Strips

8.1 cm

PTL

4.5 cm

Be Beampipe

3.5 cm

0.0 cm + Interaction Point

one layer to the next in order to facilitate local resolution of left-right ambiguity. The innermost and outermost layers are covered in 1 cm wide cathode strip pads segmented longitudinally allowing the DR to provide a precise measurement of the z coordinates at the beginning and end of most tracks.

Some particle identification is possible with the DR since the ionization energy loss (dE/dx) as particles traverse the chamber is recorded. To perform particle identification using dE/dx, we cut on the standard deviation, defined as the difference between the mean of the lowest half of the pulse heights and the expected pulse height for a given particle type divided by the expected resolution for that pulse height and number of hits. Figure 3.4 shows how effective the dE/dx measurements are for identifying particle species.
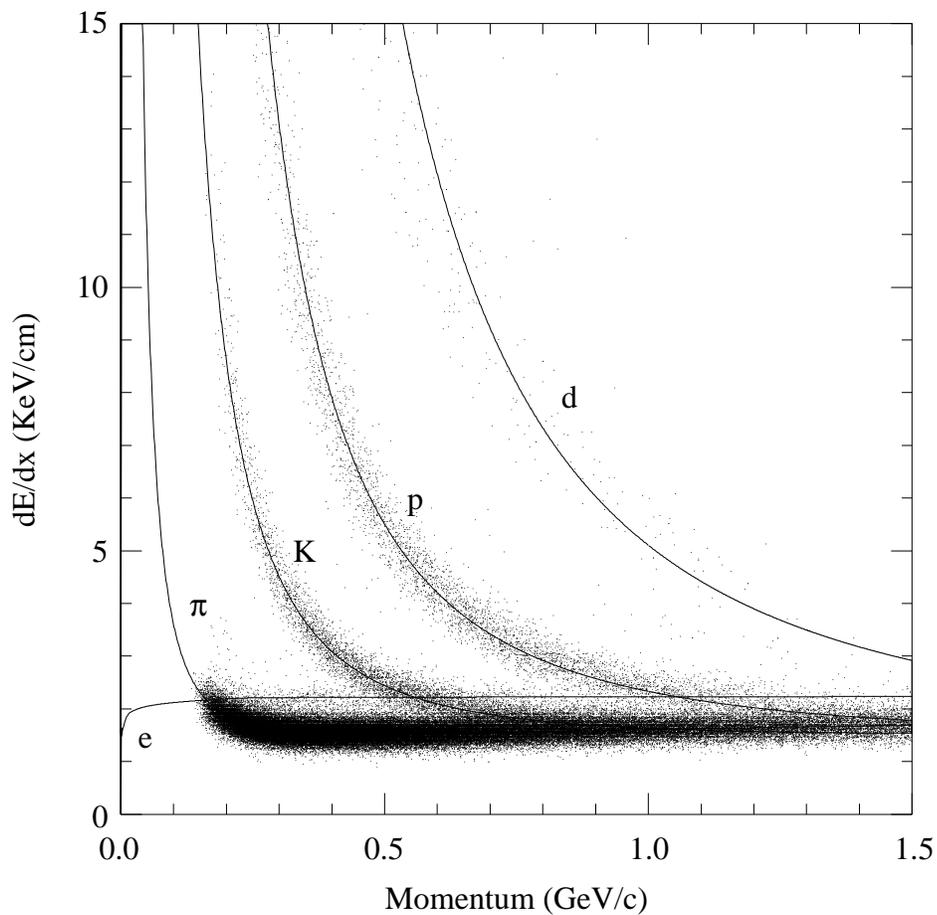
Figure 3.4: dE/dx vs. momentum of protons, π, and K.

3.2.2 Time of Flight

The Time of Flight system (TOF) provides information useful for particle identification and is the primary trigger for data recording. It measures the interval between the time of the e^+e^- collision and the time when the TOF system electronics detect the photons from the particles passing through the scintillator. Together with the momentum measurements from the curvature of charged tracks in the DR, the TOF measurements of the flight time allow us to constrain the mass of the particle.

The TOF consists of 64 fast (2.1 ns) scintillation counters about 5 cm thick $\times$ 10 cm wide $\times$ 279 cm long arranged parallel to the beam line around the outside of the DR. When a charged particle passes through the scintillator, it interacts with the heavy organic molecules embedded in it to produce a flash of light. Both ends of the scintillator are connected to photomultiplier tubes read out with a time-to-digital converter. The resolution of the system averaged over all particle types is about 160 ps.

In addition to the barrel time of flight counters, the endcap is also equipped with a TOF system, comprised of 28 wedge-shaped pieces 5 cm thick and mounted in a circle on the endcap calorimeters. They cover the radial range from 25.9 cm to 89.0 cm from the beamline, and have an intrinsic resolution of 170 ps averaged over all wedges. Together, the barrel and endcap TOF counters cover 96% of the solid angle and are capable of 2σ separation between pions and kaons at a momentum of 1.07 GeV/c. The standard deviation for a given particle type is defined as the difference in the measured and expected time of flight for that particle type divided by the estimated TOF resolution.

3.2.3 The Electromagnetic Calorimeter

While the job of the tracking system is to provide information about the momenta of charged particles, the energy and direction of neutral particles is measured using the electromagnetic calorimeter (although it can also measure the energy of some charged particles, such as electrons, aiding in their identification). Because the accurate measurement of high energy photons is a key factor in this analysis, electromagnetic calorimeter measurements are vital to a good measurement of $b \to s\gamma$.

Cesium iodide (CsI) is a scintillating crystal that produces light when particles pass through it. Through interactions with the crystal, a particle deposits some portion of its initial energy into the crystal creating scintillation light that is read out by four photo-diodes at the end of each crystal. Particles such as photons and electrons react electromagnetically with the CsI crystals and leave almost all of their energy inside the detector. Muons are heavier than electrons, so they are not stopped by bremsstrahlung processes as easily as electrons and do not deposit much of their

energy in the calorimeter. Hadrons, however, can also undergo strong interactions with the nuclei of the thallium-doped CsI. These interactions can produce secondary hadrons and cause non-localized showers. Because different particles interact differently with the matter in the calorimeter, the characteristics of the shower depend on the incident particle.

The CLEO calorimeter consists of 7800 thallium-doped cesium iodide (CsI) crystals distributed over the barrel ($\theta > 32°$) and the two endcaps ($5° < \theta < 36°$) covering 95% of the solid angle. The barrel calorimeter contains 6144 blocks arranged in 48 rows in the $z-$direction with 128 azimuthal segments in each. Photons that originate at the interaction point strike the crystal faces in the barrel at close to normal incidence. The remaining 1656 crystals are stacked inside two concentric ring-shaped aluminum holders that cover both ends of the detector. The dimensions of the crystals are approximately 5 cm $\times$ 5 cm $\times$ 30 cm long The 30 cm crystal length corresponds to 16 radiation lengths, good enough to contain the full energy of most photons and electrons produced in $\Upsilon(4S)$ processes.

Because material that blocks the CsI crystals degrades the energy resolution, the performance of the calorimeter depends on the amount of material between the interaction point and the crystals. The amount of material between the crystals and the interaction point varies with polar angle. The best resolution is achieved in the central barrel which covers about 71% of the solid angle. Here, the amount of material between the beam line and the crystals is about 18% of a radiation length whereas the amount of material between the interaction point and the endcap detectors is about 1 radiation length thanks to the drift chamber endplates along with the cables, readout boards, and support structures for the drift chamber and vertex detector. In the barrel, the photon energy resolution is 1.5% at 5 GeV, 3.8% at 100 MeV, and the angular resolution in azimuth is 3 mrad at 5 GeV, 11 mrad at 100 MeV. By contrast, the resolution in the endcap is much worse with a photon energy resolution of 2.6% at 5 GeV, 5.0% at 100 MeV and an angular resolution of 9 mrad at 5 GeV and 19 mrad at 100 MeV.

The good barrel (between 45° and 135°) offers the most precise measurement of energies, therefore we search for photons only in this region. In an attempt to squeeze more statistics out of our sample, we tried to include photons from the endcap region ($cos\theta < 0.85$), but we found that these events all had a very low weight and did not significantly increase our yield, so we chose to ignore the endcap region altogether.

3.2.4   The Magnet

The magnet provides a 1.5 Tesla axial magnetic field uniform to $\pm0.2\%$ over 95% of the solid angle in the drift chamber volume. The magnetic field is provided by a superconducting solenoid wound with two layers of 5×16 mm^2 aluminum stabilized

superconductor. It is 3.5 m in length and 2.9 m in diameter, large enough to contain the calorimeter and the tracking chambers inside the coil. The magnet return yoke is instrumented with the muon identification system, and the first two 36 cm thick layers are the main elements for the magnetic field flux return. The three layers of steel serve the dual purpose of being both the magnetic field flux return and providing interaction length for absorbing non-muon particles.

3.2.5 The Muon Chambers

Since muons are leptons, they do not interact strongly with the detector and lose only a moderate amount of energy due to ionization as they pass through the detector material. Muons do not lose much energy due to bremsstrahlung because they are much heavier than electrons. They are the only charged particles that are able to penetrate through the iron flux return yoke of the CLEO detector and are relatively easy to identify. A muon with a momentum of 2 GeV/c can reach the outermost layer of muon chambers. The muon chambers cover the polar angle range $30°$ to $150°$, or about 85% of the total solid angle.

The muon identification system consists of plastic streamer counters embedded in the iron return yoke for the magnet at depths of 36, 72, and 108 cm. Additional muon chambers are embedded in the endcaps of the detector. The thickness of the iron absorber varies from about 7.2 to 10 nuclear absorption lengths depending on the direction of the track.

The counters are approximately 5 m long and 8.3 cm wide. Each is separated lengthwise into 8 compartments, each with its own anode wire. This gives a resolution per counter of 2.4 cm. The exterior is coated with graphite to provide a cathode on three sides of each anode. One hit coordinate is obtained by reading out the anode signal, while the orthogonal coordinate is measured with external copper pickup strips that have the same width as the counters. Charge division readout is used to determine the two orthogonal space coordinates for each hit in the muon detector.

Muons are identified by extrapolating each reconstructed track from the tracking chambers into the muon detector. The path length to each muon layer is calculated in terms of nuclear absorption length. Each track is assigned the depth of the outermost unit in which it is detected. If the track's depth is less than predicted then the track is considered a non-muon.

3.2.6 Trigger and Data Acquisition

The e^- and e^+ beams at CESR are made to cross at the interaction point inside the CLEO detector. Although the beam crossings occur at rates of 2.7 MHz, the rate of production of $\Upsilon(4S)$ is only around $1-3$ per second. Other physics processes of interest such as tau, charm production, 2-photon, and QED processes contribute

at a higher rate so that the total physics event rate is around 10 Hz. The collisions occurring every 370 ns must be filtered by about a factor of a million. The filtering process is known as the trigger, a combination of hardware and software that recognizes certain event characteristics and determines whether the data comprising that event should be read out or not. The goal is to keep the trigger rate as low as possible to avoid being flooded with useless data, while at the same time keeping it high enough to capture all of the interesting physics processes.

The CLEO II and II.V trigger system is designed as a hardware system with three levels. As the data flows from the lowest level of the trigger to the highest, the decision process becomes more and more refined, and more time is available to make the decision. Each layer receives only those events that have passed the previously layer, and each uses increasingly more detailed information to make its decision.

The level zero trigger (L0) is fast and simple, reducing the 2.7 MHz crossing frequency to a rate on the order of 20 kHz, based on inputs from the electromagnetic calorimeter, the time-of-flight scintillators, and the vertex detector. If an event is accepted, it is passed along to the level one (L1) trigger, and the detector stops taking additional data until the higher-level triggers finish making the final decision. When an event is rejected or read out, the detector resumes collection of data.

The level one trigger (L1) accepts all of the events that have been passed along by the L0 trigger. It is a hardware based trigger that receives input from the ToF, the vertex detector, the calorimeter, and the DR. The goal of L1 is to reduce the input rate from L0 (which is about 20 kHz) by about 3 orders of magnitude to 25 Hz. If none of the L1 criteria have been met, the trigger logic is reset and the experiment is ready to read out again.

The Level two trigger (L2) accepts events that have passed the L1 criteria and examines information from the VD and the DR. If the L2 criteria are not met, the system is reset and readied for readout again. An event that is accepted by the L2 trigger is read out, limited only by the speed of the data acquisition system, approximately 50 Hz.

Should an event pass the L2 trigger, it will be passed along to a software level three (L3) trigger that is more sophisticated than the lower level triggers. At this stage, deadtime is no longer the driving concern, and some software tests of event quality can be done before the event is recorded for further analysis. L3 uses information from the entire detector to reject cosmic rays and interactions of the beam with the wall of the vacuum chamber or with residual gas molecules. Approximately 50% of events which pass level 2 are subsequently rejected by level 3.

The data acquisition system collects information read out by the detector once the trigger has decided that an interesting event has occurred. The data that is read out is in the form of analog signals which must be converted into digital signals before they are transferred to tape. The detector electronics is organized into an

Process	$\sigma(\text{tot})$ (nb)	$\sigma(\text{barrel})$ (nb)	$\sigma(\text{endcap})$ (nb)
$e^+e^- \to e^+e^-$	72	19	53
$e^+e^- \to \gamma\gamma$	6.2	3.7	2.5
$e^+e^- \to \mu^+\mu^-$	0.92	0.77	0.15
$e^+e^- \to \tau^+\tau^-$	0.92	0.77	0.15
$e^+e^- \to q\bar{q}$	3.5		
$e^+e^- \to B\bar{B}$	1.0		
$e^+e^- \to e^+e^-X$	2−25		
Total	∼100		

Table 3.1: Sample cross-sections for physics events at $E_{cm} = 10.58$ GeV. Assuming a luminosity of 6×10^{32} cm^{-2}s^{-1}, the expected rate for hadronic events is about 0.6 events per second [1].

array of data crates, each of which has an on-board controller containing a fast ADC and local memory for storage of pedestals and time constants.

When the trigger system detects an event, an interrupt signal is sent to the on-line computer and the crate controllers. This signal causes the analog crates to begin digitizing their data. The controllers then determine which channels satisfy the pedestal and timing constant requirements and store all valid information in data buffers for readout by the computer system. The data is then sparsified, packaged into an event and written to tape for later analysis. Some typical cross-sections for physics events at $E_{cm} = 10.58$ GeV are given in Table 3.1

3.3   Data Set

CESR runs at a center-of-mass energy of 10.58 GeV at the energy of the $\Upsilon(4S)$ resonance, just above the production threshold for $B\overline{B}$ pairs ($e^+e^- \to \Upsilon(4S) \to B\overline{B}$). It takes about 5.279 GeV to produce a single B meson, so there is no energy left over to produce anything else in addition to the meson pair in the $\Upsilon(4S)$ decays. The small amount of excess energy results in a small B momentum of about 320 MeV/c. With a lifetime of about 1.6 ps, the B meson travels only about 30μm before it decays.

The Υ resonances sit on top of an appreciable hadronic background, called "continuum". The basic physics process is the annihilation of the e^+e^- pair into a virtual photon or Z^0, which ultimately decays to a light $q\bar{q}$ pair ($u\overline{u}, d\overline{d}, c\overline{c}, s\overline{s}$).

Data is collected at two different energies, on the $\Upsilon(4S)$ resonance (On data) and off the $\Upsilon(4S)$ resonance (Off data). CESR operates at the $\Upsilon(4S)$ resonance about two thirds of the time and takes data with center-of-mass energies about 60 MeV below the $\Upsilon(4S)$ during the remaining third. In this way, CLEO collects a large sample of continuum only events in addition to the On resonance sample. Off resonance continuum data are used to understand and remove the continuum background from B decay studies.

On the $\Upsilon(4S)$ resonance, about 25% of the events are $B\overline{B}$ and about 75% are continuum. Figure 3.5 shows the 4 lowest Υ resonances. For this analysis we are running on or near the $\Upsilon(4S)$. For the remainder of this analysis, we refer to a sample of data as On resonance if its run energy is between 5.280 GeV and 5.300 GeV. If the run energy is between 5.187 GeV and 5.280 GeV or if the run energy is greater than or equal to 5.315 GeV, then we consider it Off resonance.
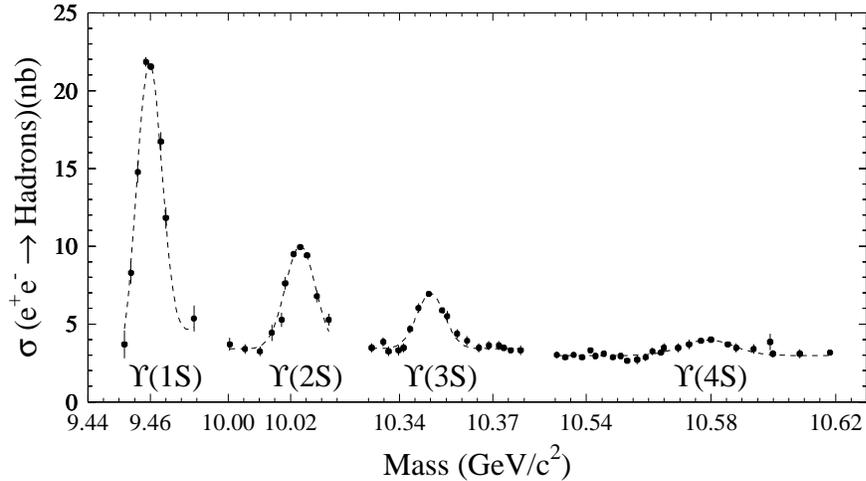
Figure 3.5: Hadronic cross section as a function of center of mass energy [5],[6],[7].

We use the full dataset collected by CLEO II and CLEO II.V. The dataset consists of a total of 9.2 fb^{-1} on-resonance and 4.6 fb^{-1} off-resonance corresponding to 9.7×10^6 $B\overline{B}$ pairs.

3.4 Monte Carlo at CLEO

In order to do an unbiased analysis, we do not determine event-selection criteria using the data sample that is used to measure the signal. It is easy to exploit statistical fluctuations in the data sample, biasing the signal upwards. Event simulations called Monte Carlo (MC) are a useful tool for evaluating selection procedures and detection efficiencies. Detailed MC simulations of the detector response to signal and background processes are used to better understand the kinematics of these decays and ultimately improve the accuracy of the measurement.

Monte Carlo events are generated in two distinct steps. The first step is to generate the desired decay process. At CLEO, we generate the particles and their 4-momenta using an event generator called QQ [32]. It uses a list of particles and their physical properties, such as mass, charge, spin, decay modes, etc., and initial conditions defined by the user to generate the final-state particles and any unstable daughters using user specified branching fractions.

The second step of Monte Carlo generation is to simulate the propagation of the decay products generated by QQ through the CLEO detector. We use the GEANT [33] based CLEOG program which simulates the CLEO detector's response to particles traversing its volume. CLEOG contains a full description of the detector including the amount and kinds of materials present and the responses of the detector elements to different particles. It takes particles from QQ and propagates them through the detector recording their trajectories and simulating the response of each detector's electronics. The output files are designed to look exactly like the raw data taken by the detector, so we can run the same analysis software on both data and MC.

CHAPTER 4

Analysis Overview

The first penguin process to be observed was the exclusive decay $B \to K^*(892)\gamma$ [34]. Although this result confirmed the existence of penguin decays, the large theoretical uncertainties in the hadronization process provided only a rough measure of the inclusive rate, the quantity of theoretical interest. CLEO's first measurement for the inclusive rate measured photons between $2.2-2.7$ GeV for 2 fb^{-1} On data and 1 fb^{-1} Off data. The result obtained was $\mathcal{B}(b \to s\gamma) = (2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$ [35]. Other experiments (BELLE [36] and ALEPH [37]) have also measured the inclusive rate for $b \to s\gamma$, but none have been able to extend the photon energy window below 2.2 GeV.

Our measurement determines $\mathcal{B}(b \to s\gamma)$ in the photon energy range $2.0 - 2.7$ GeV and uses a combination of shape analysis, pseudoreconstruction, and some new techniques for suppressing continuum . Previous analyses were limited by large backgrounds to the region of the photon energy spectrum above 2.2 GeV. However, our improved analysis technique allows us to open the photon energy window to $2.0 - 2.7$ GeV. This is the first time it is possible to extract the first and second moments of the photon energy spectrum.

4.1 Signal

The signal for $b \to s\gamma$ is a photon from a two-body decay. The B meson decays to a high energy photon with $E_\gamma \sim m_b/2 \approx 2.2$ GeV and an X_s system. The X_s system can be any kinematically allowed combination of mesons containing a strange quark, for example a K^*. Many decay channels contribute to the inclusive rate including exclusive decays like $B \to K^*\gamma$, where K^* represents one of several K^* mesons with different masses, or non-resonant decays such as $B \to K\pi\gamma$. Taking into account the momentum of the B mesons in the lab frame ($|\vec{p}_B| \approx 320$ MeV/c), simple kinematics gives an upper limit on the photon energy of 2.74 GeV. To measure the inclusive rate, we search for an excess of events with a high energy photon such that $2.0 < E_\gamma < 2.7$ GeV. Figure 6.2 shows that the photon energy distribution can have a long tail extending below 2.0 GeV to as low as 1.5 GeV. Although, we would

like to measure the spectrum over this full range $(1.5 - 2.7 \text{ GeV})$, we are limited by significant backgrounds which force us to apply a cut at $E_\gamma > 2.0$ GeV.

4.2 Background

To establish a signal for $b \to s\gamma$, it is important to understand the significant backgrounds from other processes. A background is any process other than the desired decay that exhibits the same signature as the signal and is an indistinguishable contribution to the measured experimental distribution. In the case of $b \to s\gamma$, a background is any process that can create a shower in the calorimeter with an energy between $2.0 - 2.7$ GeV. Here, we describe the sources and features of the three significant sources of background: initial state radiation (ISR), continuum $(e^+e^- \to q\overline{q}$, where $q\overline{q} = u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c})$, and $B\overline{B}$ processes. Figure 4.1 shows the sizes of the photon backgrounds due to ISR, π^0s produced in continuum processes, and $B\overline{B}$ backgrounds compared to the magnitude of the expected $b \to s\gamma$ signal.

Naïvely, we could imagine measuring the spectrum in the region between 2.3 and 2.7 GeV, since very few B decays other than $b \to s\gamma$ can produced photons in this energy range. However, a significant portion of the spectrum lies below 2.3 GeV. To extrapolate this measurement to the full phase space, a good understanding of the shape of the spectrum is necessary, so this incurs a large model dependence. Conversely, if we open up this energy window and look at the spectrum down to 1.8 GeV, the backgrounds due to $B\overline{B}$, namely $b \to c$ decays, dominate. The energy window for this analysis is $2.0 - 2.7$ GeV, a compromise between the model dependence incurred by making the window too narrow and the statistical error incurred from having to subtract large $B\overline{B}$ backgrounds.

4.2.1 Continuum and ISR

The dominant background to $b \to s\gamma$ comes from continuum processes. Since the data collected off the $\Upsilon(4S)$ resonance contain only continuum events, it is possible to estimate the size of the continuum background using the off-resonance data. The Off data produces a photon energy spectrum that represents the continuum background. This spectrum is subtracted from the corresponding distribution from the on-resonance data. There is one problem with this simple scenario: because the continuum background is so much larger than the signal contribution, the statistical error after a raw On-Off subtraction is larger than the expected signal yield!

In our case, we measure $\mathcal{B}(b \to s\gamma)$ and find there are S signal events and $B_{On} = B$ background events in the On data. We find a total of $S + B$ events in the On data with a statistical uncertainty of $\pm\sqrt{S+B}$. In the off-resonance data, we will measure B_{Off} events, where $B_{On} = B = \lambda B_{Off}$ and λ is the factor by which we scale the Off data to account for the luminosity and energy differences between
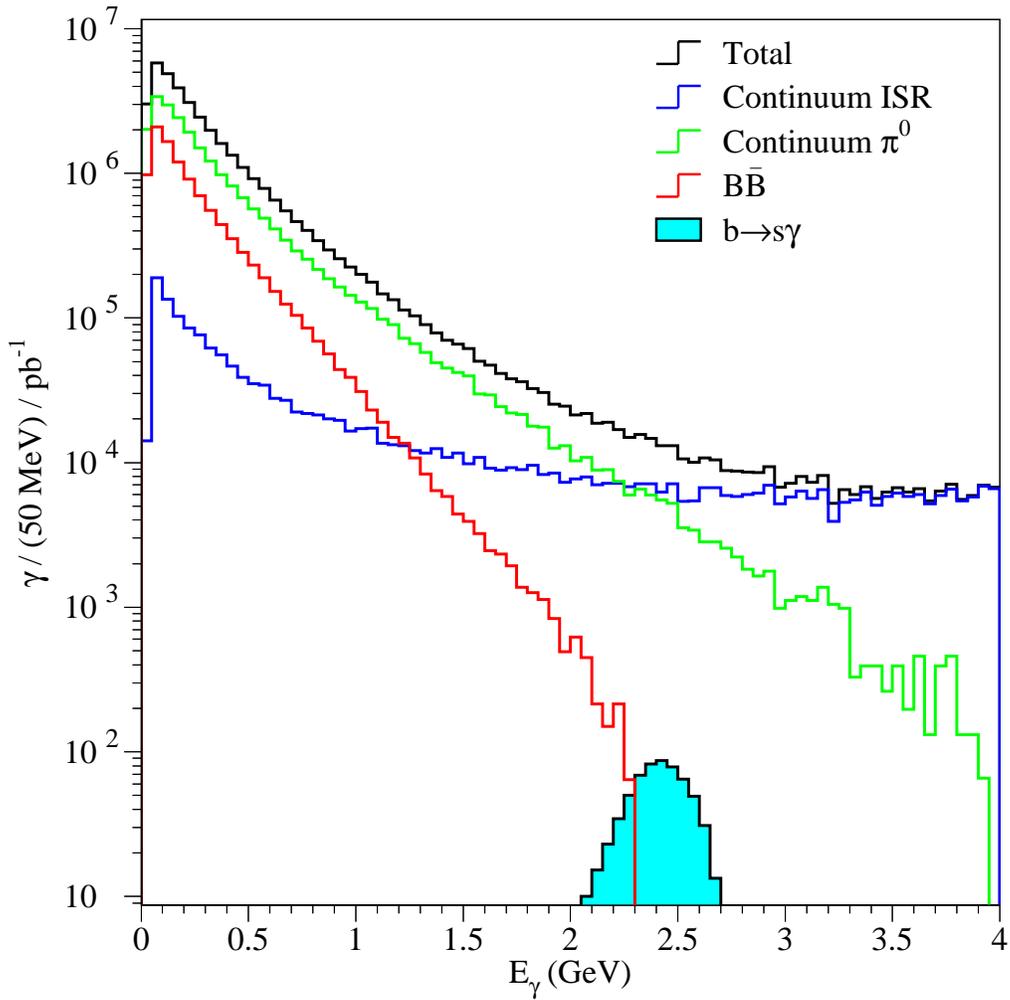
Figure 4.1: Sources of photon backgrounds include ISR, π^0s and $B\overline{B}$. These backgrounds dwarf the $b \to s\gamma$ signal. (This plot assumes that $\mathcal{B}(b \to s\gamma) = 3 \times 10^{-4}$.)

the On and Off datasets. We define the scale factor as

$$\lambda = \frac{\mathcal{L}(On)}{\mathcal{L}(Off)} \times \frac{E_{Off}^2}{E_{On}^2}.$$

The error on the measured number of Off events will be $\left(\lambda\sqrt{\frac{B}{\lambda}}\right)$. After On-Off subtraction, the overall error on the yield is

$$\Delta Y = \sqrt{S + (1 + \lambda)B}. \tag{4.1}$$

To get a statistically meaningful result[2], we will remove as much of this continuum background as possible before performing the On-Off subtraction. Once the continuum is suppressed, it is subtracted from the On-resonance spectrum. The key to this analysis is background suppression.

Reducing this continuum background is a challenge because it arises from two distinct sources. First, initial state radiation (ISR) is the process in which a photon is radiated from the e^- or e^+ before annihilation. Continuum describes events in which $e^+e^- \rightarrow q\overline{q}$ where $q\overline{q} = u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}$. In these decays, the high energy photon most often comes from the decays of π^0 or η mesons which typically decay to two photons. Figure 4.2 shows the Feynman diagrams for $q\overline{q}$ and ISR.
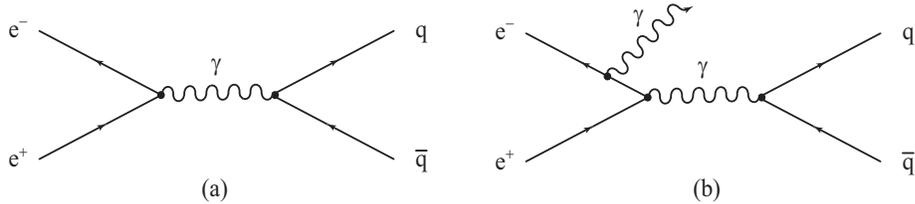
Figure 4.2: Feynman diagrams for (a) $q\overline{q}$ and (b) ISR.

Our suppression scheme will have to address two questions: how do we distinguish hard photons produced in continuum decays from signal photons and how do we eliminate them before doing the On-Off subtraction?

4.2.2 $B\overline{B}$ Backgrounds

After the continuum background is removed, the remaining backgrounds are due to $B\overline{B}$ processes. Contributions to this background come from $b \rightarrow c, b \rightarrow u$, and

---

[2]Assuming that $\mathcal{B}(b \rightarrow s\gamma) = 3.0 \times 10^{-4}$ and $\sigma_{q\overline{q}} = 3.3$ nb, $\lambda = 2.0$ and the On dataset has $\mathcal{L}_{On} = 9.165$ fb^{-1}, we expect $S = 5820$ signal events and $B = 30 \times 10^6$ background events. This gives an overall On-Off subtraction error of $\Delta Y \sim 9500$ events.

$b \to sg$ decays. Like the continuum, the majority of the background photons come from the daughters of π^0 and η decay that have slipped through π^0 and η vetoes.

These backgrounds are modeled by Monte Carlo and then subtracted. The MC sample was carefully tuned to match the data yields of π^0 and η as well as other $B\overline{B}$ decay modes that produce high energy photons. A large correction comes from including background caused by neutral hadrons, K_L^0 and $\overline{n}$. These backgrounds and the methods used to get rid of them are described in detail in Chapter 5. A significant part of the uncertainty in the result comes from the reliability of the Monte Carlo models, both for the signal and the background.

4.3   Analysis Overview

Because different events offer different types of information useful for separating continuum background from signal, we try to gather as much information about each event as possible and use all of it to judge the event as signal-like or continuum-like.

The signature for $b \to s\gamma$ is a photon between 2.0 and 2.7 GeV from the decay of a B meson. For all events with a good photon candidate, we distinguish between continuum and $B\overline{B}$ events by looking at the angular distribution of energy, or the "shape" of the event. Since B mesons are spin zero particles and are produced nearly at rest, they have very little momentum, and their decay products tend to be uniformly distributed throughout the detector giving a "spherical" event shape. The decay products of $q\overline{q}$ events, on the other hand, have a significant amount of initial momentum and decay in collimated "jets".

We can further eliminate continuum events if we determine that the photon came from a B meson decay rather than a continuum event. We try to reconstruct the $b \to s\gamma$ decay by requiring that the photon and the X_s system recoiling against it are kinematically consistent with the decay of a B meson. For each event, we try to combine the candidate photon with a kaon and $1-4$ pions, including at most one π^0. If we have more than one reconstruction, we choose the "best" combination in the event using a χ^2 based on kinematic constraints. Not all events can be reconstructed.

For all events that have been reconstructed, we search through the decay products of the non-signal B meson for a lepton. Since B mesons frequently decay to high momentum (> 1.5 GeV/c) leptons, finding a leftover lepton in the "other" B is a good indicator that this is, in fact, a $B\overline{B}$ event and not continuum. Even if the event cannot be reconstructed, it is still possible to look for a lepton by defining a likely X_s system using a high momentum π and K recoiling against the candidate photon and scanning the remaining particles for a lepton. Further information is obtained by observing whether the sign of the lepton charge agrees or disagrees with the flavor of the reconstructed B, although this test is only possible if the event is reconstructible.

Because not all events have all of this information, we separate events into four categories: non-reconstructible without a lepton, non-reconstructible with a lepton, reconstructible without a lepton, and reconstructible with a lepton.

In measuring our $b \rightarrow s\gamma$ signal, we are searching for photons produced by B decay with an energy between 1.5 and 2.7 GeV. Because the photon spectra of continuum processes dominate this region, we apply a cut at 2.0 GeV and use the continuum suppression schemes outlined above to distinguish hard photons from $B\overline{B}$ decay from those of ISR and $q\overline{q}$ processes in the region $2.0 - 2.7$ GeV. The selection criteria will therefore involve searching for hadronic events containing a high energy calorimeter shower consistent with a photon. We then identify the other tracks and showers in the event so that we may try to reconstruct the parent B and search for leptons.

4.4   Event Selection

The CLEO detector collects data from many physical processes including continuum hadron production, lepton pair production, two-photon and "junk" events like the interactions of the beam with the wall of the beam pipe or with molecules of gas in the beam pipe. There is a standard set of criteria for selecting hadronic events at CLEO [38]:

- Must contain at least three good quality charged tracks.

- The vertex of the event should be within 2 cm of the run average in x and y, and within 5 cm of the average in z. In other words, the event vertex should be near the expected collision point for the run.

- At least 20% of the total center of mass energy should be visible in the detector. "Visible energy" is defined as the sum of the energy of all charged tracks and the energy of neutral particles in the calorimeter.

- At least 65% of the center of mass energy must be visible in the calorimeter unless there are more than four charged tracks.

- If there are more than four charged tracks, then at least 15% of the center of mass energy must be visible in the calorimeter.

We employ these standard hadronic event criteria used in many CLEO analyses by requiring KLASGL = 10, a cut that has an efficiency of 99.8% for $B\overline{B}$ events [38].

40

4.5 Candidate Photon Selection

Each accepted event is searched for a high energy calorimeter shower in the barrel region with $|cos\theta| < 0.7$. This angular requirement is imposed not just because this region of the detector has the highest energy resolution (particles miss the endplate of the drift chamber and related support structures), but also because this cut reduces the background caused by photons from ISR, whose angular distribution goes as $1/sin^2\theta$.

Because charged particles also deposit energy in the calorimeter, we must determine whether the candidate shower is caused by a photon or some other particle, such as an electron, a hadron interacting strongly, or a muon ionizing in the crystals. There are five basic criteria that the candidate shower must meet before it is classified as a good photon.[3]

First, it is important that the electronics associated with any crystal in the shower be in good working order. Showers that contain crystals with noisy, dead, or shorted electronic channels are identified as noise showers and are rejected as potential candidates.

Second, we try to reject energy deposits in the calorimeter caused by charged particles. We accomplish this by attempting to match a charged track to its shower. A shower is called "matched" if a drift chamber track, projected to the calorimeter surface, is within 8 cm of a crystal contained within the shower [39].

Third, we evaluate the "shape" of the shower. Showers from different particle types have a characteristic distribution of energy among the crystals that form the shower. Showers originating from a single photon will usually involve the group of nine crystals at the center of the shower. A hadronic interaction or the overlap of two particles will form a much broader shower. To reconstruct a shower, clusters are formed by grouping together individual crystal hits. The center of the cluster is determined by evaluating the energy and position of each crystal in the cluster. To evaluate the "shape" of the shower, we define a block of 25 crystals surrounding the center of the shower and measure the ratio of the sum of the energy contained in the central nine crystals to the energy contained in all 25. We require that this ratio (E9/E25) be greater than 0.95, ensuring that the candidate shower is well-contained. In addition, we use a second characterization of the shower's shape, called "shower mass", m_{shr}. The "mass" of a shower is the invariant mass computed by assuming that the energy in each crystal was produced by a single photon. More widely distributed showers will therefore have a larger shower mass. The requirements on shower mass are energy dependent and were developed and used in the original analysis [40]. Specifically, the shower mass (m_{shr}) is given by the product of the width of the candidate shower in meters and the ratio of the candidate energy and the distance from the origin to the shower in meters. The candidate shower is

---

[3]A full list of these cuts using the standard CLEO variables is given in Appendix A.

rejected if the shower mass is greater than the value of this shower mass cut. This cut is described in greater detail in Appendix A. Extensive documentation about how this cut was derived can be found in Reference [40].

Fourth, we employ SPLITF, the splitoff rejection package [41]. Photons that hit the crystals cause a fairly well-contained shower. An electron can cause a shower similar in shape to a high energy photon. These two particles are distinguished by looking for a charged track leading to the center of the shower. Hadrons, however, interact strongly with the material of the crystal and will convert to secondary particles. These secondary hadrons may travel laterally through neighboring crystals and deposit most of their energy away from the central shower resulting in secondary showers that are split off from the primary hadronic shower. These showers are difficult to identify because there is no track pointing to this splitoff shower. The splitoff routine takes advantage of the fact that splitoff showers are more likely to lie near the parent charged shower than far away. It determines whether a cluster is likely to be a splitoff based on several parameters. The inputs to SPLITF include an array of good tracks to be used for track matching, the track types (whether the particle associated with the track is a hadron, an electron, or a muon), and the minimum shower energy allowed. We use a minimum shower energy of 25 MeV. The output of this package tells us whether we should accept or reject the given cluster.

Finally, since most of our background originates with photons from π^0s and ηs, we reject a photon as a candidate if it can form a π^0 or an η with another shower in the event. We use a π^0 mass window with 110 MeV/c^2 < $m_{\gamma\gamma}$ < 150 MeV/c^2 and an η mass window of 515 MeV/c^2 < $m_{\gamma\gamma}$ < 575 MeV/c^2. The sibling photons used for the π^0 veto must be good showers with a minimum energy of 30 MeV in the barrel and 50 MeV in the endcap. For the η veto, the sibling photons must also be good showers, but with energy requirements of E_γ > 200 MeV for all sibling photons.

In addition to these cuts, there is an energy requirement of $2.0 - 2.7$ GeV, For the purposes of measuring the spectrum, we accept all photons above 1.5 GeV and apply the energy requirement when determining the yield. If all of these conditions are met, then we have a candidate photon in this event.

To cut down on the continuum, we must learn more about the event and the other particles in it. All of the remaining cuts and techniques are used to identify and reconstruct the other particles in the event and to implement our continuum suppression schemes (shape, pseudoreconstruction, and lepton tagging).

4.6   Track Selection

Charged particles are used in the event to form other composite particles as well as to characterize the angular distribution of energy in the event. The charged

tracks that are used to reconstruct the X_s system and to characterize the "shape" of the event must pass standard track quality cuts.[4]

These tracks are required to come from the origin and must pass passing the CLEO track quality package [42]. When a charged particle passes through the drift chambers, it generates a set of hits to which the tracking programs will fit a track. The vector momentum and the error on the measurement of a charged track is computed from a fit to the reconstructed track.

To select good tracks, the track cannot have too many missing hits and is rejected if it has a momentum greater than 5.5 GeV/c. This is an unphysical momentum resulting from a bad fit to the hits. We also require that the particle that caused the track did not escape down the beam pipe. The track is constrained to come from the origin; it is allowed to originate from a point that is radially within 50 cm of the beam line and within 10 cm of the interaction point in z.

