SLAC-R-706

Studying $b \rightarrow s\,\gamma$ at BaBar Using a Fully Inclusive Method

Ruth Elisabeth Schmitz

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-706

Prepared for the Department of Energy
under contract number DE-AC03-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

UNIVERSITY of CALIFORNIA
SANTA CRUZ

STUDYING $b \to s\gamma$ AT BaBar USING A FULLY INCLUSIVE METHOD

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Ruth Elisabeth Schmitz

June 2003

The Dissertation of Ruth Elisabeth Schmitz is approved:

Professor Bruce Schumm, Chair

Professor Abraham Seiden

Professor Michael Dine

Professor Hartmut Sadrozinski

Frank Talamantes
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Studying $b \to s\gamma$ at BABAR Using a Fully Inclusive Method

by

Ruth Elisabeth Schmitz

The $b \to s\gamma$ decay represents a possible electromagnetic "loop" penguin decay of the B meson. This FCNC process is of high theoretical interest because various scenarios of new physics are expected to have contributions at the same (one loop) level as the Standard Model. The large sample of $B\overline{B}$ meson decays collected by the BaBar experiment makes a precision measurement of this rare decay possible. In conjunction with Standard Model predictions at the 10% level, it brings new physics effects into the realm of detection – or seriously constrains models that could predict them.

A fully inclusive technique is presented to study the $b \to s\gamma$ transition as a function of photon energy, using $88.5 \pm 1.0 \times 10^6$ $B\overline{B}$ meson decays collected by the BaBar experiment in the years 2000 to 2002. The expected statistical and systematic uncertainties have been fully determined which enables first comparisons with theoretical predictions and other experimental results. It also lays the basis for the determination of the inclusive branching fraction $\mathcal{B}(B \to X_s\gamma)$ and the measurement of the photon energy spectrum.

Meinen Eltern Eva und Hans-Georg

Acknowledgements

Five years ago, in Bonn, I was about to embark on an exciting journey from Bonn to Santa Cruz, from a masters to a graduate program and from the known and familiar into the unknown and new. Exciting it has been, and wonderful, and challenging, and enjoyable. It has brought me so many new experiences and friendships, has let me explore life and myself — and taught me a thing or two about physics. Too many people have impacted my life since then to mention them all. They have contributed not only to my successful completion of this work but also to the enjoyable experience the past years have been. I am grateful to them all.

First there has to be my home at UCSC, the physics department and especially SCIPP. Bruce Schumm has seen me through all the critical phases of the program and lately the writing, always understanding and supportive of my ideas — it was him who turned up just the thesis project for me that I had started to be interested in through my classes, and who made it possible for me to see it come this far. Abe Seiden never ceases to amaze me with his knowledge about everything and anything in particle physics. I have learned a lot from him. Terry Schalk always fostered the broader understanding of and excitement about physics, be it through TESLA demonstrations or great physics discussions over sour worms. Hartmut Sadrozinski never held back with advice or his connections to help me out (from convincing me to come to Santa Cruz to helping me find connections in the world of jobs). I'm grateful to have been at a place where the professors take a real interest in teaching and in their

students; Howie Haber and Josh Deutsch stand out in that respect. At SCIPP I have also learned that I am a happier person when I share an office — this must be credited to my office mates: Max Turri, Wolfgang Walkowiak, Christian Flacco, Patrick Spradlin, Michael Wilson, and Tim Beck. I would like to also thank the administrative staff here, in particular Georgia Hamel, Edward Hawking, Lesia Monson, Terry Hart, Judy Hobor — and the late Marilyn Stevens, who made the sailing through the red tape so much smoother and always had a word of encouragement.

I spent a substantial amount of time getting to and from, and working at SLAC, my "department" over the hill. I have benefited immensely from the sense of community there and all the people who created it. I could not have done this thesis without my fully-inclusive group at BABAR. It was a pleasure working with them all. It is the discussions, exchanges of ideas and the collaboration that make particle physics so exciting for me. I would like to mention some in particular. Colin Jessop was almost a second adviser to me. I have profited much from his guidance, his insisting on the larger picture, and all his ideas for how to make things better. Al Eisner's support strengthened my back and his thoroughness in thinking things through forced me to better understand our measurement and the physics behind it. Jim Libby has an ever-friendly way of helping with everything from understanding the theory to writing a CV. And finally, Erkcan is not only the person with whom I work amazingly well (and enjoyably) together but who is also simply a wonderful friend.

Graduate student life here in the US is certainly very different from the

"Doktorandendasein" in Germany. I am glad I didn't miss it! It all started with the people at Graduate Student Housing (Line, Kerstin, Beate, Adrian, Em, Jay, Eric, and also Daniel and Norm). We shared the intensity of that first year and developed coping mechanisms with ("This American Life") dinner parties, after-study reward gatherings and hikes around campus.

I am most lucky to have found the best imaginable fellow grads in Yuko and Luisa. The awesome trio has seen many changes but never ceases to nourish and celebrate its members — whether we waste the nights studying or celebrating together.

It is a memorable fact that I have lived through my five years and four apartments here with the same house-mate: Line. She and later also Patrick not only took care of me with any support needed (food, boggle, 5-for-5-for-5, ...), but moreover grew to be true family. I will miss having the good times of their support, humor and friendship immediately around me. With the Linguistics connection through them, thanks also go to Chris for taking it on himself to proof-read part of this thesis.

Oscar has been cheering me and encouraging me throughout, making me laugh and stop to see all the good things. His support and love kept me going whenever I got discouraged and his personality is truly an inspiration to me. I look forward to exploring life further with him.

I would not be here without the unfailing support from my family, in particular my parents and my brother. They not only stand behind me in any of my decisions but have always encouraged me to follow my dreams, whether those take me

close to or far away from home. It is impossible to describe just in how many ways they have contributed to this success. This work is dedicated to them.

Chapter 1

Introduction

Electroweak "penguin" loop decays of b-quarks were first proposed theoretically in 1977 by the CERN physicist John Ellis [1]. It took 16 years until first experimental evidence for these processes was seen by the CLEO experiment in their observation of the $B \to K^*\gamma$ decay [2], a manifestation of the $b \to s\gamma$ transition, the decay of a b-quark into an s-quark with the radiation of a final state photon. This verification of the existence of flavor changing neutral currents in the Standard Model was much anticipated by the particle physics community.

Theoretically, the $b \to s\gamma$ penguin decay is of high interest because the rate is sensitive to new physics in extensions of the Standard Model. This can manifest itself via the addition of new particles in the loop which may change the inclusive rate for $b \to s\gamma$ [3]. Since there are very large fragmentation uncertainties connected with the extrapolation from the $B \to K^*\gamma$ to the full $b \to s\gamma$ rate, inclusive measurements are

needed. This is relevant because the sensitivity to new physics effects is determined by the precision of the experimental measurement and that of the Standard Model theoretical calculation. The latter has been obtained from Next-to-Leading order calculations and is given as

$$\mathcal{B}(B \to X_s\gamma)_{theory} = (3.57 \pm 0.30) \times 10^{-4}.$$

Experimental measurements of the inclusive rate have been published by CLEO [4, 5], ALEPH [6], BELLE [7], and BABAR (preliminary) [8, 9]. They have been averaged [10] to

$$\mathcal{B}(B \to X_s\gamma)_{worldAvg} = (3.34 \pm 0.38) \times 10^{-4}.$$

Thus, to date the experimental measurements have been unable to seriously challenge the Standard Model calculations, because of their large errors. The largest experimental uncertainty, after limitations due to statistics, arises from restrictions on the accepted energy range of the radiated final state photon. The cuts are made to exclude the background dominated region. However, since an explicit signal model has to be assumed to extrapolate back to the full spectrum, a model-dependence is introduced to the final result, which is quite substantial in most of the literature.