4.6.1   Assigning a Particle Type to a Track

When a track's energy and momentum are computed, corrections are made to compensate for the energy losses of the particles as they pass through the material of the detector. Since these corrections depend on the mass of the particle, we need to identify the particle species that caused the track. Our strategy is to treat the track as a pion unless we find that it is better suited to some other particle. This selection process is as follows.[5]

The information provided from the time of flight (TOF) counters and the dE/dx measurements is given in the form of four variables for a given track. These are the standard deviation of the measurement from that expected for a $\pi^{\pm}, K^{\pm}, p/\overline{p}$, and $e^{\pm}$. Whether we use information from TOF, dE/dx, or both to determine the particle type depends on whether these devices are reporting reliable data.

In order for us to use time of flight information to identify the particle, the track must pass through the TOF counters in the barrel. These counters give the cleanest measurement because they are not obstructed by endcap material. We also require that the time of flight information is of good quality, and that the detector was not misbehaving during this run. Likewise, if the CLEO dE/dx subroutines determine that there is no dE/dx information available for a given track, then we do not use dE/dx to identify the particle. Once we have determined which device has reliable data, we calculate sigmas based on the available TOF and dE/dx information for each particle hypothesis (e, μ, π, K, p) for this track.

We determine the charge of the particle by examining the direction of curvature of the track. To assign a particle type to a given track, we determine if the specific

---

[4]For those familiar with CLEO variables, a list of track selection cuts is given in Appendix A.

[5]CLEO readers are encouraged to read Appendix A which describes the particle identification process in greater detail using CLEO variables.

energy loss measured for this track is consistent with a proton hypothesis. Then, we see if it is consistent with a kaon hypothesis and finally a pion hypothesis. If a particle is consistent with a π, then it will be considered a π regardless of the other hypotheses.

Electrons can be distinguished from hadrons and muons by taking several quantities into account. First, the amount of energy left in the calorimeter compared to the momentum of the track matched to that shower. Muons deposit relatively little of their overall energy in the calorimeter. Hadrons sometimes interact strongly leaving some of their energy in the calorimeter. Electrons, however, tend to deposit all of their energy in the crystals. To a good approximation, in the energy range that we are concerned with, the electron's energy is nearly equal to the electron's momentum. This means that E/p is usually close to one for electrons while it is much less than one for hadrons and muons. Information from dE/dx is also useful in distinguishing between electrons, muons, and hadrons. The REID electron identification package combines these and other variables using a likelihood fit to produce a variable that discriminates between electrons and hadrons at various momenta in different regions of the detector. We use the REID output to tell us if a given track is an electron.

Muon identification relies on the fact that only muons can penetrate through one or more layers of the iron return yoke to reach the muon chambers within. A track is considered a muon if there are muon hits everywhere along the track that muon hits were expected and if at least two layers hit in the muon chambers [43].

Once we have determined whether the given track can be identified as a muon or electron, we override all previous particle type settings. First we check to see if it has been classified as an electron. If yes, we assign it an electron particle type. Next, we check to see if it has been classified as a muon. If yes, then we override all previous type settings and call it an electron.

4.7   Shower selection

It is important to this analysis to identify photons that have come from π^0s, ηs and other neutral particles. These neutral particles leave showers in the calorimeter. We use these showers to perform our π^0/η veto with the photon candidate and in the pseudoreconstruction of the event. In order to be considered a good shower caused by a neutral particle, a cluster in the calorimeter must have an energy greater than 50 MeV within the barrel or the endcap. It cannot be matched to a charged track, must be well-contained, and must pass the splitoff rejection package. In addition, the cluster must not be associated with crystals that contain noisy, dead, or shorted electronic channels. If the neutral fulfills these requirements, then we call it a photon and add it to our list of good neutrals.

4.8 Building Intermediate Particles

Once we have raw tracks and showers, we can start combining them to form π^0s, ηs and K_s^0s. These composite particles are later used to pseudoreconstruct the B meson in the event. We reconstruct these particles in the following modes: $\pi^0 \to \gamma\gamma$, $\eta \to \gamma\gamma$, and $K_s^0 \to \pi^+\pi^-$.

We build π^0s using a two-photon invariant mass window of 110 MeV/c^2 < $m_{\gamma\gamma}$ < 150 MeV/c^2. Likewise, to reconstruct an η, we accept combinations of two photons with an invariant mass between 515 MeV/c^2 < $m_{\gamma\gamma}$ < 575 MeV/c^2. The photons are taken from the list of neutral showers described above. The K_s^0 is formed by combining two good tracks identified as a pair of oppositely charged pions as described above and accepting them as a K_s^0 if their invariant mass is between 486 MeV/c^2 < $m_{\pi^+\pi^-}$ < 510 MeV/c^2. We record the number of good reconstructions that are found, the quality of the reconstructions, and which two tracks or showers were used to create the composite particle. These particles are then combined with the unused tracks and showers to form a full list of particles for the event.

The analysis machinery itself goes through these selection criteria in the following order. First, it looks for good tracks (according to the track selection criteria) and assigns a particle ID to all of the tracks passing the cuts. Next, it looks for good showers and creates separate lists for each of the following categories: all showers (from which we later select potential candidates), showers fulfilling the proper energy requirements in the barrel (for π^0 veto), showers passing the energy criteria in the endcap (for π^0 veto), and showers passing the energy requirements for the η veto. Next, we build intermediate particles such as π^0s, ηs, and K_s^0s and add them to a full list of particles in the event. Finally, we look for photon candidates by looping over all showers and determining whether or not any of these showers fit the candidate criteria. We veto a potential candidate, if it combines with any other shower in the event to form a π^0 or an η. Should a photon pass the π^0/η veto, we proceed to the next part of the analysis: continuum suppression.

4.9 Combining Event Information Using a Neural Net

To distinguish signal from continuum background, we combine shape information, reconstruction information, and lepton information into a single discriminatory variable that distinguishes between signal and continuum using a neural network.

A neural network is a combination of simple functions called nodes which can be trained to perform pattern detection functions [44]. A set of input variables is fed into one end of the network and the decision-making algorithm inside outputs a single variable that represents the extent to which the input variables exhibit a certain signal characteristic. Details about the original neural net can be found in reference [40].

A neural network must be trained so that application of a set of inputs produces the desired set of outputs. Our nets are trained on signal and continuum Monte Carlo, generated expressly for this purpose. Not all events have the same types of information, so we optimize a net for each of the four event categories, then run each event through the appropriate one. Each net has its own set of inputs and internal architecture. In the end, we take all the input variables we have and combine them using the appropriate neural net to form a single net output that is the ultimate arbiter of how signal-like or continuum-like the event is. In all cases, this output, which I will generically call r, tends toward $+1$ for signal events and -1 for continuum background events. Usually, r refers specifically to the output of the shape net, but for the sake of simplicity in discussing weights, in Section 4.10, I will use r generically to denote the output of all of the four neural nets.

4.10   Weights

Given that the net output, r, tends toward $+1$ for signal events and -1 for continuum background events, we could choose a "cut" on this variable, and label all events above this cut as signal and below it as background. However, we would be ignoring information contained in the distribution of r. For example, accepted events near the r cut boundary are more likely to be background than are accepted events far from the boundary.

One way to quantitatively extract the information on both sides of the cut is to fit the measured distribution of r values to the *expected* distribution of r values for signal. We use the equivalent procedure in which we weight each event according its neural net output value. We weight each event according to its r value, with a weight given by

$$w_i = \frac{S_i(r)}{S_i(r) + (1 + \lambda)B_i(r)}$$

where $S_i(r)$ is the expected signal yield in bin i of the r distribution, $B_i(r)$ is the expected background yield in that bin, and λ is the luminosity scale factor. We define the weight to give the minimum statistical error on the yield, $\Delta Y / Y$. After the On-Off subtraction described in Section 4.2.1, our expected error is

$$\frac{\Delta Y}{Y} = \frac{\sqrt{S + (1 + \lambda)B}}{S}$$

and the optimum weight per event is $1/(expected\ error)^2$. The weights are normalized such that a distribution of n events will have a weighted sum of n if they follow the expected signal r distribution exactly.

We convert the neural net output of each of the four nets, r, to the weight, w, which is the probability that a given point in r space will be a signal rather than a background event. The event weighting function is obtained from signal and

continuum Monte Carlo samples that are different from the samples used to estimate signal efficiency. Using these samples, we determine the expected precision for the analysis at any r value. If the MC makes poor estimates of the distribution of our net output, then the choice of weights will not be optimum, and our statistical errors will ultimately be larger than they would have been had our MC been better. With this method, events in regions of low background are given more weight than events in regions of high background.

Once each event has a weight, the branching ratio is determined by summing the weights for all events. We define the efficiency as the weighted sum of a sample of signal events divided by the number of events in the sample.

CHAPTER 5

Backgrounds

In the previous chapter, we described the major backgrounds to $b \rightarrow s\gamma$: continuum (ISR and $q\overline{q}$) and $B\overline{B}$. This chapter describes how we eliminate them.

5.1 Continuum Suppression

In the previous chapter, we outlined a three-pronged continuum suppression scheme (shape, pseudoreconstruction, lepton tagging) that uses all the information available in an event to characterize it as signal-like or continuum-like. The real work, however, is in finding the variables that will separate the signal from background.

5.1.1 Shape Analysis

The first continuum suppression scheme is based on the difference in "shape" between signal and continuum events. Figure 5.1 shows a typical $B\overline{B}$ event in which the tracks and showers are uniformly distributed throughout the detector. Figure 5.2 by contrast, shows a $q\overline{q}$ event in which the products of the decay are more narrowly collimated. This kind of topology can be described with the set of eight suppression variables whose effectiveness for separating between signal, $q\overline{q}$ and ISR is shown in Figures 5.3 and 5.4 [40].

- R_2

  Qualitatively, R_2 is a measure of the isotropy of an event. It is the ratio of two Fox-Wolfram moments ($R_2 = H_2/H_0$), and tends toward 0 for spherical $B\overline{B}$ decays, and tends toward 1 for the jet-like continuum events [45]. It is defined as follows:

  $$R_2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |\vec{p}(i)||\vec{p}(j)|P_2(cos\theta_{i,j})}{\sum_{i=1}^{n} \sum_{j=1}^{n} |\vec{p}(i)||\vec{p}(j)|}$$

48
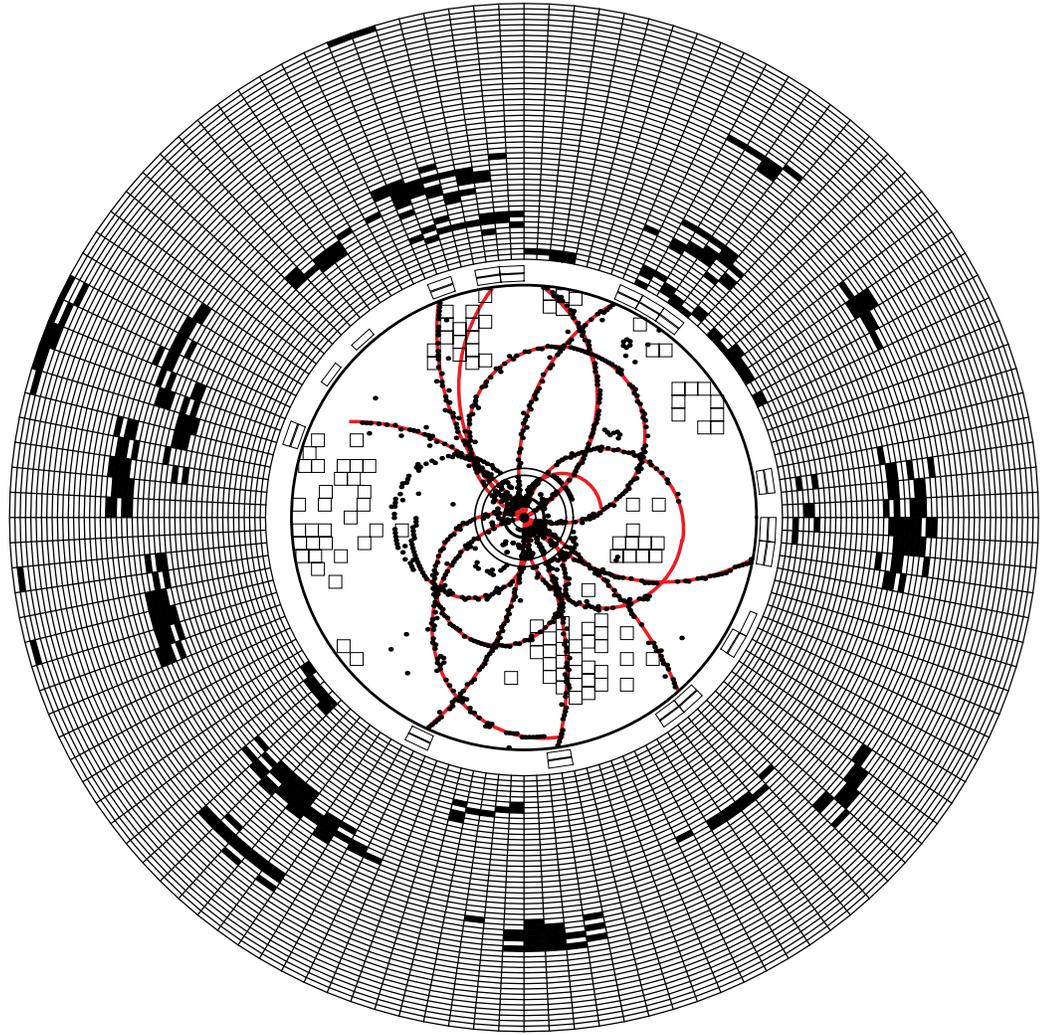
Figure 5.1: A typical $B\overline{B}$ event with a "spherical" topology.
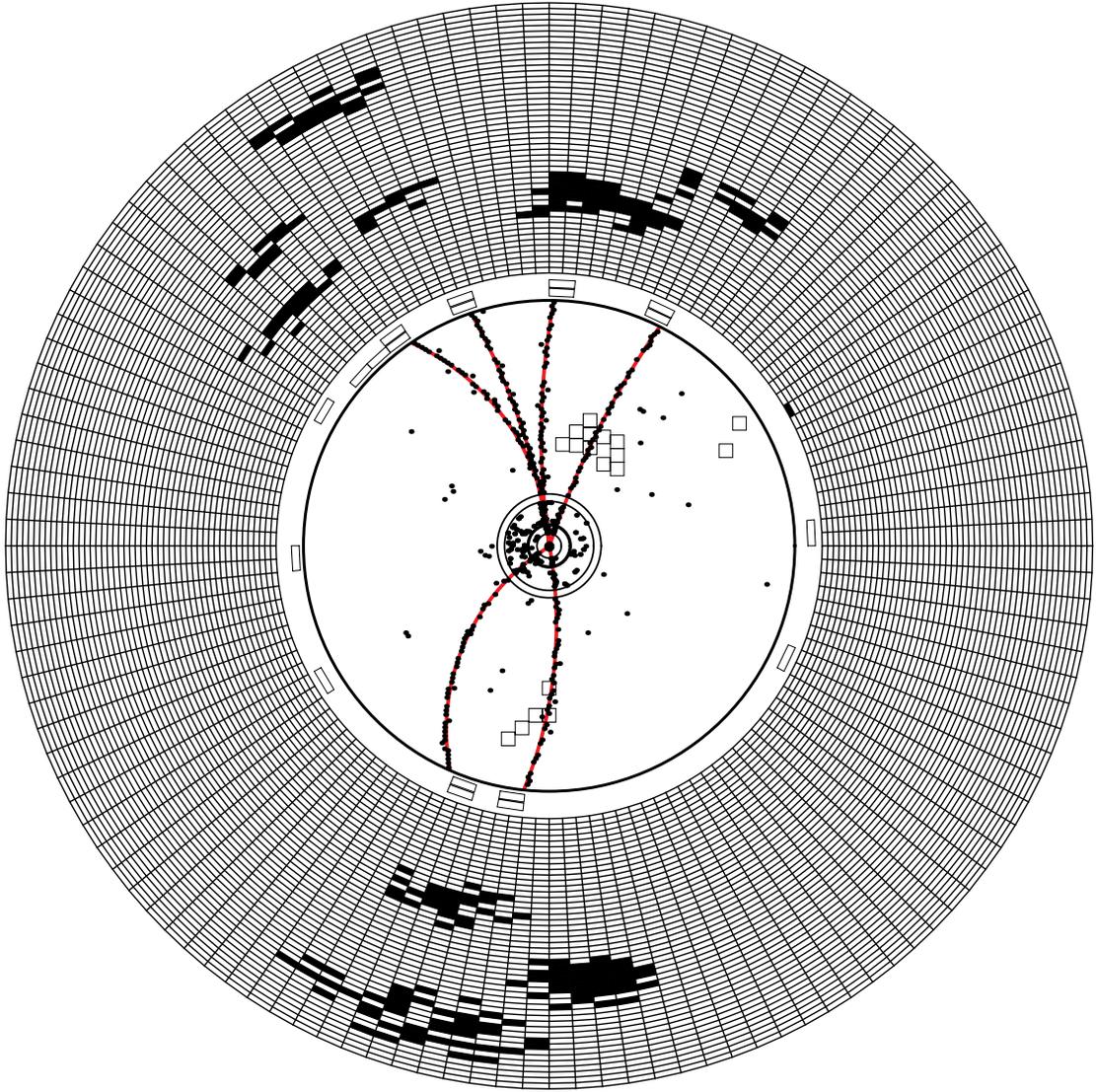
Figure 5.2: A typical $q\overline{q}$ event with a "jetty" structure.

where n is the number of particles in the particle list, i,j is a pair of non-identical particles, $\vec{p}(i), \vec{p}(j)$ are the momenta of particles i and j respectively, $cos\theta_{i,j}$ is the cosine of the opening angle between the ith and j^{th} particles, and $P_2(cos\theta)$ is the Legendre polynomial of order two. R_2 is effective at separating signal events from $q\overline{q}$. Since ISR events are a little more spherical in shape than typical $q\overline{q}$ events, R_2 is not very effective at separating signal from ISR.

- R_2'

  We calculate R_2 in the primed frame, or the frame of reference recoiling against the radiated photon. Since ISR events are just $q\overline{q}$ events recoiling against the radiated photon, this variable is especially powerful in separating signal from ISR. Having lost the energy of the photon, an ISR event in the primed frame has the same shape as a reduced energy continuum event. By transforming to the primed frame, we transform the event to a jetty $q\overline{q}$ event.

- $S_\perp$

  This variable is the sum of the magnitudes of components of momenta perpendicular to the photon for all particles more than 45° from the photon axis divided by the sum of the magnitudes of the momenta of all particles except the photon:
  $$S_\perp = \frac{\sum |\vec{p}(j)| sin\theta_{j\gamma}}{\sum |\vec{p}(j)|}$$
  The summation runs over all selected tracks and showers, j. In the numerator, the forward-backward cone about the direction of the photon is excluded such that $|cos\theta_{j\gamma}| > 0.707$. For a perfect two jet event, $S_\perp$ would be zero while for spherical $B\overline{B}$ events it tends toward 1. $S_\perp$ provides some discriminatory power between the relatively spherical $b \rightarrow s\gamma$ events and the jetty $q\overline{q}$ events and offers some discrimination between signal and ISR.

- $cos\theta'$

  This variable is also calculated in the primed frame. It is the cosine of the angle between the photon and the thrust axis of the rest of the event calculated in the frame recoiling against the photon. The thrust axis of the event is the direction for which the sum of the magnitudes of the longitudinal components of momenta is a maximum. In signal events, the photon will tend to line up along this thrust axis, whereas, in ISR events, there is no correlation between the photon direction and the direction of the rest of the event.

- **20° and 30° forward and backward cones**

  The cone energy sums act as a virtual calorimeter mapping the energy flow of the event which varies depending on whether the event is signal, continuum,
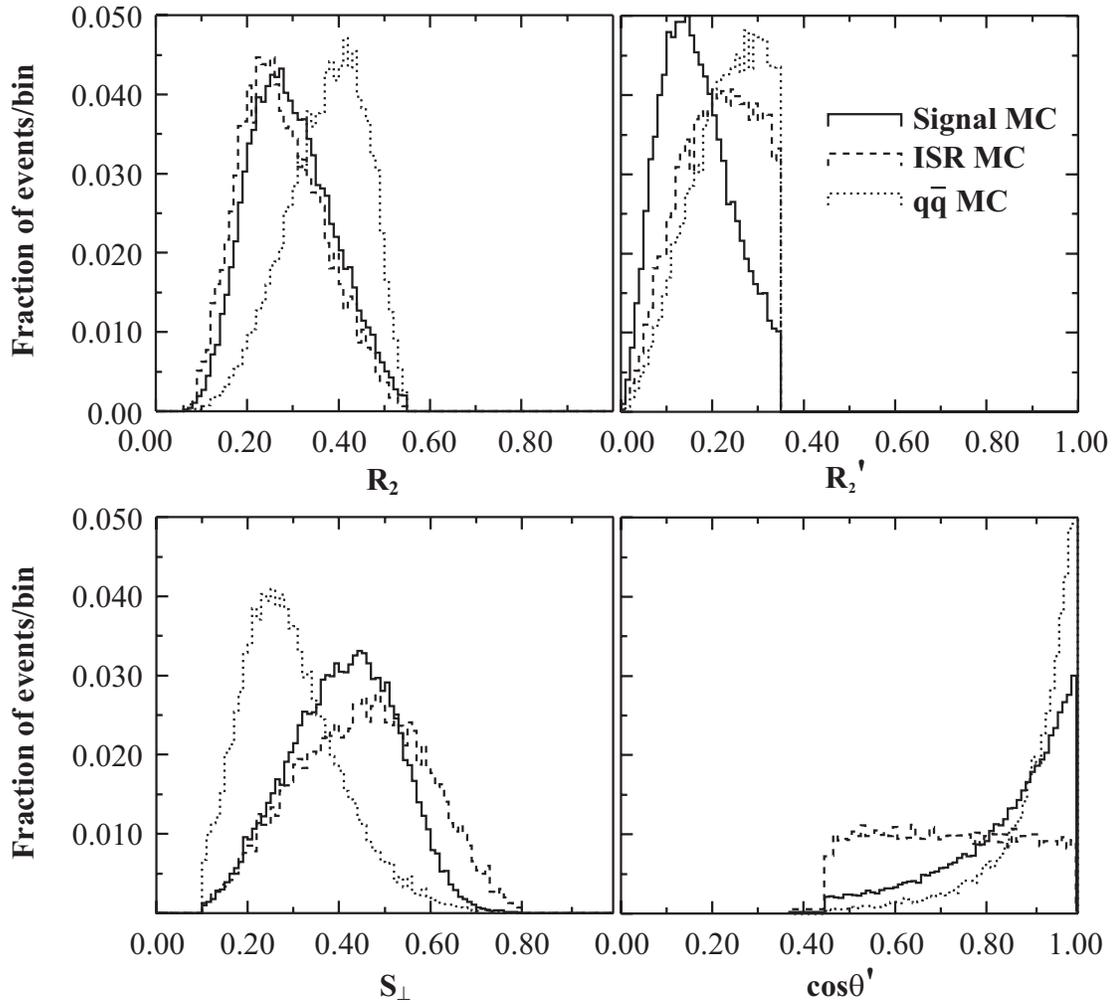
Figure 5.3: Angular distribution of signal, ISR, and $q\bar{q}$ for four of the shape variables: R_2, R_2', $S_\perp$, and $cos\theta'$. Loose shape cuts (see Appendix B) have been applied.

or ISR. These four variables are calculated by summing the energy (excluding the photon) of particles within the 20° and 30° cones surrounding the photon. For signal events, we expect a small amount of energy in the forward cones (those in the direction of the γ) and a moderate amount of energy in the backward cones. In $q\overline{q}$ events, on the other hand, we expect a large amount of energy in both the forward and the backward cones. For ISR events, we only expect a small amount of energy in both forward and backward cones. Figure 5.4 shows the energy in the forward and backward cones for signal, ISR, and $q\overline{q}$.

When calculating the shape variables, we use all of the particles in the event that came from the interaction point. This list includes all of the tracks, showers, and composite particles that meet the criteria listed in the previous chapter and excludes the candidate photon.

In the "shape" analysis, these eight variables are combined using a neural net to form one discriminatory variable, called r, that distinguishes between signal and continuum background based solely on the shape of the input event. Figure 5.5 shows the distribution of r for signal and continuum MC.

We perform the shape analysis for ALL events that have a good photon candidate. For events that are not reconstructible and do not have a leftover lepton, the final weight is determined from this r value alone. For events that are also capable of being reconstructed or for events that have a leftover lepton, we use r along with other variables as an input to other neural nets.

5.1.2   Pseudoreconstruction

Further continuum suppression is achieved if we determine that the transition photon and the recoiling X_s system are kinematically consistent with the decay of a B meson. We suppress the continuum background by reconstructing some subset of the particles in the event into a $B \rightarrow X_s\gamma$ candidate. For each event, we consider only one combination of particles; the combination that best satisfies some $B \rightarrow X_s\gamma$ decay hypothesis. A fully inclusive analysis is achieved by using a large number of exclusive decay modes. For each event, we try to reconstruct the X_s system recoiling against the photon. The X_s system can be composed of one kaon (which can be a K_s^0) and 1-4 pions including at most one π^0. We try to reconstruct the B meson in one of the modes listed in Table 5.1.

As with the shape variables, we combine particles from the list of tracks, showers, and composite particles that satisfy the criteria listed in the previous sections. We keep track of the list of B candidates and the quality of their reconstruction.

The reason that this is called "pseudoreconstruction" instead of reconstruction is because it is not important for us to get the reconstruction correct in every detail. Since we are only using the reconstruction as a continuum suppression technique,
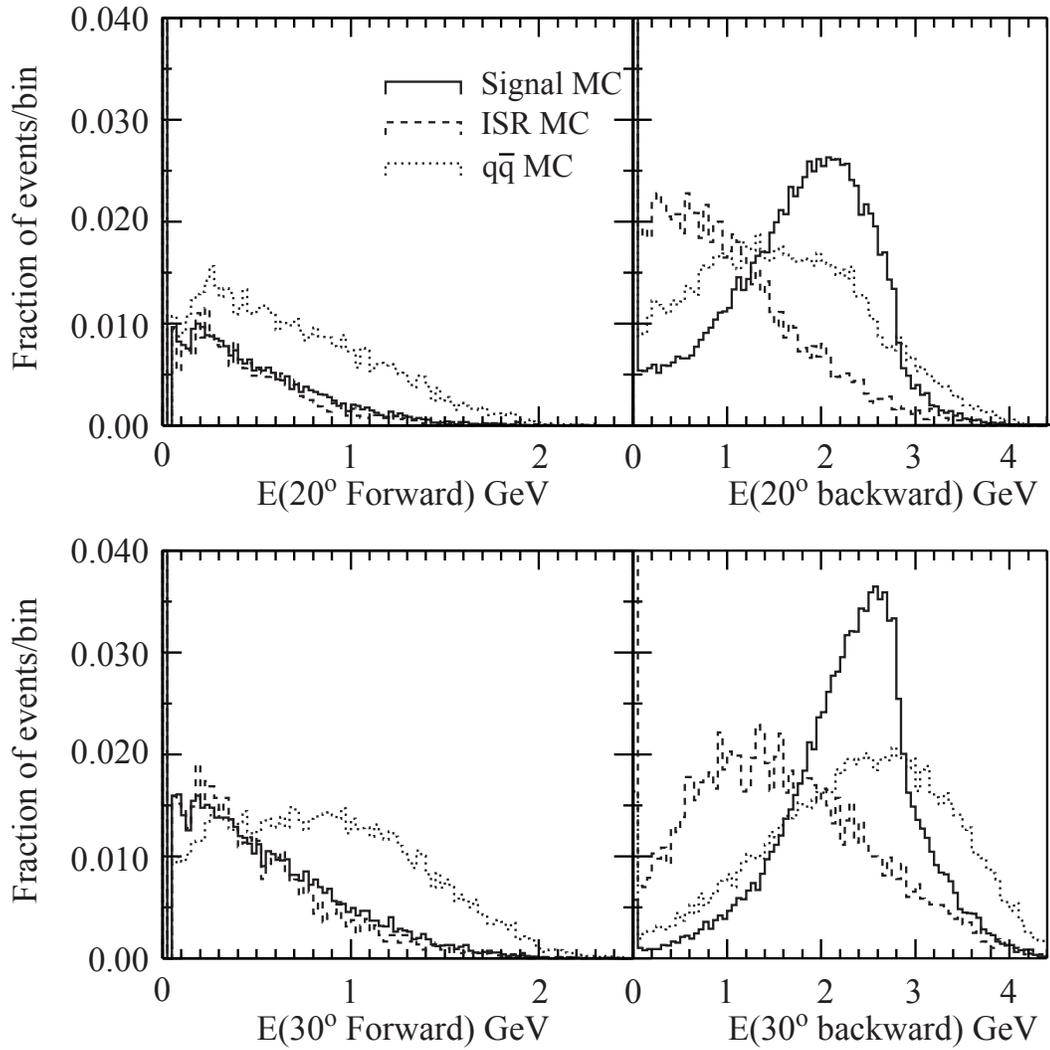
Figure 5.4: Energy in forward and backward cones.
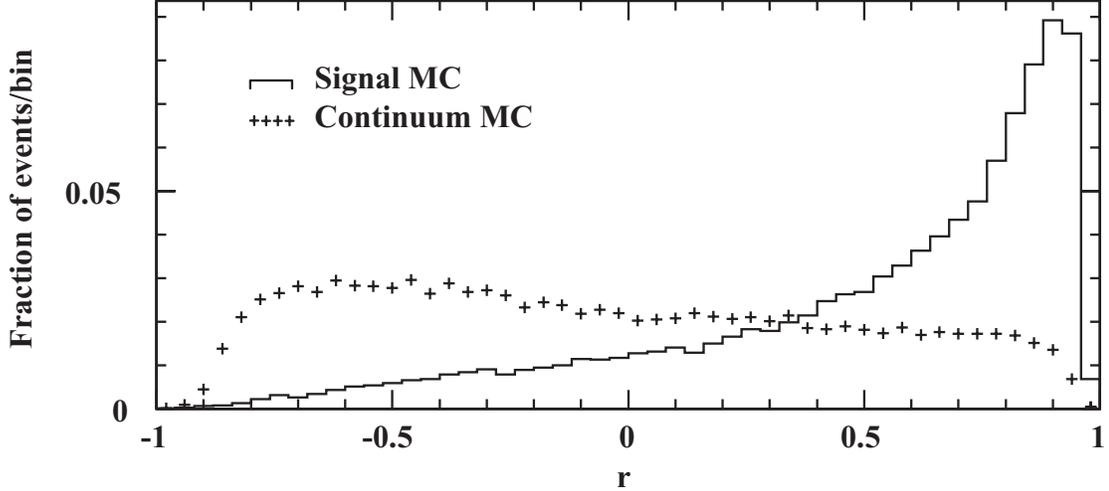
Figure 5.5: r distribution for shape analysis done on signal MC and continuum MC. Although the shape analysis is done for all events, this particular distribution comes from events that were reconstructed without a lepton leftover. Signal events peak strongly towards 1 whereas continuum tends toward -1.

our only concern is to determine whether the reconstruction is *consistent* with a B meson decay.

For each set of candidate particles that can be combined to form a B meson, we obtain the total energy, E, and the total momentum, P for all of the particles in the reconstruction. Then, we calculate the energy difference, ΔE, beam-constrained mass, M, and mass difference, ΔM for this combination of particles. These variables are defined as follows:

$$\Delta E = E - E_{beam}$$

where E is the total energy of the combination and E_{beam} is the beam energy. This kinematic requirement makes use of our knowledge that the energy of the parent B meson must be equal to the beam energy, E_{beam}. When computing the invariant mass of a set of particles assumed to come from the decay of a B meson, we can impose the constraint $E_B = E_{beam}$. For the correct combination of candidate particles, ΔE will be zero, within measurement error. The beam-constrained mass of the system is written as

$$M = \sqrt{E_{beam}^2 - P^2}.$$

Because the energy resolution on the beam is better than the resolution of the reconstructed B meson, this gives a more precise value for the mass of the B meson

55

Pseudoreconstructed modes	
$\overline{B}^0 \to \gamma K^- \pi^+$	$B^0 \to \gamma K^+ \pi^-$
$\overline{B}^0 \to \gamma K^- \pi^+ \pi^0$	$B^0 \to \gamma K^+ \pi^- \pi^0$
$\overline{B}^0 \to \gamma K^- \pi^+ \pi^+ \pi^-$	$B^0 \to \gamma K^+ \pi^- \pi^+ \pi^-$
$\overline{B}^0 \to \gamma K^- \pi^+ \pi^+ \pi^- \pi^0$	$B^0 \to \gamma K^+ \pi^- \pi^+ \pi^- \pi^0$
$\overline{B}^0 \to \gamma K_s^0 \pi^0$	$B^0 \to \gamma K_s^0 \pi^0$
$\overline{B}^0 \to \gamma K_s^0 \pi^+ \pi^-$	$B^0 \to \gamma K_s^0 \pi^+ \pi^-$
$\overline{B}^0 \to \gamma K_s^0 \pi^+ \pi^- \pi^0$	$B^0 \to \gamma K_s^0 \pi^+ \pi^- \pi^0$
$\overline{B}^0 \to \gamma K_s^0 \pi^+ \pi^+ \pi^- \pi^-$	$B^0 \to \gamma K_s^0 \pi^+ \pi^- \pi^+ \pi^-$
$B^- \to \gamma K^- \pi^0$	$B^+ \to \gamma K^+ \pi^0$
$B^- \to \gamma K^- \pi^+ \pi^-$	$B^+ \to \gamma K^+ \pi^+ \pi^-$
$B^- \to \gamma K^- \pi^+ \pi^- \pi^0$	$B^+ \to \gamma K^+ \pi^+ \pi^- \pi^0$
$B^- \to \gamma K^- \pi^+ \pi^- \pi^+ \pi^-$	$B^+ \to \gamma K^+ \pi^+ \pi^- \pi^+ \pi^-$
$B^- \to \gamma K_s^0 \pi^-$	$B^+ \to \gamma K_s^0 \pi^+$
$B^- \to \gamma K_s^0 \pi^- \pi^0$	$B^+ \to \gamma K_s^0 \pi^+ \pi^0$
$B^- \to \gamma K_s^0 \pi^- \pi^+ \pi^-$	$B^+ \to \gamma K_s^0 \pi^+ \pi^+ \pi^-$
$B^- \to \gamma K_s^0 \pi^- \pi^+ \pi^- \pi^0$	$B^+ \to \gamma K_s^0 \pi^+ \pi^+ \pi^- \pi^0$

Table 5.1: The following modes for B^+/B^- and $B^0/\overline{B}^0$ are used to pseudoreconstruct a $B \to X_s \gamma$ decay.

candidate. In analyses of hadronic decays, this technique gives a mass resolution of $\sigma_M \approx 4$ MeV/c^2, about a factor of ten better than the resolution without the constraint.

We then calculate the mass difference $\Delta M = M - M_B$ where M_B is the PDG mass of the B meson (5.279 GeV/c^2). Again, if we have the correct set of candidate particles, ΔM will be zero within measurement error. We construct the variable

$$\chi_B^2 = \left(\frac{\Delta E}{\sigma_E}\right)^2 + \left(\frac{\Delta M}{\sigma_M}\right)^2.$$

We use $\sigma_E = 40$ MeV and $\sigma_M = 4$ MeV/c^2 and define the reconstruction quality as $\chi_B^2 = (\Delta M/0.004)^2 + (\Delta E/0.040)^2$, which reflects how well the kinematic properties of the combination of particles represent a B meson. The distributions of ΔE and ΔM for all reconstructed events from a sample of signal MC are given in Figure 5.6. Ultimately, we use two variables to help with continuum suppression:

- χ_B^2
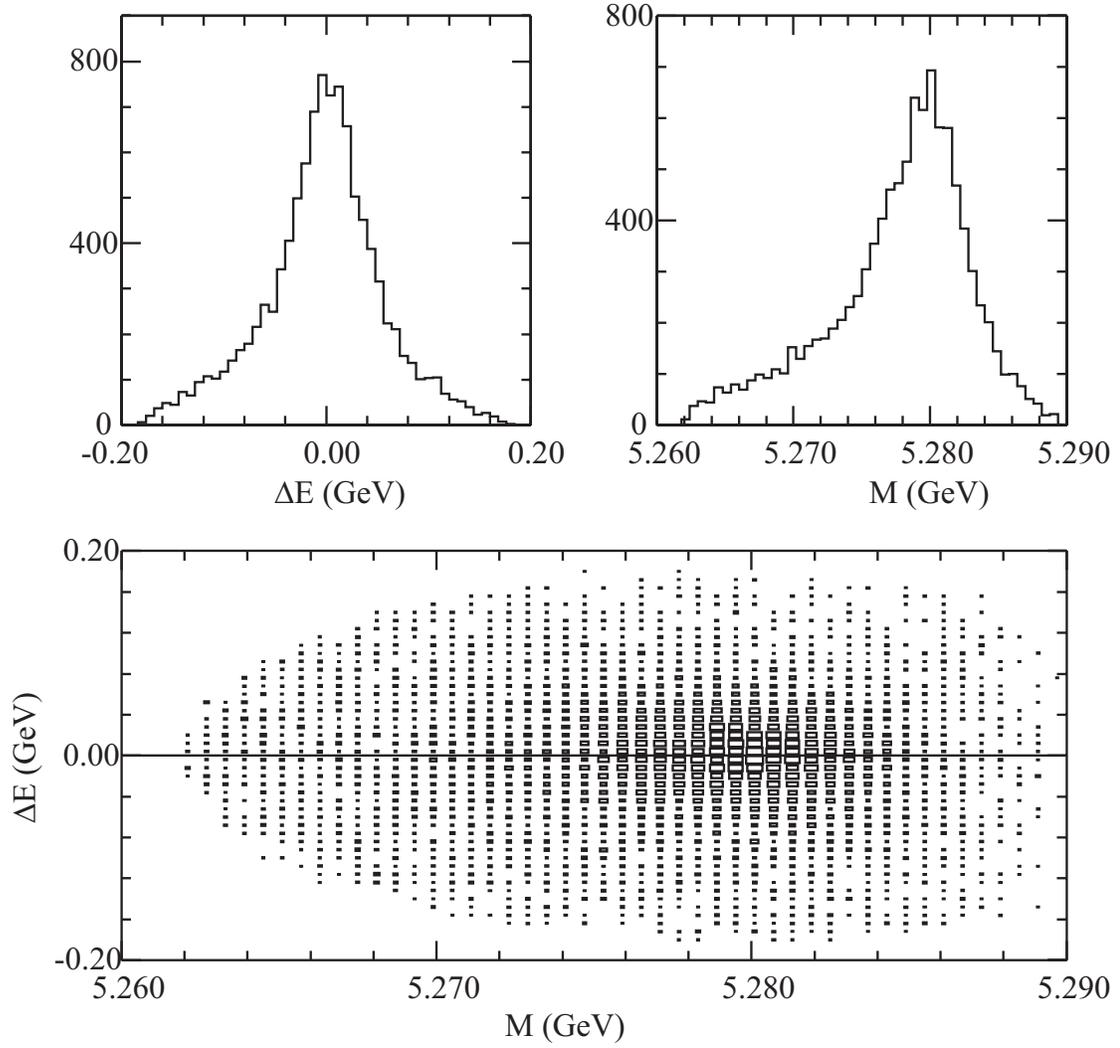
Figure 5.6: Distribution of ΔE and beam-constrained mass for reconstructed events from signal MC. The plot of ΔE vs. M shows that the signal is concentrated around $\Delta E = 0$ and $M = 5.280$ GeV/c^2 with an intrinsic width of about 40 MeV in ΔE and 4 MeV/c^2 in M.

This variable is defined above. A valid reconstruction must have $\chi_B^2 < 20$, and the best B reconstruction is the one with the lowest χ_B^2. As a general rule, the χ_B^2 values for signal are smaller than those for continuum. Figure 5.7 gives the distribution for $log(\chi_B^2)/4$.[6]

- $|cos\theta_{tt}|$

    The variable $|cos\theta_{tt}|$ exploits the fact that the decay angles of the two B mesons in each event are uncorrelated. In jetty continuum events, the thrust axes of the best combination of particles and the rest of the event tend to be aligned, whereas these axes are randomly distributed in phase space for signal events. $|cos\theta_{tt}|$ is the cosine of the angle between the thrust axis of the particles used to form the best candidate B and the thrust axis of the remaining particles in the event. Since there is no angular correlation between the decay products of the two B mesons in a $B\overline{B}$ event, $|cos\theta_{tt}|$ is expected to be flat. For continuum events, the two thrust axes will tend to align so $|cos\theta_{tt}|$ will peak near 1 as shown in Figure 5.7

5.1.3   Lepton Tagged Events

The final handle we have on continuum suppression is the presence of a lepton in the event. $B\overline{B}$ events decay semileptonically to an electron or a muon about 20% [1] of the time. These leptons tend to be stiff with much of the spectrum from primary semileptonic decays above 1.0 GeV/c [46]. We can use this characteristic of B meson decays as a continuum suppression tool.

Once we have identified our best B meson candidate by choosing the reconstruction with the lowest χ^2, we search through the leftover particles for the highest momentum lepton (electron or muon) from the non-signal B. In reconstructed events, we also look at the sign of the lepton. For example, if our signal B were a B^+, then the other B would be a B^- containing a b quark with charge $-1/3$. If this B^- were to decay semileptonically (i.e., $b \rightarrow cl\nu$), then the lepton would also have a negative charge. Having a "hard" lepton, a lepton with $P_l \gtrsim 1.0$ GeV/c in an event increases the probability that the event is signal. We gain information by looking at whether the sign of the lepton is in agreement with the sign of the reconstruction.

The angle between the highest momentum lepton and the photon is also indicative of whether this event is continuum or $B\overline{B}$. Because continuum events are jetty, there is an angular correlation between the direction of the photon and the direction of the lepton, so the cosine of this angle will peak at ± 1. There is no such correlation in $B\overline{B}$ decays so the distribution of the cosine of this angle is flat.

---

[6]The log of this variable is used as the neural net input because the nets converge more quickly if the input variables are of similar range.
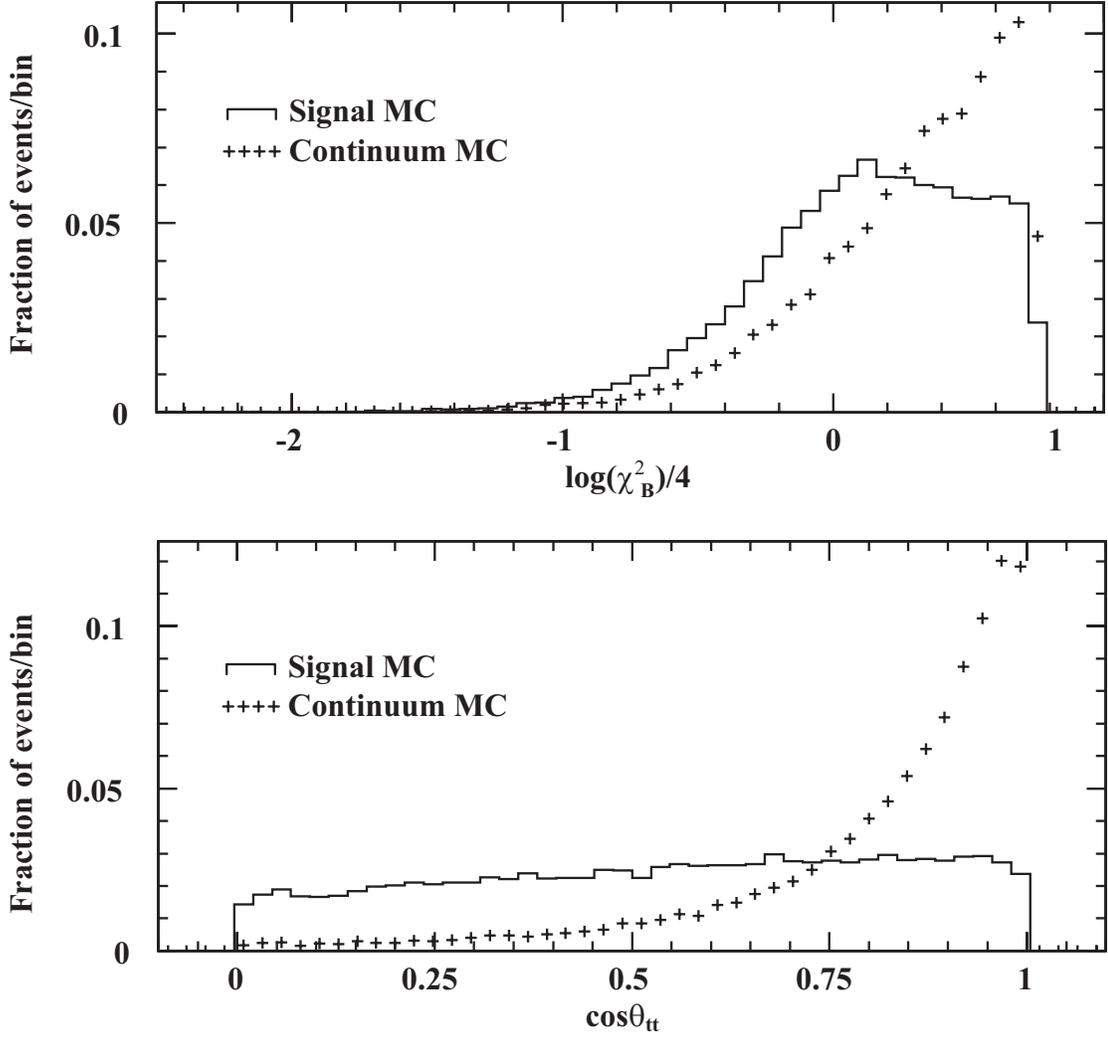
Figure 5.7: The top plot gives the distribution of $\log(\chi^2_B)/4$ for signal MC and continuum MC. The lower plot gives the distribution of $|cos\theta_{tt}|$, the cosine of the angle between the candidate photon and the thrust axis of the particles in the "other" B plotted for signal MC and continuum MC. These plots are for events that were reconstructed but had no lepton in the non-signal B.

For the highest momentum lepton in the non-signal B, we record the charge, momentum, and the angle between it and the candidate photon. Lepton tagged events are shown in Figure 5.8 and use the following variables in the neural net:

- **Signed lepton momentum**

  The neural net input is given by the expression

  $$P_l \times q_l \times q_b$$

  where P_l is the absolute value of the momentum of the highest momentum lepton in the "other" B, q_l is the charge of this lepton, and q_b is the charge of the quark in the B meson.

- **Signed photon-lepton angle**

  The neural net input is given by the expression

  $$\theta_{l\gamma} \times q_l \times q_b$$

  where $\theta_{l\gamma}$ is the angle between the hardest lepton in the non-signal B and the candidate photon.

Recall that we search for leptons in the event even if we have not reconstructed a B meson. To do this, we perform a crude B reconstruction by combining a couple of high momentum tracks recoiling against the photon. For each good track identified as a π or a K, we find the component of its momentum opposite to the photon candidate. Then, we choose the two tracks with the highest momentum opposite the photon. If both of those tracks are identified as pions, then we look for the kaon with the highest momentum opposite the photon and add it as an additional daughter. This becomes the one and only B reconstruction for this event. We do not treat this event as pseudoreconstructed, but we use these daughters to define the particles that are in the signal B. We then look in the leftover particles of the "other" B for a lepton and determine the variables described above, unsigned (we do not multiply by the charge of the lepton or the charge of the b quark). If we are unable to perform even this crude reconstruction, then the event is of such poor quality that we do not use it and throw it out.

5.1.4   Event Categories

All events with a candidate photon are sorted into one of four event categories depending on whether we are able to reconstruct the event and whether we are able to find a lepton in the non-signal B. For all events, we perform the shape analysis. Next, we attempt to reconstruct the signal B, and the event is either reconstructible or not. The event is considered reconstructed if it has at least one B candidate with $\chi^2_B < 20$. Finally, an event either has a lepton in the non-signal B or it does not (independent of whether it is reconstructible). The four event categories are
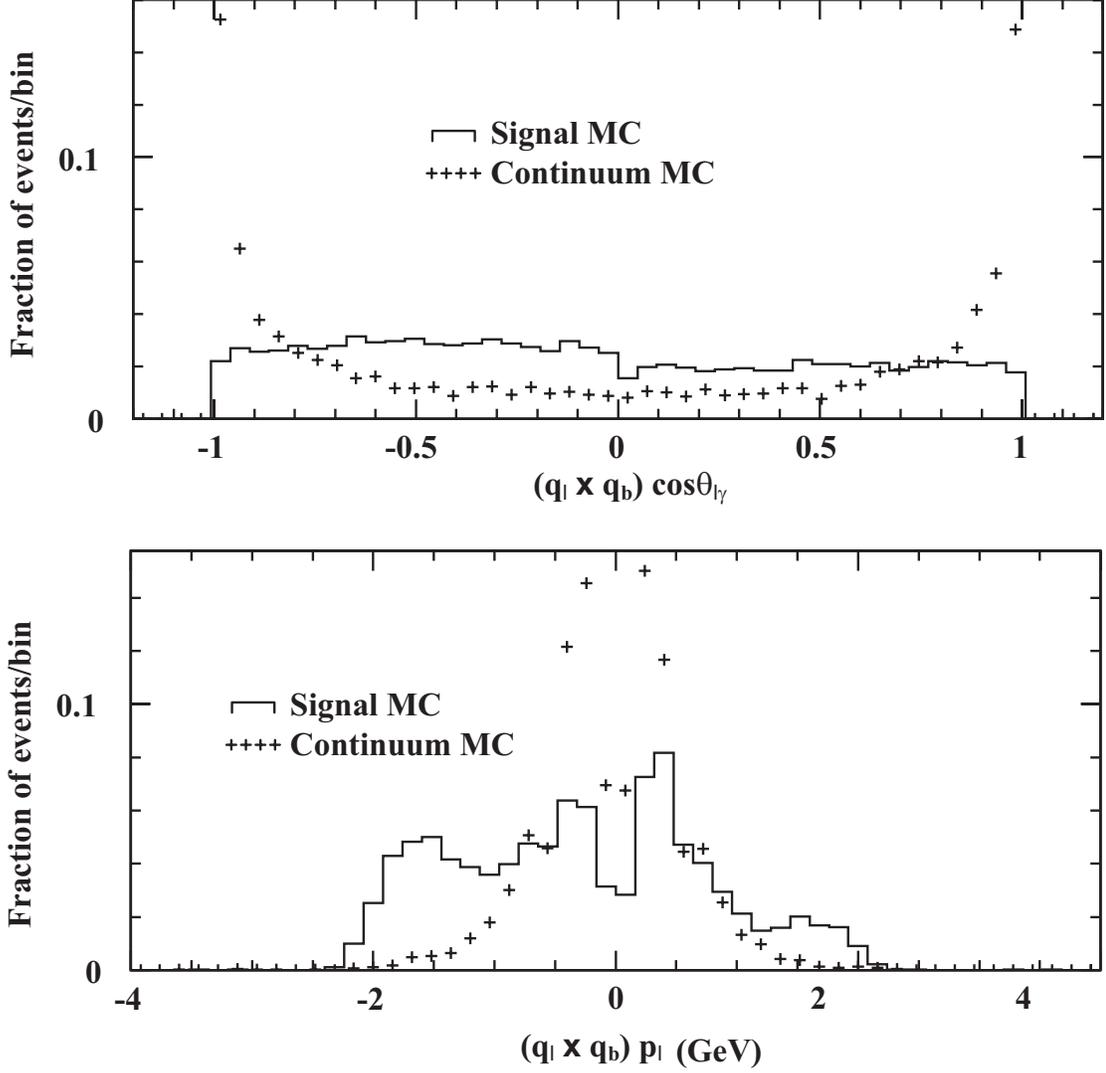
Figure 5.8: The top plot shows the distribution (for signal MC and continuum MC) of $|cos\theta_{l\gamma}|$ where $\theta_{l\gamma}$ is the angle between the highest momentum lepton in the non-signal B and the candidate photon. The quantity is signed by the charge of the lepton and the charge of the b quark ($q_l \times q_b$). The bottom plot is the distribution of the momentum of the highest momentum lepton for signal MC and continuum MC. The quantity is signed by the charge of the lepton and the charge of the b quark in the reconstruction ($q_l \times q_b$). These plots are for events that were reconstructed and had a leftover lepton.

- Pseudoreconstructed, with lepton

- Pseudoreconstructed, no lepton

- Not pseudoreconstructed, with lepton

- Not pseudoreconstructed, no lepton

Pseudoreconstructed		Not Pseudoreconstructed	
44%		56%	
Lepton	No lepton	Lepton	No lepton
10%	34%	15%	41%

Table 5.2: The percentage of signal MC events that fall into each of the 4 event categories.

The four event classes determine what information is available for making the best possible distinction between signal and continuum. We optimize a neural net for each of the four categories and run each event through the appropriate one. The inputs to each neural net are the variables described in Sections 5.1.1, 5.1.2, and 5.1.3. Before combining all of these variables using the neural net, we apply some loose shape cuts described in Appendix B to remove obvious background. Table 5.1.4 summarizes the four neural nets and the number and names of their input variables.

5.1.5   Continuum Subtraction

We developed a continuum suppression scheme to eliminate as much of the continuum background as possible by taking advantage of key differences between signal and continuum events. We combined this information using several neural nets and assigned a weight to each event. This weight is higher for signal events than continuum events.

To measure the On−Off subtracted yield, we use the full CLEO dataset (4s2 − 4sT) totaling 9.2 fb^{-1} on-resonance data and 4.6 fb^{-1} off-resonance data. We use the Off data to model the remaining continuum background and subtract its contribution from the On data. The On-Off scale factor that corrects for the difference in cross section and event shape which stems from the energy dependence is

$$\lambda = \frac{\mathcal{L}(On)}{\mathcal{L}(Off)} \times \frac{E_{Off}^2}{E_{On}^2} = 1.99.$$

Event Class	# of Inputs	Input Variables
Shape Analysis (not PR, no lepton)	8	R_2, R'_2, $S_\perp$, $|cos\theta'|$
PR, no lepton	3	r, χ^2_B, $|cos\theta_{tt}|$
PR, with lepton	5	r, χ^2_B, $|cos\theta_{tt}|$, $(q_l \times q_b) \times \theta_{l\gamma}$, $(q_l \times q_b) \times |\vec{p}_l|$
Not PR, with lepton	3	r, $\theta_{l\gamma}$, $|\vec{p}_l|$

Table 5.3: Summary of the four neural nets used in this analysis, one for each event category. PR stands for pseudoreconstructed events.

From the On data sample, we find a yield (summed weights) of 1861.7±16.5 between 2.0 and 2.7 GeV. Subtracting the continuum and including the error due to On−Off subtraction bias, we determine the On−Off subtracted yield to be 1101.4 ± 11.6. After this subtraction, we are left with $b \to s\gamma$ signal and $B\overline{B}$ backgrounds as shown in Figure 5.9.

On−Off Subtraction Bias

The yield is already corrected for the difference in cross section which results from the On−Off energy difference (λ). However, this correction does not account for differences between the On and Off continuum events that arise because of the different center of mass energies of the On and Off samples. For example, assuming that the On and Off events have the same multiplicity, then the average particle energy in On events will be higher by a factor of E_{On}/E_{Off}. This means that the candidate photons from the On continuum background are shifted upwards relative to candidate photons from the Off. Because the photon spectrum decreases rapidly as a function of energy, this introduces an under-subtraction. By scaling the energy of the candidate photons from the Off data by E_{On}/E_{Off}, we prevent this under-subtraction. Nevertheless, there remain smaller effects that we do not compensate for directly.

Because On continuum events are produced at a higher energy, they will have a "jettier" shape and be rejected more effectively by continuum suppression cuts. Also, the minimum energy allowed for a sibling photon used to veto candidates from π^0s or ηs is independent of run energy, making it more likely that a candidate from the On data will be rejected. Since the On−Off energy difference is small, we expect that these effects will also be small. We scale the energy of the candidate photon by E_{On}/E_{Off} to compensate, but we do not adjust the minimum energy for photon siblings used in the π^0/η veto and we do not adjust any of the neural net
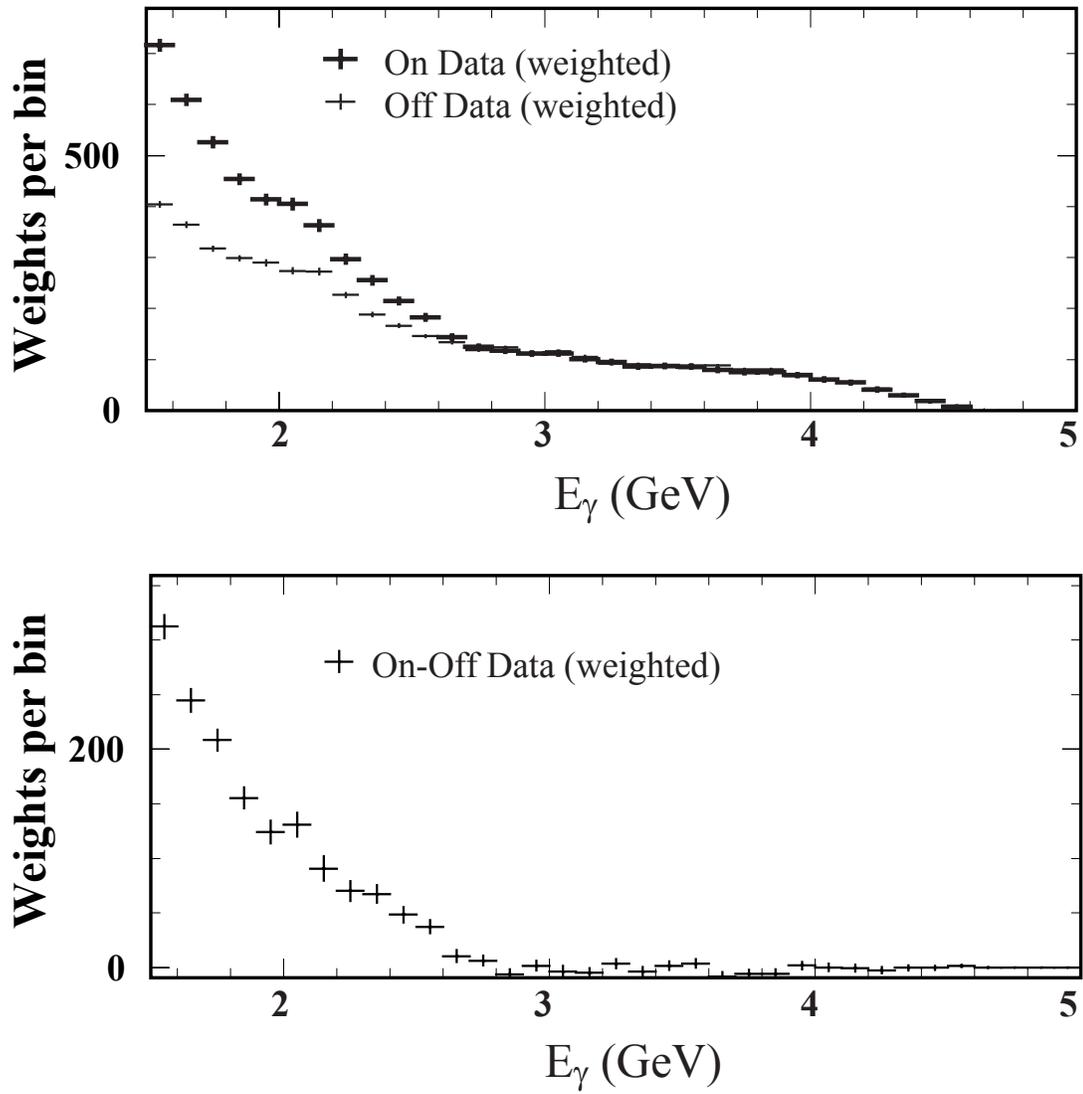
Figure 5.9: The upper plot shows the photon energy spectrum from on-resonance data (weighted) and off-resonance data (weighted). The lower plot is the On−Off subtracted spectrum from data.

inputs. Instead, we measure the effect of not making these corrections and apply a correction factor with a conservative error at the end. The yield in Off data is decreased by 0.5% of itself to account for the On−Off subtraction bias. All reports of the On−Off yield include the correction for this bias. The error is not included in the statistical error but is included in the systematic error.

5.2 B Backgrounds

Backgrounds due to $B\overline{B}$ decays are small relative to continuum processes, but they are non-negligible. These decays become especially important in the lower energy region, $1.5 - 2.2$ GeV, where the $b \to c$ backgrounds increase significantly. This section addresses the issue of estimating and subtracting the backgrounds due to B decays.

While the standard CLEO MC simulation provides a good model of the underlying particle reactions and the detector response, some processes of importance to this analysis are incorrectly modeled or absent. We start with a default MC sample, and use it to estimate the background due to $B\overline{B}$ decays.[7]

Ultimately, we subtract the estimated $B\overline{B}$ background yields from our On−Off subtracted spectrum to obtain the final $b \to s\gamma$ yield. The bottom plot of Figure 5.10 shows the photon energy spectrum that results after subtracting the default $B\overline{B}$ MC from the On−Off data. There is a huge excess in the lower energy region below 2.0 GeV, an indication that the default $B\overline{B}$ MC alone cannot accurately estimate the backgrounds due to B decays. To correct the MC, we must add the missing decays that could cause a high energy shower in the calorimeter and we must carefully match the detector response and yields in MC to actual yields from data.

5.2.1 π^0/η Corrections

As with continuum background, the B decay backgrounds are predominantly from π^0 and η decays that survive our veto. The photons enter the analysis in spite of our π^0/η vetoes either because we have not detected their sibling photon, or because they do not reconstruct well enough with their sibling to form a π^0 or an η. Table 5.4 shows the sources of photons above 1.5 GeV in a small MC sample of generic $B\overline{B}$ decays. Since these are the largest sources of backgrounds, our analysis depends on the MC estimates of the $B \to \pi^0 X_s$ and $B \to \eta X_s$ yields. By comparing the π^0 and η spectra in data and MC, we reduce our dependence on the MC prediction of our background.

---

[7]CLEO readers: Our default MC is not to be confused with the CLEO generic $B\overline{B}$ MC which contains only $b \to c$ decays. The default MC consists of generic $b \to c$ MC, $b \to u$ MC [47], and some charmless hadronic MC [48].

Photon Energy Spectrum - Uncorrected
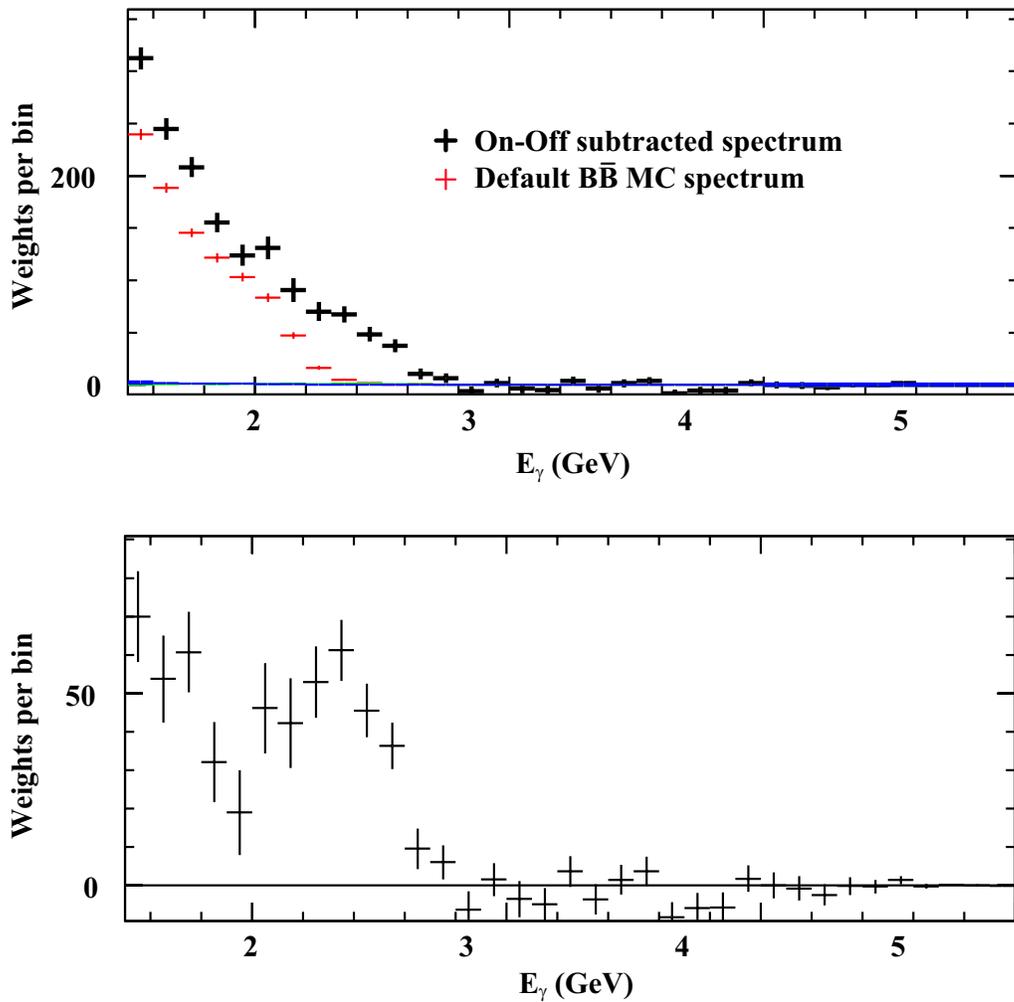
Figure 5.10: The upper plot shows the photon energy spectrum for all events. In this plot, the upper curve is the On−Off subtracted spectrum (data) while the lower curve is the default $B\overline{B}$ MC prediction for the B backgrounds. The bottom plot shows the spectrum that results if you subtract the two curves in the top plot. The default $B\overline{B}$ MC shows poor agreement below 2.0 GeV.

Source	# of γ from src		# passing π^0 veto	# passing η veto	# passing both	
TOTAL	**66583**		**19127**	**60757**	**16690**	
#γ from π^0	59773	90%	13147	55742	12269	74%
#γ from η	4477	7%	3948	2875	2551	15%
#γ from η'	274	0.4%	235	257	222	1%
#γ from $e^\pm$	1646	2%	1436	1516	1325	8%
#γ from ω	349	0.5%	309	308	273	2%
#γ from ψ, ψ', oth	64	0.1%	52	59	50	0.3%

Table 5.4: Sources of high energy photons in $B\overline{B}$ decays. Also included in the table is the number of photons that pass the π^0 and η vetoes applied individually and applied together. The percentages are the number of photons from a given source out of the total number of photons in the sample, before and after the vetoes are applied.

Our strategy is to perform the $b \rightarrow s\gamma$ analysis using π^0s and ηs rather than photons as the candidate particles. The rates for $B \rightarrow \pi^0 X_s$ and $B \rightarrow \eta X_s$ are measured from the data using exactly the same method as is used for $B \rightarrow \gamma X_s$, but without the π^0 and η vetoes. We accept as candidates those events with π^0s reconstructed between $110 - 150$ MeV/c^2 or ηs reconstructed between $515 - 575$ MeV/c^2. Using this method, events that are reconstructed as $B \rightarrow \gamma\gamma X_s$ are subject to the same biases as events reconstructed as $B \rightarrow \gamma X_s$ allowing us to observe the π^0/η spectrum in the same region of phase space that we use for the $b \rightarrow s\gamma$ spectrum.

To form π^0/η candidates, we select showers based on the same criteria listed in Chapter 4 and look for combinations of two photons that fall within the π^0 or η mass window selecting candidates with $1.5 < E_{\pi^0/\eta} < 5.0$ GeV for $|cos\theta| < 0.7$. We also require that the candidates not be matched to a charged track.

We compute the eight shape variables, attempt to reconstruct a $B \rightarrow \pi^0/\eta X_s$ event, and search for leptons in the non-signal B. As before, events are split into their four event classes and neural net outputs are computed for each. Finally, we obtain the weighted sum of the $B \rightarrow \pi^0 X_s$ and $B \rightarrow \eta X_s$ events in each 100 MeV bin between 1.5 and 2.7 GeV.

Random combinations of two photons in the event can fall within the allowed mass window for the π^0/η even if these two photons do not really come from a π^0/η decay. The $\pi^0 X_s$ and ηX_s yields are obtained from the raw $\gamma\gamma X_s$ yields by fitting the $\gamma\gamma X_s$ spectrum to a gaussian plus a simple background shape and removing

the combinatoric background in each 100 MeV bin between $1.5 - 2.7$ GeV for the On$-$Off data and default $B\overline{B}$ MC. In this way, we determine the true $B \to \pi^0 X_s$ and $B \to \eta X_s$ yields for On$-$Off Data and default $B\overline{B}$ MC in each of the 100 MeV energy bins between 1.5 and 2.7 GeV. Figure 5.11 shows the fraction of π^0 candidates that are true π^0s as a function of π^0 energy for the On$-$Off data. Figures 5.12 and 5.13 show the On$-$Off subtracted π^0 and η spectra in data and in MC. In each energy bin, we determine the ratio of data to MC, the factor by which we must adjust the MC to make it agree with the On$-$Off data.

Fraction of π^0 candidates that are true π^0s vs. $\mathbf{E_{\pi^0}}$
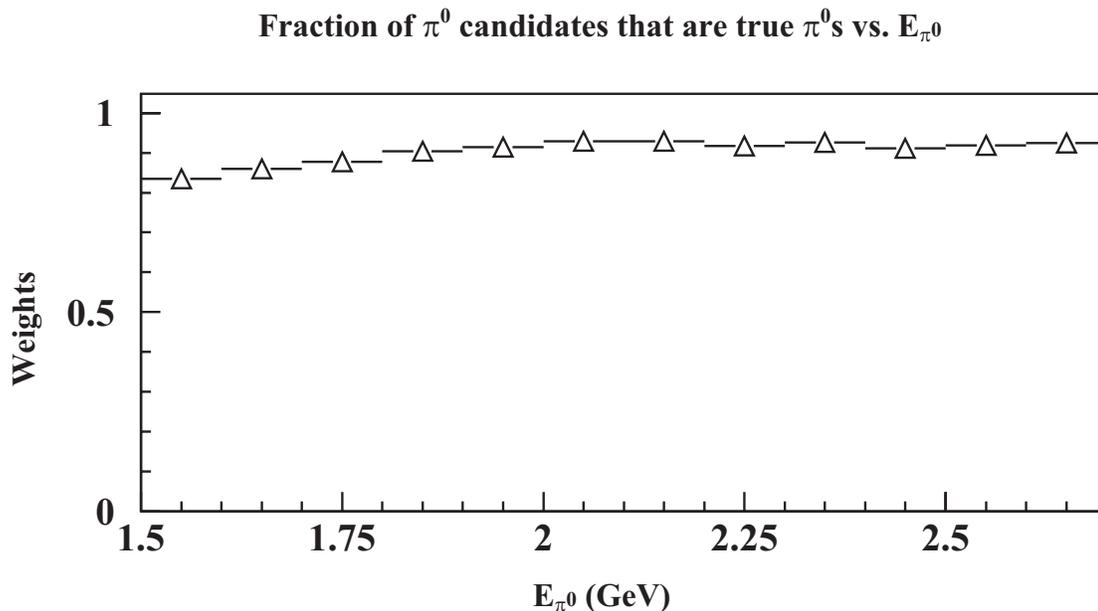
Figure 5.11: This figure shows the fraction of π^0 candidates that are true π^0s as a function of π^0 energy. This is the result for On data only.

We translate the deficiency of the π^0/η spectrum to a correction in the photon energy spectrum by running the standard analysis on the default $B\overline{B}$ MC sample and searching for candidate photons as before. For each candidate photon, we determine whether it came from a π^0/η and if so, record E_γ and $E_{\pi^0/\eta}$. By doing this, we are examining the sample of photons coming from π^0s and ηs that have slipped through our vetoes, exactly the background we are trying to correct. Then, for each candidate photon at a given energy, E_γ, we look at the energy of the π^0/η and scale the number of candidate photons at the energy E_γ by the value of the
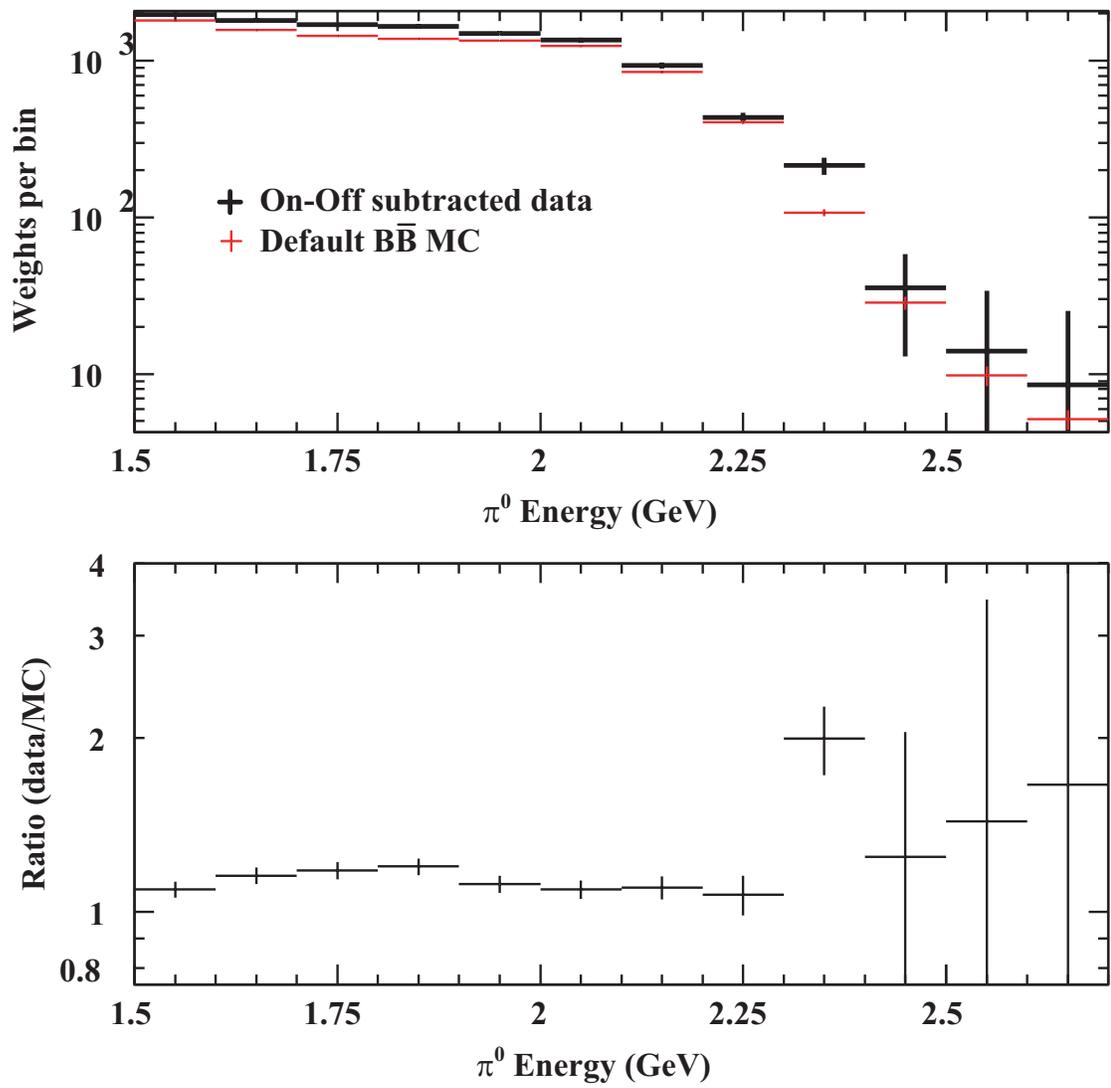
Figure 5.12: The top plot shows the π^0 spectra in On−Off subtracted data (upper curve) and in MC (lower curve). The bottom plot shows the ratio of data to MC as a function of π^0 energy.