In this thesis a measurement of $b \to s\gamma$ is presented that is designed to use 88.5 million $B\overline{B}$ decays taken with the BABAR detector at the Stanford Linear Accelerator Center in the years from 2000 to 2002. It supersedes the previous fully-inclusive BABAR measurement [9] and is designed to be done in bins of the photon

energy, so that both a branching fraction and a spectral shape can be obtained for the $b \to s\gamma$ process. The data are kept "blind", meaning that the signal region of the photon energy spectrum is not looked at while any analysis considerations are made using simulated events only. It is only "unblinded" once all analysis decisions have been made, all errors determined and the strategy is approved. The blinding methodology is used because it avoids experimental bias. This thesis only includes the blinded results up to the full determination of the experimental uncertainties.

A fully-inclusive technique is used here in which the model dependence is substantially reduced. The idea behind the analysis is then to exploit the large available statistics that the BABAR data set offers, by applying a rigorous selection that reduces the background drastically. The use of a good photon selection with especially tuned π^0 and η vetoes, requirements on the event shape (combined into a Fisher linear discriminant), and the use of lepton-tags (the requirement of a high-momentum lepton in the event) are the main ingredients for separating the signal from the background. Although the resulting overall efficiency is quite low, the BABAR data set is large enough that the overall statistical error is not a concern. A dedicated selection and study of different sources of background provides critical constraints from data on the production rate of π^0s and ηs yielding the necessary systematic precision. The gains in the reduction and understanding of the background and the small model dependence yield a total error on the branching fraction that is significantly smaller than for any other experimental measurement and compatible with the precision of the theoretical

prediction.

This document is organized as follows: First the theoretical foundations are laid in Chapter 2, which gives a brief introduction into relevant parts of the Standard Model and its possible extensions before focusing on the theoretical description of the $b \to s\gamma$ branching fraction and the photon energy spectrum. The chapter ends with a description of the signal model that is prescribed for the simulation of signal events.

With this understanding of the process, Chapter 3 introduces the environment in which the data for this measurement are taken: the accelerator and the BABAR detector. A summary of the detector and event selection and the compilation of all available data sets ends this chapter.

Next, an overview over the analysis strategy as used here is given, introducing the main sources of backgrounds and comparing different possible approaches to inclusive measurements. It also tabulates and discusses the available experimental numbers. This can be found in Chapter 4.

All details of the selection of signal events are described in Chapter 5.

The remaining background from $B\overline{B}$ decays must be subtracted off using Monte Carlo simulation. It is found not to model the data very well and must be tuned to data for a better match. Chapter 6 describes all procedures and considerations that go into this.

Backgrounds from continuum events that survive the analysis selection are eliminated from the sample by subtracting off-resonance data. General comparisons

of subtracted data and Monte Carlo expectations are given in Chapter 7.

The study and evaluation of the systematic uncertainties that affect the measurement are discussed in Chapter 8.

All errors on the new measurement are then fully determined, which is the result of this thesis. The data still have to be unblinded. However, the error alone already allows for an evaluation of the performance and influence of the measurement. This and general conclusions are presented in Chapter 9.

Appendices A to C with additional material conclude the thesis.

Chapter 2

Theoretical Foundations

2.1 The Standard Model

This section portrays the Standard Model (SM), an extremely successful, albeit incomplete theory of particle physics. It is the generally accepted model in the field, but the hope is that it will be superseded within the next decade. It is known to be an effective theory of fundamental interactions at energies up to ~ 1 TeV. In that range it makes powerful predictions that experiments to date have not been able to prove wrong.

The SM describes the current view of how the physical world (matter, molecules, atoms) is composed of fundamental (*i.e.*, indivisible, point-like) building blocks, and how the latter are held together by one or more of three fundamental forces: electromagnetic, weak and strong, which are mediated by the exchange of force-carrier

particles. The SM unifies the electromagnetic and weak forces in one theoretical description, although their effects appear as originating from two different interactions, because their strengths differ by about three orders of magnitude at commonly observed energies. It is still an open question if and at what energy scale the strong force can be included with the electroweak force in one unified description. Gravity is currently not included in the Standard Model (one strong indication that the SM can, at best, be a subsumed by a more comprehensive theory). Luckily, for typical processes in particle physics, gravitational effects are completely negligible.

Here, only elements of the Standard Model structure are described that are of relevance for the subject of this thesis, namely the decay of a b-quark into an s-quark with the radiation of a final state photon, $b \to s\gamma$. A more complete treatment of the Standard Model can be found in many textbooks, *e.g.* [11].

Elementary particles are characterized by several properties called quantum numbers, like the electric charge. They allow for a grouping into different categories and also determine which forces govern the particle's interactions. Classification by spin, *e.g.* , results in fermions with half-integer spin and bosons with integer spin. This also determines their statistical behavior.

All particles with electric charge interact electromagnetically through the coupling to photons. The latter are massless and electrically neutral so they cannot interact amongst themselves.

All particles carry the "charge" of the weak interaction: weak isospin. There-

fore all particles can interact weakly. The weak force is mediated by the massive vector-bosons W^+, W^- and Z^0 and is responsible for, *e.g.* , the beta decay of certain nuclei in nature. It is also a main part of the $b \to s$ decay studied in this thesis.

Fundamental particles carrying the charge of the strong interaction are called quarks, but not all fundamental mass-particles carry it. Those that don't are called leptons. The so called "color" charge occurs in the three possible states red, yellow, and blue.[1] The strong force is therefore also called color force and the theory that describes the interactions it mediates is known as Quantum Chromo-Dynamics (QCD). The eight carriers of the strong force, the gluons, are massless and themselves doubly color-charged (in any of the combinations of colors or of anticolors). This way they couple quarks of different color, but they can also interact with each other.

No colored particles have ever been observed separately, but always in color-neutral bound states, called hadrons. The outward neutralization of the color charge of a composite particle is achieved by combining a quark and with an antiquark of the corresponding (complementary) anti-color, resulting in a composite meson — or by combining three quarks, one of each color, making a baryon. This requirement of color-neutrality of the strong force arises from the peculiar way the gluon exchange force behaves at different distances: at small distances it barely has an effect, leaving the quarks quasi-free; with growing separation, though, the strong coupling constant α_s grows so rapidly that the single quarks experience a *confinement* into the hadron.

[1] This naming of the colors is somewhat arbitrary, but there is a certain analogy to chromatics where there are also three primary states that add together to neutral (white).

Therefore, strong interaction effects are present and may have to be taken into account for any particle observed in the laboratory, a fact that complicates the theoretical calculation of many processes, the decay studied in this thesis being one example.

Ordinary matter is composed of two types of quarks, up (u) and down (d), which are the constituents of protons and neutrons (these in turn make up the cores of all atoms in different combinations), and one type of lepton, the electron, which makes up every atom's outer shell and determines the chemical properties of the elements.

All in all there are six quarks (really 18 under the strong interaction where all quarks can carry one of the three colors) and six leptons presently described by the SM as the fundamental building blocks of matter. The ones not found in regular matter are heavier and not stable under human living conditions. They can only be created in very high-energetic environments, *e.g.* , in collisions in particle accelerator laboratories, where their decay, although extremely rapid, can teach about their properties and interactions — and even about the early stages of the universe and its evolution. Conditions in the lab today are thought to recreate the ones less than 10^{-34} sec after the Big Bang.

In the representation of the electroweak symmetry group, the six left-handed leptons are divided into three doublets (pairs), also called generations.[2] The upper members of the doublets have a weak isospin of $+1/2$ and are electrically neutral,

[2]Why there are exactly three generations is another one of the puzzling questions that the SM does not answer. All experimental indications to date support the formulation of the SM with the existence of three and only three generations.

whereas the lower members carry a weak isospin of -1/2 and an electric charge of -1.

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

The diagram above shows the left-handed components of the leptons (indicated by the "L"), which are doublets under rotation in weak isospin space. The right-handed components are singlets. Therefore, parity (P) transformations between left-handed and right-handed components do not lead to equivalent systems under the action of the weak force; this is referred to as parity-violation.