69
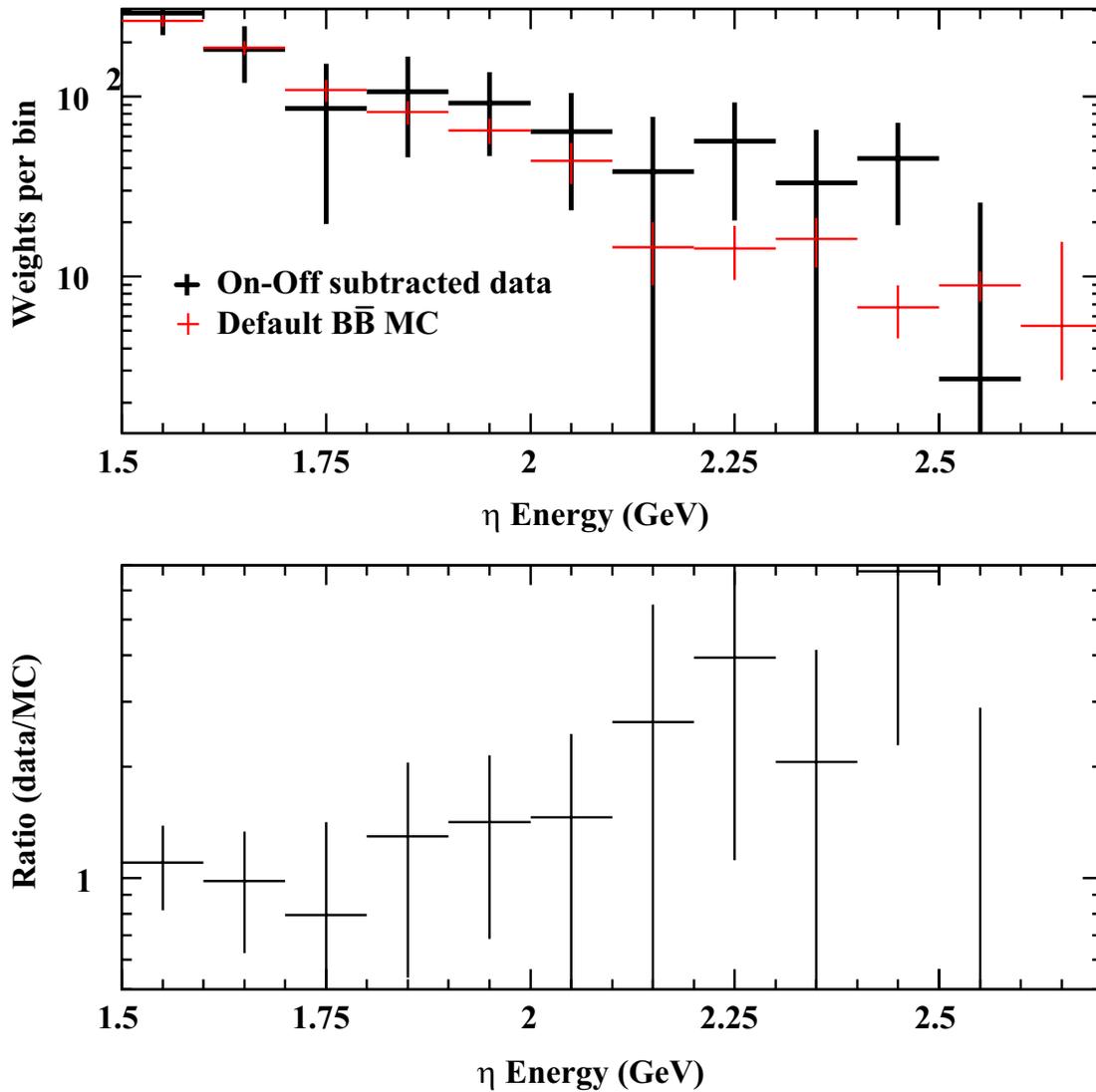
Figure 5.13: The top plot shows the η spectra in On−Off subtracted data (upper curve) and in MC (lower curve). The bottom plot shows the ratio of data to MC as a function of η energy.

Energy Range	1.50 − 1.80	1.80 − 2.00	2.00 − 2.20
π^0 Correction	58.32 ± 2.16	24.63 ± 1.60	12.27 ± 1.41
η Correction	11.34 ± 4.05	16.11 ± 2.46	23.06 ± 3.01

Energy Range	2.00 − 2.70	2.10 − 2.70	2.20 − 2.70
π^0 Correction	18.39 ± 1.55	10.68 ± 1.06	6.16 ± 0.66
η Correction	34.28 ± 3.39	19.26 ± 2.13	11.21 ± 1.56

Energy Range	2.70 − 3.00	3.00 − 4.00	4.00 − 5.00
π^0 Correction	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
η Correction	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

Table 5.5: Size of π^0 and η correction for a range of energy bins.

ratio of data to MC in that π^0/η energy bin. We apply the π^0 and η corrections to the MC and obtain the spectrum shown in Figure 5.14 for B backgrounds. The size of the individual π^0 and η corrections is given in Table 5.5 and the total correction from π^0s and ηs is listed in Table 5.6.

5.3   Neutral Hadrons

In the previous section, we found that our default MC sample was deficient in the number of π^0s and ηs present at certain energies. In general, MC can differ from the real (measured) data in two ways, either in the numbers and momentum of the particles generated or in the simulation of the interactions of these particles with the detector.

Even after the π^0/η corrected $B\overline{B}$ MC has been subtracted from the On−Off data, the resulting photon energy spectrum, shown in Figure 5.14, has an excess of photons between 1.5 and 1.8 GeV. At first, we believed that this excess was due to the fact that we were missing some physics process in our default $B\overline{B}$ MC sample that produced a large number of photons with an energy between 1.5 − 1.8 GeV. We estimated the effect of adding several modes (Appendix C) not included in the original $B\overline{B}$ MC sample and concluded that the background was not from photons at all, but from neutral hadrons that caused showers in the calorimeter that could be mistaken for photons. Although our particle generator, QQ, was doing a decent job in predicting the numbers and momentum of the particles, CLEOG was incorrectly modeling the behavior of the detector. This caused us to incorrectly estimate the

Photon Energy Spectrum - with π^0/η Correction

+ **On-Off subtracted spectrum**
+ **Default B$\overline{\text{B}}$ MC spectrum corrected for π^0/η yield**
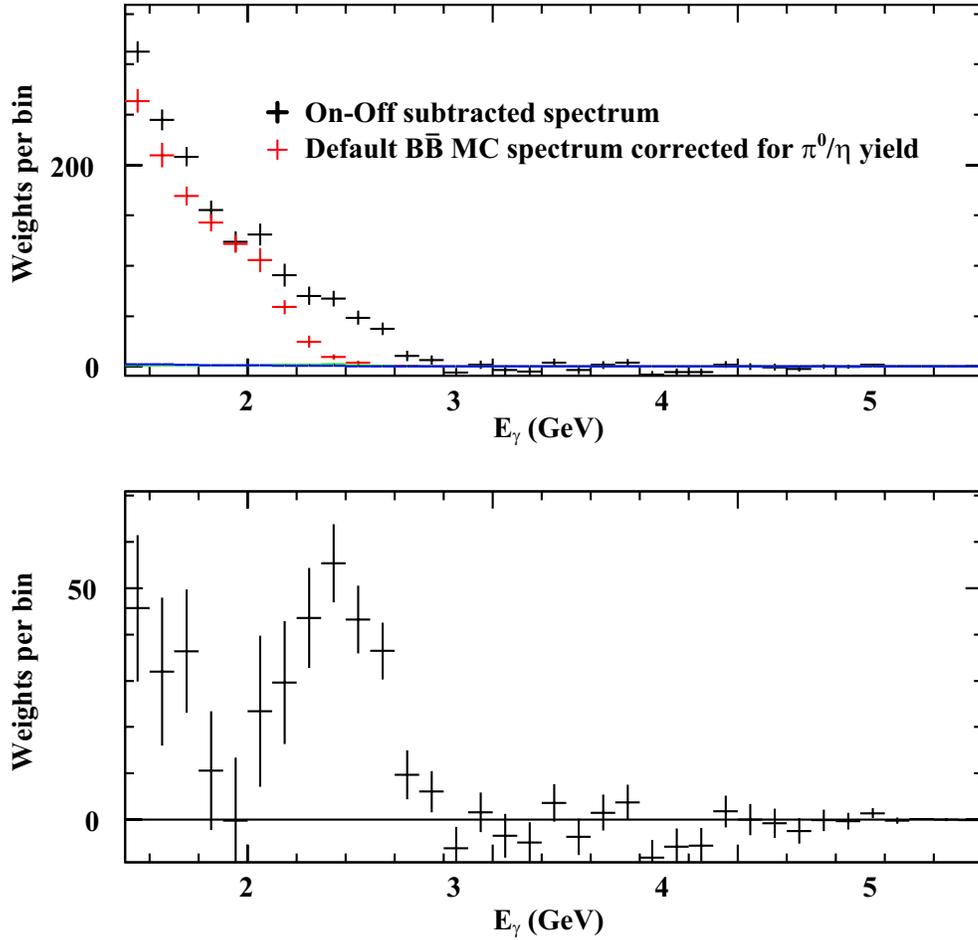
Figure 5.14: The top plot shows the weighted E_γ spectra for all events. The upper curve is On−Off data, and the lower curve is the photon energy spectrum from the π^0/η corrected $B\overline{B}$ MC. We subtract the corrected $B\overline{B}$ MC from the data and obtain the lower plot. The signal region $(2.0 - 2.7 \text{ GeV})$ looks very nice, but the region of the spectrum below 2.0 GeV, where we expect our $b \rightarrow s\gamma$ yield to be small, still looks inconsistent with zero. There are still more corrections that need to be made to the MC.

photon yield between $1.5 - 1.8$ GeV in the default $B\overline{B}$ MC. Consequently, we developed a technique that reduced our dependence on MC for the purpose of estimating the background from neutral hadrons.

The neutral particles that we can detect in the crystals are photons (γ), long-lived neutral kaons (K_L), and anti-neutrons ($\overline{n}$). K_L and $\overline{n}$ showers are usually rejected by our E9/E25 cut (E9/E25 > 0.95). Photons react electromagnetically with the CsI crystals and create a well-contained shower with an E9/E25 value that is close to 1. Anti-neutrons annihilate in the crystals, and annihilation showers tend to be farthest from E9/E25 = 1. K_Ls that hit the crystals will be subject to a number of inelastic nuclear interactions. Given the length of the crystals and the nuclear interaction length of K_L in CsI, a K_L will enter the calorimeter and interact somewhere within the CsI crystal about 56% of the time.[8] Depending on where it decays, it can leave some, all, or none of its energy in the crystal. In addition, a K_L can also regenerate to a K_s and then decay to two π^0s somewhere within the crystal, leaving a shower similar to a photon.

We compare the E9/E25 distribution in On−Off data to the E9/E25 distribution in $B\overline{B}$ MC, the distributions of which are shown for $1.5 < E_\gamma < 1.8$ GeV in Figure 5.15. Although the data and MC are in good agreement for E9/E25 > 0.9, at values of E9/E25 below 0.9 the agreement is dismal. We have our smoking gun.

The data rises with E9/E25 while the MC remains relatively flat (excluding the spike due to photons at $0.95 - 1.00$). Figure 5.16 shows the E9/E25 distributions for γ, K_L, and $\overline{n}$ from $B\overline{B}$ MC. From the E9/E25 distributions shown in Figure 5.15, the "flatness" in the MC must be caused by some combination of the $\overline{n}$ and K_L component. This comparison leads us to believe that CLEOG is not simulating the E9/E25 distribution of the neutral hadrons correctly. Either the data has more K_L and less $\overline{n}$ than the MC or the true $\overline{n}$ E9/E25 distribution is peaked much higher in E9/E25 than it is in MC. For the branching fraction, we don't care which it is.

We assume that the true E9/E25 distribution for neutral hadrons is a linear combination of the three MC distributions (γ, $\overline{n}$, K_L) pictured in Figure 5.16. We correct the E9/E25 distribution by fitting the distribution for γ, K_L, and $\overline{n}$ in $B\overline{B}$ MC to the On−Off subtracted data in four separate photon energy bins: $1.5 - 1.8$ GeV, $1.8 - 2.0$ GeV, $2.0 - 2.2$ GeV, and $2.2 - 2.7$ GeV. We fit the E9/E25 distribution in four bins between $0.8 <$ E9/E25 < 1.0 and obtain the factors by which the MC E9/E25 distribution in a given energy bin must be adjusted to agree with the data. We correct the E9/E25 distribution of the MC and obtain a predicted background yield from K_L and $\overline{n}$ to the photon energy spectrum.

We eliminate the contribution from γ using the knowledge that the $0.95 - 1.00$ bin contains only photons and that the $0.9 - 0.95$ photon bin contains 4% of the

---

[8]The CsI crystals at CLEO are about 30 cm in length and the nuclear interaction length of a K_L in CsI is 36.9 cm. Therefore, the probability that a K_L will interact within a crystal is given by $P_{int} = 1 - e^{-\frac{x}{\lambda}} = 1 - e^{-\frac{30}{36.9}} = 56\%$.
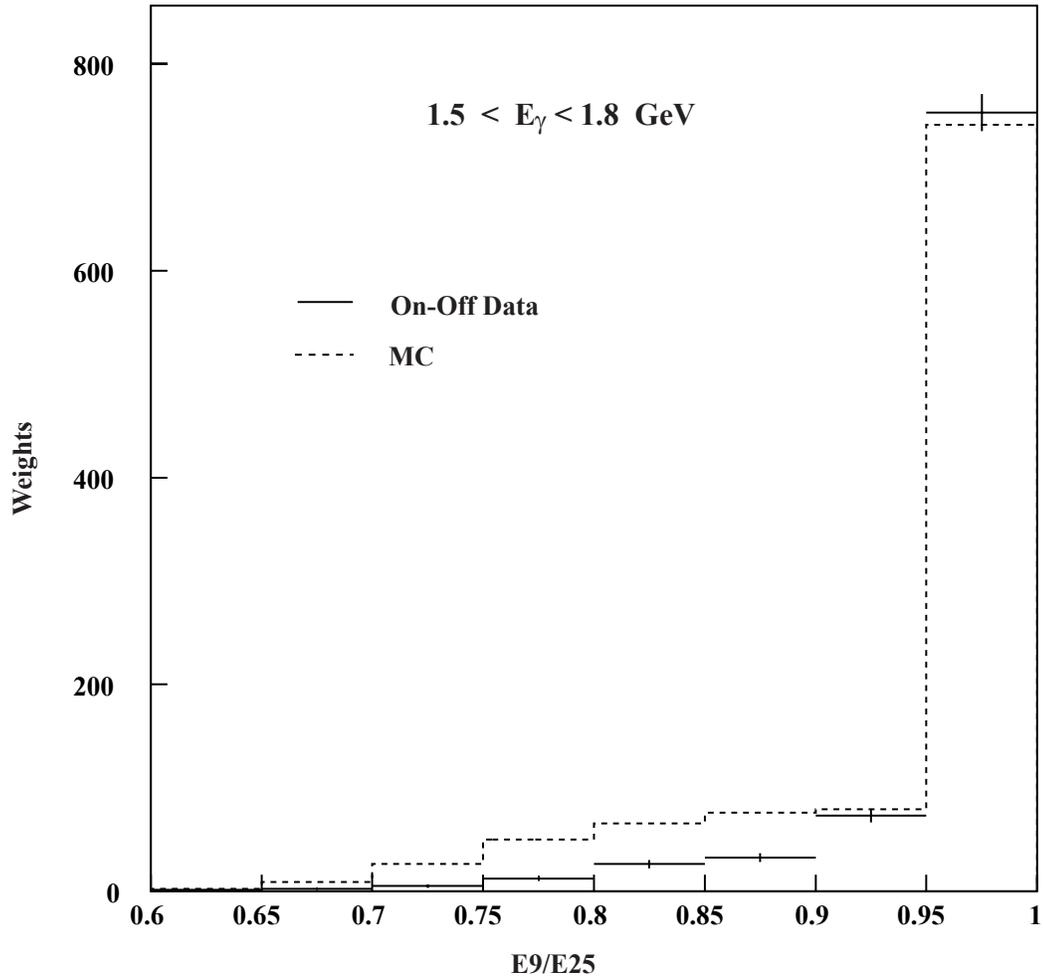
Figure 5.15: Comparison of E9/E25 distributions for On−Off subtracted data and $B\overline{B}$ MC for 1.5 GeV $< E_\gamma <$ 1.8 GeV. The Monte Carlo normalization is arbitrary, and the plots were made without applying any shape cuts or shower mass cuts on the showers.
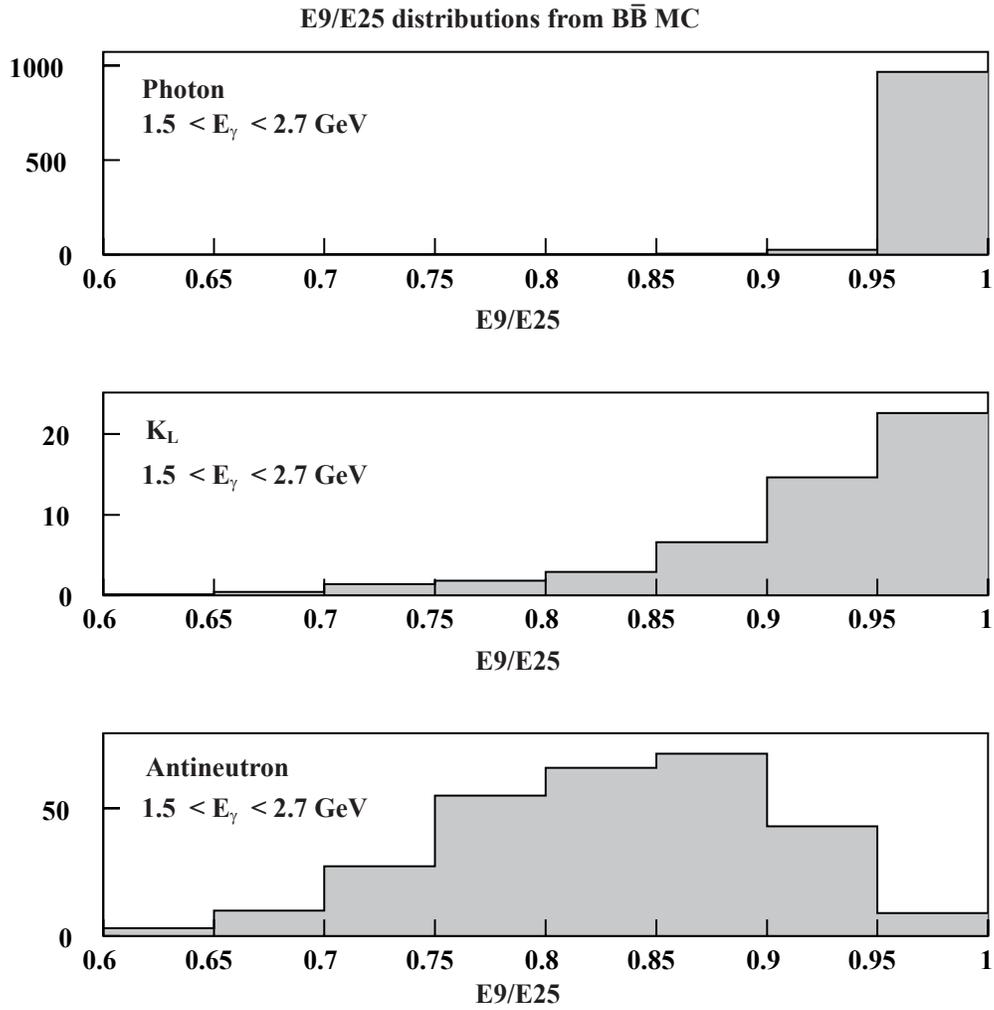
Figure 5.16: E9/E25 distributions from $B\overline{B}$ MC for γ, K_L, and $\overline{n}$ for the full photon energy range, $1.5 - 2.7$ GeV.

$0.95 - 1.00$ bin. The K_L and $\overline{n}$ contribute 65.8 ± 25.17 weights to the $1.5 - 1.8$ GeV bin and no weights to the $2.0 - 2.7$ GeV bin. The $B\overline{B}$ MC predictions, and the "measured" K_L and $\overline{n}$ background yields can be found in Table C.4 in Appendix C.

5.3.1 The Many Small Contributions

Several modes capable of producing high energy photons were not included in our default $B\overline{B}$ MC sample. We evaluated the contributions of these modes to our yield, and found that most of them were very small. Nevertheless, we correct for them in the analysis and describe them in Appendix C for completeness.

Although we call this analysis $b \to s\gamma$, we actually make no distinction in the code between $b \to s\gamma$ and $b \to d\gamma$. We are sensitive to both and will make a correction to the yield to remove the $b \to d\gamma$ contribution in Chapter 7.

5.3.2 The Control Region: $1.5 - 2.0$ GeV

Our best efforts to determine the $B\overline{B}$ backgrounds is shown in Figure 5.17. There is a clear $b \to s\gamma$ peak and little evidence of other $B\overline{B}$ processes. We must now estimate how well we have determined the $B\overline{B}$ background to determine a systematic error on the subtraction. For this, we use the $1.5 - 2.0$ GeV region as a control region.

Some of the $b \to s\gamma$ signal will spill over into the $1.5 - 2.0$ GeV region, and Kagan and Neubert claim that 91.5% of the $b \to s\gamma$ spectrum lies above 2.0 GeV [19]. As a reasonable approximation, we take 4% as the fraction lying between 1.5 and 2.0 GeV with 4.5% lying below 1.5 GeV. If we have done the background subtractions correctly, the sum of the weights between $1.5 - 2.0$ GeV minus 4% of the sum of the weights between 2.0 and 2.7 GeV should be zero.

Our $B\overline{B}$ subtraction gives us a difference of 19.83 ± 42.29 weights, an excess of 0.47σ above zero. If we subtract 5% more $B\overline{B}$ background, we have -30.46 ± 42.29 weights, a deficit of 0.72σ. If we subtract 5% less $B\overline{B}$ background, we have 70.11 ± 42.29 weights for an excess of 1.66σ. We conclude from this that our $B\overline{B}$ subtraction is reasonable, bringing the control region to within 0.5σ of zero. Therefore, a $\pm 5\%$ variation in the $B\overline{B}$ subtraction is a good estimate for a 1σ systematic error in the subtraction.

After subtracting the continuum and $B\overline{B}$ backgrounds, the yield for $b \to s\gamma$ between $2.0 - 2.7$ GeV is measured to be 233.61 ± 31.17 weights, where the error is statistical and includes the error on the π^0/η and neutral hadron corrections to the background. The final yields are summarized in Table 5.6.
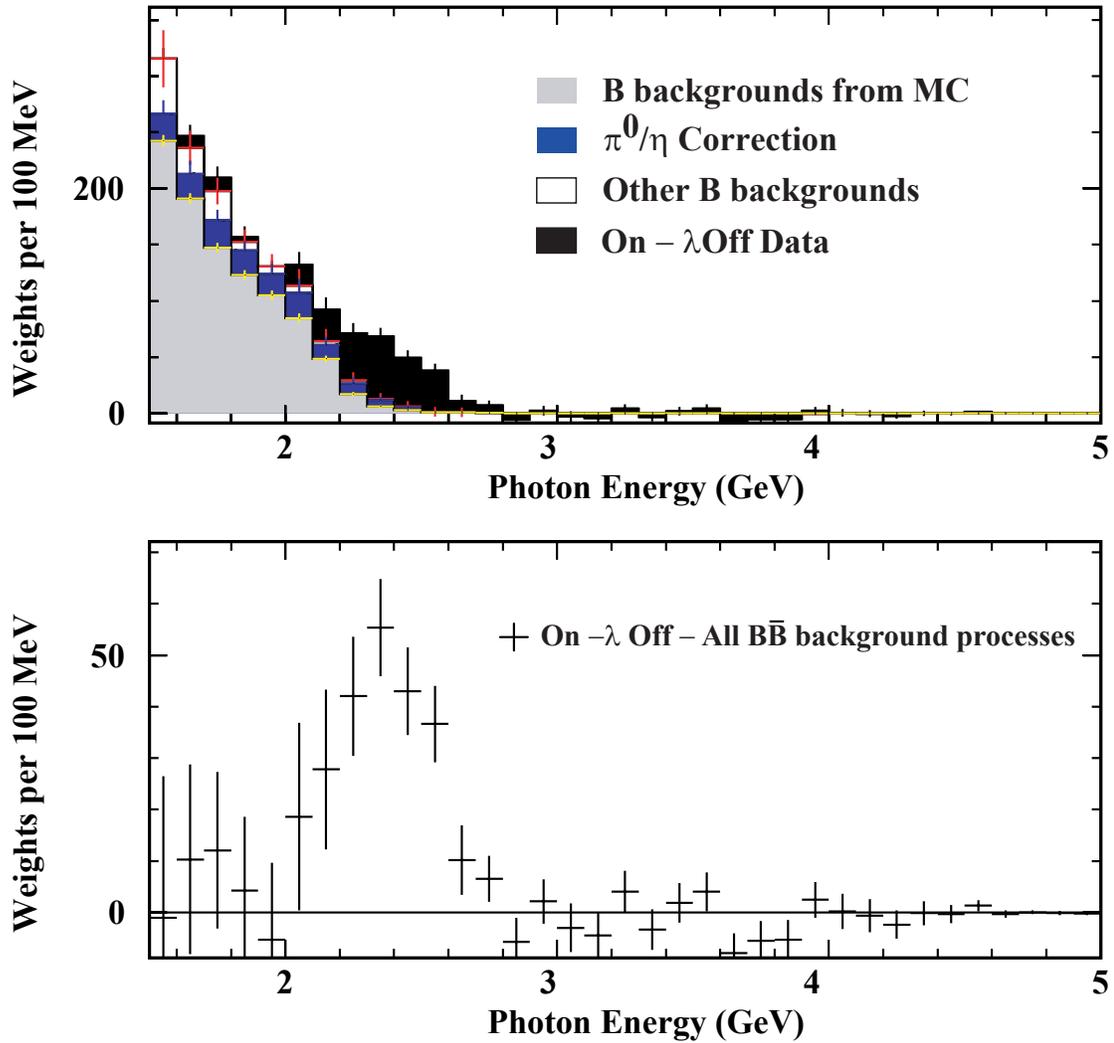
Figure 5.17: The upper plot is an overlay of the various contributions to the final yield. The largest contribution of the four is MC prediction for the B backgrounds. Next, is the π^0/η correction, followed by the "other" B backgrounds. The greatest contribution is the On−Off subtracted yield. The lower plot shows the yield after all of the background processes are subtracted from the On−Off yield.

Energy Range	1.50 − 1.80	1.80 − 2.00	2.00 − 2.20
ON	1851.34 ± 13.72	867.47 ± 10.38	767.98 ± 11.84
$\lambda \times$OFF	1080.43 ± 11.18	585.19 ± 9.43	543.70 ± 10.36
ON$-\lambda \times$OFF	**770.91 ± 17.70**	**282.28 ± 14.02**	**224.28 ± 15.74**
$b \to c$	573.65 ± 8.07	224.61 ± 5.83	130.41 ± 5.22
non-$b \to c$ had	2.57 ± 0.17	1.64 ± 0.15	1.56 ± 0.16
$b \to u$ semilep	4.86 ± 0.24	1.94 ± 0.18	1.26 ± 0.15
π^0/η Correction	70.33 ± 17.49	40.74 ± 10.82	35.34 ± 12.91
Other bkds	98.16 ± 25.42	14.45 ± 9.14	9.30 ± 11.48
MC bkd Sum	**749.57 ± 31.89**	**283.39 ± 15.32**	**177.88 ± 18.05**
Yield	**21.34 ± 36.48**	**-1.11 ± 20.76**	**46.40 ± 23.95**

Energy Range	2.00 − 2.70	2.10 − 2.70	2.20 − 2.70
ON	1861.65 ± 16.45	1456.71 ± 14.04	1093.67 ± 11.42
$\lambda \times$OFF	1399.40 ± 15.31	1126.77 ± 13.68	855.69 ± 11.26
ON$-\lambda \times$OFF	**462.25 ± 22.47**	**329.94 ± 19.60**	**237.98 ± 16.04**
$b \to c$	153.89 ± 5.64	70.46 ± 3.75	23.47 ± 2.14
non-$b \to c$ had	5.41 ± 0.28	4.71 ± 0.27	3.85 ± 0.23
$b \to u$ semilep	2.07 ± 0.17	1.39 ± 0.15	0.80 ± 0.08
π^0/η Correction	52.62 ± 14.57	29.90 ± 9.25	17.30 ± 6.77
Other bkds	14.65 ± 14.90	8.49 ± 12.47	5.35 ± 9.49
MC bkd Sum	**228.64 ± 21.59**	**114.95 ± 15.98**	**50.77 ± 11.85**
Yield	**233.61 ± 31.17**	**214.99 ± 25.29**	**187.21 ± 19.94**

Energy Range	2.70 − 3.00	3.00 − 4.00	4.00 − 5.00
ON	353.74 ± 4.43	867.47 ± 7.14	215.69 ± 3.60
$\lambda \times$OFF	350.38 ± 6.34	884.36 ± 10.43	217.01 ± 5.16
ON$-\lambda \times$OFF	**3.36 ± 7.74**	**-16.89 ± 12.64**	**-1.32 ± 6.29**
$b \to c$	0.13 ± 0.06	0.00 ± 0.00	0.04 ± 0.04
non-$b \to c$ had	0.11 ± 0.04	0.00 ± 0.00	0.00 ± 0.00
$b \to u$ semilep	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.01
π^0/η Correction	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
Other bkds	0.17 ± 0.25	0.35 ± 0.57	1.13 ± 0.96
MC bkd Sum	**0.43 ± 0.26**	**0.35 ± 0.57**	**1.18 ± 0.96**
Yield	**2.93 ± 7.74**	**-17.25 ± 12.65**	**-2.50 ± 6.36**

Table 5.6: Yields for a range of energy bins. The yield in Off data has been decreased by 0.5% of itself to account for the On−Off subtraction bias.

CHAPTER 6

Signal Efficiency

After subtracting the continuum and $B\overline{B}$ backgrounds, the yield for $b \to s\gamma$ between $2.0 - 2.7$ GeV is measured to be 233.61 ± 31.17 weights, where the error is statistical. Statistical errors are uncertainties in the measurement that come from fluctuations in the data and depend on the number of events in the sample. In the limit of an infinite number of events, the statistical fluctuations would disappear and the statistical error would go to zero. However, a measurement is also subject to systematic biases and uncertainties which affect the measured results. These uncertainties may derive from detector issues, from the tracking and shower reconstruction programs, or from the analysis procedure itself.

In this analysis, Monte Carlo simulations (MC) are used to estimate the $B\overline{B}$ backgrounds and to calculate the signal detection efficiency. We count on the MC to accurately model the physical process we are trying to measure, and we rely on detector simulation to perform particle detection with the correct efficiencies and smearing. Any inaccuracy in the MC in either of these two areas will lead to systematic uncertainties. This chapter describes how we model the $b \to s\gamma$ process, how the signal efficiency is determined, and how systematic uncertainties are assigned.

6.1   Modeling $b \to s\gamma$

To model the decay $b \to s\gamma$, we use the spectator model of Ali and Greub, which includes gluon bremsstrahlung and higher-order radiative effects, in order to determine the kinematic acceptance of photons in the window from 2.0 to 2.7 GeV [49]. In this particular model, the b quark is assumed to decay independently of the $\overline{u}$ or $\overline{d}$ spectator quark in the B meson. The two input parameters to the spectator model are the average b-quark mass, $\langle m_b \rangle$, and the momentum of the quarks in the B meson frame, given by a Gaussian distribution whose width is defined as the Fermi momentum, p_F.

Bin #	Decay Mode	Mass Range (MeV/c^2)
1	$B \to K^*(890)\gamma$	$M_{X_s} \leq 1.08$
2	$B \to K^*(1270)\gamma$	$1.08 < M_{X_s} \leq 1.35$
3	$B \to K^*(1430)\gamma$	$1.35 < M_{X_s} \leq 1.56$
4	$B \to K^*(1680)\gamma$	$1.56 < M_{X_s} \leq 1.73$
5	$B \to K^*(1780)\gamma$	$1.73 < M_{X_s} \leq 1.91$
6	$B \to K^*(2045)\gamma$	$1.91 < M_{X_s} \leq 2.15$
7	$B \to K^*(2250)\gamma$	$2.15 < M_{X_s} \leq 2.35$
8	$B \to K^*(2450)\gamma$	$M_{X_s} > 2.35$

Table 6.1: Mass bins used for efficiency estimation with the spectator model. The upper (lower) limit for each mass range is exactly halfway between the mass of the K^* mode in question and the mass of the next heaviest (lightest) K^* resonance.

Since $b \to s\gamma$ is a two-body decay, the photon energy (in the rest frame of the B meson) and the recoil mass M_{X_s} are related by the expression

$$E_\gamma^{B\ rest\ frame} = \frac{M_B}{2} - \frac{M_{X_s}^2}{2M_B}$$

where M_B is the mass of the B meson. A model needs to specify either the photon energy distribution *or* the recoil mass distribution. As an example, the photon energy distributions and X_s mass distributions for several values of the spectator model parameters are shown in Figure 6.1. The parameter $\langle m_b \rangle$ varies the mean while p_F varies the width of the photon energy spectrum.

Using this model, we extract spectra for the mass of the strange system (X_s) and for the energy of the emitted photon in the lab frame for 24 ($\langle m_b \rangle$, p_F) pairs. We have used two approaches to hadronize the $s\bar{q}$ pair. In the first approach, we assume that the system recoiling against the photon hadronizes into the nearest spin-1 kaon resonance, and chop the X_s mass spectrum into several exclusive modes listed in Table 6.1. We generated MC for all eight of these exclusive modes. In the second approach, we let JETSET hadronize the $s\bar{q}$ resonance to give the Ali-Greub X_s mass distribution for all 24 sets of ($\langle m_b \rangle$, p_F) parameters.

We use these MC samples to obtain the efficiency for each of the 24 models and for each of the two hadronization methods. From this data, we calculate our final efficiency and systematic errors.
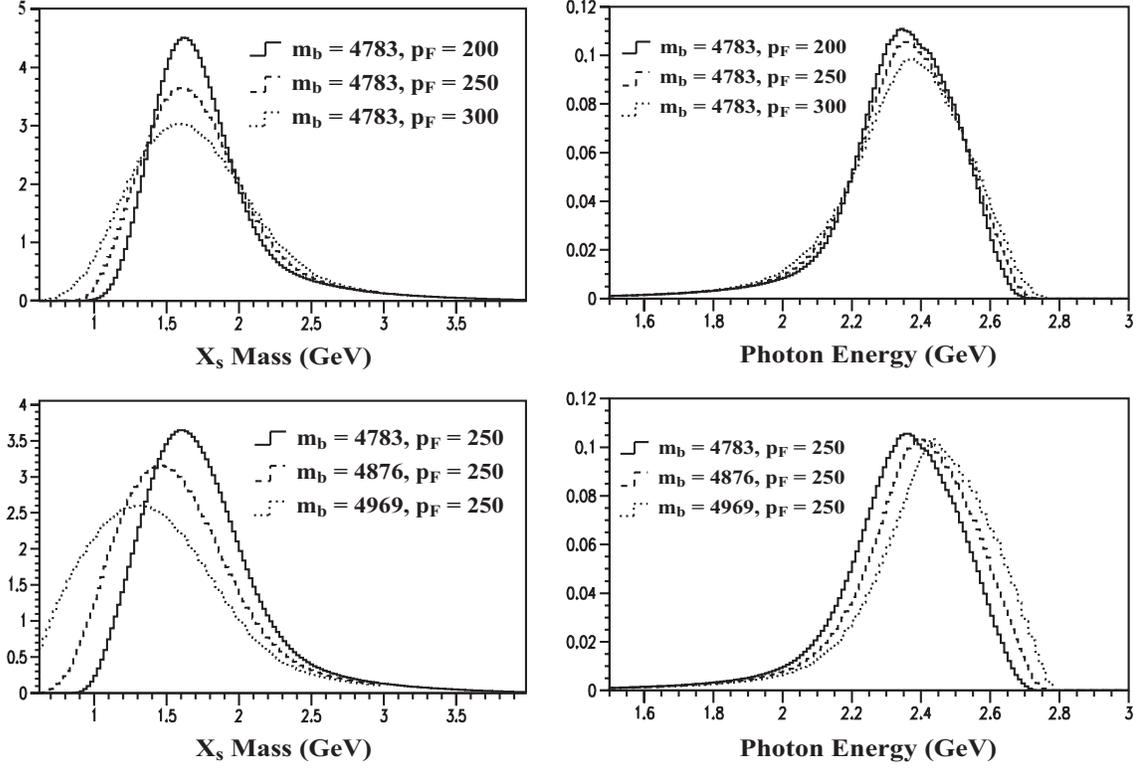
Figure 6.1: The left plots show recoil mass distributions from Ali and Greub in the rest frame of the B meson for several values of the spectator model parameters. The plots on the right show photon energy distributions from Ali and Greub. The upper plots show that the width of the recoil mass spectrum is determined mainly by p_F. The lower plots show that the peak of the mass distribution is controlled by $\langle m_b \rangle$. The fraction of photons in the $2.0 - 2.7$ GeV window is mainly sensitive to $\langle m_b \rangle$ and to a lesser extent p_F.

6.2   Definition of Weighted Efficiency

In a typical sample of $b \rightarrow s\gamma$ MC signal events, about half pass all of the analysis cuts, including a photon energy cut. However, because we are weighting our candidate events, this efficiency is not what we need for determining the branching fraction. Instead, we define a *weighted* efficiency, ϵ_w. For a given Monte Carlo signal sample, we sum the weights of the events passing the analysis cuts and divide by the total number of events in the sample. This average weight per signal event is our weighted efficiency, ϵ_w.

When we calculate the weighted efficiency, we find that the efficiency is higher for events with a low X_s mass than for events with a high X_s mass. Using our knowledge of the dependence of ϵ_w on X_s mass, we adjust the weight for each event based on the reconstructed X_s mass for that event, ultimately reducing the dependency of the efficiency on the X_s mass. The method of adjusting the weights for each event is described in Appendix D.

To calculate our efficiency, we sum these adjusted or "flattened" weights and divide by the total number of events in the sample to obtain the flattened weighted efficiency, ϵ_{flat}. Since our final yields are based on the sums of the flattened weights, all references to yields, weighted sums, efficiencies, ϵ, or weighted efficiencies, refer to ϵ_{flat} unless otherwise stated.

To obtain an efficiency for the inclusive process from the exclusive modes, a weighted sum is performed over the K^* modes, with the efficiency for each exclusive mode weighted by the fraction of the total rate expected per respective mass bin. The mass-bin weights (not to be confused with *event* weights) obtained from the Ali-Greub spectator model are listed in Table 6.2. The mass-weighted efficiency is obtained for several sets of model parameters to help us determine a systematic uncertainty to the overall efficiency due to hadronization effects. We refer to this as our "K^* sum" sample or method. The inclusive efficiencies resulting from the sum of exclusive efficiencies in Table 6.3 weighted by the mass-bin weights given in Table 6.2 are shown in Table 6.5.

6.2.1   Sources of X_s mass dependence

The spectator model affects our answer through the efficiency. The efficiency for $b \rightarrow s\gamma$ can be divided into two separate components. The first component is the fraction of $b \rightarrow s\gamma$ events in which the photon lies between 2.0 and 2.7 GeV. The second component is the probability that we will detect an event, given that its photon is between 2.0 and 2.7 GeV. The efficiency needed to compute the branching fraction for $b \rightarrow s\gamma$ is the product of these two components.