It was observed that for anti-particles the opposite seemed to be true: right-handed anti-particles are the doublets in the weak isospin space, the left-handed ones are the singlets. Naively one can then expect the combination of charge conjugation (C) and parity operation (P) to yield a final system that is equivalent to the original one (it would be CP symmetric). In particular, this results in the expectation of matter and anti-matter behaving exactly alike. This symmetry was found to be broken for neutral Kaons where differences on the 2% level were discovered. CP violating effects and clues about their origin are also expected in decays of neutral B mesons — the main incentive for building B Factory experiments like BABAR and BELLE. The former of which yielded the data set for the present analysis.

The six (18) quarks are eigenstates of the strong interaction. They are also affected by the weak force and, like the leptons, can be grouped into equivalent weak isospin doublets. Quarks carry fractional electrical charges (with respect to the electron or the proton charge, which is set as ∓ 1), the upper members of the doublets

carry $+2/3$, the lower ones $-1/3$.

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

Quark eigenstates of the weak interaction (as shown in the diagram above) are not simultaneously mass eigenstates.[3] Instead each weak eigenstate can be derived as a superposition of the three mass eigenstates that carry the same electric charge (and vice versa for relating a mass eigenstate to superpositions of the weak ones). Thus the lower quark states in the isospin grouping above are marked with a prime, indicating that they are the particular superpositions of the mass eigenstates d, s, and b that couple to the u, c and t mass eigenstates under weak interactions.[4] This mixing of weak eigenstates, *i.e.*, the coupling between the upper and lower quarks, is expressed in the Cabibbo-Kobayashi-Maskawa [12] matrix, or CKM matrix:

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

With the mixing that is described by the above matrix, quarks can decay under the weak force directly into other quarks, even across generations, as in Figure 2.1. This also occurs in the loop of the $b \to s$ decay which is studied in this thesis.

In the Standard Model, the CKM matrix is unitary. One consequence of this

[3] Mass eigenstates are eigenstates to the combination of the strong and the electromagnetic interaction. They combine the three possible color states of each quark type in one description.

[4] Either the upper or the lower states in the doublets have to be expressed this way, choosing the lower ones is merely convention. It is also purely convention to denote the mass eigenstates *without* and the weak eigenstates *with* a prime.

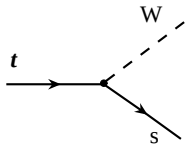

Figure 2.1: Feynman diagram for a direct weak quark decay across generations.

is that, *e.g.* , s- and $\bar{d}$-quarks cannot annihilate into a Z^0, because the contributions from the different weak eigenstates that make up the mass eigenstates will cancel. The same is true for the *direct* decay of b- to s-quarks with Z^0 emission, a fact which is relevant for this thesis, because it requires that the $b \rightarrow s\gamma$ transition studied here proceed through a second order process, which influences the rate (see Section 2.3.1).

With the above ingredients, the Standard Model correctly predicted a wealth of particle and decay properties, some of them even before the discovery of the corresponding phenomena. No measurement to date has been able to seriously challenge this picture. Yet, there are large questions that are left open in the Standard Model: the origin of masses, *e.g.* , and the corresponding breaking of the electroweak symmetry, the unification of all forces in the model, and the inclusion of gravity.

The Standard model does include one proposal for electroweak symmetry breaking: the Higgs mechanism [13]. In this scenario the electroweak symmetry which assumes the same (zero) mass for all (fundamental) particles is broken through interaction with an additional field: the Higgs field. It must be present locally at any time, *i.e.*, everywhere at any time, and it couples to all elementary particles, giving

them mass. It also interacts with the carriers of the weak force. In the theoretical description of that process the carriers of the new interaction are absorbed into the weak force carriers W^+, W^-, and Z^0 themselves, making them massive and leaving only one extra, neutral Boson, the Higgs particle, to be discovered within the framework of the Standard Model. Since the Higgs particle interacts with its own field it acquires mass, and the field is characterized by a non-zero Vacuum Expectation Value (VEV) that minimizes its potential and can be connected to the Higgs mass. Unfortunately the mechanism cannot make numerical predictions for these parameters. It simply offers a mechanism (or idea) of how to generate masses. Proposed extensions of the SM predict up to two Higgs-doublets with both electrically neutral and charged Higgs bosons. Experimentally, none of them have been found yet, which puts strict limitations on their expected masses and the mechanisms that implement them.

In this context theorists and experimentalists are now searching intensely for something — anything — that disagrees with the predictions of the SM. To this end precision measurements of predicted Standard Model parameters are being done in as many different ways as possible, but also searches for particles and effects that are not predicted by the Standard Model are being conducted at present and planned for the future. Tight experimental constraints on these effects yield valuable hints to constructing extensions of the SM, which may offer a more fundamental view and understanding of the compositness of matter. Somewhere in between lies the approach of measuring processes that are predicted as rare in the Standard Model, especially if

extensions of the SM predict otherwise. If the process is still measurable, it can be a very sensitive probe for non-SM effects, although interpretations rely on compatible accuracy in both experimental and predicted results. One example of exactly this approach, the measurement of the transition between a b-quark and an s-quark with radiation of a final state photon, $b \to s\gamma$, is the topic of this thesis. It is studied with a fully inclusive analysis designed to keep the total error on the result as small as possible.

2.2 A Standard Model Extension: SuperSymmetry

Amongst the many different scenarios that have been developed as possible extensions of the Standard Model, SuperSymmetry (SUSY) is expected to show the largest effects in measurements of $b \to s\gamma$. Conversely, experimental $b \to s\gamma$ measurements are thought to have very large constraining power on SUSY assumptions. Therefore a very brief overview shall be given here.

SuperSymmetry extends the group of symmetry transformations that are present in the Standard Model (there they are mostly external symmetries, *i.e.*, space-time related) by an internal symmetry of each particle: the spin. The Standard Model separation and different treatment of fermions and bosons is reversed in SuperSymmetry by the introduction of a super-partner for each fundamental particle: SM fermions (quarks and leptons) each acquire a bosonic partner (squarks and sleptons); SM bosons are assigned fermionic superpartners (their names are the same as those of the SM

bosons, but with the suffix "-ino" attached, *e.g.* , gluon – gluino, W – wino). For the new particle–superpartner pairs there are now no spin distinctions anymore.

SuperSymmetry attractively solves several problems that appear in the Standard Model. One is the *naturalness problem* of the SM Higgs mass being divergent unless highly tuned (up to 22 decimal places), which seems "unnatural". The same divergences also give rise to the *hierarchy problem* of why the electroweak scale (several hundred GeV) and the scale of unification of all forces (called Planck scale, about 10^{19} GeV) are so different. The idea of SuperSymmetry also comes about in other new physics scenarios, for example, it is a natural postulation in string theories that started from the very different point of extra dimensions.

Obviously, SuperSymmetry must be a broken symmetry at energy scales that are typically measured in experiments, because all experimental results so far show no direct evidence for SUSY partners of any of the measured SM particles. There is a multitude of model approaches for how SuperSymmetry can be achieved — and broken. Only the simplest one is discussed here. See Reference [14] for more examples. This simplest SUSY model is often called the Minimal SuperSymmetric Standard Model (MSSM) because it has minimal particle content. This means it has only three generations of spin-1/2 quarks and leptons, like the Standard Model. Yet, it still doubles the number of Standard Model fundamental particles through the introduction of their super-partners. It works with two Higgs doublets, one charged and one neutral. There are then two Higgs fields, one coupling (and giving mass) to

the up-type quarks, the other to the down-type ones. But even in this simplest model there are 105 additional parameters - masses and couplings of all particles. It is a huge task to constrain all these parameters. But this is necessary for arriving at calculable consequences of the model. Experimental results that are SUSY-model independent provide the best constraints. The suppression of flavor changing neutral currents in the measured mixing of the neutral kaon states, K^0 and $\bar{K^0}$ mixing into K^0_L and K^0_S, is one good example. In general, exploiting the mixing of weak eigenstates as described in the CKM matrix yields a lot of information for this purpose. However, experimental constraints are not sufficient and theoretical assumptions must be made to arrive at more tangible models. Specific scenarios implement different ways to mediate the breaking of SuperSymmetry. Within them, hypotheses and assumptions can be made that constrain the multitude of parameters further.

For example, the simplest model uses a gravity-mediated SUSY breaking. It assumes a unification of the coupling constants and gaugino masses at the very high energy of the unification scale (M_{Planck}). Gauginos are the super-partners to the Ws, gluons and neutral gauge particles (Z^0 and γ) of the Standard Model. At the unification scale also the sfermions and Higgs bosons are required to be degenerate (*i.e.*, to have equal mass). Renormalization to lower (accessible) energies is then necessary, but the parameter space is very much constrained. Because it includes the simplest approach to supergravity, this scenario is called mSUGRA (minimal SUperGRAvity). Its parameter space is spanned by only five parameters: $m_{1/2}$, m_0, A_0, $\tan\beta$, and

$sign(\mu)$. Where $m_{1/2}$ is the degenerate mass of the three gauginos (spin $\frac{1}{2}$) and m_0 is the degenerate mass of the sfermions (spin 0), both taken at the unification scale. A_0 is the three scalar coupling of the Higgs to sferminons (analog to the Yukawa coupling) and $\tan\beta$ is the ratio of the vacuum expectation values for the up- and down-type Higgs particles, and μ is related to the mass of the higgsino.

2.3 The Decay $b \to s\gamma$

The remainder of this chapter discusses one specific process with sensitivity to new physics: the decay $b \to s\gamma$. First an introduction to the process is given from the point of view of the Standard Model. Then the theoretical deduction of a branching fraction for this decay is explained in Section 2.3.1. The calculation of a spectral shape follows in Section 2.3.2. Finally, in Section 2.4, the model used to represent this process is introduced.

The decay $b \to s\gamma$ proceeds via a flavor-changing neutral current (FCNC). It is called a radiative electroweak penguin[5] process and is schematically represented by the diagram in Figure 2.2. The quark (initially a b-quark) emits and then re-absorbs a W boson, changing flavor with each interaction while a photon is radiated from any of the charged particles. The intermediate, virtual quark can be either of t, c, or u, but

[5]The term "penguin" decay for this process was coined by CERN theorist John Ellis in 1977, after he had lost a bet obliging him to use the word "penguin" in his next paper [1]. Under this pressure (and, he admits, the influence of illicit substances) it occurred to him that the Feynman diagrams for the FCNC processes he was working on resembled those antarctic animals. See appendix A for a Feynman diagram that expresses this notion better.

the largest contribution numerically comes from the charm quark. The second quark in the B meson can be approximated (at first order) as not participating in the b-quark decay. It is therefore called a spectator quark and is not included in the diagram. Of course, it *does* participate in the fragmentation of the final hadronic system, a fact that will be important later.

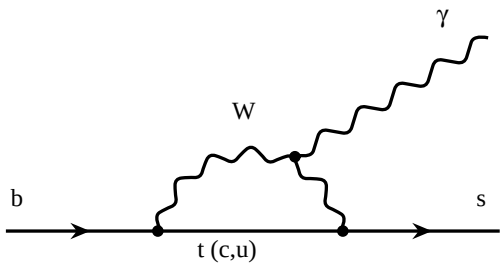

Figure 2.2: One of the leading order Feynman diagrams for the SM $b \to s\gamma$ transition.

This one loop decay is the lowest order the Standard Model allows for the $b \to s$ transition since the unitarity of the CKM matrix forbids this transition to occur through a direct $b \to s$ decay without loops ("tree"-level process). Being a second order process it occurs far less frequently than tree-level processes, yet it is measurable due to the absence of a competing tree-level decay.

These rare B decays are of considerable interest because of their good sensitivity to new physics. New high-mass virtual particles might enter the loop and contribute to the decay amplitude at the same order as the process in the SM. This can happen either constructively or destructively compared to the Standard Model

contribution, thus either raising or lowering the rate expected there [3]. In supersymmetric models, for example, both particles in the loop can be replaced with their SUSY partners, a wino and a squark, and/or a new fundamental boson, such as a charged Higgs and a (s)quark (see Fig. 2.3). Consequently, experimental results on $b \to s\gamma$ can either discover new physics, or provide tight constraints on extensions of the Standard Model [15, 16].

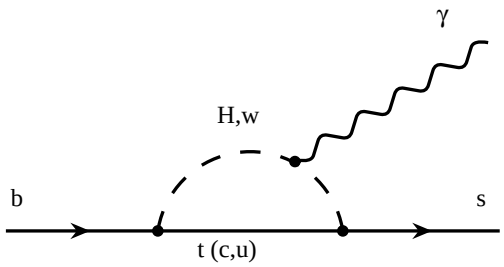

Figure 2.3: Feynman diagram for additional processes contributing to the $b \to s\gamma$ transition in extensions of the Standard Model.

Furthermore, while the photon energy spectrum cannot constrain new physics models, it does give information about the mass, m_b, and the momentum, p_F, of the b-quark inside the B meson. Both of these parameters can be used in the extraction of $|V_{ub}|$ from measurements of charmless semi-leptonic B decays, allowing for an independent, additional approach for the $|V_{ub}|$ measurement.

The CLEO experiment observed the first penguin $b \to s\gamma$ process in 1993: the exclusive decay $B \to K^*\gamma$ [2]. It has since been measured by the BELLE and BABAR experiments also, most precisely in [18]. Unfortunately, due to high QCD

fragmentation uncertainties, it is impossible to deduce the inclusive $b \to s\gamma$ rate from this exclusive analysis. Instead, the measurement has to include (as much as possible) the full array of possible hadronic final states. One approach to that end, a fully inclusive method, is presented in this thesis.

A word of caution is necessary regarding the final quark in the $b \to s\gamma$ transition. Similar to all up-type quarks (t, c, and u) contributing in the loop, all down-type quarks lighter than the b can occur as the final quark: although the main interest is in a final s-quark, a d-quark is also possible. The decay leading to it is fully equivalent to the $b \to s$ transition, and with the fully inclusive experimental technique presented here the two cannot be distinguished. (Therefore, really the sum of $b \to s\gamma$ and $b \to d\gamma$ is measured!) But since the coupling between the final d- and intermediate c- (or t-)quark is a factor of $(|V_{cd}|/|V_{cs}|)^2 \approx 5\%$ (or $(|V_{td}|/|V_{ts}|)^2 \approx 4\%$) [19] smaller than that between c (or t) and s, the final contribution is expected to be smaller by that fraction. To extract the pure $b \to s\gamma$ contribution, 5% have to be subtracted from the total measured rate.

2.3.1 The Branching Fraction for $b \to s\gamma$

Standard Model calculations of the branching fraction for $B \to X_s\gamma$ work in a framework of heavy-quark effective theory (HQET) and operator product expansions (OPE). They rely on parton-hadron duality to predict the inclusive decay width of $B \to X_s\gamma$ from the parton level decay width of $b \to s\gamma$. This relies on the fact that

the energy released by the decay of the b-quark into the final state is large compared to the QCD scale, Λ_{QCD}. Then the final hadronic states can be mostly non-resonant (with a smooth hadronic mass spectrum) and need not consist dominantly of sharp local resonances. In this case the QCD short-distance decay of the b-quark can be factored from the subsequent long-distance fragmentation and hadronization of the resulting s-quark system. This factorization is known as local parton-hadron duality. It is expected to be accurate to better than 10% [20].

The parton level calculation can be performed perturbatively

$$\Gamma[b \to X_s^{parton}\gamma] = \Gamma[b \to s\gamma] + \Gamma[b \to s\gamma g] + \ldots$$

with the contribution of higher-order diagrams being the dominant uncertainty. The electroweak decay (*i.e.*, the interaction with the charm quark and the W boson) is fully contained in the leading order calculation (apart from electroweak corrections, which are small). For the following discussion of QCD and QED effects in the OPE it will be useful to draw a modified Feynman diagram as in figure 2.4 to visualize the processes.