Both the fraction of events falling between $2.0 - 2.7$ GeV and our ability to detect an event, given that the photon lies between $2.0 - 2.7$ GeV, are dependent on the X_s mass. Evidence that the fraction of events between $2.0 - 2.7$ GeV has a mass

$< m_b >$	p_F	Fraction in Mass Bin(%)							
MeV/c^2	MeV/c	1	2	3	4	5	6	7	8
4597	390	2.24	9.04	15.47	11.76	16.34	18.02	11.68	15.44
	440	4.47	11.06	15.13	10.61	14.43	16.22	11.20	16.88
	260	0.31	6.86	19.64	16.48	21.25	18.28	8.14	9.04
	310	1.93	10.39	19.22	14.08	18.05	17.01	8.98	10.33
4690	360	4.13	13.11	18.11	12.19	15.59	15.69	9.33	11.84
	410	7.42	14.60	16.58	10.63	13.66	14.39	9.35	13.36
	200	0.23	9.85	27.51	19.27	19.57	12.10	4.52	6.94
	250	1.47	14.77	24.45	15.64	17.29	13.16	5.71	7.52
4783	300	5.28	17.13	21.21	13.01	15.06	13.16	6.68	8.47
	350	10.19	17.59	18.30	10.99	13.19	12.68	7.33	9.73
	430	17.25	15.65	14.70	8.90	11.22	11.92	7.96	12.39
	200	2.75	23.15	29.01	14.66	13.03	7.96	3.38	6.04
4876	250	9.19	22.97	23.13	12.29	12.52	9.27	4.23	6.40
	300	15.76	20.78	19.01	10.49	11.62	9.99	5.15	7.19
	350	20.42	18.06	16.26	9.30	10.96	10.42	6.04	8.54
	200	16.99	28.75	22.51	9.87	8.50	5.49	2.63	5.26
4969	250	24.38	23.02	18.44	9.15	9.14	6.92	3.32	5.63
	270	25.27	21.44	17.51	9.03	9.46	7.54	3.74	6.01
	140	33.77	31.58	16.06	5.14	3.99	3.03	1.90	4.52
5072	160	35.81	28.01	16.10	5.75	4.57	3.26	1.94	4.57
	180	35.60	25.66	16.31	6.45	5.39	3.73	2.80	4.79
	80	59.88	21.00	6.35	2.39	2.45	2.33	1.57	4.03
5164	90	57.82	21.87	7.10	2.52	2.54	2.40	1.61	4.14
	100	54.99	22.91	8.14	2.75	2.69	2.52	1.68	4.32

Table 6.2: Mass bin weights for recoil mass bins used in the inclusive efficiency estimate for various spectator model inputs.

Decay Mode	ϵ (%)	ϵ_w (%)	ϵ_{flat} (%)
$B^- \to \gamma K^{*-}(890)$	91.44	4.33	3.68
$B^- \to \gamma K^{*-}(1270)$	99.97	5.20	4.86
$B^- \to \gamma K^{*-}(1430)$	100.00	4.35	4.29
$B^- \to \gamma K^{*-}(1680)$	99.98	3.83	4.15
$B^- \to \gamma K^{*-}(1780)$	100.00	3.79	4.26
$B^- \to \gamma K^{*-}(2045)$	99.99	2.87	3.65
$B^- \to \gamma K^{*-}(2250)$	98.29	2.55	3.43
$B^- \to \gamma K^{*-}(2450)$	78.40	1.82	2.51
$\overline{B}^0 \to \gamma \overline{K}^{*0}(890)$	91.79	6.26	4.86
$\overline{B}^0 \to \gamma \overline{K}^{*0}(1270)$	99.98	5.15	4.85
$\overline{B}^0 \to \gamma \overline{K}^{*0}(1430)$	100.00	4.96	4.83
$\overline{B}^0 \to \gamma \overline{K}^{*0}(1680)$	99.96	4.28	4.71
$\overline{B}^0 \to \gamma \overline{K}^{*0}(1780)$	100.00	4.01	4.55
$\overline{B}^0 \to \gamma \overline{K}^{*0}(2045)$	99.99	3.46	4.52
$\overline{B}^0 \to \gamma \overline{K}^{*0}(2250)$	98.43	2.38	3.15
$\overline{B}^0 \to \gamma \overline{K}^{*0}(2450)$	78.69	1.71	2.32
$B^- \to \gamma X_s^{*-}$	49.93	3.95	4.06
$\overline{B}^0 \to \gamma \overline{X}_s^{*0}$	45.61	4.12	4.20

Table 6.3: The efficiencies listed for each decay mode are ϵ, the fraction of the generated signal events that have a photon between 2.0 and 2.7 GeV with no other cuts applied, ϵ_w, the unflattened weighted efficiency calculated by summing the weights of the events that pass all of our analysis cuts divided by the total number of generated signal events, and ϵ_{flat}, the sum of the flattened weights of the events that pass our cuts divided by the total number of generated signal events. Included in this table are efficiencies for both exclusive and inclusive decays. Modes denoted by X_s indicate a fictitious $s\overline{q}$ resonance hadronized by JETSET [2].

dependence can be seen within each column of Table 6.3 where ϵ rises slightly as the K^* mass increases above 890 MeV/c^2 and more photons are included in the energy window. The efficiency decreases for heavier masses as more photons with energies less than 2.0 GeV are excluded from the signal window. The mass dependence from the fraction of photons falling between 2.0 and 2.7 GeV is illustrated in Figure 6.2.

The X_s mass dependence of our ability to detect an event with a photon between $2.0 - 2.7$ GeV results from a difference in the shape of the X_s decay. In Table 6.3, the weighted efficiencies (both flattened and unflattened) decrease with increasing X_s mass. Part of this dependence of ϵ_w and ϵ_{flat} on X_s mass comes from the fact that as X_s mass increases, there are an increasing number of photons below 2.0 GeV. The other part of the dependence comes from a difference in the shape of the X_s decay. A lighter resonance, such as $K^*(890)$ recoiling against a higher energy photon will produce X_s decay products that are collimated in a narrow jet matching the signature for $b \rightarrow s\gamma$ very well. A heavier resonance, however, will usually have lower momentum and broader decays that look less like the "typical" $b \rightarrow s\gamma$ events. These heavier K^* events will be weighted less heavily than the lighter K^* events, contributing to the dependence of ϵ_w and ϵ_{flat} on X_s mass. Notice that ϵ_{flat} drops much less steeply than ϵ_w, evidence that flattening the weights has the desired effect of decreasing the model dependence of our efficiency. The mass dependence of the net output variable r keys on the shape differences between the lightest and heaviest K^* masses and is shown in Figure 6.3.

6.3   Efficiency

We have generated two samples of MC, the "K^* sum" sample and the JETSET sample. Now, we turn to calculating the inclusive efficiency for both of these samples. For each of the $(\langle m_b \rangle, p_F)$ pairs, we determine the efficiency, which we take to be the sum of flattened weights for events with $2.0 < E_\gamma < 2.7$ GeV as measured in the laboratory frame divided by the total number of $b \rightarrow s\gamma$ events generated with photon energy above 2.0 GeV in the B rest frame ($E_\gamma^{B\ rest\ frame} > 2.0$ GeV). The final inclusive efficiencies for all $(\langle m_b \rangle, p_F)$ pairs are listed in Table 6.4 for the JETSET sample and Table 6.5 for the K^* sum sample.

Once we have an efficiency for each of the 24 $(\langle m_b \rangle, p_F)$ pairs for both our JETSET sample and the K^* sum sample, we determine which values of $(\langle m_b \rangle, p_F)$ best represent our data and find the efficiency associated with these values. We determine the best-fit values and the efficiency separately for the two MC samples (JETSET and K^* Sum) and use the difference between the results to determine the systematic error due to model dependence, hadronization effects and $B\overline{B}$ subtraction. Previous measurements used external means to determine the best $(\langle m_b \rangle, p_F)$, but the present measurement of the photon energy spectrum is good enough that we can let our own data pick the best values of these parameters. It is important to

E_γ Distribution for Several K* Modes
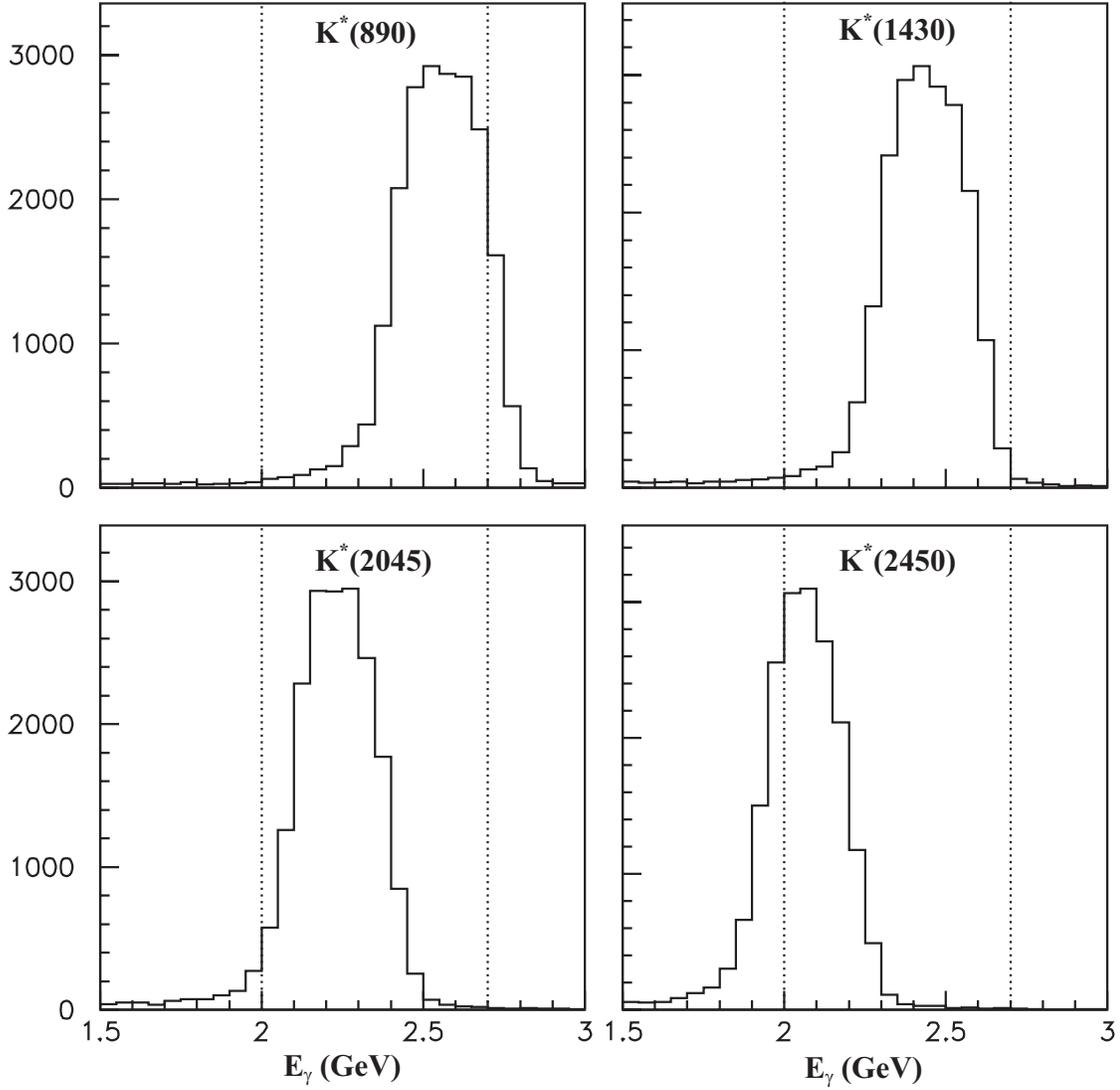
Figure 6.2: Photon energy distributions for several $B \to K^*\gamma$ modes from the lightest K^* mass, $B \to K^*(890)\gamma$, to the heaviest, $B \to K^*(2450)\gamma$, and two intermediate masses, $B \to K^*(1430)\gamma$ and $B \to K^*(2045)\gamma$, boosted in the lab frame of reference. The vertical lines indicate our acceptance window, $2.0 - 2.7$ GeV. The change in the fraction of photons between 2.0 and 2.7 GeV is a source of mass dependence.

Neural Net Output Distribution for Several K* Modes
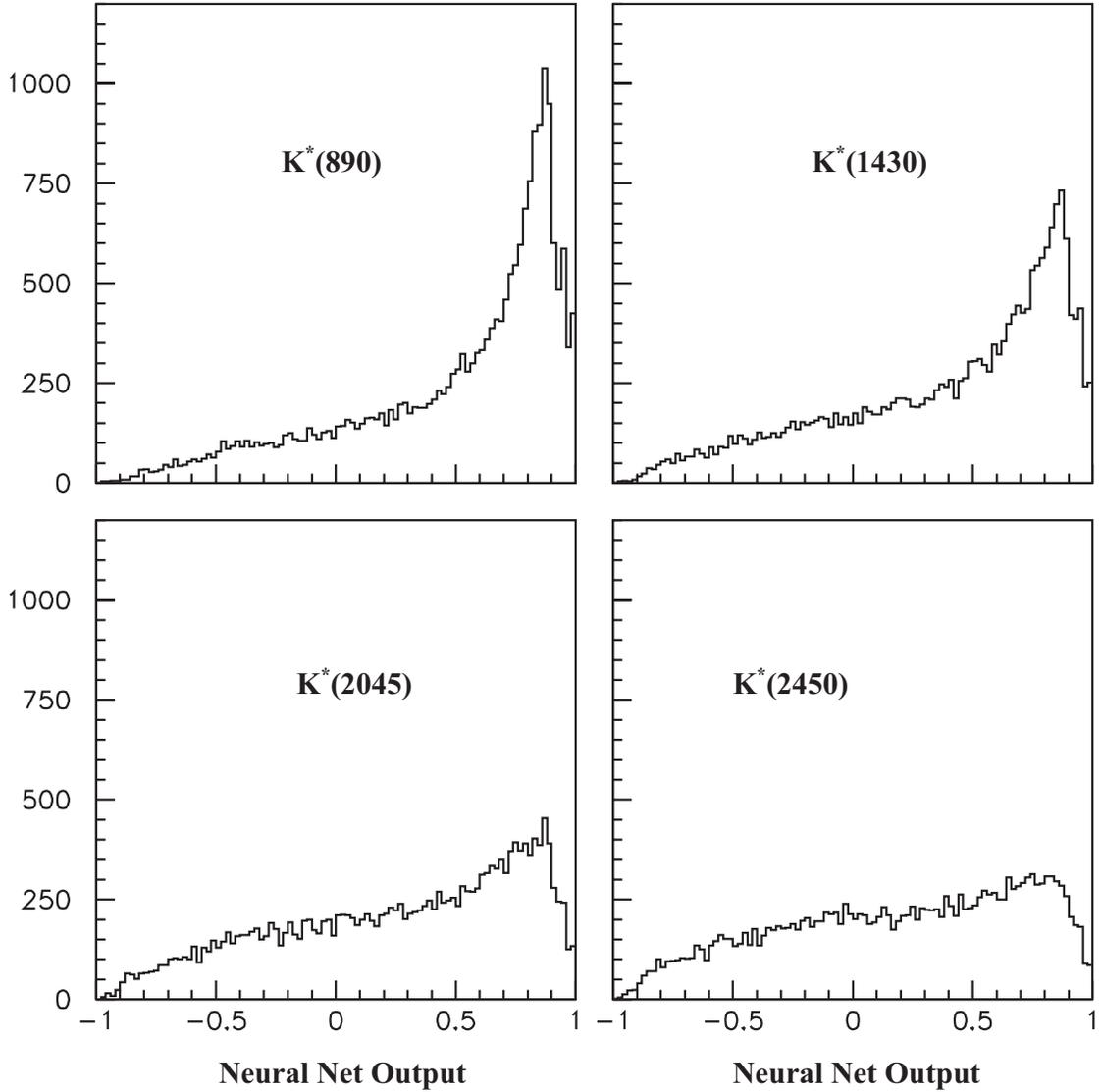
Figure 6.3: Neural net output distributions for several $B \to K^*\gamma$ modes from the lightest K^* mass, $B \to K^*(890)\gamma$, to the heaviest, $B \to K^*(2450)\gamma$, and two intermediate masses, $B \to K^*(1430)\gamma$ and $B \to K^*(2045)\gamma$, boosted in the lab frame of reference. The lighter masses tend to have higher values of the neural net output than the heavier masses. We interpret this to mean that the lighter masses are more signal-like than the heavier masses, a source of model dependence.

realize that we do not attribute any physical meaning to $(\langle m_b \rangle, p_F)$ in doing this. We just take them as two correlated parameters that vary the mean and width of the photon energy spectrum (recall Figure 6.1).

We fit our measured spectrum over the range $2.0 - 2.8$ GeV to each of the photon energy spectra for the 24 spectator models in both the K^* sum method and the JETSET method. When fitting our data spectrum to the MC generated spectrum, we need to take into account the fact that our measured spectrum will look different if we subtract a different amount of $B\overline{B}$ background. We previously concluded that a good estimate for a 1σ error on the $B\overline{B}$ subtraction was a $\pm 5\%$ variation in the $B\overline{B}$ background. Because of this $\pm 5\%$ uncertainty on the $B\overline{B}$ subtraction, we actually have three measured photon energy spectra to fit: a nominal, a $+5\%$, and a -5%. For this reason, we fit three measured spectra with each of the 24 MC spectra for both the JETSET and the K^* sum samples. The only fitting parameter in these fits is the overall scale. Table 6.4 shows the three χ^2 (nominal, $+5\%$, and -5%) for each of the 24 $(\langle m_b \rangle, p_F)$ pairs for the JETSET sample. Table 6.5 gives the same information for the K^* sum sample.

The best-fit values of $\langle m_b \rangle$ and p_F correspond to the minimum of our χ^2 distribution. For any increase in the χ^2 of the fit over the minimum a 1σ error ellipse is defined over the variables $\langle m_b \rangle$ and p_F. For the best $(\langle m_b \rangle, p_F)$ pair, we obtain the statistical errors by determining $(\langle m_b \rangle, p_F)$ for one unit increases in the χ^2. Knowing the best fit value of $(\langle m_b \rangle, p_F)$ and the 1σ error ellipse, and the relation between $(\langle m_b \rangle, p_F)$ and efficiency, we can find the efficiency and the error in the efficiency for the six fits. These are given in Table 6.6.

By projecting the ellipse down to the $\langle m_b \rangle$ and p_F axes, we determine the 1σ errors on these parameters. The best-fit points and error ellipses for fits to the nominal, $+5\%$, and -5% measured photon energy spectra are shown in Figure 6.4. This figure shows that there is a negative correlation between the variables $\langle m_b \rangle$ and p_F. As p_F increases from its best-fit value, the value of $\langle m_b \rangle$ that minimizes our distribution decreases in proportion to the correlation between the two parameters and how far p_F moves from its best-fit value.

Our best-fit values correspond to $\langle m_b \rangle = 4706 \pm 116$ MeV/c^2, $p_F = 438 \pm 152$ MeV/c for JETSET and $\langle m_b \rangle = 4740 \pm 100$ MeV/c^2, $p_F = 374 \pm 142$ MeV/c for the K^* sum model. We take our efficiency for $b \rightarrow s\gamma$ to be $\epsilon = 4.12\%$, the average of the two nominal efficiencies. It is important to state at this point that this is NOT the final efficiency for $b \rightarrow s\gamma$ because as yet, it is uncorrected for several known differences between the data and our MC. The uncorrected efficiency is listed in the final row of Table 6.6.

$\langle m_b \rangle$	p_F	χ^2_{nom}	$\chi^2_{+5\%}$	$\chi^2_{-5\%}$	$\langle E_\gamma \rangle$	$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$	ϵ_{flat}	$\epsilon^{00}_{flat} - \epsilon^{+-}_{flat}$
4597	390	6.95	7.59	6.55	2.311	0.0169	0.040	0.0025
4597	440	4.08	4.62	3.78	2.320	0.0194	0.040	0.0026
4690	260	9.03	9.28	9.02	2.326	0.0115	0.041	0.0022
4690	310	5.70	5.96	5.69	2.332	0.0143	0.041	0.0023
4690	360	3.43	3.65	3.46	2.339	0.0169	0.041	0.0025
4690	410	2.17	2.30	2.30	2.346	0.0196	0.041	0.0028
4783	200	6.13	5.87	6.66	2.356	0.0096	0.042	0.0017
4783	250	4.25	4.03	4.73	2.359	0.0120	0.042	0.0019
4783	300	2.77	2.54	3.26	2.365	0.0150	0.042	0.0023
4783	350	2.44	2.16	2.99	2.371	0.0181	0.042	0.0029
4783	430	3.85	3.47	4.50	2.378	0.0226	0.041	0.0040
4876	200	4.51	3.74	5.56	2.392	0.0106	0.043	0.0015
4876	250	4.20	3.45	5.25	2.396	0.0135	0.043	0.0024
4876	300	4.88	4.11	5.94	2.401	0.0166	0.043	0.0034
4876	350	5.93	5.17	6.98	2.401	0.0198	0.042	0.0042
4969	200	9.42	8.11	11.03	2.431	0.0121	0.044	0.0029
4969	250	10.45	9.16	12.04	2.433	0.0151	0.043	0.0043
4969	270	10.27	9.03	11.80	2.431	0.0163	0.043	0.0046
5072	140	21.09	19.18	23.32	2.473	0.0109	0.044	0.0049
5072	160	21.08	19.19	23.29	2.473	0.0118	0.044	0.0054
5072	180	19.96	18.14	22.10	2.469	0.0128	0.044	0.0056
5164	80	36.89	34.50	39.60	2.511	0.0104	0.043	0.0086
5164	90	35.94	33.58	38.62	2.510	0.0105	0.043	0.0085
5164	100	34.72	31.96	36.90	2.506	0.0109	0.043	0.0082

Table 6.4: For each of the 24 models ($\langle m_b \rangle$, p_F pairs), as hadronized by JETSET, we give the efficiency, the neutral-charged efficiency difference, the first and second moments of the photon energy spectrum, and the χ^2 of the fits to the measured photon spectrum, nominal and with B background increased and decreased by 5%.

$\langle m_b \rangle$	p_F	χ^2_{nom}	$\chi^2_{+5\%}$	$\chi^2_{-5\%}$	$\langle E_\gamma \rangle$	$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$	ϵ_{flat}	$\epsilon^{00}_{flat} - \epsilon^{+-}_{flat}$
4597	390	5.69	6.55	5.06	2.299	0.0207	0.040	0.0032
4597	440	3.70	4.50	3.14	2.306	0.0233	0.040	0.0031
4690	260	6.12	6.56	5.93	2.319	0.0158	0.042	0.0038
4690	310	3.92	4.35	3.74	2.324	0.0183	0.041	0.0036
4690	360	2.54	2.93	2.40	2.328	0.0209	0.041	0.0034
4690	410	1.91	2.24	1.84	2.333	0.0236	0.041	0.0034
4783	200	3.77	3.66	4.15	2.350	0.0144	0.043	0.0038
4783	250	2.84	2.76	3.20	2.352	0.0164	0.043	0.0036
4783	300	2.08	1.99	2.45	2.356	0.0191	0.042	0.0036
4783	350	2.22	2.09	2.63	2.360	0.0221	0.042	0.0038
4783	430	3.88	3.70	4.33	2.362	0.0265	0.042	0.0041
4876	200	3.91	3.25	4.87	2.385	0.0150	0.044	0.0034
4876	250	3.90	3.25	4.84	2.388	0.0177	0.043	0.0038
4876	300	4.84	4.18	5.79	2.391	0.0206	0.043	0.0043
4876	350	5.93	5.30	6.84	2.389	0.0235	0.042	0.0046
4969	200	9.13	7.91	10.64	2.423	0.0164	0.044	0.0041
4969	250	10.34	9.15	11.82	2.423	0.0190	0.043	0.0049
4969	270	10.13	9.00	11.55	2.419	0.0200	0.043	0.0049
5072	140	20.08	18.28	22.20	2.463	0.0152	0.044	0.0051
5072	160	20.15	18.36	22.24	2.462	0.0158	0.044	0.0054
5072	180	19.08	17.36	21.10	2.457	0.0166	0.043	0.0055
5164	80	34.52	32.26	37.10	2.497	0.0144	0.043	0.0074
5164	90	33.37	31.14	35.92	2.494	0.0147	0.043	0.0072
5164	100	31.76	29.58	34.26	2.490	0.0151	0.043	0.0069

Table 6.5: For each of the 24 models ($\langle m_b \rangle$, p_F pairs), as hadronized by the K^* sum method, we give the efficiency, the neutral-charged efficiency difference, the first and second moments of the photon energy spectrum, and the χ^2 of the fits to the measured photon spectrum, nominal and with B background increased and decreased by 5%.
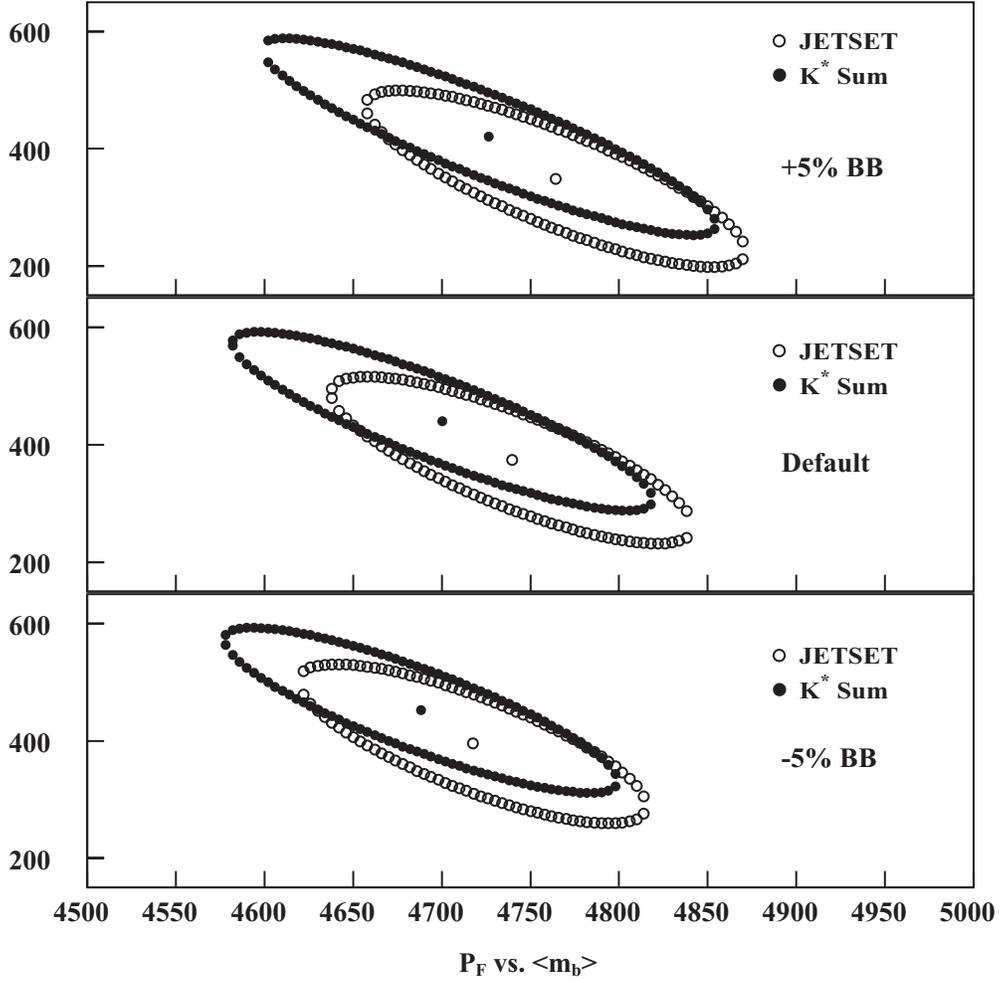
Figure 6.4: The best-fit points and error ellipses, in $\langle m_b \rangle - p_F$ space, of the fits to the photon energy spectrum. Top to bottom, the panels are $+5\%$ subtraction, nominal subtraction, and -5% subtraction.

Measured Spectrum	$\langle m_b \rangle$	p_F	Efficiency (%)
	JETSET		
Nominal	4706 ± 116	438 ± 152	4.08 ± 0.16
$+5\%$ $B\overline{B}$ Background	4727 ± 126	421 ± 167	4.10 ± 0.17
-5% $B\overline{B}$ Background	4688 ± 110	452 ± 141	4.06 ± 0.15
	K^* sum		
Nominal	4740 ± 100	374 ± 142	4.15 ± 0.13
$+5\%$ $B\overline{B}$ Background	4764 ± 106	349 ± 151	4.18 ± 0.14
-5% $B\overline{B}$ Background	4717 ± 96	395 ± 136	4.13 ± 0.12
Final Efficiency	$4.12 \pm 0.14 \pm 0.04 \pm 0.02$		

Table 6.6: Shown in this table are the best-fit values of $\langle m_b \rangle$, p_F to the Ali-Greub model as well as the efficiency and error in the efficiency for the six fits of the three measured spectra (nominal and $\pm 5\%$) and the two MC samples (K^* sum and JETSET).

6.4   Systematic Studies

This analysis relies on MC to determine the efficiency of the $b \to s\gamma$ process. In this section, we describe the studies that quantified the uncertainty on the efficiency due to MC simulation and detector modeling.

6.4.1   Model Dependence and Hadronization Effects

To obtain the error due to the model dependence of the $b \to s\gamma$ process, we take the average of the two nominal efficiency errors listed in Table 6.6 to get ± 0.14. This represents the uncertainty due to the spectator model inputs (how well we know $\langle m_b \rangle$ and p_F).

For hadronization of $s\overline{q}$ into X_s, we used two approaches, represented by our our two MC samples, to determine the uncertainty on the efficiency due to hadronization effects. The K^* sum method assumed that the recoil system always hadronized into a nearby resonance. In reality, non-resonant contributions (represented by the JETSET sample) are present, and may have different final-state multiplicities than the resonant contributions. We take the error due to uncertainty in the hadronization of the recoiling system hadronization to be a weighted average of the difference between the JETSET and K^* sum efficiencies. The error due to hadronization is ± 0.04.

Finally, the error on the efficiency due to the uncertainty in $B\overline{B}$ subtraction is given by half the difference between the average efficiencies for +5% ($\epsilon_{avg} = 4.14\%$) and −5% ($\epsilon_{avg} = 4.09\%$), or ±0.02.

6.4.2   Uncertainty due to $f_\pm/f_{00}$

For this measurement, the decay rate is summed over neutral and charged B decays. However, there is currently an 8% uncertainty in the ratio of B^+B^- to $B^0\overline{B}^0$ resulting from $\Upsilon(4S)$ decays. Standard error propagation shows that the uncertainty in the efficiency due to the uncertainty in $f_\pm/f_{00}$ is

$$\delta\epsilon = |\epsilon_{+-} - \epsilon_{00}| \frac{\delta(f_{+-}/f_{00})}{(1 + f_{+-}/f_{00})^2} \tag{6.1}$$

The recent CLEO measurement of the relative branching fraction of the $\Upsilon(4S)$ to charged and neutral B meson pairs concludes that $f_{+-}/f_{00} = 1.04 \pm 0.08$ [50]. The values $f_{+-}/f_{00} = 1.04$ and $\delta(f_{+-}/f_{00}) = 0.08$ imply $\delta\epsilon = |\epsilon_{+-} - \epsilon_{00}| \times 0.02$.

From Tables 6.4 and 6.5, we see that the $B^0\overline{B}^0 - B^+B^-$ efficiency difference, $\epsilon_{flat}^{00} - \epsilon_{flat}^{+-}$, is typically +0.38% for the K^* sum MC sample and +0.31% for the JETSET MC sample for relevant values of $\langle m_b \rangle$ and p_F. We conservatively use the larger of these and obtain an error in efficiency of $\delta\epsilon = \pm0.008\%$ out of $\epsilon = 4.12\%$, a relative error of ±0.2%.

6.4.3   Detector Modeling

In addition to possible deficiencies in the event generator, the signal efficiency estimate is subject to uncertainties associated with the accuracy of the detector simulation (CLEOG). For the MC to agree with the data, it must agree in a number of aspects of particle detection. Some examples of where the MC could stray from the data are the resolution of momentum measurement, the efficiency of charged particle detection, and the number and distribution of photons. Discrepancies in any one of these ways will affect the smearing of the photon energy which will affect our measurement of the spectrum and all of the values derived from it.

These aspects of particle detection have been studied within the CLEO collaboration and the MC code was tuned to reproduce the measured efficiency but only to the data-measured precision. We have investigated the impact that these aspects of the MC have on our measurement by changing them within the data-measured precision. For example, the charged particle detection efficiency is known to about 1%. By hand, we can randomly throw out 1% of the tracks, and then proceed with the analysis as usual. We make all measurements with the altered data and thereby find the uncertainty in our measured quantities due to the uncertainty in the part of the detector simulation that we altered. Each aspect of simulation that we change

is called a "knob"; finding the systematic error due to detector simulation is then a knob-turning study.

It is necessary to make a few assumptions for knob-turning studies. For the most part, the knobs were turned only one way and we assume that the errors are symmetric. Implicit in the knob-turning study is the assumption that the response to the knob turn is linear so that one turn is equal to half of the uncertainty of two turns. In most cases, we over-turn the knobs to exaggerate the effect and then scale back the systematic uncertainty by the appropriate amount.

We define the "Nominal" efficiency as the reference point for the knob-turning study. It is the efficiency (defined as sum of flattened weights between $2.0 - 2.7$ GeV and with $|cos\theta| < 0.7$ divided by total number of $b \rightarrow s\gamma$ events generated) for the MC with no change. We take the average of the efficiency calculated from the K^* sum sample and the efficiency calculated from the JETSET MC sample to be our nominal efficiency: $\epsilon = 4.08\%$. This value is not the same as the central value that we eventually obtain as our final efficiency because an earlier version of the code was used to look at a different sample of events. However, as long as we determine the percentage difference of each turn of the knob relative to the nominal, we can apply the results to the final efficiency regardless of what that nominal value actually is. The sum total of the systematic error obtained from the knob-turning studies is $\pm 0.057\%$ out of a nominal efficiency of $\epsilon = 4.08\%$. The combined knob-turn error of ± 0.057 out of 4.08 is a relative error of 1.4%. The results of the knob-turning studies are summarized in Table 6.7, and a full description of these studies is given in Appendix E.

6.4.4 Efficiency for finding candidate photons

There is an uncertainty involved in our efficiency for finding and keeping the high energy cluster. To evaluate the difference between this efficiency for data and MC, we use radiative bhabhas embedded in hadron events. The radiative bhabha component gives us a high energy cluster from a photon that we can study in a hadronic environment.

In both samples, we selected photons between 1.5 and 3.5 GeV and required that $|cos\theta| < 0.7$. We ran on the radiative bhabha events before embedding them in the hadronic events (pre-embedded photons). If the hard photon passed both the energy and the angular cuts, then we searched for it in the embedded sample (post-embedded photon). The results of this search are given in Table 6.8.

There were two main reasons why some photons were not found in the embedded event. First, some photons were thrown out because they formed a type-1 or type-2 match with a track. These were due to the overlap in the calorimeter of energy due to a charged particle and our signal photon. We found that this accidental overlap of photons with charged particles is higher in rate for MC by about 1.5%.

Knob	Description	Turn	Scale (%)	ϵ (%)	Systematic Error
Nominal	No Change. MC taken as is.	—	—	4.082	—
CHINEFF	Charged track inefficiency. Tracks cut randomly with a 1% probability per track.	2%	±1/2	3.952	±0.033
PHINEFF*	Photons are cut randomly with a probability of 2% per photon.	2%	±1	3.960	±0.021
NOTMNG	TRKMNG is not used.	Off	±1/5	4.233	±0.030
NOSPLITF	Splitoff packages is not used.	Off	±1/5	4.223	±0.028
MISTRA	Error in track momentum measurement is increased by 10%.	10%	±10%	4.077	±0.006
MISPHO	Error in photon measurement is increased by 10%.	10%	±10%	4.077	±0.005
Total Systematic Error					±0.0057

Table 6.7: This table is a summary of the results of the knob-turning study. Shown for each of the knobs are: the amount the knob was turned, the scaling factor by which we multiply the turn to get a $\pm 1\sigma$ systematic, the efficiency calculated with the full (unscaled) turn, and the systematic error obtained. *For PHINEFF, the photon inefficiency knob, the 2% loss of signal photons is included in the efficiency column, but not in the systematic error column. The final row represents the sum of all the systematic errors contributed by the knob-turns, a total of ± 0.057 out of 4.08, which is a $\pm 1.4\%$ relative error.

# of γ after cut	Monte Carlo	
	Pre-Embed	Post-Embed
$1.5 < E_\gamma < 3.5$, $\cos\theta < 0.7$	1593.31	1489.27
E9/E25	1452.99 (97.56%)	1437.00 (96.49%)
Shower Mass	1440.40 (96.72%)	1417.33 (95.17%)
Overall Efficiency	88.95%	

# of γ after cut	Data	
	Pre-Embed	Post-Embed
$1.5 < E_\gamma < 3.5$, $\cos\theta < 0.7$	1551	1476
E9/E25	1413 (95.73%)	1409 (95.46%)
Shower Mass	1395 (94.51%)	1386 (93.90%)
Overall Efficiency	89.36%	

Table 6.8: Each column gives the number of photons that passes each cut. Cuts are applied in succession. The percentages given are the number of events passing the cut out of the total number of photons passing both the energy and the angular cuts in the *post*-embedded sample. The overall efficiency is the percentage obtained by dividing the number of photons that pass all the cuts in the post-embedded sample by the total number of photons that were found in the pre-embedded sample.

Second, the efficiency of the E9/E25 and shower mass cuts were different in MC and data. Since, propagating a large number of lower energy showers through the crystals slows down the simulation tremendously, the MC halts shower development below 1 MeV for electromagnetic showers in the calorimeter. The consequence of this "cutoff energy" is that data showers tend to be broader than simulated showers in MC. This difference shows itself in the distributions of our two shower shape variables which tend to be higher in MC than in data.

The efficiency for finding photons in data is 95.16% while in MC, the efficiency is 93.47%, a difference of 1.69%. Once a photon has been found in the pre- and post-embedded samples, the efficiency for keeping it after the E9/E25 and shower mass cuts have been applied is 93.90% in data and 95.17% in MC, a difference of −1.27%. The overall efficiency for finding and keeping candidate photons is 89%. The overall efficiency difference between the data and the MC is small at 0.41% and can be neglected. We take the relative error on the photon finding and keeping efficiency by combining the two differences in quadrature and dividing by the overall efficiency, obtaining a relative uncertainty of 2.4%.

6.4.5 Correcting the TOF Radiation Length

The description of the time of flight system in CLEOG is incorrect [51]. Standard CLEOG uses a different radiation length for the scintillator than the correct radiation length. Code has been developed to to correct this error. While the correction will affect lower energy showers most, we check what it does to high energy photons.

We generated two samples of MC, one using the standard CLEOG and the other using the corrected CLEOG. We generated events of 2 GeV single photons within the good barrel ($\cos\theta < 0.7$), and examined these samples for differences. The overall difference between the number of photons passing all analysis cuts in the two MC samples is small, about 12 events out of 7444. We conclude that the correction to the TOF material in CLEOG does not affect high energy photons.

6.4.6 Errors from Imperfect Modeling of the "Other B"

Features of the "other B" will influence our analysis by affecting the distribution of energy in the event, hence the output of our shape variable neural net. In addition, the modeling of the "other B" will also influence pseudoreconstruction and our ability to find leptons. Incorrectly modeling the number or quality of the leptons in the non-signal B will have a negative impact on our analysis since we use leptons in the "other B" to distinguish between signal and continuum. Our efficiency also depends on the number of charged particles and π^0s in the "other B", so a careful comparison of the charged and neutral multiplicity in data and MC needs to be made.