Currently the full next-to-leading order (NLO) terms have been calculated [16, 21] and work on the next order is underway (see below). The framework for this calculation is based on Operator Product Expansions (OPE) of an effective Hamiltonian [22]:

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{ts}^* V_{tb} \sum_{i=1}^{8} C_i(\mu) O_i + \ldots$$

where G_F is the Fermi constant, V_{ij} are elements of the CKM matrix, O_i are the relevant operators, and $C_i(\mu)$ are the Wilson coefficients. The OPE works analogously to Fermi theory: particles with masses greater than m_b (*i.e.*, top and W) are represented by the effective Hamiltonian and a point interaction. The operators O_i represent the complete set of possible space-time structures of the effective currents (behaviors) that arise when the vertex is collapsed as shown in Figure 2.4. All the higher-order, short-

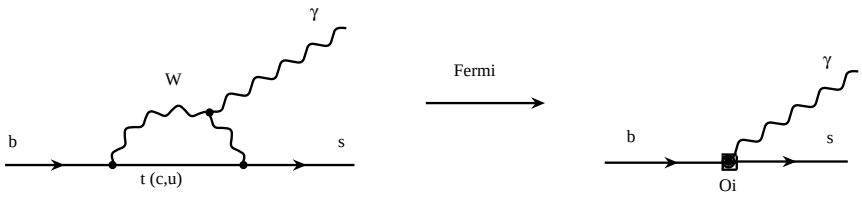

Figure 2.4: Visualization of the effects of the operator product expansion for the leading order diagram for the $b \to s\gamma$ decay. The weak decay loop is contracted to a point-like vertex (analogous to Fermi theory). This way the effects of the Operators can be demonstrated more easily. Here the radiation of the final state photon through the operator O_7 is shown.

distance QCD and electroweak effects within the vertex (*e.g.* hard gluon exchange between the quark lines in Fig 2.2) are parameterized by the Wilson coefficients. The latter effects enhance the $B \to X_s\gamma$ rate by about a factor of three [23]. Use of appropriate Renormalization Group Equations yields a resummation of large logarithms of M_W^2/m_b^2 that are present in the $b \to s\gamma$ amplitude. The operators O_i include part of the longer-distance physics; the rest must be evaluated as non-perturbative corrections outside of the spectator model. The dominant contribution comes from the electro-

magnetic operator O_7 which contains $F_{\mu\nu}$ and couples the photon to the quark lines (see the right side of Figure 2.4). Extra gluons are produced by the chromomagnetic operator O_8 which contains $A^a_{\mu\nu}$. Operators O_1 – O_6 are four-quark operators; they produce diquark loops, the most important of which (charm loops) will be discussed below.

With electroweak corrections under control (cited as less than 3% in [21]), one can then deduce the actual branching fraction BF$[B \to X_s\gamma\ ]$. This is done by relating the width and branching fraction for $B \to X_s\gamma$ to those for $B \to X_c e\bar{\nu}$ and to the width for $B \to X_u e\bar{\nu}$:

$$
\begin{aligned}
& BF[B \to X_s\gamma] \\
&= BF[B \to X_c e\bar{\nu}]_{exp} \left(\frac{\Gamma[B \to X_u e\bar{\nu}]}{\Gamma[B \to X_c e\bar{\nu}]}\right)_{th} \left(\frac{\Gamma[B \to X_s\gamma]}{\Gamma[B \to X_u e\bar{\nu}]}\right)_{th} \\
&\simeq BF[B \to X_c e\bar{\nu}]_{exp} \left(\frac{\Gamma[b \to X_u^{parton} e\bar{\nu}]}{\Gamma[b \to X_c^{parton} e\bar{\nu}]}\right)_{NNLO} \left(\frac{\Gamma[b \to X_s^{parton}\gamma]}{\Gamma[b \to X_u^{parton} e\bar{\nu}]}\right)_{NLO} \\
&\quad + \begin{pmatrix} \text{non} - \text{perturb.} \\ \text{corrections .} \end{pmatrix}
\end{aligned}
$$

This procedure avoids uncertainties from the CKM matrix elements and an otherwise strong dependence on the b-quark mass, as well as poor convergences (see [21]). The X_u/X_c ratio of the widths is theoretically known up to next-to-next-to-leading order (NNLO) and the NLO ratio for X_s/X_u can be calculated perturbatively. The non-perturbative corrections turn out to be quite small ($\mathcal{O}(\Lambda^2_{QCD}/m_b^2)$). They are calculated, in analogy to the treatment of semi-leptonic B decays, by applying operator product expansions in a framework of HQET. This introduces a dependence of the

final rate on the standard HQET parameters m_b, λ_1, and λ_2 [24], with $\lambda_{1,2} \approx \Lambda^2_{QCD}$. Conceptually, λ_1 is an expression of the kinetic energy of the b-quark inside the B meson (it is thus related to the b's momentum inside the B) and λ_2 parametrizes the mass difference between the B and B^* mesons.

The most precise theoretical treatment to date includes the full next-to-leading order (NLO) QCD terms including up to three-loop diagrams, all relevant electroweak corrections, and an extensive study of non-perturbative effects. Much effort [25, 26, 16] went into all these areas recently. It is summarized in [21].

Gambino and Misiak [16] brought all these pieces together in 2001 for the most precise calculation to date, yielding $BF[B \to X_s\gamma]_{SM} = 3.73 \times 10^{-4}$ for the full photon energy range, E_γ. Although, in the light of the charm loop uncertainties discussed above, the more important number they find is:

$$BF[B \to X_s\gamma\,(E_\gamma > 1.6\ \text{GeV})]_{SM} = (3.57 \pm 0.30) \times 10^{-4}.$$

This relatively precise expectation (less than 10% error) allows for a meaningful comparison between the Standard Model and new experimental results at the expected total significance of about 10%. It can yield tighter constraints on extensions of the Standard Model, since the latter use the same methods as outlined here (but of course with an effective Hamiltonian that includes different Wilson coefficients). Some examples for the constraints that the $b \to s\gamma$ measurements can pose on new physics will be given at the end of this section.

The above Standard Model expectation agrees with the current experimen-

tal world average of: $BF[B \to X_s\gamma\ (E_\gamma > 1.6\ \text{GeV})]_{exp} = (3.22 \pm 0.37) \times 10^{-4}$, with the photon energy range of extrapolation in the experiments corrected to the range motivated above. For details about the experimental results that enter this number, see Section 4.4.

The next step, a NNLO calculation, is presently underway. It is expected to fix the renormalization scheme for the charm loop contributions discussed below. However, non-perturbative effects, especially those connected to the same charm loops, also need to be systematically evaluated or brought under control — a difficult task, since a new method needs to be developed.

The main sources of uncertainty in both the perturbative treatment and the non-perturbative corrections that are included above arise from two-loop charm quark diagrams with soft virtual gluons (see Fig 2.5). Factorization of short and long distance contributions is not sufficient here, because the relatively large c-quark mass results in $2m_c/m_b$ being not small.

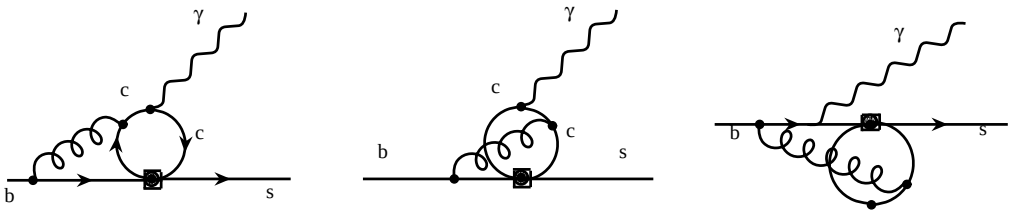

Figure 2.5: Charm quark contributions to the matrix elements of four-quark operators.

The exact treatment of the non-local parts of these diagrams[6] can change the predicted $b \to s\gamma$ rate by a significant fraction, depending on the assumed mass of the charm quark. Since the the c-quark in the loop is never on-shell, it is wrong to assume the pole-mass for the c-quark (as is done for real quarks). Instead, similar to "running" (*i.e.*, energy-dependent) coupling constants, the charm mass in the loops has to be treated as a running constant. To determine its exact value, the charm mass would have to be evaluated at a scale where c- and b-quarks are renormalized. To define this renormalization scheme, a calculation at the next order of precision (3 loop level, NNLO) would need to be performed. This is presently not available, leaving large renormalization scheme ambiguities. The latest calculations use the $\bar{M}S$ scheme and $m_c = (0.22 \pm 0.04)m_b$, which leads to an error of 6% for charm mass effects. The selected mass is, as expected, well below the upper mass limit of charm pair-production in a b-quark decay: $m_c = \frac{1}{2}m_b$. As an illustration of the importance of this point, consider that varying m_c between the (extreme) points of $m_c = 0.22m_b$ and $m_c = \frac{1}{2}m_b$ changes the predicted $b \to s\gamma$ rate by as much as 35%. A more physical variation between the m_c pole mass (as used in earlier calculations [27]) and $m_c = 0.22m_b$ results in a rate enhancement of 11%.

At low photon energies, intermediate real $c\bar{c}$ states such as ψ or ψ' significantly enhance the rate (experimentally they are usually treated as background and subtracted off). One expects the largest contributions from virtual $c\bar{c}$ pairs in this

[6]Buras and Misiak define as non-local those parts that cannot be removed off-shell by finite local counterterms.

region, too. They enter both the experimental measurements and the perturbative and non-perturbative theoretical calculations. This has one immediate consequence for experimental work. In measurements of the $b \to s\gamma$ rate, a cut is made on a minimum accepted photon energy, mostly to avoid the overwhelming background from $b \to c$ transitions in this region (see Chapter 6). A secondary effect of a cut is that it removes the region where all charm loop effects result in the largest uncertainties. Extrapolating back to the full E_γ range through the use of theoretical predictions of the cut-out part can then re-introduce those effects. Clearly this is not a desirable option. In this situation there are two ways to proceed: quoting an experimental result for the region above the cut-off only (this would make it hard to compare results from different experiments with potentially different cut-offs) or extrapolating to only a part of the total E_γ range. Using only the range above $E_\gamma = 2.0$ GeV allows in only about 1% uncertainty from the perturbative treatment of charm pairs, but significantly increases the dependence on the signal model (see Sec. 2.4), whereas extrapolation down to $E_\gamma = 1.6$ GeV, as suggested in [16][7], yields an uncertainty of about 1.7% from perturbative charm loop calculations (only slightly more) while including more than 95% of the total signal and thus reducing the related model error. Quoting experimental results extrapolated down to 1.6 GeV as well as the numbers within the measured region (only above the applied cut) seems advisable and both

[7]The value of 1.6 GeV is arbitrarily chosen, it might as well be 1.5 or 1.7 GeV, but it is important that it stays below about 1.8 GeV to avoid model dependence and above about 1.4 GeV to avoid the charm uncertainties.

approaches can be considered for the analysis presented here.

Figures 2.6 and 2.7 demonstrate the sensitivity of the $b \to s\gamma$ rate to new physics: they highlight constraints on new physics models that are imposed by the experimental average of the rate and its calculation in extended SM scenarios. The lower bound on the charged Higgs mass from the $B \to X_s\gamma$ measurement, as seen in Figure 2.6, uses the experimental world average from early 2001, which does not include the 2001 CLEO measurement. (The experimental results and the techniques that were used to obtain them are described in more detail in Section 4.4.) The absolute lower limit, at 99% confidence level, is $M_H > 350$ GeV from the $B \to X_s\gamma$ measurement, as indicated in the plot. The great constraining power of the $B \to X_s\gamma$ measurement as compared to direct searches is obvious in this scenario.

The bound for the excluded light blue region in Figure 2.7 comes from the experimentally measured branching ratio for $B \to X_s\gamma$. The calculation uses the average of the numbers measured in 2001 by CLEO [5] and BELLE [7]. Here again, there are strong constraints from the $B \to X_s\gamma$ measurements, as can be seen from the significant part of the open parameter space that is excluded.

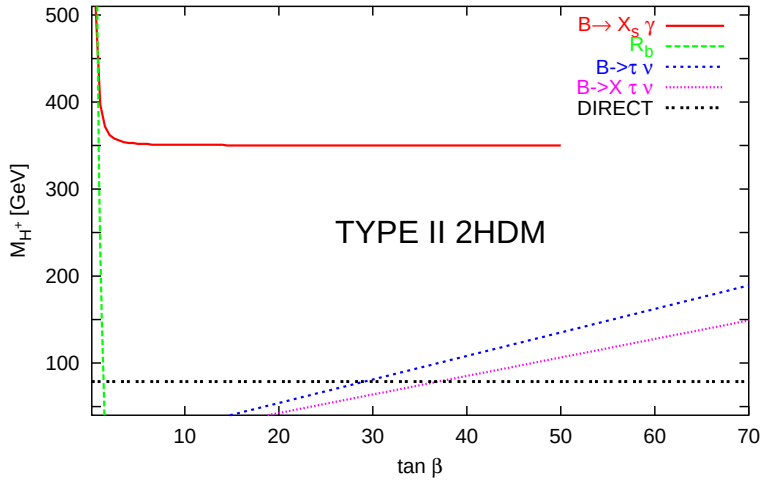

Figure 2.6: Direct and indirect bounds on the minimum mass of the charged Higgs boson as a function of $\tan\beta$ in a two Higgs-doublet model scenario (*e.g.* the MSSM with no specified SUSY breaking mechanism). M_{H^+} values below the curves are excluded. The plot and calculation of the $B \to X_s\gamma$ bound is taken from [16].

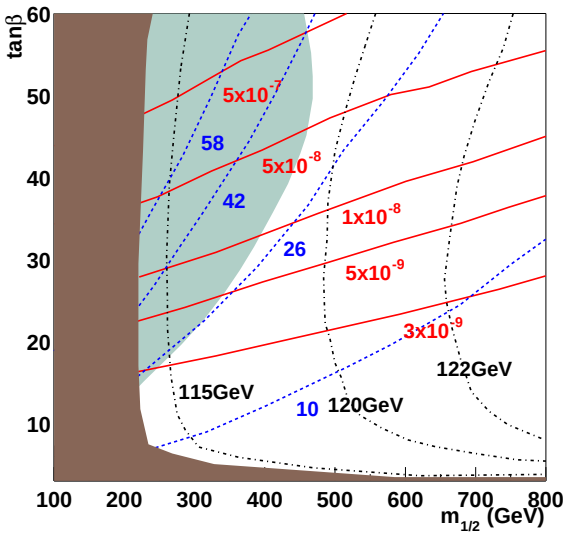

Figure 2.7: Parameter space of two mSUGRA elements ($m_{1/2}$, $\tan\beta$) for $m_0 = 300$ GeV and $A_0 = 0$. The light blue region is excluded by the measured $B \to X_s\gamma$ branching ratio and the brown (dark) region is excluded by the mass bounds from direct searches for the Higgs and SUSY particles. The colored curves are contour plots for the SUSY anomalous magnetic moment of the muon a_μ^{SUSY} (blue curves, $\times 10^{-10}$), for the branching ratio for $B_s \to \mu^+\mu^-$ (red curves), and for different values of the Higgs mass (black curves). The plot is taken from [28].

2.3.2 The Photon Energy Spectrum

If the strong force were to be "switched off", so that quarks did not fragment or were confined, then in the b-quark rest frame the two monochromatic decay products, s and γ, would be observed, each with momenta of about $m_b/2$. In reality the strong force causes the b-quark to be confined into a B meson and the s-quark to fragment into a complicated spectrum of X_s states. The interaction of the b-quark with the spectator quark in the meson causes a scale dependence in m_b, and confinement leads to the b-quark having an average momentum inside the B meson. This results in a smearing of the E_γ spectrum such that $< E_\gamma > \approx \ m_b/2$ and the width (second moment) is related to the momentum, p_F, of the b-quark.