We have investigated the possible errors in efficiency from modeling of the "other B" in three ways. In the first, we examined a mode with similar topology to $b \to s\gamma$ events and treated a high momentum $(1.5 - 2.2 \text{ GeV/c})$ muon as if it were the high energy signal photon and compared On-Off subtracted data and $B\overline{B}$ MC. In the second, we did a knob-turning study of the number of secondary leptons. In the third investigation, we determined the dependence of efficiency on the multiplicity of the "other" B. Finally, while doing the $B \to X\mu\nu$ study, we noticed that there was poor agreement between data and MC on the electron momentum spectrum. This turned out to be a flaw in the way we identified electrons in MC. We compared electron spectra and MC to find the uncertainty in electron identification efficiency.

$B \to X\mu\nu$

Our MC prediction of the detection efficiency relies on the following three elements: 1) The topology of the $b \to s\gamma$ events produced by our event generator, 2) The topology of generic B decays as produced by our event generator, and 3) The simulation of the events by CLEOG. By using a well understood B decay mode with

a topology similar to $b \to s\gamma$, we can test the accuracy of the generated topology for the "other B" and the accuracy of the detector simulation. To do this, we compare the MC and data efficiencies for $B \to X\mu\nu$. Treating a high momentum muon as the signal photon tests the correctness of the modeling of those features of the "other B" that influence r, the output of the shape variable neural net. This study will not address features of the "other B" that might influence pseudoreconstruction or the leftover lepton, but the other two studies will.

We select events containing a μ with momentum between 1.5 and 2.2 GeV/c. The topology of these events is similar to $b \to s\gamma$ events – a high energy particle recoiling against the other products of the B decay.[9] Once the event passes the initial requirements, we treat the μ just as we did the candidate photon in the analysis and compute the eight shape variables and the shape net output, r. Our goal is to determine whether there is a difference in the features of the "other B" between data and MC. Since this particular method is sensitive to the energy distribution of the event, we compare the r distributions of data and MC and quantify the difference. Finally, we translate this shift in r into an efficiency correction.

For the data, we used 9.2 fb^{-1} On data and 4.6 fb^{-1} Off corresponding to 9.7 million $B\overline{B}$ pairs and carried out the usual On$-$Off subtraction to find the r spectrum from B decays. For the MC, we used a sample of 3.8 million $B\overline{B}$ pairs. We then looked at the r-distribution for On-Off subtracted data and MC split into the four different classes of events listed in Table 6.9. Although the distributions were very similar in shape, the MC distributions were shifted to positive r relative to the data.

The overall shift in the r distribution was 0.03 units of r, however, the high-r region ($r > 0.5$) appeared shifted by only 0.006 units of r. We computed the means of the three r-distributions: all events, events with $r > 0$, and events with $r > 0.5$. Then, we calculated the differences in the means of the r distributions between MC and data, $\langle r \rangle_{MC} - \langle r \rangle_{Data}$. The results are given in Table 6.9. The change in the means with the cuts of $r > 0$ and $r > 0.5$ underestimates the shifts because of the fixed lower edge. For a linearly falling spectrum, the shift is three times the change in mean.

What effect does this shift in r have on our efficiency? Weights increase with increasing r, and the vast majority of the summed weight comes from events with $r > 0.5$. The difference that will be most significant will be the MC-Data difference for events with $r > 0.5$. We take the difference to be 0.015 ± 0.015. We shift r in MC down by that amount and find the efficiency changes by a factor of 0.97 ± 0.03. We apply this correction to the final efficiency.

---

[9]For CLEOns: to select good muon, we required DPTHMU ≥ 5 and QUALMU $= 0$.

Event Class	Difference in r Distribution		
	All Events	$r > 0$	$r > 0.5$
All Events	$+0.047 \pm 0.003$	$+0.010 \pm 0.002$	$+0.004 \pm 0.002$
PR, w/ lepton	$+0.066 \pm 0.036$	-0.008 ± 0.023	-0.009 ± 0.015
PR, no lepton	$+0.111 \pm 0.009$	$+0.015 \pm 0.006$	$+0.003 \pm 0.004$
Not PR, w/ lepton	$+0.014 \pm 0.007$	-0.001 ± 0.006	$+0.003 \pm 0.005$
Not PR, no lepton	$+0.045 \pm 0.003$	$+0.010 \pm 0.003$	$+0.004 \pm 0.002$

Table 6.9: The differences, $\langle r \rangle_{MC} - \langle r \rangle_{Data}$ for the four event class types and for the three r-distributions: all r values, events with $r > 0$, and events with $r > 0.5$. The errors given are rough estimates. PR stands for pseudoreconstructed.

SECLEP - Varying the number of secondary leptons

The second investigation is a knob-turning study on the fraction of events having secondary leptons. It was motivated by a finding that seemed to indicate that the default $B\overline{B}$ MC contained perhaps as much as 20% too large a secondary lepton contribution [47]. Such an overestimate would affect our analysis most significantly in those events with a leftover lepton. Since these events usually receive the heaviest weight, the overestimated secondary lepton contribution could significantly increase the efficiency of the MC over the data.

To determine the effect of having additional secondary leptons in the "other B", we ran on our usual samples of $b \to s\gamma$ MC (K^* Sum and JETSET) and separated them into events containing one or more secondary leptons and events containing no secondary leptons. A secondary lepton is defined as any lepton produced by the decay chain $b \to c \to l$. We found that in the parent sample, 20% of the events had one or more secondary leptons and 80% of the events contained no secondary leptons. We calculated the efficiency for the no-secondary sample and secondary sample separately. We found that the no-secondary sample had an efficiency of 4.02% while the secondary sample had an efficiency of 4.33%. The combination of the 80%/20% mix is 4.08%. A reassuring aspect of this result is that the secondary sample has a significantly higher efficiency than the nominal even though it represents a small fraction of the small sample. This is what we would expect given that having a lepton in the event makes our neural net believe the event is more likely to be $B\overline{B}$ than continuum and thus weight it more heavily.

To determine the effect of increasing the percentage of secondary leptons in our sample, we exaggerated the percentage of secondary leptons in our sample (originally at 20%) by $\pm 25\%$ of itself. The no-secondary/secondary mix of 85%/15% gives an efficiency of 4.07%, and the 75%/25% mix gives an efficiency of 4.10%. For a $\pm 25\%$

variation in the secondary lepton component of the $b \to s\gamma$ MC sample, the efficiency changes by $\pm 0.016\%$ out of 4.08%. This gives a relative error of $\pm 0.4\%$ which is negligible.

Efficiency as a Function of Multiplicity

Since the Monte Carlo prediction for our efficiency depends on how well we model the number of charged particles and π^0s in the "other" B, we investigated the effect of these decay features of the "other" B on our efficiency.

We used a fully simulated sample of $b \to s\gamma$ signal Monte Carlo with about 60,000 events, $\langle m_b \rangle = 4783$ MeV/c^2, and $p_F = 430$ MeV/c. Since the efficiency of events from charged B mesons differs from the efficiency of neutral B mesons, it is possible that the multiplicity will also differ between events from charged Bs and neutral Bs, therefore we performed the entire analysis separately for events coming from charged B mesons and events coming from neutral B mesons. Because the topology and multiplicity of a given $b \to s\gamma$ event varies with the presence or absence of primary and secondary leptons in the "other" B, we look at each of the following decay categories of the "other" B individually: 1) no semileptonic decay, 2) primary semileptonic decay, 3) secondary semileptonic decay, 4) both primary and secondary semileptonic decays.

To determine particle multiplicity, we counted the number of charged and neutral QQ particles in the "other" B. We defined the charged multiplicity as the number of $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, $p^\pm$, and charged hyperons that come from the decay of the "other" B as well as the number of $\pi^\pm$ from K_s^0 decays and the number of $e^\pm$ and $\mu^\pm$ from the semileptonic decays of charmed mesons. Similarly, we defined neutral multiplicity as the number of π^0s and half the number of γs (not from π^0 decay) that come from the decay of the "other" B plus the number of π^0s from K_s decay. The decay multiplicity is based on QQ information and does not include any particles produced through interactions with the detector or the material surrounding it.

For each of the four categories listed above, we determined the overall mean of the charged (neutral) multiplicity distribution. We then considered the events above the overall mean as their own category, the "high" charged (neutral) multiplicity region. Likewise, the events below the mean became the "low" charged (neutral) multiplicity region. We then calculated the mean and determined the efficiency for the events that fell into these regions. We calculated the efficiency in the usual way by dividing the flattened weighted sum of all events in a given category by the number of events in that category that had a signal photon with generated (QQ) energy in the rest frame of the signal B meson greater than 2.0 GeV.

We are interested in the change in efficiency per unit change in charged (neutral) multiplicity for each of the four decay categories. We get $\Delta\epsilon$ by subtracting ϵ_{low} from ϵ_{high}. A similar prescription is used to get ΔMult: ΔMult $=$ Mult$_{high}-$Mult$_{low}$. By

		Charged Multiplicity	Neutral Multiplicity
"Other" B decay	Fraction	$\Delta\epsilon/\Delta mult$	$\Delta\epsilon/\Delta mult$
No semileptonic	0.62	0.262 ± 0.026	0.067 ± 0.033
Primary	0.18	0.114 ± 0.194	0.288 ± 0.239
Secondary	0.16	0.171 ± 0.065	0.090 ± 0.085
Both	0.04	-0.425 ± 0.535	2.421 ± 0.695
Weighted sum		$+0.195 \pm 0.044$	$+0.199 \pm 0.056$

Table 6.10: This table contains the fraction of the total sample that has primary leptons, secondary leptons, both primary and secondary leptons, or no semileptonic decays. For the corresponding category of "other B" decay, we give the change in efficiency per unit change in charged multiplicity and unit change in neutral multiplicity. Also shown is the fraction of total events that fall into the given category.

dividing these two quantities, we get the change in efficiency per unit change in multiplicity, $\Delta\epsilon/\Delta$Mult. The 50/50 average of the quantity $\Delta\epsilon/\Delta$Mult for charged and neutral B mesons is shown in Table 6.10 for both charged and neutral multiplicities.

We determined the unit change in efficiency per unit change in multiplicity for each of the four decay categories, and the results for the four categories were combined in a weighted sum based on the fraction of the total number of events falling into each of the four categories. The unit change in efficiency per unit change in multiplicity for charged multiplicity is 0.195 ± 0.044. For neutral multiplicity, the unit change in efficiency per unit of multiplicity is 0.199 ± 0.056.

To find the difference in charged (neutral) multiplicity between Monte Carlo and data, we compared our results for the mean charged particle multiplicity of the "other" B from Monte Carlo, to the CLEO result for data [52]. We measured a charged multiplicity mean of 5.090 for all events. The CLEO result for data is 5.36 $\pm$ 0.09. We are low by 0.27 ± 0.09. There is no CLEO result for the π^0 multiplicity so we assume that the data has half as much neutral as charged with the same error, or 0.135 ± 0.09 more units of π^0 than Monte Carlo.

Overall, our Monte Carlo multiplicity is low compared to the measured multiplicity for data. This means we must correct the efficiency upward by a factor that is the overall difference in multiplicity between data and Monte Carlo multiplied by the change in efficiency per unit multiplicity. For charged multiplicity, we must correct the efficiency upward by $0.195 \times (0.27 \pm 0.09) = 0.053 \pm 0.018$. For neutral multiplicity, we must correct the efficiency upward by $0.199 \times (0.135 \pm 0.09) = 0.027 \pm 0.018$. Added up, this gives an overall efficiency correction of 0.080 ± 0.040. Given the

overall efficiency from this sample of events, $(4.158 \pm 0.049)\%$, we conclude that the efficiency must be shifted upward by 2.0% of itself. In the end, our $b \to s\gamma$ efficiency must be multiplied by a factor of (1.02 ± 0.01).

REID efficiency

While doing the $B \to X\mu\nu$ study, we noticed a discrepancy between data and MC in the r-distribution for event classes with a leftover lepton. We found that there was poor agreement between data and MC on the electron momentum spectrum that stemmed from the way we identified electrons with REID. Originally, we identified electrons using REID in combination with dE/dX and ToF. Any track within 3σ of the electron hypothesis and not within 3σ of π, K, p, or μ hypotheses was treated as an electron. Given the bad agreement between data and MC, we dropped this approach and insisted that for a track to be called an electron, REID had to say it was an electron. We then decided to compare the electron momentum spectra for data and MC and quantify the difference.

Comparing data and MC electron spectra in the non-good barrel region ($\cos\theta > 0.7$), we noticed a spike in data at $0.6 - 0.8$ GeV/c that we attributed to kaon fakes, which occur with $\sim 30\%$ faking probability at this momentum in the non-good barrel region when ToF is not used [53]. Rather than include electron fakes (negligible except in this narrow interval) in the MC, we instead chose to impose a momentum requirement of $\vec{p} > 0.8$ GeV/c for electrons in the non-good barrel region.

Given these two changes, the remaining differences between data and MC are covered by a $\pm 5\%$ uncertainty in electron identification efficiency as well as the knob-turning study of secondary leptons described above.

6.4.7 Corrections to our MC Sample

We apply several corrections to our MC measured efficiency based on known differences between the data and our MC sample. These corrections are listed below and summarized in Table 6.12.

- **TOF**

  Based on a study of TOF calibration [53], TOF information is not used in $\sim 25\%$ of the runs. In MC however, we use TOF in all the runs. We correct for this by measuring the difference in efficiency with and without TOF. Based on the results of this study, we scale our efficiency by 1.0043.

- **MIXING and** f_{+-}/f_{00}

  Our original MC sample did not have the right mixing parameter of $\chi_d = 0.174$ [1] and did not have the correct charged to neutral ratio given by the recent CLEO result [50]. To correct for this, we first computed our MC efficiency for

charged, neutral-unmixed, and neutral-mixed events separately. From that, we were able to determine the corrections to our standard sample. Based on this study, we scale our efficiency by 0.9985±0.0008 to correct the amount of mixing and by 0.999±0.002 to correct the value of f_{+-}/f_{00}.

- **Mix of CLEO II and CLEO II.5 MC**

  Our MC did not have the right mix of CLEO II and CLEO II.5 MC, therefore we computed our efficiency for the two data-types separately and determined a correction factor. Based on this study, we scale our efficiency by 0.995±0.002.

- **REID Efficiency**

  Our REID code mistakenly had 100% efficiency for electrons in MC. We measured our dependence on the REID efficiency in MC over a range of REID efficiency values from 25% to 100% efficient. We then fit this to a line, and took the true REID efficiency to be 90±5%. Based on this study, we scale our efficiency by 0.983 ± 0.009.

6.5 Efficiency Summary

From the best-fit to the Ali-Greub spectator model, we determined the efficiency for $b \to s\gamma$ for photons with generated energy greater than 2.0 GeV to be 4.12×10^{-2}. This is only the starting point.

The final corrected efficiency is obtained by multiplying the starting efficiency by all of the MC Corrections listed in Table 6.12. These include the corrections due to r distribution shift, multiplicity, candidate vetoing, ToF, Mixing, Dataset, REID, and f_{+-}/f_{00}. Together, these MC corrections combine to form an overall multiplicative factor of 0.9549. Multiplying the starting efficiency, 4.12% by this factor, we obtain the final efficiency, $\epsilon = 3.93\%$.

The final error on the efficiency is obtained by adding all of the errors in quadrature. There are three types of errors to be dealt with: absolute error, relative error, and error on the MC corrections. Added in quadrature, the three absolute errors listed in Table 6.11 combine to from the error due to modeling of the signal B, $\pm 0.151 \times 10^{-2}$. The relative errors are reported as a percentage of the efficiency. To get the magnitude of the error, we multiply each relative error by the starting efficiency of $\epsilon = 4.12 \times 10^{-2}$. The error on the MC corrections is also multiplied by the starting efficiency to calculate the magnitude of the error on the efficiency due to the shift. The relative errors and the errors from the MC corrections, when added in quadrature, give the total error due to modeling of the "other" B and detector simulation, $\pm 0.179 \times 10^{-2}$.

The starting efficiency, the multiplicative shifts, and the errors (added in quadrature) are given in Table 6.12. The final corrected efficiency and total error for $b \to s\gamma$

Absolute Errors		
Item	Relative Error on ϵ	Error on ϵ
$\langle m_b \rangle$, p_F	—	0.00144
Hadronization	—	0.00037
$B\overline{B}$ subtraction	—	0.00025
Total error from signal B Modeling		**± 0.00151**

Relative Errors		
Item	Relative Error on ϵ	Error on ϵ
Secondary leptons	0.004	1.65×10^{-4}
MC knob-turning	0.014	5.76×10^{-4}
Candidate finding	0.024	9.88×10^{-4}
Sum of relative errors		**± 0.00116**

Relative Errors from MC Corrections		
Item	Relative Error on ϵ	Error on ϵ
r Distribution	0.03	0.124×10^{-4}
Multiplicity	0.010	4.12×10^{-4}
Candidate vetoing	0.005	2.06×10^{-4}
ToF	0.0000	0.0000
Mixing	0.0008	0.31×10^{-4}
Dataset	0.0020	0.82×10^{-4}
REID	0.0085	3.50×10^{-4}
f_{+-}/f_{00}	0.002	0.82×10^{-4}
Sum of errors from MC Corrections		**± 0.00137**

Total error from other B Modeling and detector simulation	0.179×10^{-2}

Table 6.11: Summary of absolute and relative errors. When added in quadrature, the absolute errors give the total error on the efficiency from modeling of the signal B. We add the errors on the efficiency due to the relative errors to find the total error from detector simulation and modeling of the other B.

for generated $E_\gamma > 2.0$ GeV is

$$\epsilon = (3.93 \pm 0.23) \times 10^{-2}.$$

Starting efficiency for $E_\gamma > 2.0$ GeV	4.117×10^{-2} (from best fit)

Errors		
Item	Error	Note
$\langle m_b \rangle$, p_F	0.144×10^{-2} (absolute)	Variation in ϵ over Ali-Greub models within 1 unit of χ^2 of best-fit.
Hadronization	0.037×10^{-2} (absolute)	JETSET vs. K^* Sum
$B\overline{B}$ subtraction	0.025×10^{-2} (absolute)	ϵ change from fitting yield vs. fitting yield with $\pm 5\%$ $B\overline{B}$ backgrounds.
Secondary leptons	0.4×10^{-2} (relative)	Effect on ϵ from varying the # of secondary leptons.
MC knob-turning	1.4×10^{-2} (relative)	Effect on ϵ from varying MC.
Candidate finding	1.4×10^{-2} (relative)	Error from uncertainty in efficiency for finding the hard γ.

Shifts - MC Corrections		
Item	Corr. Factor	Note
r distribution shift	0.97 ± 0.03	Data/MC differ in the distribution of r.
Multiplicity	1.020 ± 0.010	Data/MC differ in event multiplicity.
Candidate Vetoing	0.985 ± 0.005	Data/MC differ in probability of false signal γ veto.
TOF	1.0043 ± 0.0000	MC always uses TOF; data rejects TOF in runs known to be poorly calibrated.
Mixing	0.9985 ± 0.0008	Some of our MC had an incorrect amount of mixing.
Dataset	0.9950 ± 0.0020	The ratio of CLEO II to CLEO II.5 events in our MC sample does not match the data.
REID	0.983 ± 0.0085	Our REID efficiency in MC was wrong.
f_{+-}/f_{00}	0.999 ± 0.002	Our charged to neutral mix in MC differed from recent CLEO f_{+-}/f_{00} result.

Corrected efficiency for generated $E_\gamma > 2.0$ GeV	3.931×10^{-2}
Error due to signal-B modeling	0.151×10^{-2}
Error due to other-B modeling and detector simulation	0.179×10^{-2}
Total error on efficiency	0.234×10^{-2}

Table 6.12: Summary of final efficiency determination.

CHAPTER 7

Results

The yield from $b \to s\gamma$, the number of B mesons in the dataset, and the detection efficiency of the analysis are the three quantities required to compute $\mathcal{B}(b \to s\gamma)$. Additional information can be extracted from the shape of the photon energy spectrum. The first and second moments can be related to HQET parameters which can then be used to determine several CKM matrix elements. This chapter will collect all of the numbers to obtain the branching fraction and present the final photon energy spectrum with its first and second moments.

7.1 Branching Fraction

The branching ratio for $b \to s\gamma$ is given by

$$\mathcal{B}(b \to s\gamma) = \frac{Y}{\epsilon \times 2 \times \mathcal{L} \times \sigma}$$

where Y is the yield, ϵ the efficiency, $\mathcal{L}$ the luminosity, and σ is the $B\overline{B}$ cross section. The factor of two arises because the luminosity times the cross section gives the number of $B\overline{B}$ pairs.

The yield for a set of events is given by their flattened weighted sum. Table 7.1 summarizes the yield in three photon energy intervals and the section number where the details can be found. To measure the yield from $b \to s\gamma$, we obtain the yield from the On data and subtract the yield from the background processes. Our final yields are for events in which the candidate photon is between 2.0 and 2.7 GeV, although other energy ranges are listed for comparison and as a check of the results.

The measured values for the relevant quantities are

- $Y = 233.6 \pm 31.2 \pm 13.4$ weights

- $\epsilon = (3.93 \pm 0.15 \pm 0.17)\%$

- $N_{B\overline{B}} = \mathcal{L} \times \sigma = (9.7 \pm 0.2) \times 10^6$ events

Source	1.5 − 2.0 (GeV)	2.0 − 2.7 (GeV)	3.0 − 3.5 (GeV)	Reference Reference
ON	2718.8 ± 17.2	1861.7 ± 16.5	1083.2 ± 8.0	4.2.1
$\lambda \times$OFF	1665.6 ± 14.6	1399.4 ± 15.3	1101.4 ± 11.6	5.1.5, 4.2.1
$b \to c$	798.3 ± 10.0	153.9 ± 5.6	0.04 ± 0.04	Table 5.6
charmless hadronic	4.2 ± 0.2	5.4 ± 0.3	0.0 ± 0.0	Table 5.6
$b \to u$	6.8 ± 0.3	2.1 ± 0.2	0.01 ± 0.01	Table 5.6
π^0/η Correction	111.1 ± 20.6	52.6 ± 14.6	0.0 ± 0.0	5.2.1
Other B Bkds	112.6 ± 27.0	14.7 ± 14.9	1.5 ± 1.1	5.3, C
Yield	20.3 ± 42.0	233.6 ± 31.2	-19.8 ± 14.2	Table 5.6

Table 7.1: Yields (summed flattened weights) with statistical errors for three photon energy intervals. Given are the yields on the $\Upsilon(4S)$ resonance, scaled off resonance yields, and estimated backgrounds from B decay processes other than $b \to s\gamma$ and $b \to d\gamma$. The yield represents the $b \to s\gamma$ plus $b \to d\gamma$ signal. Note that the yield in Off data has been decreased by 0.5% of itself to account for the On−Off subtraction bias. The error on this bias is not included in the statistical error, but is included in the systematic error.

We use the full CLEO dataset (4s2 − 4sT) totaling 9.2 fb^{-1} on resonance and 4.6 fb^{-1} off resonance corresponding to 9.7×10^6 B mesons, a value determined from cross section measurements for each dataset [54],[55]. We take a 2% error on this number based on a conservative estimate of the fits to the cross section. The first error on the yield is statistical, the second is systematic from the 5% uncertainty on the $B\overline{B}$ subtraction and a 0.5% uncertainty on the continuum subtraction. The first error on the efficiency is from the model dependence of the $b \to s\gamma$ decay and the second error is from detector simulation and model dependence of the decay of the other B. The systematic errors on the efficiency include the uncertainties in the modeling of the signal and non-signal B mesons, the uncertainty due to hadronization, the systematic error on the $B\overline{B}$ subtraction, and an uncertainty due to the simulation of the detector performance including photon-finding, track-finding, and shower/track momentum resolutions.

Based on these numbers, we obtain branching fraction for photons with $E_\gamma > 2.0$ GeV:

$$\mathcal{B}(b \to (s+d)\gamma) = (3.06 \pm 0.41 \pm 0.26) \times 10^{-4}.$$

To obtain the branching fraction for $b \to s\gamma$ alone extrapolated to the full spectrum, we must apply two "theory" corrections.

7.1.1 $b \to d\gamma$

Although we call this analysis $b \to s\gamma$, we actually make no distinction in the code between $b \to s\gamma$ and $b \to d\gamma$. We are sensitive to both and apply a correction to the yield to remove the $b \to d\gamma$ contribution to the branching fraction.

Using a model for $b \to d\gamma$ that is similar to $b \to s\gamma$ we find that the efficiency for $b \to d\gamma$ is the same as that for $b \to s\gamma$. Although pseudoreconstruction favors $b \to s\gamma$ over $b \to d\gamma$, the shape variables and the photon energy spectrum favor $b \to d\gamma$, so the two effects pull the efficiency in opposite directions and effectively cancel. The expectation from the SM is that the $b \to d\gamma$ and $b \to s\gamma$ branching fractions are in the ratio $|V_{td}/V_{ts}|^2$ [1]. We take $|V_{td}/V_{ts}| = 0.20 \pm 0.04$ [1] which gives a downward correction to the branching fraction of $(4.0 \pm 1.6)\%$. After this correction, the value of the branching fraction for photons with energy in the B rest frame above 2.0 GeV is $\mathcal{B}(b \to s\gamma)_{E_\gamma > 2.0} = (2.94 \pm 0.40 \pm 0.25) \times 10^{-4}$.

7.1.2 Extrapolating to Full Spectrum

Finally, we extrapolate the branching fraction to the full spectrum. The fraction of $b \to s\gamma$ decays with photon energies above 2.0 GeV is sensitive to the b quark mass and Fermi momentum which means that the fraction depends on the model that is used to simulate the decay. We use the value of $0.915^{+0.027}_{-0.055}$ from Kagan and Neubert as the fraction of the spectrum above 2.0 GeV [19]. To extrapolate the branching fraction to the full energy range, we divide by 0.915. With these corrections, we get the final branching fraction for $b \to s\gamma$ over all energies:

$$\mathcal{B}(b \to s\gamma) = (3.21 \pm 0.43 \pm 0.27^{+0.018}_{-0.010}) \times 10^{-4}$$

where the first error is statistical, the second is systematic, and the third (asymmetric) error is from the extrapolation to the full spectrum. The result is in good agreement with the Standard Model predictions. The SM prediction of Chetyrkin, Misiak, and Münz is $(3.28 \pm 0.33) \times 10^{-4}$ [18]. Recently, Gambino and Misiak argue that the charm-loop contribution is larger than previously thought and obtain a value of $(3.74 \pm 0.30) \times 10^{-4}$ [22]. The result is also in good agreement with the CLEO collaboration's previously measured result of $(2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$ [35].

7.2 Photon Energy Spectrum

The final photon energy spectrum with all of the corrections applied is shown in Figure 7.1. Useful information can be obtained from the first and second moments of the photon energy spectrum. To a good approximation, the mean of the photon energy distribution, $\langle E_\gamma \rangle$ is equal to half of the b quark mass, m_b. Likewise, the mean square width of the photon energy spectrum, $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2$, depends on the momentum of the b quark within the B meson, p_F.
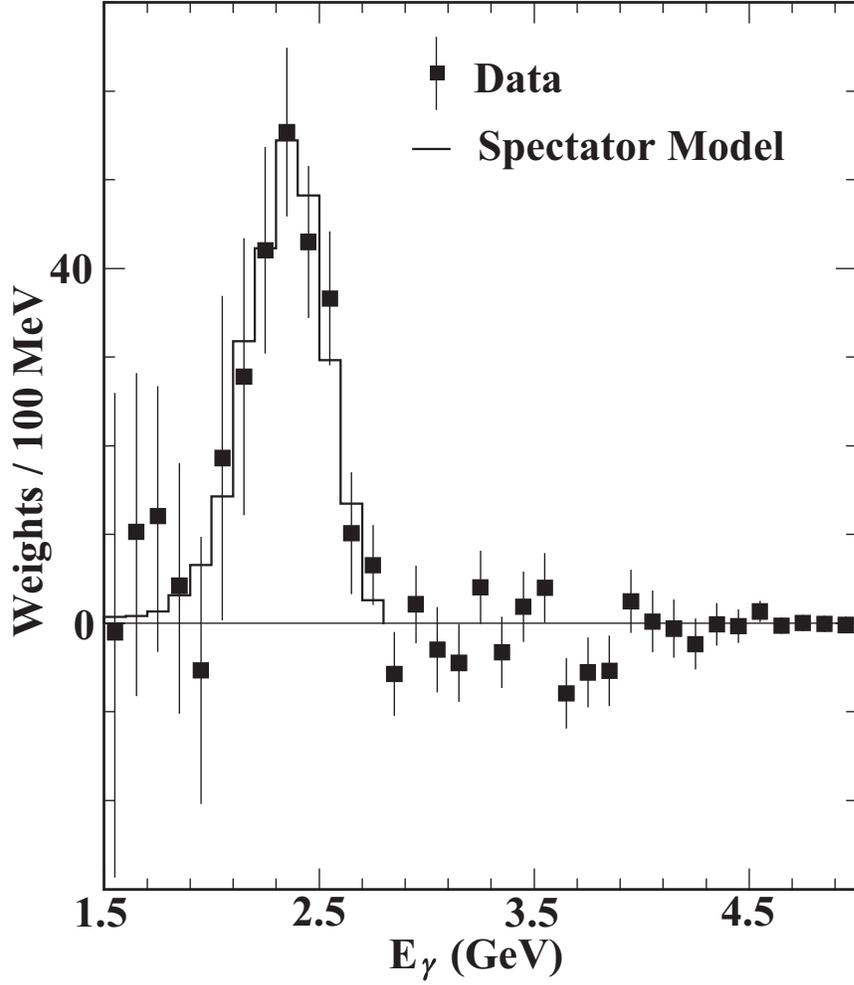
Figure 7.1: The observed laboratory frame photon energy spectrum (weights per 100 MeV) for On−Off data and with B backgrounds subtracted. This represents the $b \to s\gamma + b \to d\gamma$ signal. Superimposed is the spectrum from MC simulations of the Ali-Greub spectator model with parameters $\langle m_b \rangle = 4690$ MeV/c^2 and $p_F = 410$ MeV/c.

7.2.1   Extraction of Moments

We have two methods for calculating the first and second moments of the spectrum for $E_\gamma > 2.0$ GeV in the rest frame of the B meson. In the first method (which we creatively call "Method 1"), we correct the moments from the raw, measured spectrum for the energy dependence of the efficiency, calculate moments from that spectrum, and apply several corrections to obtain the moments of the spectrum in the rest frame of the B meson. In the second method, "Method 2", we use the best-fit values of $(\langle m_b \rangle, p_F)$ and obtain an expression for the idealized moments as functions of $\langle m_b \rangle$ and p_F by expanding about the best-fit point.

Method 1

In the first method, we start by correcting the moments from the raw, measured spectrum for the energy dependence of the efficiency. Next, we obtain the moments from the efficiency-corrected spectrum, and apply several corrections to the first and second moments.

There is only one correction to be made for the first moment – scaling it down by the ratio of the energy of the B meson in the lab frame to the B meson mass, 5.29/5.28. This takes the moment from the lab frame to the rest frame of the B meson. For the second moment, we need to correct for several effects that will broaden the width of the measured spectrum. From the second moment, we subtract the mean square widths due to the bin width $(100 \text{ MeV}/\sqrt{12})^2$, Doppler broadening $(\langle E_\gamma \rangle \vec{p}_B/\sqrt{3} M_B)^2$, and photon energy resolution $(0.026 \langle E_\gamma \rangle)^2$.

Finally, we apply an empirical correction obtained from the Monte Carlo (both K^* sum and JETSET samples). For each of the 48 $b \rightarrow s\gamma$ models, corresponding to the 24 $(\langle m_b \rangle, p_F)$ pairs times two methods of hadronization, K^* sum and JETSET, we run the MC through the analysis code to obtain a "measured" spectrum based on flattened weighted events with detector resolution and detector efficiency effects included. We extract the true first and second moments from the MC obtained at the generator level for $E_\gamma > 2.0$ GeV in the B meson rest frame. We compare the true moments (from MC) to the "measured" first and second moments (from MC).

The empirical correction is the difference between the true moment and the corrected moment. We apply the empirical correction associated with the best of the 48 $b \rightarrow s\gamma$ models, chosen by minimizing a χ^2 determined by the measured first (M_1) and second (M_2) moments. The χ^2 is calculated by taking the difference between the measured data and MC moments and normalized by the statistical error of the data:

$$\chi^2 = \frac{(M_{1_{Data}} - M_{1_{MC}})^2}{\sigma(M_{1_{Data}})^2} + \frac{(M_{2_{Data}} - M_{2_{MC}})^2}{\sigma(M_{2_{Data}})^2}$$

Table 7.2 summarizes the results for the first and second moments from Method 1.

Spectra	Raw Moments	True Moments	Empirical Correction (JETSET)	Empirical Correction (K^* Sum)
	$\langle E_\gamma \rangle$	$\langle E_\gamma \rangle$		
Nominal	2.347 ± 0.034	2.343	2.349 ± 0.034	2.344 ± 0.034
$+5\%$	2.359 ± 0.037	2.355	2.366 ± 0.037	2.361 ± 0.037
-5%	2.336 ± 0.031	2.332	2.338 ± 0.031	2.333 ± 0.031
	$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$	$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$		
Nominal	0.0299 ± 0.0064	0.0191	0.0224 ± 0.0065	0.0242 ± 0.0065
$+5\%$	0.0281 ± 0.0076	0.0173	0.0196 ± 0.0077	0.0214 ± 0.0077
-5%	0.0313 ± 0.0055	0.0206	0.0238 ± 0.0056	0.0256 ± 0.0056

Table 7.2: Results for First and Second Moments from Method 1. Given are raw moments (from data) as initially calculated, the true moments obtained from the best-fit MC, and the moments after the "empirical" correction, with empirical corrections taken from JETSET and K^* sum determinations.

Method 2

In the second method, we take our best fit Monte Carlo model, in which $\langle m_b \rangle$ and p_F are determined from a fit to the measured photon energy spectrum, and obtain the moments given by that model. This method of fitting the measured spectrum to the MC spectrum and obtaining the 1σ error ellipse has already been described in Section 6.3. Only one step remains: obtaining the first and second moments given the best fit value of ($\langle m_b \rangle$, p_F).

We obtain the true (QQ level) first and second moments for all 48 $b \rightarrow s\gamma$ MC models for $E_\gamma > 2.0$ GeV. This results in a set of points that define $\langle E_\gamma \rangle$ and $\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$ as functions of $\langle m_b \rangle$ and p_F. We fit these points to obtain the functions for the first and second moments. Since we only need the expression to be valid in the region bounded by the error ellipse, linear terms are sufficient to obtain an accurate expression for the moments. Using the spectrum obtained from the On$-$Off data, we obtain the best-fit $\langle m_b \rangle$ and p_F, as described in Section 6.3.

Given these expressions for the first and second moments, the central value of each is taken at the value of the best-fit point (obtained from data), and the statistical error on each is taken as half the difference between the high and low value as one travels around the error ellipse. Results for the first and second moments from Method 2 are given in Table 7.3.

Spectrum	JETSET	K^* Sum
	$\langle E_\gamma \rangle$	
Nominal	2.350 ± 0.031	2.344 ± 0.029
+5%	2.356 ± 0.033	2.352 ± 0.031
-5%	2.345 ± 0.030	2.338 ± 0.029
	$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle$	
Nominal	0.0217 ± 0.00700	0.0226 ± 0.0066
+5%	0.0211 ± 0.0076	0.0216 ± 0.0069
-5%	0.0222 ± 0.0065	0.0234 ± 0.0063

Table 7.3: Results for First and Second Moments from Method 2.

Combining the Results

The central value is taken to be the average of the 4 nominal numbers such that $\langle E_\gamma \rangle = 2.346$ GeV and $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2 = 0.0226$ GeV^2. The statistical errors are taken to be the straight average of the statistical errors given on the 4 nominal numbers, or ± 0.032 GeV for the first moment, and ± 0.0066 GeV^2 for the second moment.

The first and second moments of the spectrum also need to be corrected for the $b \rightarrow d\gamma$ component. We calculated the first and second moments for $b \rightarrow d\gamma$ and $b \rightarrow s\gamma$ MC samples and found that the first moment of $b \rightarrow d\gamma$ is higher than the $b \rightarrow s\gamma$ first moment, therefore we correct the first moment down by 0.001 ± 0.005 GeV. The second moment of a mixed sample of 95% $b \rightarrow s\gamma$ MC and 5% $b \rightarrow d\gamma$ is larger than pure $b \rightarrow s\gamma$, so we correct the second moment down by $(7 \pm 4) \times 10^{-5}$ GeV^2.

The systematic error includes the following contributions:

- $B\overline{B}$ **subtraction** The systematic error is taken to be half the straight average of the four differences between -5% and $+5\%$, or ± 0.010 GeV for the first moment, $\pm(1.44 \times 10^{-3})$ GeV^2 for the second moment.

- **Uncertainty in the On-Off subtraction bias** The error is determined by creating spectra with 2.5% over-subtraction and 2.5% under-subtraction, exaggerating the effect of this uncertainty by a factor of 5. We calculate the moments from the two spectra using Method 1, and take 1/10 of their difference as the systematic error. In this way, we find an error of $\pm(1.1 \times 10^{-3})$ GeV for the first moment, and $\pm(9.1 \times 10^{-4})$ GeV^2 for the second moment.

- **Hadronization** The systematic error is half the straight average of the two differences between the nominal K^* sum and JETSET samples. It is $\pm(2.9 \times 10^{-3})$ GeV for the first moment and $\pm(6.5 \times 10^{-4})$ GeV^2 for the second moment.

- **Method** The systematic error due to method determined from the absolute values of the difference between Method 1 and Method 2 for all three spectra. The error is taken to be half the average of the 6 differences and is $\pm(2.8 \times 10^{-3})$ GeV for the first moment and $\pm(6.5 \times 10^{-4})$ GeV^2 for the second moment.

- $b \to d\gamma$ **Correction** The error on the correction to the first moment is $\pm(5 \times 10^{-3})$ GeV and to the second moment is $\pm(4 \times 10^{-5})$ GeV^2

The final values for our first and second moments are as follows (first error is statistical, second is systematic):

$$\langle E_\gamma \rangle = 2.35 \pm 0.03 \pm 0.01 \; GeV$$

$$\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2 = 0.023 \pm 0.007 \pm 0.002 \; GeV^2$$

CHAPTER 8

Interpretation of Results

8.1 Standard Model Implications

Because the measurement and theoretical prediction for $\mathcal{B}(b \to s\gamma)$ are in such good agreement, there is not much room for new physics. Physics beyond the SM affects the branching fraction and manifests itself in two ways. First, extensions to the Standard Model contribute additional Feynman diagrams (at the high energy scale) and modify the values of the Wilson coefficients in the effective theory (low energy scale). Second, new operators can appear that are not present in the Standard Model calculation of the branching fraction. There are over 100 references in SPIRES[10] to the original $b \to s\gamma$ paper, most of them involving the implications of the branching fraction measurement on beyond-SM physics [23].

Our measurement of the branching fraction can be used to extract the magnitudes of the relevant Wilson coefficients, $\mathcal{C}_{7,8}$. These values for the Wilson coefficients can be compared to various model predictions, and constraints can be placed on the parameters of several extensions to the Standard Model, for example, supersymmetry, left-right symmetric models, and multi-Higgs models [19]. To give the reader an idea of what is possible with this measurement, we list several models that have, in the past, used $\mathcal{B}(b \to s\gamma)$ to constrain them:

- Two Higgs doublet Model

  In the "Type II" Two Higgs doublet Model [56],[57], one doublet gives mass to the up-type quarks, while the other gives mass to the down type quarks. The presence of these charged Higgs particles would increase the branching fraction through $\mathcal{C}_7$ and $\mathcal{C}_8$. Using their prediction for $\mathcal{B}(b \to s\gamma)$ for the Two Higgs Doublet Model II, Gambino and Misiak have ruled out the possibility of a charged Higgs boson lighter than 350 GeV [22].

---

[10]http://www.slac.stanford.edu/spires/hep

- Left-right supersymmetric model

  This model introduces new neutral and charged gauge bosons, Z_R and W_R, as well as a right-handed gauge coupling, g_R. Because the $b \to s\gamma$ decay can be mediated by left-handed and right-handed W bosons as well as other supersymmetric particles, it is sensitive to the parameters of this model. Although the model contains too many parameters to allow for a precise restricion on any single one, some general constraints can be obtained [58],[59],[60].

- Minimal Supersymmetric Model (MSSM)

  The MSSM introduces a large number of superpartners of the Standard Model particles. These supersymmetric partners can contribute to the $B \to X_s\gamma$ branching fraction. Since the experimental value for the branching fraction agrees so well with the theoretical predictions, limits can be placed on the contributions of certain combinations of these supersymmetric partners [61],[62].

- Constraining Fourth Generation with $B \to X_s\gamma$

  If fourth generation fermions exist, the new quarks would enhance the $B \to X_s\gamma$ branching fraction. Together with a measurement of the CP asymmetry in $B \to X_s\gamma$, it is possible to constrain the parameter space of the fourth generation [63],[64].

8.2   HQET Parameters

Measuring the inclusive photon energy spectrum allows us to learn about the motion of the b quark inside the B meson. This knowledge leads to a better understanding of QCD, especially in the way that QCD affects the lepton spectrum in $b \to ul\nu$ decays. The moments of the $B \to X_s\gamma$ spectrum can be used to experimentally determine several nonperturbative HQET parameters which determine the quark pole mass and kinetic energy.

Inclusive semileptonic B decays have been calculated using an operator product expansion (OPE). The OPE and the parton model predictions are equal in the heavy quark limit, $m_B \to \infty$. Observables can be written as the parton model expectation plus nonperturbative corrections, expansions in inverse powers of the B meson mass, M_B.

At order $1/M_B$, the nonperturbative parameter $\overline{\Lambda}$ appears, and at order $1/M_B^2$, two more parameters, λ_1 and λ_2 enter. These parameters may be thought of as the energy of the light quark and gluon degrees of freedom $(\overline{\Lambda})$, the average momentum-squared of the b quark $(-\lambda_1)$, and the energy of the hyperfine interaction of the spin of the b quark with the light degrees of freedom (λ_2/M_B). The parameter $\lambda_2 = 0.128$ GeV^2 is determined from the $B^* - B$ mass difference. The parameters ρ_1, ρ_2, $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$, and $\mathcal{T}_4$ appear at order $1/M_B^3$ and are estimated, from dimensional

$$\langle E_\gamma \rangle = \frac{M_B}{2}\Big[1 - .385\frac{\alpha_s}{\pi} - .620\beta_0\big(\frac{\alpha_s}{\pi}\big)^2 - \frac{\bar\Lambda}{M_B}\big(1 - .954\frac{\alpha_s}{\pi} - 1.175\beta_0\big(\frac{\alpha_s}{\pi}\big)^2\big)$$
$$- \frac{13\rho_1 - 33\rho_2}{12M_B^3} - \frac{\mathcal{T}_1 + 3\mathcal{T}_2 + \mathcal{T}_3 + 3\mathcal{T}_4}{4M_B^3} - \frac{\rho_2 \mathcal{C}_2}{9M_B M_D^2 \mathcal{C}_7} + \mathcal{O}(1/M_B^4)\Big]. \qquad (8.1)$$

$$\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle = M_B^2\Big[\frac{-\lambda_1}{12M_B^2} + \big(0.00815\frac{\alpha_s}{\pi} + 0.01024\beta_0\big(\frac{\alpha_s}{\pi}\big)^2\big)$$
$$\frac{-\bar\Lambda}{M_B\big(0.05083\frac{\alpha_s}{\pi} + 0.05412\beta_0\big(\frac{\alpha_s}{\pi}\big)^2\big)}$$
$$- \frac{2\rho_1 - 3\rho_2}{12M_B^3} - \frac{\mathcal{T}_1 - 3\mathcal{T}_2}{12M_B^3} + \mathcal{O}(1/M_B^4). \qquad (8.2)$$

considerations, to be $\sim (0.5\mathrm{GeV})^3$. The expressions for the moments of the photon energy spectrum in $B \to X_s\gamma$, for $E_\gamma > 2.0$ GeV to order $\beta_0\alpha_s^2$ (where β_0 is the one-loop QCD β function) and $1/M_B^3$ are given in Equations 8.1 and 8.2.

The expression for the first moment converges rapidly in the $1/M_B$ expansion. Taking the values of $\mathcal{T}_i$ and ρ_2 to be $0.0 \pm (0.5~\mathrm{GeV})^3$ and the value of ρ_1 to be $1/2(0.5~\mathrm{GeV})^3 \pm 1/2(0.5~\mathrm{GeV})^3$, and taking $\lambda_2 = 0.128~\mathrm{GeV}^2$ the expression for the first moment combined with our measurement for the first moment defines a band in $\bar\Lambda - \lambda_1$ space and is shown in Figure 8.3. The calculation of the moments contains an error from the measurement as well as an error from the theoretical expression, in particular from the neglected $1/M_B^3$ terms, and the uncertainty of the scale to use for α_s, which we take to range from $m_b/2$ to $2m_b$. We use $\alpha_s(m_b) = 0.220$, $\alpha_s(m_b/2) = 0.275$, $\alpha_s(2m_b) = 0.176$, $\beta_0 = (33 - 2n_f)/3$ and $n_f = 4$. We take the B meson mass to be $M_B = 5.280$ GeV, $\mathcal{C}_2 = 1.11$, and $\mathcal{C}_7 = -0.32$.

The expression for the second moment converges slowly in $1/M_B$, so we do not use the second moment for the extraction of expansion parameters. It is shown in Figure 8.1 only to illustrate the size of the error.

Using the equation for the first moment and the parameter values listed in Equation 8.1 and 8.2, we obtain $\bar\Lambda = 0.35 \pm 0.08 \pm 0.10$ GeV, where the first error is from the experimental error in the determination of the first moment, and the second error is from the theoretical expression, in particular from the neglected $1/M_B^3$ terms, and the uncertainty of the scale to use for α_s.

Only $\bar\Lambda$ can be determined from the $b \to s\gamma$ spectrum alone. Although in theory we could extract λ_1 using the second moment, in practice, the errors are too large to allow an accurate determination of this parameter. However, it is possible to obtain λ_1 by using the first moment of the hadronic mass-squared distribution in $b \to cl\nu$ decays. CLEO recently measured the first and second moments of the distribution in the hadronic mass-squared in the inclusive semileptonic decay $b \to cl\nu$ for leptons restricted to the region $P_l \geq 1.5$ GeV/c [65]. These moments are also related to the
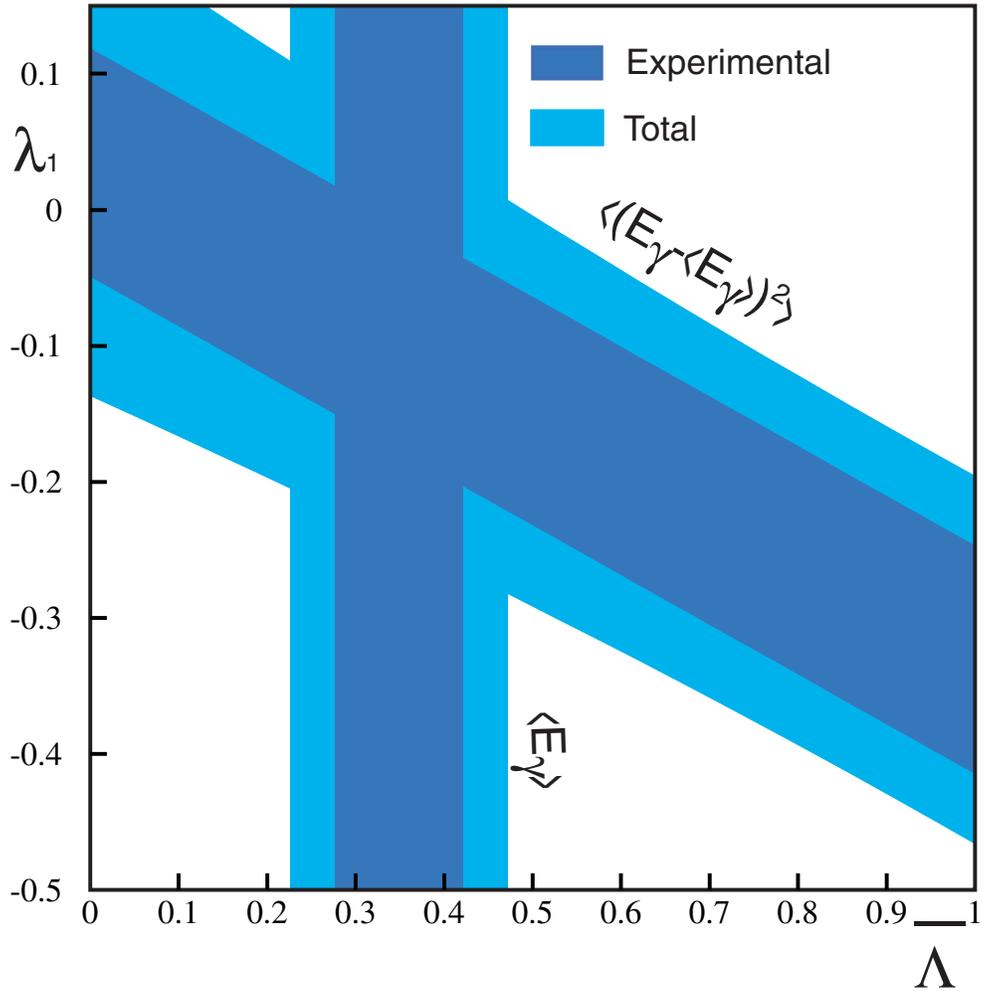
Figure 8.1: Bands in $\overline{\Lambda} - \lambda_1$ space allowed by our measurements of $\langle E_\gamma \rangle$ and $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2$. The dark bands indicate the error from the measurement while the light bands include the errors from the theory.

$$\frac{\langle M_X^2 - \bar{M}_D^2 \rangle}{M_B^2} = [0.0272\frac{\alpha_s}{\pi} + 0.058\beta_0\left(\frac{\alpha_s}{\pi}\right)^2 + 0.207\frac{\bar{\Lambda}}{M_B}\left(1 + 0.43\frac{\alpha_s}{\pi}\right)$$
$$+0.193\frac{\bar{\Lambda}^2}{M_B^2} + 1.38\frac{\lambda_1}{M_B^2} + 0.203\frac{\lambda_2}{M_B^2} + 0.19\frac{\bar{\Lambda}^3}{M_B^3} + 3.2\frac{\bar{\Lambda}\lambda_1}{M_B^3} + 1.4\frac{\bar{\Lambda}\lambda_2}{M_B^3}$$
$$+4.3\frac{\rho_1}{M_B^3} - 0.56\frac{\rho_2}{M_B^3} + 2.0\frac{\mathcal{T}_1}{M_B^3} + 1.8\frac{\mathcal{T}_2}{M_B^3} + 1.7\frac{\mathcal{T}_3}{M_B^3} + 0.91\frac{\mathcal{T}_4}{M_B^3} + \mathcal{O}(1/M_B^4)]. \quad (8.3)$$

$$\frac{\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle}{M_B^4} = [0.00148\frac{\alpha_s}{\pi} + 0.0025\beta_0\left(\frac{\alpha_s}{\pi}\right)^2 + 0.027\frac{\bar{\Lambda}}{M_B}\frac{\alpha_s}{\pi} + 0.0107\frac{\bar{\Lambda}^2}{M_B^2} - 0.12\frac{\lambda_1}{M_B^2}$$
$$+0.02\frac{\bar{\Lambda}^3}{M_B^3} - 0.06\frac{\bar{\Lambda}\lambda_1}{M_B^3} - 0.129\frac{\bar{\Lambda}\lambda_2}{M_B^3} - 1.2\frac{\rho_1}{M_B^3} + 0.23\frac{\rho_2}{M_B^3} - 0.12\frac{\mathcal{T}_1}{M_B^3} - 0.36\frac{\mathcal{T}_2}{M_B^3}$$
$$+\mathcal{O}(1/M_B^4)]. \quad (8.4)$$

parameters $\overline{\Lambda}$ and λ_1. The expressions for the hadronic mass moments in $B \rightarrow X_c l \nu$ to order $\beta_0 \alpha_s^2$ and $1/M_B^3$ subject to the restriction $P_l \geq 1.5$ GeV/c are given in Equations 8.3 and 8.4. These expressions are only approximate, but are believed to be good to $\pm 50\%$. The values of λ_2, ρ_1, ρ_2, α_s, and $\mathcal{T}_1 - \mathcal{T}_4$ are taken to be the same as before and the bands in $\overline{\Lambda} - \lambda_1$ space are shown in Figure 8.2.

The expression for the first moment converges rapidly in the $1/M_B$ expansion, and the second moment converges slowly in $1/M_B$. Consequently, the second moment is not used to define a band in $\overline{\Lambda} - \lambda_1$ space. It is shown in Figure 8.2 as an illustration of the size of the error.

Even though we cannot trust the second moment of either individual measurement to give an accurate extraction of λ_1, we can use the two first moments together to obtain this parameter. The intersection of the two bands from the first moments determines $\overline{\Lambda}$ and λ_1, and is shown in Figure 8.3 along with a $\Delta\chi^2 = 1$ error ellipse. The inner bands of Figure 8.3 represent the error bands from the measurement while the light grey extensions include the errors from the theory. The values obtained are

$$\overline{\Lambda} = 0.35 \pm 0.08 \pm 0.10 \; GeV,$$

$$\lambda_1 = -0.236 \pm 0.071 \pm 0.0078 \; GeV^2.$$

Here, the first error is from the experimental error on the determination of the two moments, and the second error is from the theoretical expression. Specifically, the theory error on $\overline{\Lambda}$ comes from $(1/M_B)^3$, ± 0.0039, and from the scale uncertainty in α_s, ± 0.092. The theory error on λ_1 also comes from $(1/M_B)^3$, ± 0.0065, and from the scale uncertainty in α_s, ± 0.041. The answers derived by using the first and second moments from both spectra to determine these parameters differ little both in central values and and in errors.
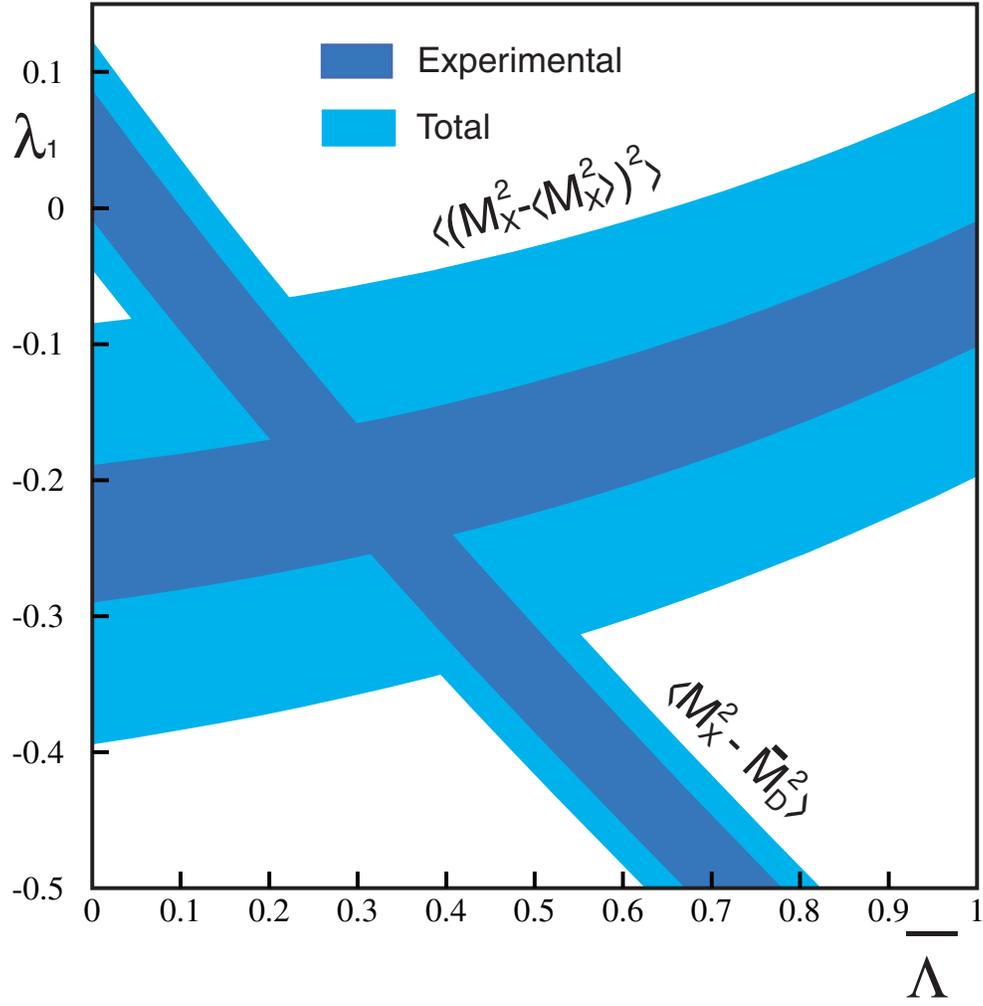
Figure 8.2: Bands in $\overline{\Lambda} - \lambda_1$ space defined by $\langle M_X^2 - \overline{M}_D^2 \rangle$ and $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$, the measured first and second moments of hadronic mass-squared. The inner bands indicate the error bands from the measurement. The light grey extensions include the errors from the theory.
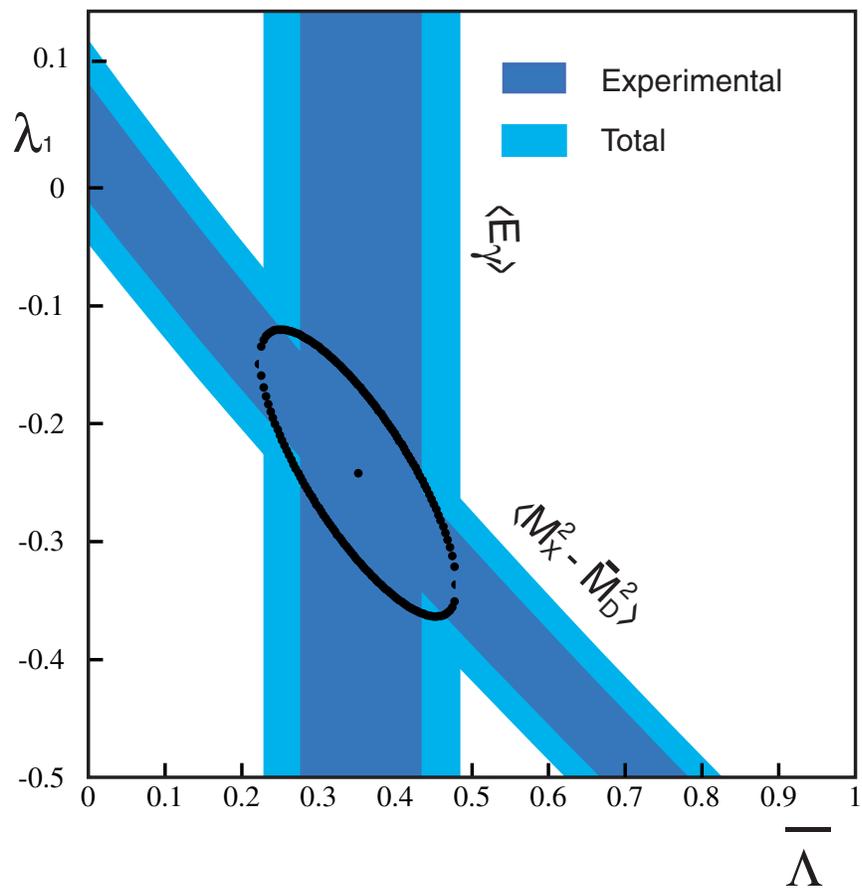
Figure 8.3: Bands in $\overline{\Lambda} - \lambda_1$ space defined by $\langle M_X^2 - \overline{M}_D^2 \rangle$, the measured first moment of hadronic mass-squared, and $\langle E_\gamma \rangle$, the first moment of the photon energy spectrum in $b \to s\gamma$.

$$\Gamma_{sl} = \frac{G_F^2 |V_{cb}|^2 M_B^5}{192\pi^3} 0.3689 [1 - 1.54\frac{\alpha_s}{\pi} - 1.43\beta_0(\frac{\alpha_s}{\pi})^2 - 1.648\frac{\bar{\Lambda}}{M_B}(1 - 0.87\frac{\alpha_s}{\pi})$$
$$-0.946\frac{\bar{\Lambda}^2}{M_B^2} - 3.185\frac{\lambda_1}{M_B^2} - 0.02\frac{\lambda_2}{M_B^2}$$
$$-0.298\frac{\bar{\Lambda}^3}{M_B^3} - 3.28\frac{\bar{\Lambda}\lambda_1}{M_B^3} + 7.997\frac{\bar{\Lambda}\lambda_2}{M_B^3} - 6.153\frac{\rho_1}{M_B^3} + 7.482\frac{\rho_2}{M_B^3} - 7.4\frac{\mathcal{T}_1}{M_B^3}$$
$$+1.491\frac{\mathcal{T}_2}{M_B^3} - 10.41\frac{\mathcal{T}_3}{M_B^3} - 7.482\frac{\mathcal{T}_4}{M_B^3} + \mathcal{O}(1/M_B^4)] \ . \tag{8.5}$$

8.2.1 $|V_{cb}|$

Given the values of $\bar{\Lambda}$ and λ_1, we can improve the determination of the CKM matrix element $|V_{cb}|$ from the measured $B \to X_c l\nu$ semileptonic width. The expression for the semileptonic width to order $\beta_0\alpha_s^2$ and $1/M_B^3$ is given in Equation 8.5.

For the experimental determination of Γ_{sl}, we use: $M_B = 5.280$ GeV, $\mathcal{B}(B \to X_c \ell\nu) = (10.39 \pm 0.46)\%$ [66], $\tau_{B^{\pm}} = (1.548 \pm 0.032)$ ps [1], $\tau_{B^0} = (1.653 \pm 0.028)$ ps [1], $f_{+-}/f_{00} = 1.04 \pm 0.08$ [50], giving $\Gamma_{sl} = (0.427 \pm 0.020) \times 10^{-10}$ MeV.

Combining the measured semileptonic width with the theoretical expression for it, and using the determination of $\bar{\Lambda}$ and λ_1 from the first moments, we find

$$|V_{cb}| = (4.04 \pm 0.09 \pm 0.05 \pm 0.08) \times 10^{-2},$$

where the errors are from experimental determination of Γ_{sl}, from experimental determination of $\bar{\Lambda}$ and λ_1, and from the $1/M_B^3$ terms and scale uncertainty in α_s, in that order [65]. This gives a determination of $|V_{cb}|$ from inclusive processes, with an accuracy of $\pm 3.2\%$. Breaking the theory errors down, $\pm(3.5 \times 10^{-4})$ comes from $(1/M_B)^3$ and $\pm 7.2 \times 10^{-4})$ from scale uncertainty in α_s.

Other important parameters, such as the CKM matrix element V_{ub} can be expressed in terms of the HQET parameters $\bar{\Lambda}$ and λ_1. By measuring the lepton yield in the endpoint region of the $b \to u l\nu$ spectrum and combining it with the information obtained from the shape of the $b \to s\gamma$ spectrum, one can determine the CKM matrix element V_{ub} without incurring large model dependence [67],[19].

There are two problems associated with the interpretation of measured inclusive properties, one associated with the convergence of the expansion, and another with the validity of the assumptions underlying the expansion. The inclusive observables are expansions in powers of $1/M_B$ and at each order, more nonperturbative parameters appear. By order $1/M_B^2$ there are three parameters and at order $1/M_B^3$ another six parameters, ρ_1, ρ_2, $\mathcal{T}_1 - \mathcal{T}_4$. Without good estimates for the additional parameters we must rely on the rapid convergence of the expansion. The other problem is the validity of the assumption of parton-hadron duality implicit in this approach. This error is unquantifiable and not included in our estimates. The experimental

determination of $\overline{\Lambda}$ and λ_1 with several different methods is necessary to support the validity of parton-hadron duality.

8.3   Conclusion

We have measured the branching fraction and photon energy spectrum of $b \to s\gamma$ for photons with $E_\gamma > 2.0$ GeV, a significant improvement over previous measurements [35]. The branching fraction, $\mathcal{B}(b \to s\gamma) = (3.21 \pm 0.43 \pm 0.27^{+0.018}_{-0.010}) \times 10^{-4}$, is in good agreement with SM predictions and can be used to place limits on beyond-SM physics. We used the moments of the photon energy spectrum ($\langle E_\gamma \rangle = 2.35 \pm 0.03 \pm 0.01$ GeV and $\langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2 = 0.023 \pm 0.007 \pm 0.002\,\mathrm{GeV}^2$) to extract the HQET parameters $\overline{\Lambda}$ and λ_1 ($\overline{\Lambda} = 0.35 \pm 0.08 \pm 0.10$ GeV and $\lambda_1 = -0.236 \pm 0.071 \pm 0.0078$ GeV^2), which can be used to extract V_{cb}. In addition, the shape of the spectrum can also be used in combination with the lepton yield in the endpoint region of the $b \to ul\nu$ spectrum to obtain V_{ub} without large model dependence. This is very useful for doing precision CKM measurements.

Future efforts at CLEO will be aimed at reducing the systematic error especially in the $1.5 - 1.8$ GeV region. Improvements to this part of the spectrum will come at a small sacrifice in yield, but will allow us to more precisely measure the shape of the spectrum, something which is drawing significant attention among theorists because of the potential for precisely determining V_{ub} and constraining the Standard Model.

APPENDIX A

Track/Photon Selection and Particle Identification

A.1   Track Selection

The charged tracks used in the shape analysis and pseudoreconstruction must pass the following standard track quality cuts:

- Pass TRKMNG

- KINCD ≥ 0

- $|PQCD| \leq 5.5$ GeV/c

- $0.0 < |CZCD| < 0.9$

- $0.0 < |DBCD| < 0.01$

- $0.0 < |Z0CD| < 0.05$

- $0.0 < |RESICD| < 0.001$

- Must be able to assign a reasonable particle type to the track.

There is a second list of tracks with the same requirements above except that the track is allowed to fail Z0CD and/or DBCD. These are tracks from neutral particles that decayed in flight so their tracks do not come from the origin. This second track list, used to form composite particles such as $K_s^0 \to \pi^+ \pi^-$, must fulfill the following requirements:

- Pass TRKMNG

- KINCD ≥ 0

- $|PQCD| \leq 5.5$ GeV/c

- $0.0 < |CZCD| < 0.9$

- $0.0 < |RESICD| < 0.001$

- Must be able to assign a reasonable particle type to the track.

These tracks are required to come from the origin and must pass passing the CLEO track quality package, TRKMNG (TRKMAN the Next Generation) [42]. When a charged particle passes through the drift chambers, it generates a set of hits to which the tracking programs will fit a track. Unfortunately, more than one track can be generated from a single set of hits generated by one particle. The purpose of TRKMNG is to appropriately discard the extra tracks so that one particle results in only one track that is the best description of its matching particle. The requirement that KINCD be greater than or equal to zero flags the track as good based on the CLEO tracking package. It means that the track did not have too many missing hits or bad residuals and it is not believed to be the back end of a curler. Eliminating tracks with PQCD greater than 5.5 GeV/c is equivalent to rejecting tracks with an unphysical momentum resulting from a bad fit to the hits. CZCD is the z-direction cosine and a value of 0.9 ensures that the particle that caused the track did not escape down the beam pipe. DBCD is the signed radial distance between the interaction point and the beam line. Z0CD is the z-coordinate at that point, the point of closest approach to the origin. Together these last three requirements constrain the track to come from the origin; they allow the track to originate from a point that is radially within 50 cm of the beam line and within 10 cm of the interaction point in z. RESICD is the mean square residual per degree of freedom for the hits assigned to the track. It is the RMS distance between the fit track and the position of the track's wire chamber hits and is a measure of the quality of the fit.

A.2 Photon Selection

The full list of cuts used to select candidate photons is given below. The description of these cuts use CLEO-specific vocabulary and is intended for the reader who is well-versed in CLEO analysis and wants the nitty gritty.

A.2.1 Shower selection

We use these showers to perform our π^0/η veto with the photon candidate and in the pseudoreconstruction of the event. In order to be considered a good shower caused by a neutral particle, a cluster in the calorimeter must fulfill the following requirements:

- $E_\gamma > 50$ MeV

- $|cos\theta| < 0.95$

- Must not be matched to a charged track (no type-1 or type-2 match)

- Unfolded E9/E25 (LP5SH) ≥ 0.95

- Must pass BADSH requirement

- Must pass SPLITF

A.2.2 Candidate Photon Selection

The following is a full list of the cuts used to select a candidate photon:

- 1.5 GeV $< E_\gamma <$ 5.0 GeV

- $|cos\theta| < 0.7$

- Must not be matched to a charged track (no type-1 or type-2 match)

- E9/E25 (LP5SH) ≥ 0.95

- (LP2SH $\times E_{candidate}$/RADSH) $\leq 0.035 + (E_{candidate} - 0.5) \times 0.028$

- BADSH $= 0$

- Must pass SPLITF

- Must pass π^0 and η vetoes

The shower mass cut, which requires that $m_{shr} < 0.035 + E_{candidate} - 0.5) \times 0.028$, was evaluated for the original analysis and details about how it was evaluated can be found in reference [40].

A candidate photon is rejected if it can be combined with another shower in the event to form a π^0 with $110 < m_{\gamma\gamma} < 150$ MeV/c^2. Requirements on the sibling photon for the π^0 veto are as follows:

- $E_\gamma > 30$ MeV for $|cos\theta| < 0.708$

- $E_\gamma > 50$ MeV for $0.708 < |cos\theta| < 0.95$

A candidate photon is rejected if it can be combined with another shower in the event to form an η with $515 < m_{\gamma\gamma} < 575$ MeV/c^2. Requirements on the sibling photon for the η veto are as follows:

- $E_\gamma > 200$ MeV for all sibling photons

A.3 Particle Identification

Our strategy is to treat the track as a pion unless we find that it is better suited to some other particle. This selection process is as follows.

The information provided from the time of flight (TOF) counters and the dE/dx measurements is given in the form of four variables for a given track. These are the standard deviation of the measurement from that expected for a $\pi^{\pm}, K^{\pm}, p/\overline{p}$, and $e^{\pm}$. Whether we use information from TOF, dE/dx, or both to determine the particle type depends on whether these devices are reporting reliable data. If a track fails any of the conditions below, we do not use time of flight information to identify the particle:

- $|CZCD| < 0.7$

  This requirement ensures that the tracks pass through the TOF counters in the barrel. These counters give the cleanest measurement because they are not obstructed by endcap material.

- TFIDQL $\neq 8$

  TFIDQL is the quality word returned by the CLEO time of flight analysis package. A value of 8 indicates that the pulse-height from a photo-tube used in the TOF measurement for this track has been saturated. This value assures that the time of flight information is of good quality.

- NTUBTF $= 2$

  A value of NTUBTF $= 2$ means that the time of flight measurement was determined from the barrel.

- Adam Lyon's ISRUNTOFBAD routine determines that the TOF info for this run is bad [68].

- If all four of the time of flight standard deviations (e, μ, π, K, p) are greater than three, then the TOF values are flaky, and the time of flight information for this track is ignored.

If a track fails the following condition, we do not use dE/dx information:

- IQALDI $\neq 0$

  IQALDI is the quality of the track as reported by the CLEO dE/dx subroutines. A value of zero means that there is no dE/dx information available for that track.

Once we have determined which device has reliable data, we calculate sigmas based on the available TOF and dE/dx information. If the track has no reliable

information from either time of flight or dE/dx, then it won't get any type at all. If only dE/dx is good, then we calculate the chi-square probability per degree of freedom for each hypothesis ($e/\mu/\pi/K/p$) for this track. This is then converted to a sigma. If only the time of flight is good, then we consider it a bad track and don't give it any type at all. If both dE/dx and the TOF information are good, then we use both pieces of information by combining them to form an effective sigma, one for each of the four particle type hypotheses.

A.3.1   Identifying hadrons

The charge of the particle is determined from the the direction of curvature of the track (the sign of PQCD). To assign a particle type to a given track, we examine the the sigmas for the hypotheses in the following order:

- If $\sigma_{proton} \leq 3.0$ then particle is a proton

- If $\sigma_K \leq 3.0$ then particle is a kaon

- If $\sigma_\pi \leq 3.0$ then particle is a pion

In each case the particle is identified as the hypothesis that is within 3σ of the expected value. If a particle is within 3σ of a π, then it will be a π no matter what the other sigmas are. Next, we override these hadron type settings if the particle passes our muon or electron id criteria.

A.3.2   Identifying leptons

We use the Rochester electron ID package (REID) to determine whether a given track is an electron or not. The REID package combines E/p, dE/dx, E9/E25, and other variables using a likelihood fit to produce a variable that discriminates between electrons and hadrons at various momenta in different regions of the detector.

The following conditions must be met if a track is to be called an electron:

- $|PQCD| \geq 600$ MeV/c

  Since REID works best for stiffer tracks, we do not attempt an electron iden-
  tification if the track has a momentum of less than 600 MeV/c.

- IRANGE $\neq 0$ and IRANGE $\neq -2$

  IRANGE is a variable that indicates where in the calorimeter the track hit:
  good barrel, bad barrel, overlap barrel, good endcap, bad endcap. A value
  of -2 means that the track went into the bad endcap while a value of zero
  indicates that there was no hit at all in the calorimeter for this track. We only
  attempt an electron identification if the track went into the barrel or the good
  endcap.

Next, we look at the LIKERB variable in the good and bad barrel and for different track momenta to determine whether the track is likely to be an electron. A track is called an electron if any of the following sets of conditions is true:

- Good Barrel (IRANGE = 1)

  If a track hits the calorimeter in the good barrel and has $|PQCD| > 250$ MeV/c, we consider it an electron if LIKERB > 3.

- Not good barrel (IRANGE = 1, 2, 3, -1)

  - If a track hits the calorimeter outside the good barrel and has a momentum in the range 0.8-1.8 GeV/c ($|PQCD| > 0.8$ GeV/c and $|PQCD| < 1.8$ GeV/c, then it is called an electron if LIKERB > 0.6.

  - If a track hits the calorimeter outside the good barrel and has a momentum greater than 1.8 GeV/c ($|PQCD| > 1.8$ GeV/c), then it is an electron if LIKERB > 0.4.

Muon identification relies on the fact that only muons can penetrate through one or more layers of the iron return yoke to reach the muon chambers within. A track is considered a muon if the following conditions are satisfied:

- MUQUAL = 0

  MUQUAL is the muon track quality flag. A value of zero implies that there are muon hits everywhere along the track that muon hits were expected.

- DPTHMU > 5

  DPTHMU is the maximal depth in absorption lengths at which the track correlates to good quality hits in the muon detector. A good quality muon requires at least two layers hit in the muon chambers. A DPTHMU of five is 75% efficient for muons [43].

Once we have determined whether the given track can be identified as a muon or electron, we override all previous particle type settings. First we check to see if it has been classified as an electron. If yes, we assign it an electron particle type. Next, we check to see if it has been classified as a muon. If yes, then we override all previous type settings and call it an electron.

Finally, we determine alternate particle types for each of these tracks. For example, if a particle was previously called a pion, then we assign the next likely particle type (proton or kaon).

APPENDIX B

Loose Shape Cuts

Cuts 1 and 2 are applied as soon as we have a list of candidate photons. Cut 3 requires only that the eight shape variables are calculated. Once the shape variables are determined, cut 3 is applied and events are rejected. Cuts 4 and 5 require the use of all of the particles in the event and are not applied until after we have attempted to pseudoreconstruct the event. Cut 5 is applied after the shape variables have been calculated, after the neural net values are calculated, and after the final-judgment variables are determined. Here is a list and a description of these loose shape cuts.

- **CUT 1:** The event must have at least one candidate photon.

- **CUT 2:** We select hadronic events by requiring KLASGL = 10. We reject events that have a track with momentum greater than 2.5 GeV/c. (Since B mesons are produced nearly at rest, they will rarely contain a track with such high momentum whereas continuum events often will). We also do not allow more than one good shower over 2.4 GeV.

- **CUT 3:** Loose cuts on shape variables. These cuts were defined in the previous analysis [40], and have the virtue of cutting out very little signal while removing some background.

  - $0.0 < R_2 < 0.55$
  - $0.1 < S_\perp < 0.8$
  - $0.0 < R_2' < 0.35$
  - $0.45 < cos\theta' < 1.0$
  - $0.5 < R_2 + S_\perp < 1.0$
  - $-0.7 < R_2 - S_\perp < 0.7$

- **CUT 4:** If we cannot reconstruct any B meson candidate at all, even by performing the trivial reconstruction in which we identify the highest momentum

π and K recoiling against the photon as a B candidate, then we discard the event. Only 1 in 200 events fails to make a trivial reconstruction.

- **CUT 5:** This cut is used to throw out pathological events in which we cannot calculate the shape variables. A handful of these events were investigated and found to be cases in which we tried to compute R_2 with only two tracks that were almost back-to-back. Only about 1 in 3000 events fails this cut.

APPENDIX C

Other Backgrounds

Several modes capable of producing high energy photons were not included in our default $B\overline{B}$ MC sample. We describe them here and summarize their contributions to each energy bin in Tables C.4 and C.5. In general, our procedure for obtaining a correction was to generate a sample of MC for the missing process, select high energy photons and run the $b \to s\gamma$ analysis code on the sample to obtain the number of weights contributed by the missing process.

C.1 FSR Subtraction

Generic $b \to c$ MC does not include final state radiation, photons radiated from the B meson prior to decay. We generated a sample of generic $B\overline{B}$ MC events and required that the event have a photon with energy above 1.5 GeV and that its parent be a B meson.

C.2 η' Yields

Our default MC sample underestimated the η' yield. Since $\eta' \to \gamma\gamma$, a discrepancy in the η' yield will correspond to a discrepancy in the photon yield. We generate additional MC so that the sum of our exclusive η' modes in the MC add up to the measured inclusive rate for η', 6×10^{-4} for $2.0 < P_{\eta'} < 2.7$ GeV/c [69].