The shape of the energy spectrum expected from the high-energy photon in $b \to s\gamma$ decays is not sensitive to new physics [26], yet it yields important information for the study of this radiative penguin process. First of all, experimentally the spectrum provides an alternative method of determining the total $b \to s\gamma$ decay rate through a fit to the spectrum, even with a minimum cut imposed on the photon energy. This aids in reaching a model-independent interpretation of the measured rate. A theoretical understanding and prediction of the spectral shape is a necessity for the experimental extrapolation of the rate (measured in a restricted energy range) to a broader phase space. Secondly, since a theoretical approach to the photon energy spectrum must include the effects of fragmentation in the hadronic system (a circumstance that was treated only as a correction for the calculation of the rate) the predicted spec-

trum will depend on m_b and λ_1 as parameters. They can then be extracted from the study of the experimentally measured spectrum. The theoretical method works in a framework of HQET. The first moment of the photon spectrum, $\langle 1 - 2E_\gamma/m_B \rangle$, is expressed in terms of the HQET parameter $\bar{\Lambda} \approx m_B - m_b$, where m_B is the mass of the B meson. It allows the extraction of the b-quark pole mass, m_b, *i.e.*, the mass of the b-quark bound into the B meson (including all one-loop effects). In addition, it is expected that an accurate measurement of the first moment can determine the size of the non-perturbative contributions to the $\Upsilon(1S)$ mass (the lowest lying $b\bar{b}$ bound state), for corrections which cannot be absorbed into the b-quark pole mass [29]. The second moment, $\langle (1-2E_\gamma/m_B)^2 \rangle$, determines the parameter λ_1 (sometimes also called $-\mu_\pi^2$) which in turn determines the motion of the b-quark in the B meson.

Theorists first tried to use information from semi-leptonic $b \to c$ decays to evaluate these HQET parameters and predict the photon spectrum for $b \to s\gamma$ [30]. This treatment was found not to be theoretically justified, because, as mentioned earlier, the charm mass is large enough to make $2m_c/m_b$ not a small number. In this case, factorization is not sufficient. On the other hand, the transition $b \to s$ is certainly heavy-to-light and can be factorized. (The same is true for $b \to u$, which makes the use of information from the $b \to s\gamma$ spectrum valid in the extraction of $|V_{ub}|$ from $B \to X_u l\nu$ decays.) Instead, the current approach is based on a description of the b-quark motion inside the B meson which determines the shape of the photon spectrum to a large degree. HQET incorporates the b-quark momentum (expressed

in the parameter λ_1) by resummation of an infinite set of higher-order corrections into a shape function. The latter is very similar to the parton distribution (structure) functions in deep inelastic scattering.

The b-quark motion effects have been calculated in a NLO analysis in [26]. The resulting shape function is a universal characteristic of the inclusive decay spectra of the B meson, for decays into light-mass final state partons. The spectrum is obtained as a convolution of this shape function with the parton-level spectrum. The results found by Kagan and Neubert [26] are used in the signal model for this analysis, as described in Section 2.4. The shape calculations are far less accurate than the corresponding ones for the rate, mostly because fragmentation issues now play a large role and a phenomenological description is necessary. In addition, the current approach is only valid away from the high-energy endpoint of the spectrum, where quark-hadron duality breaks down due to the presence of the exclusive $K^*(892)$ (see also Section 2.4). High uncertainties and corrections to the spectrum are then expected at best, they are avoided by replacing that region in the shape function with the Breit-Wigner shape representing the $K^*(892)$.

To relate the b-quark pole mass and λ_1 (which expresses the b momentum) to the measured spectrum, both the theoretically predicted spectrum and its integral (the (partial) $b \to s\gamma$ rate) are expanded perturbatively in α_s, the strong interaction coupling constant that enters the equation through the Wilson coefficients. The expansion is known to first order and the second order has been estimated in [29, 31].

Using the theoretical description of both the α_s-expanded spectrum and its integral, the following expression is obtained for the determination of $\bar{\Lambda}$ from the experimentally measured average energy [29]:

$$\begin{aligned}\langle 1 - 2E_\gamma/m_B\rangle_{exp} = \\ & B_0(\delta) + B_1(\delta)\frac{\alpha_s}{\pi} + B_2(\delta)\beta_0(\frac{\alpha_s}{\pi})^2 \\ - \quad & \frac{m_B - m_b}{m_B}\cdot\left\{C_0(\delta) + C_1(\delta)\frac{\alpha_s}{\pi} + C_2(\delta)\beta_0(\frac{\alpha_s}{\pi})^2\right\}\end{aligned}$$

with $\delta = 1 - E_\gamma^{min}/m_B$. B_i and C_i are combinations of the coefficients used in the parametrization of the spectrum and its integral. They are calculated and tabulated [32]. $\beta_0 = 11 - 2\,n_f/3$, where n_f is the number of flavors with higher mass than (in this case) the b-quark (one).

For the relationship between the measured variance (second moment) of the photon spectrum and λ_1 $(-\mu_\pi^2)$ a very similar expression is calculated in [29].

2.4 The Signal Model for $b \to s\gamma$

The full X_s system in $b \to s\gamma$ decays consists both of a number of K^* resonances, given in Table 2.1, and of non-resonant modes, including mostly kaons and pions with appropriate net strangeness. A signal model is required to describe this system. It must include a well-defined theoretical description of the spectral shape, so that an extrapolation into the experimentally unmeasured part of the spectrum is possible.

In the rest frame of the B ($\overline{B}$) decaying to $X_s\gamma$, the photon energy (E_γ) spectrum is equivalent to the hadronic mass (m_{X_s}) spectrum through the kinematic relation:

$$E_\gamma = \frac{m_B^2 - m_{X_s}^2}{2m_B}$$

For the present analysis, the spectral shape predicted by Kagan and Neubert (KN) [26] is used. It predicts a smooth hadronic mass spectrum. This approach works, even in the presence of $K^*\gamma$ resonances, because the higher resonances are closely spaced and wide enough that they overlap (see Table 2.1).

State X_s	M_{X_s} [GeV]	Γ_{X_s} [MeV]	E_γ [GeV]	Fraction [%]
$K\,(n\pi)$	≥ 0.629	continuum	≤ 2.60	62.6 ± 13.6
$K^*(892)$	0.894	50	2.56 ± 0.01	16.8 ± 6.4
$K_1(1270)$	1.273	90	2.49 ± 0.02	4.3 ± 1.6
$K_1(1400)$	1.402	174	2.45 ± 0.05	2.1 ± 0.9
$K^*(1410)$	1.412	227	2.45 ± 0.06	4.1 ± 0.6
$K_2^*(1430)$	1.428	103	2.45 ± 0.03	6.2 ± 2.9
$K_2(1580)$	1.580	110	2.40 ± 0.03	1.7 ± 0.4
$K_1(1650)$	1.650	150	2.38 ± 0.05	1.7 ± 0.6
$K^*(1680)$	1.714	323	2.36 ± 0.10	0.5 ± 0.2
$K_2(1770)$	1.773	186	2.34 ± 0.06	——

Table 2.1: The accessible resonances for the X_s system in the $B \to X_s\gamma$ decay. The mass and width are given, and the corresponding E_γ value. The last column contains Veseli and Olsson's [33] predictions for the fraction of the total branching ratio that each state contributes.