C.3 ω Yields

Decays such as $\omega \to \pi\gamma$ can produce photons with a high enough energy to be considered a background to the analysis. Several ω producing processes are not accurately represented in our default MC. These are given in Table C.3 with the branching fraction/upper limit with which they were included in our MC. From our ω MC sample, we selected events containing a photon with an energy of at least 1 GeV and obtained the number of weights contributed to the background.

Mode	Branching Fraction
$\overline{B} \to D^{*+}\omega\pi^-$	0.29×10^{-2}
$B \to D^{*0}\omega\pi^-$	0.45×10^{-2}
$\overline{B} \to D^+\omega\pi^-$	0.28×10^{-2}
$B \to D^0\omega\pi^-$	0.41×10^{-2}
$\overline{B}^0 \to D^0\omega$	5.1×10^{-4} (Upper Limit)
$\overline{B}^0 \to D^{*0}\omega$	7×10^{-4} (Upper Limit)
$B^- \to D^{(*)-}\omega$	0
$\omega \to \pi\gamma$	8.5%

Table C.1: Branching fractions for the recently observed $B \to D^{(*)}\omega\pi$ decays [3],[4].

C.4 Radiative ψ Decays

The ψ meson is composed of a $c\overline{c}$ quark pair and can decay to lighter mesons by radiating a photon. Some of these transitions can produce a photon spectrum that could contribute above 1.5 GeV. The default $b \to c$ MC does *not* include any radiative decays of the ψ meson, so we generate a Monte Carlo sample that includes several two-body and three-body decays of the ψ meson. The modes we included and their branching fractions are listed in Tables C.2.

The sum of the branching fractions of the modes included in our MC is $1.94\% \approx 2\%$. The $\psi \to \gamma e^+ e^-$ mode is not included in our MC because the photon spectrum it produces is too soft to contribute to the region above 1.5 GeV. We split up the contributions into two-body and three-body modes and list the results in Table C.5.

C.5 $a_1 \to \pi\gamma$ and $\rho \to \pi\gamma$

The default MC does *not* include the radiative decays of a_1, ρ, and several others. We generated the decays $a_1^\pm \to \pi^\pm\gamma$ with a branching fraction of 2×10^{-3}, $\rho^\pm \to \pi^\pm\gamma$ with a branching fraction of 5×10^{-4}.

C.6 Non-$B\overline{B}$ decays of the $\Upsilon(4S)$

Only a small fraction of $\Upsilon(4S)$ decays are expected to decay into non-$B\overline{B}$ final states. For the $\Upsilon(4S)$, some of these decays will be cascades to lower $b\overline{b}$ bound states, which may then decay to ggg or $gg\gamma$. The decays to e^+e^-, $\mu^+\mu^-$, and $\tau^+\tau^-$ will be relatively small in comparison, as will the decays of $\Upsilon(4S) \to b\overline{b} \to gg$, $g\gamma$.

Mode	Branching Fraction
$\psi \to \eta(1440)\gamma$	0.30%
$\psi \to \eta'\gamma$	0.43%
$\psi \to f_4\gamma$	0.27%
$\psi \to f_2\gamma$	0.14%
$\psi \to f_2'\gamma$	0.086%
$\psi \to \eta\gamma$	0.083%
$\psi \to f_1'\gamma$	0.065%
$\psi \to f_1\gamma$	0.067%
$\psi \to \gamma e^+ e^-$	0.88%
$\psi \to \gamma\omega\omega$	0.16%
$\psi \to \eta\pi\pi$	0.67%
$\psi \to \gamma\pi^+\pi^-\pi^0\pi^0$	0.83%
$\psi \to \gamma\pi^+\pi^+\pi^-\pi^-$	0.28%

Table C.2: Branching fractions for several two-body and many-body decay modes of the ψ meson [1].

To estimate this background, we generated some MC and found that the ratio of $\Upsilon(4S) \to gg\gamma$ to $\Upsilon(4S) \to ggg$ is about 2.1%.

We checked the literature to see how our results in MC (both the rates of $\Upsilon(4S) \to gg\gamma$, ggg and the photon energy spectrum these decays produced) compared to measured results. A measurement of the direct photon spectrum in $\Upsilon(1S)$ decays found the ratio of the decay rates from $gg\gamma$ and ggg in data to be $(gg\gamma)/(ggg) = 2.75\%$ rather than the 2.1% that our MC predicts [70]. Second, the study showed that the photon energy spectrum in data did not match the spectrum produced by our MC. We scale the MC to take these differences into account [71].

C.7 Noise from Random Crossings

A possible source of background to $b \to s\gamma$ is the overlap of a hadronic event and stale energy clusters in the calorimeter from a few crossings earlier. These random events represent, on average, the background level readings in the detector that are superimposed on the events of our dataset. These backgrounds include random noise in the electronics, cosmic rays, beam-wall and beam-gas interactions, and stale information in the electronics leftover from a previous collision.

The $B\overline{B}$ MC simulates the noisy conditions in the detector by embedding random events in the MC. We searched for hard photons in the "noise" files used in the $B\overline{B}$

	"Noise" clusters passing analysis cut	PR	Not PR
Total # of events in sample	4826	—	—
# of events passing loose cuts	2593	54	2539
Average weight	—	0.064	0.037
Average weight (all events)		0.037	

Table C.3: This table lists the number of events with an embedded 2.45 GeV "noise" shower that passed each of our analysis cuts. PR stands for Pseudoreconstructed.

MC and found that none of the "noise" events pass our shower selection cuts, so we conclude that the clusters caused by noise would probably not make it into our analysis and would contribute no weight. In addition, we compare random trigger data to the "noise" files and find that either the "noise" file is deficient in high energy clusters or the random trigger data has an above-average number of high energy clusters.

To estimate the effect of having a high energy noise cluster in our $B\overline{B}$ MC, we created a 2.45 GeV cluster at a fixed angle in the good barrel and embedded it in generic $B\overline{B}$ MC events and did the rest of the analysis. The number of events that passed each stage of the analysis and the average weight given to each event category are listed in Table C.7.

The "noise" file that has actually been used in the MC suggests a considerably lower noise level than do the random trigger events we have processed. We take the claimed background as half of that listed, with an error equal to the claimed background. The final background is given in Table C.4.

C.8   Sneak-Through Electrons

It is possible that electrons from semileptonic B decay or from any $B\overline{B}$ events are a source of background to $b \rightarrow s\gamma$. Electrons deposit almost all of their energy in the calorimeter, and we usually veto them by requiring that a candidate cluster not be matched to a track in the drift chamber. Occasionally, electrons can cause a shower in the calorimeter and escape the track match veto. If the number of electrons that sneak through in data and MC is different, then this is a potential background that needs to be corrected. These electrons fall into one of two (admittedly broad) categories: 1) The electron radiates, and it is a γ that enters the crystals, and 2) The electron itself enters the crystals.

135

The first category is properly included in the generic $B\overline{B}$ MC *if* the amount of material in the MC is the same as the amount of material in the detector. A comparison of the amount of material in CLEO II MC compared to the actual detector concluded that the accuracy was good to $\pm 10\%$ indicating that we can take the MC as is for our central value [72].

The second category deals with clusters caused by electrons that actually enter the crystals, but fail to find a type-1 or type-2 match. An electron may "sneak through" and be considered a candidate cluster if its track is not found, poorly measured overall, or poorly measured in z. We developed code to identify sneak-through electrons in which the track is well-measured in $r-\phi$ but poorly measured in z, and we make a correction based on differences between the yields of these electrons in On-Off subtracted data and $B\overline{B}$ MC. For tracks that are not found or poorly measured, we use TOF information to derive the remaining electron contamination to our data and to our MC. The corrections are listed in Table C.5.

Energy Range (GeV)	1.5 − 1.8	1.8 − 2.0	2.0 − 2.2
Other B modes	30.58 ± 1.00	13.45 ± 0.72	6.97 ± 0.56
Missed electrons	4.95 ± 2.82	1.79 ± 1.72	2.07 ± 1.34
FSR	14.92 ± 1.40	3.60 ± 0.76	1.65 ± 0.60
Random	0.38 ± 0.38	0.13 ± 0.13	0.20 ± 0.37
K_L^0 in MC	13.96 ± 1.22	3.00 ± 0.60	1.27 ± 0.40
$\overline{n}$ in MC	4.52 ± 0.57	1.52 ± 0.38	0.32 ± 0.13
K_L^0 and $\overline{n}$ "measured"	65.82 ± 25.17	0.00 ± 8.89	0.00 ± 11.36
Total	98.16 ± 25.42	14.45 ± 9.14	9.30 ± 11.48

Energy Range (GeV)	2.0 − 2.7	2.1 − 2.7	2.2 − 2.7
Other B modes	8.95 ± 0.63	4.41 ± 0.43	1.98 ± 0.29
Missed electrons	5.03 ± 1.51	3.81 ± 1.09	2.96 ± 0.70
FSR	1.68 ± 0.60	0.22 ± 0.15	0.03 ± 0.02
Random	0.70 ± 0.70	0.60 ± 0.65	0.50 ± 0.59
K_L^0 in MC	1.33 ± 0.41	0.41 ± 0.26	0.06 ± 0.06
$\overline{n}$ in MC	0.38 ± 0.13	0.14 ± 0.05	0.06 ± 0.04
K_L^0 and $\overline{n}$ "measured"	0.00 ± 14.77	0.00 ± 12.40	0.00 ± 9.44
Total	14.65 ± 14.90	8.49 ± 12.47	5.35 ± 9.49

Energy Range (GeV)	2.7 − 3.0	3.0 − 4.0	4.0 − 5.0
Other B modes	0.14 ± 0.02	0.46 ± 0.02	0.17 ± 0.01
Missed electrons	-0.16 ± 0.17	-0.45 ± 0.45	0.00 ± 0.00
FSR	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Random	0.19 ± 0.19	0.35 ± 0.35	0.96 ± 0.96
K_L^0 in MC	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$\overline{n}$ in MC	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
K_L^0 and $\overline{n}$ "measured"	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Total	0.17 ± 0.25	0.35 ± 0.57	1.13 ± 0.96

Table C.4: B backgrounds other than the corrected B decay Monte Carlo given in Table 5.6. The "other B modes" listed here are further broken down in Table C.5. Note "K_L^0 and $\overline{n}$ in MC" in this table is *subtracted* from the backgrounds, since it is replaced by "K_L^0 and $\overline{n}$ measured."

Energy Range (GeV)	1.5 − 1.8	1.8 − 2.0	2.0 − 2.2
$a_1 \to \pi\gamma$	3.11 ± 0.36	2.31 ± 0.32	1.88 ± 0.27
$\rho \to \pi\gamma$	0.99 ± 0.11	0.76 ± 0.19	0.46 ± 0.16
η' 1	0.55 ± 0.04	0.11 ± 0.02	0.01 ± 0.00
η' 2	0.17 ± 0.02	0.01 ± 0.00	0.00 ± 0.00
radiative ψ (2-body)	5.69 ± 0.39	2.89 ± 0.28	1.81 ± 0.26
a1 D^{**}	0.81 ± 0.04	0.50 ± 0.04	0.11 ± 0.01
non-$\overline{B}^0 \Upsilon(4S)$	0.04 ± 0.01	0.05 ± 0.01	0.05 ± 0.01
radiative ψ (3-body)	3.91 ± 0.42	0.83 ± 0.29	0.03 ± 0.01
$D^{(*)}\omega$	15.31 ± 0.72	5.99 ± 0.46	2.62 ± 0.38
Total	$\mathbf{30.58 \pm 1.00}$	$\mathbf{13.45 \pm 0.72}$	$\mathbf{6.97 \pm 0.56}$

Energy Range (GeV)	2.0 − 2.7	2.1 − 2.7	2.2 − 2.7
$a_1 \to \pi\gamma$	2.45 ± 0.32	1.44 ± 0.26	0.56 ± 0.17
$\rho \to \pi\gamma$	0.72 ± 0.21	0.49 ± 0.19	0.26 ± 0.13
η' 1	0.02 ± 0.01	0.01 ± 0.00	0.00 ± 0.00
η' 2	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
radiative ψ (2-body)	2.47 ± 0.31	1.32 ± 0.23	0.65 ± 0.18
a1 D^{**}	0.11 ± 0.01	0.02 ± 0.01	0.00 ± 0.00
non-$\overline{B}^0 \Upsilon(4S)$	0.22 ± 0.02	0.19 ± 0.01	0.16 ± 0.01
radiative ψ (3-body)	0.03 ± 0.01	0.01 ± 0.01	0.00 ± 0.00
$D^{(*)}\omega$	2.95 ± 0.39	0.92 ± 0.16	0.35 ± 0.08
Total	$\mathbf{8.95 \pm 0.63}$	$\mathbf{4.41 \pm 0.43}$	$\mathbf{1.98 \pm 0.29}$

Energy Range (GeV)	2.0 − 2.7	2.1 − 2.7	2.2 − 2.7
$a_1 \to \pi\gamma$	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$\rho \to \pi\gamma$	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
η' 1	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
η' 2	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
radiative ψ (2-body)	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
a1 D^{**}	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
non-$\overline{B}^0 \Upsilon(4S)$	0.13 ± 0.01	0.46 ± 0.02	0.17 ± 0.01
radiative ψ (3-body)	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$D^{(*)}\omega$	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Total	$\mathbf{0.14 \pm 0.02}$	$\mathbf{0.46 \pm 0.02}$	$\mathbf{0.17 \pm 0.01}$

Table C.5: Yields in various energy bins for specific B modes that we added to the MC. Totals from this table are the "other B modes" in Table C.4.

APPENDIX D

Flattening the Weights

Table 6.3 shows how many photons lie between 2.0 and 2.7 GeV for the eight different K^* modes. The fraction of events with photons between 2.0 and 2.7 GeV drops as the mass increases. The the shape of the X_s decay affects how heavily the event will be weighted. A lighter resonance such as $K^*(890)$ recoiling against a higher energy photon will produce X_s decay products that are collimated in a narrow jet matching the signature for $b \to s\gamma$ very well. A heavier resonance, however, will usually have a lower momentum and broader decays that look less like the "typical" $b \to s\gamma$ events. Heavier K^* events will be weighted less heavily than the lighter K^* events, contributing to the dependence of the efficiency on X_s mass. If our measured values of the weight depend on the X_s mass, then our overall efficiency will depend on the X_s mass distribution as well. A dependence on the X_s mass implies a dependence on the parameters of the model, hence model dependence.

There is a way out of this predicament. If we know how the weights depend on X_s mass, then we can adjust the weight value for each reconstructed event based on its X_s mass. Once we have made this correction, we expect that the average weight for reconstructed events will be independent of X_s mass. Our overall efficiency will be less dependent on the model parameters that determine the X_s mass distribution.

Knowing that the average weight depends on X_s mass in the simple way shown in Figure D.1, we adjust the weight value for each event based on its X_s mass. Our procedure is to fit the dependence in Figure D.1 to a straight line, and then scale the weight value for each reconstructed event based on its measured X_s mass. Specifically, the weight is scaled by

$$\frac{0.15}{0.22 - 0.082 \times (M - 0.8)}.$$

The straight line that fits the dependence of the X_s mass is also shown in Figure D.1 along with its inverse. For events that are not reconstructible, we cannot apply this correction procedure and are stuck with the mass-dependence contributed by those events. Once we have made the correction, we expect that the average weight for reconstructed events will be independent of X_s mass to the extent that we have

good measures of X_s mass. The effect of this correction is shown in Figure D.2. This should then translate into our overall efficiency being less dependent on the model parameters that determine the X_s mass distribution. This is demonstrated in Table 6.3.
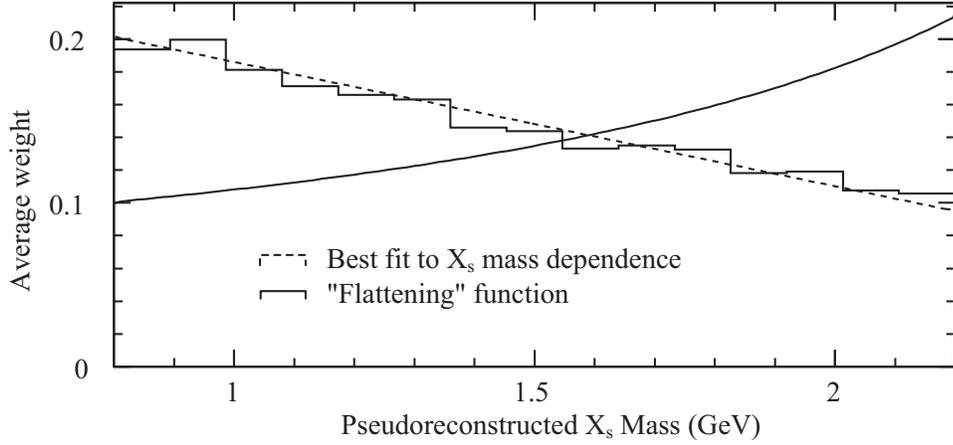
Figure D.1: The average weight (w) as a function of the reconstructed X_s mass for events that were pseudoreconstructed. The average w (the effective efficiency) value decreases as X_s mass increases, introducing a model dependence. The dotted line is a fit to the X_s mass dependence, while the rising line (solid) is the inverse of the dotted line. The dotted line is the factor by which we scale w to "flatten" the weights for a given X_s mass.

We will refer to the adjust weight value as w_{flat}. We can choose any weight we want for each event without bias, but as we stray from the optimum value, we will pay a price in statistical precision. Based on MC, we expect that using w_{flat} will result in an error on $\mathcal{B}(b \rightarrow s\gamma)$ to be 0.21×10^{-4}.
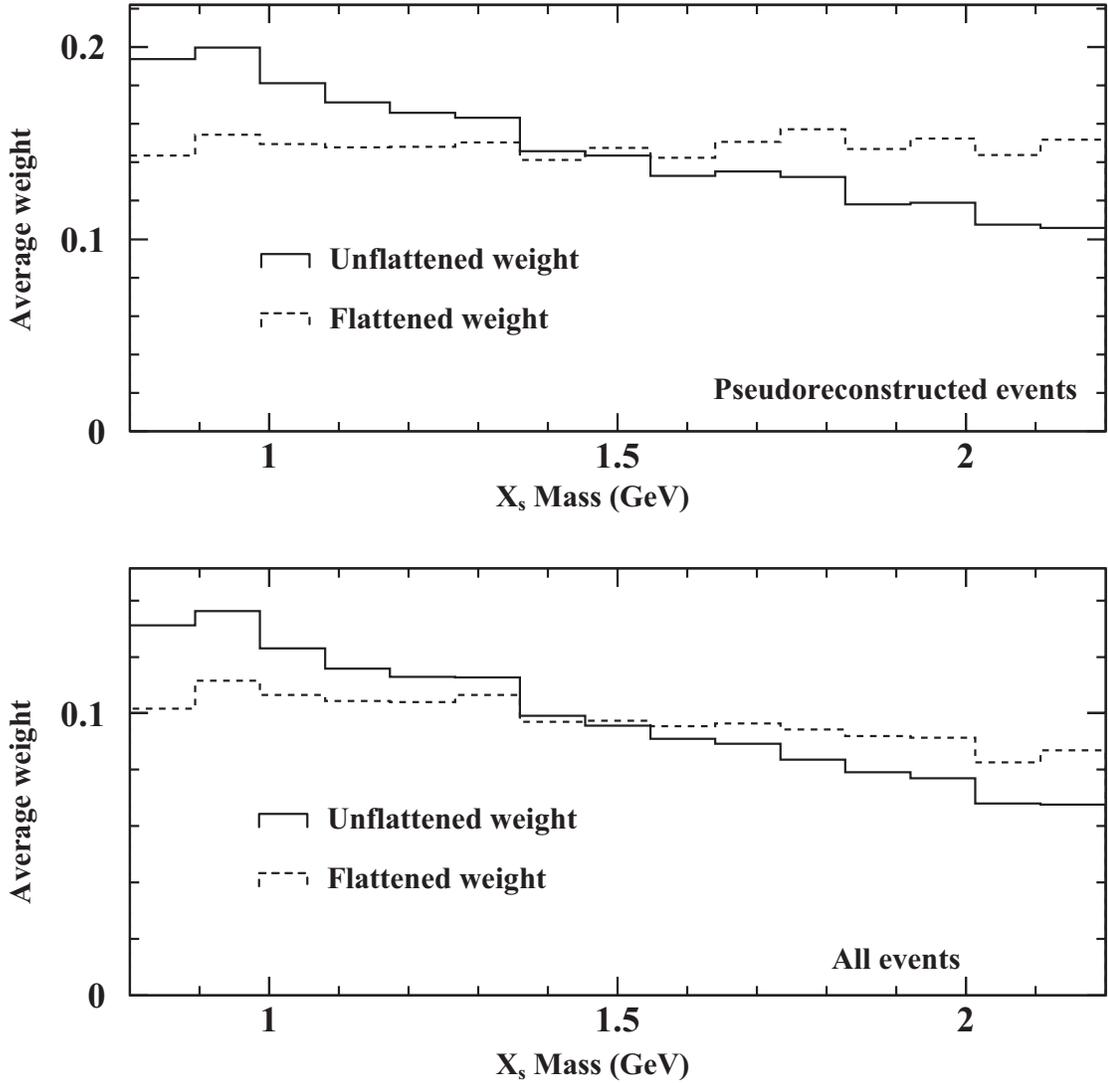
Figure D.2: The upper plot shows the average value of w as a function of generated X_s mass for reconstructed events. Similarly, the lower plot shows $\langle w \rangle$ as a function of generated X_s mass for *all* events. The flattening works very well for the reconstructed events (top plot), but the effect is diluted with the addition of non-reconstructible events in the lower plot. This procedure leaves some mass dependence, but it is significantly reduced.

APPENDIX E

Knob-Turning Studies

Here, we describe the knob-turning studies we did to evaluate the systematic error due to detector modeling. We define the "Nominal" efficiency as the reference point for the knob-turning study. It is the efficiency (defined as sum of flattened weights between $2.0 - 2.7$ GeV and with $|\cos\theta| < 0.7$ divided by total number of $b \rightarrow s\gamma$ events generated) for the MC with no change. The central value for the inclusive efficiency is taken to be $m_b = 4783$ MeV/c^2 and $p_F = 350$ MeV/c. We take the average of the efficiency calculated from the K^* sum sample and the efficiency calculated from the JETSET MC sample to be our nominal efficiency: $\epsilon = 4.08\%$. This value is not the same as the central value that we eventually obtain as our final efficiency, but as long as we determine the percentage difference of each turn of the knob relative to the nominal, we can apply the results to the final efficiency regardless of what that nominal value actually is. The sum total of the systematic error obtained from the knob-turning studies is $\pm 0.057\%$ out of a nominal efficiency of $\epsilon = 4.08\%$. The combined knob-turn error of ± 0.057 out of 4.08 is a relative error of 1.4%.

E.1   CHINEFF - Charged Track Inefficiency

This is a study of the charged particle detection efficiency, determining the effect of not detecting an existing charged particle track. The uncertainty in tracking efficiency is expected to be about 1%. The loss of tracks will affect the event's energy distribution and hence the shape variables. In addition, it will also affect our ability to accurately pseudoreconstruct the event and find leftover leptons. To test the impact of charged track inefficiency, we discard at random 2% of the tracks from the standard set of signal MC and measure the change in efficiency and obtain a $b \rightarrow s\gamma$ efficiency of $\epsilon = 3.95\%$. The difference between this efficiency and the nominal is 0.131%. Since we have doubled the effect of this knob by pitching tracks with a probability of 2% instead of 1%, we cut this number in half for a systematic error of ± 0.033.

E.2   PHINEFF - Photon Inefficiency

This is a study of the photon identification efficiency. About 2% of the time, we do not detect an existing high energy cluster. This will affect not only our ability to find and identify candidate photons, but secondarily, it will change the energy distribution of the event, affecting our shape variables. Since we need to find both the candidate and sibling clusters in order to perform the π^0/η veto, the shower inefficiency will also decrease the efficiency of the π^0/η vetoes increasing the likelihood that a π^0/η daughter will sneak through the veto and fake a candidate photon. To measure the effect of this photon inefficiency on our $b \to s\gamma$ efficiency, showers are thrown out randomly with a 2% probability per shower giving a $b \to s\gamma$ efficiency of $\epsilon = 3.96\%$. The difference between this efficiency and the nominal is 0.12%. We expect our signal efficiency to drop by 2% due to the fact that the random photon we pitch could be our candidate photon. Subtracting 2% of the nominal (0.08%) from the efficiency difference gives an overall change of 0.041. This adds a systematic error of ± 0.021.

E.3   NOTMNG - Tracking Package not used

The purpose of the second generation track quality package, TRKMNG, is to get rid of bad tracks and to associate each particle with one and only one track without vetoing too many good tracks. Since it is not always possible to accomplish these goals with perfect accuracy, we determine whether using TRKMNG biases our $b \to s\gamma$ efficiency in some way by running the analysis without using the TRKMNG track quality package. Without TRKMNG, we obtain a $b \to s\gamma$ efficiency of $\epsilon = 4.23\%$. This differs from the nominal by 0.1509. Since turning TRKMNG off entirely is a drastic change, we take the systematic error to be $\pm 1/5$ of the difference, or ± 0.030.

E.4   NOSPLITF - Splitoff Package not used

The purpose of the splitoff package, SPLITF, is to reject splitoff showers caused by hadrons interacting strongly in the crystals. We need to correctly identify the sources of the splitoff showers in the calorimeter or else we may falsely assume that they are candidate photons. We run the analysis with the SPLITF package turned off and obtain a $b \to s\gamma$ efficiency of $\epsilon = 4.22\%$. This differs from the nominal by 0.140. Not using the splitoff package at all is a significant change and we believe that it overestimates the effect by at least a factor of 5. We take the systematic error to be $\pm 1/5$ of the difference, or ± 0.028.

E.5   MISTRA - Error in Track Momentum Measurement

The track momentum in our QQ generated MC is smeared out by CLEOG in such a way as to match the resolution of the momentum measurement in data. To determine how sensitive we are to the magnitude of this smearing, we run the analysis again and increase the error in track momentum measurement by 10% of itself. We obtain a $b \rightarrow s\gamma$ efficiency of $\epsilon = 4.08\%$. This differs from the nominal by 0.0058. We take the 10% track momentum measurement error to be a conservative estimate of the systematic error, so the systematic error is ± 0.006.

E.6   MISPHO - Error in Photon Momentum Measurement

The photon momentum in our QQ generated MC is smeared out by CLEOG in such a way as to match the resolution of the momentum measurement in data. We test our sensitivity to the magnitude of this smearing by running the analysis again and increasing the error in shower momentum measurement by 10% of itself. We obtain a $b \rightarrow s\gamma$ efficiency of $\epsilon = 4.08\%$. This differs from the nominal by 0.0049. We take the 10% shower momentum measurement error to be a conservative estimate of the systematic error, so the systematic error is ± 0.005.

E.7   RNENRG - Effect of shifts in E_{beam}

The uncertainty in the beam energy can contribute to uncertainty in the efficiency. Poor knowledge of E_{beam} will affect our pseudoreconstruction as well as our neural net analysis. To quantify the effect that this uncertainty has on our efficiency, we have made shifts in the beam energy, generating signal MC with one value of E_{beam} and analyzing it with a shifted value. The results of these shifts is summarized in Table E.1.

The most conspicuous feature of the data is the linear behavior. We expected that the data would form a parabola about 0 MeV, with the efficiency dropping as we moved farther away from the actual value of E_{beam}. Instead, we saw a linear behavior, a change of 0.0018 per 1 MeV, and no evidence for quadratic behavior. We write the efficiency as a function of change in E_{beam}:

$$\epsilon(\Delta E_{beam}) = 4.0825 + 0.0018\Delta E_{beam} \pm 0.0001(\Delta E_{beam})^2.$$

We assume that the energy corrections bring the beam energy to within $\Delta E_{beam} = \pm 1/2$ MeV of the correct value, giving us an error to efficiency of ± 0.0009. We also believe that the mean square width of the beam is $(\Delta E_{beam})^2 = 9$ MeV^2, giving us ± 0.0009. Combining these two gives a total error from beam energy of ± 0.0012 out of $\epsilon = 4.0825$, a $\pm 0.03\%$ error.

Energy Shift	Efficiency, ϵ (%)	$\delta\epsilon$ (%)
+3 MeV	4.0872	+0.0047
+2 MeV	4.0861	+0.0036
+1 MeV	4.0855	+0.0030
0 MeV	4.0825	0.0000
−1 MeV	4.0818	−0.0007
−2 MeV	4.0795	−0.0030
−3 MeV	4.0764	−0.0061

Table E.1: The nominal value for the efficiency is listed in the middle of this table with a shift of 0 MeV. We shifted the value of E_{beam} up and down in 1 MeV increments up to ± 3 MeV. The resulting efficiency and change in efficiency from the nominal is shown.

BIBLIOGRAPHY

[1] Particle Data Group, D. E. Groom *et al.*, *Review of Particle Physics*, Eur. Phys. J. **C15**, 1 (2000).

[2] T. Sjöstrand and M. Bengtsson, *JETSET 7.3*, Comput. Phys. Commun. **433**, 367 (1987).

[3] R. Ayad *et al.*, *Observation of $B \to D^* \pi^+ \pi^- \pi^- \pi^0$*, CLEO Note CBX 00-16, (Internal Report).

[4] M. Artuso *et al.*, *Spin-parity Analysis of an $\omega \pi^-$ Enhancement Seen in $\overline{B} \to D \omega \pi^-$*, CLEO Note CBX 00-31, (Internal Report).

[5] E. Rice *et al.*, *Search for Structure in $\sigma(e + e- \to hadrons)$ Between $\sqrt{s} = 10.34$ GeV and 11.6 GeV*, Phys. Rev. Lett. **48**, 906 (1982).

[6] R. Giles *et al.*, *The Total Cross-Section for Electron-Positron Annihilation into Hadron Final States in the Upsilon Energy Region*, Phys. Rev. D **29**, 1285 (1984).

[7] D. Besson *et al.*, *Observation of New Structure in the $e^+ e^-$ Annihilation Cross-Section Above $B\overline{B}$ Threshold*, Phys. Rev. Lett. **54**, 381 (1985).

[8] D. Griffiths, *Introduction to Elementary Particles* (John Wiley and Sons, Inc., 1987).

[9] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104**, 254 (1956).

[10] C. S. Wu *et al.*, *An experimental test of parity conservation in β-decay*, Phys. Rev. **105**, 1413 (1957).

[11] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[12] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[13] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[14] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (John Wiley and Sons, Inc., 1984), p. 91.

[15] P. Renton, *Electroweak Interactions - An Introduction to the Physics of Quarks and Leptons* (Cambridge University Press, 1990).

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

[17] M. A. Shifman, *Theory of Pre-Asymptotic Effects in Weak Inclusive Decays*, 1994, talk given at Workshop on Continuous Advances in QCD, Minneapolis, MN.

[18] K. Chetyrkin, M. Misiak, and M. Münz, *Weak radiative B-meson decay beyond leading logarithms*, Phys. Lett. B **400**, 206 (1997).

[19] A. Kagan and M. Neubert, *QCD Anatomy of $B \to X_s\gamma$ Decays*, Eur. Phys. J. **C7**, 5 (1999).

[20] A. Falk, M. Luke, and M. Savage, *Nonperturbative Contributions to the Inclusive Rare Decays $B \to X_s\gamma$ and $B \to X_s l^+ l^-$*, Phys. Rev. D **49**, 3367 (1994).

[21] A. Falk, *The Heavy Quark Expansion of QCD*, hep-ph/9610363 (1996).

[22] P. Gambino and M. Misiak, *Quark Mass Effects in $\overline{B} \to X_s\gamma$*, Nucl. Phys. B **611**, 338 (2001).

[23] SPIRES lists over 500 citations to the previous CLEO $b \to s\gamma$ branching fraction paper (ref. [35]), most of them theory papers addressing beyond-SM issues.

[24] http://www.chess.cornell.edu/Facility/Highlights/macro_cryst.htm.

[25] K. M. Reinish *et al.*, *Structure of Reovirus Core Determined via X-ray Crystallography*, Nature **404**, 960 (2000).

[26] S. Desgreniers *et al.*, *Studying Dens Materials by High Pressure X-ray Diffraction*, Phys. Rev. B **59**, 8467 (1999).

[27] E. B. Blum *et al.*, *Performance of the Cornell High Intensity Linac Injector*, CESR Note CBN 83-8, (Internal Report).

[28] M. Sands, *The Physics of Electron Storage Rings*, SLAC Note SLAC-121, (Internal Report).

[29] CLEO Collaboration, Y. Kubota *et al.*, *The CLEO-II Detector*, Nucl. Instrum. Meth. **A320**, 66 (1992).

[30] T. S. Hill, *The CLEO II Silicon Vertex Detector*, Nucl. Instrum. Meth. **A418**, 32 (1998).

[31] F. Sauli, *Principles of Operation of Multiwire Proportional And Drift Chambers*, (1977), 1977.

[32] P. Avery *et al.*, *QQ, A Monte Carlo Generator*, CLEO Note CSN 85-212, (CLEO Internal Report).

[33] R. Brun *et al.*, CERN DD/EE/84-1, 1987.

[34] R. Ammar *et al.*, *Evidence for penguins: First observation of $B \to K^*(892)\gamma$*, Phys. Rev. Lett. **71**, 674 (1993).

[35] CLEO Collaboration, M. S. Alam *et al.*, *First Measurement of the Rate for the Inclusive Radiative Decay $b \to s\gamma$*, Phys. Rev. Lett. **74**, 2885 (1995).

[36] K. Abe *et al.*, *A measurement of the branching fraction for the inclusive $B \to X_s\gamma$ decays with Belle*, Phys. Lett. **B511**, 151 (2001).

[37] R. Barate *et al.*, *A measurement of the inclusive $b \to s\gamma$ branching ratio*, Phys. Lett. **B429**, 169 (1998).

[38] F. Morrow, *Global Event Classification Primer*, CLEO Note CBX 90-5, (Internal Report).

[39] B. Heltsley, *CDCC Track-Shower Matching*, CLEO Note CBX 92-108, (Internal Report).

[40] J. Ernst, Ph.D. thesis, University of Rochester, 1995, (Unpublished).

[41] L. Gibbons, *Exclusive $B \to \pi/\rho/\omega l\nu$ via Neutrino Reconstruction*, CLEO Note CBX 95-6, (Internal Report).

[42] S. Roberts, L. Gibbons, and E. Thorndike, *Trkman the Next Generation: The quest for track quality*, CLEO Note CBX 96-103, (Internal Report).

[43] B. K. Heltsley, *Muon Identification Efficiencies*, CLEO Note CBX 92-117, (Internal Report).

[44] P. D. Wasserman, *Neural Computing, Theory and Practice* (Van Nostrand Reinhold, 1989), p. 14.

[45] F. Morrow, *Shape Analysis of CLEO-II Events*, CLEO Note CSN 91-316, (Internal Report).

[46] R. Wang and R. Poling, *Measurement of the B Meson Semileptonic Branching Fraction*, CLEO Note CBX 93-71A, (Internal Report).

[47] A. Warburton, Private Communication.

[48] F. Wuerthwein, MC generated by F. Wuerthwein.

[49] A. Ali and C. Greub, *Inclusive Photon Energy Spectrum in Rare B Decays*, Phys. Lett. B **259**, 182 (1991).

[50] CLEO Collaboration, J. P. Alexander *et al.*, *Measurement of the Relative Branching Fraction of $\Upsilon(4s)$ to Charged and Neutral B-Meson Pairs*, Phys. Rev. Lett. **86**, 2737 (2001).

[51] B. Valant-Spaight and R. Patterson, *Time of Flight Monte Carlo Improvements for Recompress Data*, CLEO Note CBX 98-9, (Internal Report).

[52] CLEO Collaboration, G. Brandenburg *et al.*, *Charged track multiplicity in B meson decay*, Phys. Rev. D **61**, 72002 (2000).

[53] A. Lyon, J. Ernst, and E. Thorndike, *CP Asymmetry in $b \rightarrow s\gamma$*, CLEO Note CBX 00-47, (Internal Report).

[54] D. Besson, *The $\Upsilon(4s)$ Cross-Section at CLEO*, CLEO Note CBX 92-23, (Internal Report).

[55] D. Besson, *$\Upsilon(4s)$ Cross-Section, reprise*, CLEO Note CBX 92-75, (Internal Report).

[56] T. G. Rizzo, *$b \rightarrow s\gamma$ in the two-Higgs-doublet models*, Phys. Rev. D **38**, 820 (1988).

[57] V. Barger, J. L. Hewett, and R. J. N. Phillips, *New constraints on the charged Higgs sector in the two-Higgs doublet models*, Phys. Rev. D **41**, 3421 (1990).

[58] M. Frank and S. Nie, *$b \rightarrow s\gamma$ in the left-right supersymmetric model*, hep-ph/0202154 (2002).

[59] J. L. Hewett, *B Physics Beyond the Standard Model*, hep-ph/9803370 (1998).

[60] S. Rigolin, *An alternative approach to $b \rightarrow s\gamma$ in the unconstrained MSSM*, hep-ph/0204169 (2002).

[61] W. de Boer *et al.*, *The $b \rightarrow X_s\gamma$ rate and Higgs boson limits in the Constrained Minimal Supersymmetric Model*, hep-ph/0102163 (2001).

[62] L. Everett *et al.*, *Alternative approach to b → sγ in the uMSSM*, hep-ph/0112126 (2001).

[63] C.-S. Huang *et al.*, *The B → X_s ℓ^+ ℓ^- and B → X_s γ decays with the fourth generation*, hep-ph/9911203 (1999).

[64] L. Solmaz, *Constraining Fourth Generation with B → X_s γ*, hep-ph/0204016 (2002).

[65] CLEO Collaboration, D. Cronin-Hennessy *et al.*, *Hadronic Mass Moments in Inclusive Semileptonic B Meson Decays*, Phys. Rev. Lett. **87**, 251808 (2001).

[66] CLEO Collaboration, B. Barish *et al.*, *Measurement of the B Semileptonic Branching Fraction with Lepton Tags*, Phys. Rev. Lett. **76**, 1570 (1996).

[67] S. Anderson, Ph.D. thesis, University of Minnesota, 2002, (Unpublished).

[68] A. Lyon and E. Thorndike, *Study of ToF Performance*, CLEO Note CBX 00-37, (Internal Report).

[69] CLEO Collaboration, T. Browder *et al.*, *Observation of High Momentum η′ Production in B Decay*, Phys. Rev. Lett. **81**, 1789 (1998).

[70] N. Hancock and D. Besson, *Determination of α_s from a measurement of the direct photon spectrum in Υ(1s) decays*, CLEO Note CBX 95-54, (Internal Report).

[71] R. D. Field *et al.*, *The Direct Photon Spectrum in the Inclusive Decays of the Upsilon*, Phys. Lett. B **133**, 248 (1983).

[72] Kevin Sparks and Rusty Malchow did the work. Brian Heltsley did similar work for determining the absolute luminosity.