The table also shows the theoretical predictions by Veseli and Olsson [33] for the amount that each resonance contributes to the total spectrum. Large errors have to be assumed on these fractions, due to the large hadronization uncertainties

in prediction exclusive final states. Nonetheless, it can clearly be deduced from this table that the non-resonant contributions account for the majority of the spectrum. A smooth, *i.e.*, non-resonant parametrization in m_b and λ_1 is then a good description. In Kagan and Neubert's ansatz these two variables are used as parameters. In addition, their model makes use of the available experimentally measured shape of the E_γ spectrum to add additional constraints. This is achieved through the use of a spectral shape function. Various functional forms were investigated and compared to the experimental data points and an exponential shape was found to match best. The dependence on the parameter m_b is found to be much stronger than on λ_1. Therefore different values of m_b should be investigated and the ratio of $\lambda_1/\bar{\Lambda}$ should be kept constant, effectively determining λ_1 for any value of m_b chosen. Appendix B gives more information about this.

The lower plots in Figure 2.8 show how the E_γ spectrum varies for different choices of these parameters in the Kagan and Neubert model. This will give rise to an additional "theoretical" error on the experimental result in the extrapolation to the full range.

The left plot in Figure 2.9 shows the corresponding (and equivalent) m_{X_s} spectrum for this model. Obviously, it does not include the expected sharp peak of the lowest-lying K^* resonance, the $K^*(892)$, in the lowest m_{X_s} region. The model ansatz fails here, because the narrow $K^*(892)$ stands by itself, separated from the next higher resonance, $K_1(1270)$, and the rest of the resonant spectrum.

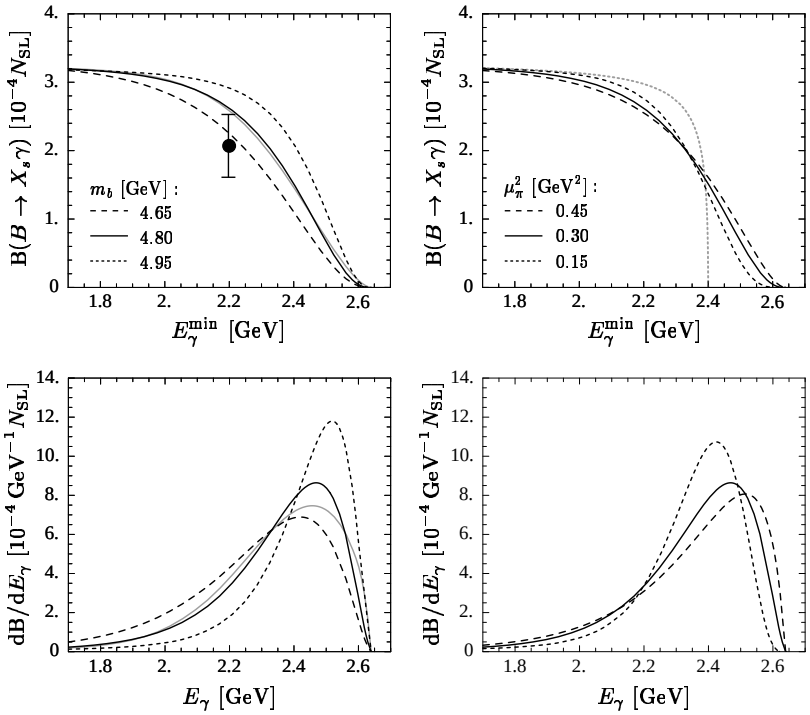

Figure 2.8: The model dependence of the E_γ spectrum in the B meson rest frame, taken from [26]. The spectra, $d\mathcal{B}/dE_\gamma$ (bottom), are shown for different choices of b-quark mass (left plots) and Fermi momentum (right plots). Also shown are the integrals of the spectra ($\mathcal{B}(B \to X_s\gamma)$) as a function of the lower bound of integration, E_γ^{min}. The data point represents the first CLEO measurement [4].

Therefore, KN argue that the smooth m_{X_s} spectrum should be modified in the $K^*(892)$ region, *i.e.*, below some cut-off m_{cutoff}:

- The m_{X_s} spectrum for $m_{X_s} > m_{\mathrm{cutoff}}$ is given by the spectral function ansatz from reference [26].

- The m_{X_s} spectrum for $m_{X_s} < m_{\mathrm{cutoff}}$ is replaced by a relativistic Breit-Wigner distribution for the $K^*(892)$, normalized to equal the same fraction of the total

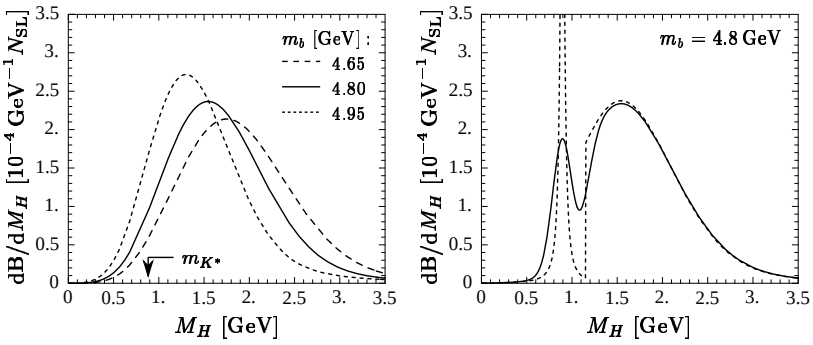

Figure 2.9: The m_{X_s} spectra taken from reference [26] ($m_H = m_{X_s}$). The spectra equivalent to Figure 2.8 for different m_b choices (left); and one of the m_{X_s} spectra modified to include the $K^*(892)$ resonance (right). The dashed curve shows the spectrum with an abrupt cut and the replacement of the $K^*(892)$, the solid curve shows the same with an experimentally more realistic energy smearing of about 100 MeV.

spectrum that was cut from the non-resonant part. (Note that the $K^*\gamma$ Breit-Wigner contribution continues above m_{cutoff}.)

This is illustrated in the right plot of Figure 2.9.

This prescription can be implemented by choosing a fixed m_{cutoff}. 1.1 GeV is the value implied by a different, preliminary, BABAR measurement of the X_s spectrum [8], with a semi-inclusive approach (see Section 4.4). Each choice of model parameters then yields a different fraction of $K^*(892)$ in the total rate. Alternatively, a fixed fraction of $K^*(892)$ might be chosen — *e.g.* , taking the experimentally measured branching fraction for $B \to K^*\gamma$ divided by the theoretically predicted branching fraction for the inclusive process — and adjust m_{cutoff} to satisfy the prescription. Measurements of the exclusive decay $B \to K^*\gamma$ have been performed by BABAR [18], CLEO [2], and BELLE [17]. BABAR's yields the most precise result. It

finds the branching fraction in the region $m_{X_s} < 1.1\,\mathrm{GeV}$ to be:

$$\mathcal{B}(B^{\pm} \to K^{*\pm}(892)\gamma) = (4.02 \pm 0.52(\mathrm{stat.}) \pm 0.22(\mathrm{syst.})) \times 10^{-5}$$

which is the average of the published values $\mathcal{B}(B^{\pm} \to K^{*\pm}(892)\gamma) = (3.83 \pm 0.62 \pm 0.22) \times 10^{-5}$ for the charged and $\mathcal{B}(B^0 \to K^{*0}(892)\gamma) = (4.23 \pm 0.40 \pm 0.22) \times 10^{-5}$ for the neutral mode. Table 2.2 shows the resulting cutoffs and $K^*(892)$ fractions for the range of m_b choices suggested in reference [26].

m_b	$K^*(892)$ fraction for $m_{\mathrm{cutoff}} = 1.1\,\mathrm{GeV}$	m_{cutoff} for $K^*(892)$ fraction $= 0.117$
4.65 GeV	0.103	1.14 GeV
4.80 GeV	0.161	1.02 GeV
4.95 GeV	0.273	0.90 GeV

Table 2.2: Choices of m_{cutoff} and $K^*(892)$ fraction for Kagan-Neubert model for several values of the b-quark mass m_b. These values satisfy the prescription [26] that the discarded integral of the X_s spectrum below m_{cutoff} equals the $K^*(892)$ contribution.

This measurement relies upon the signal model primarily to compute the efficiency. For this purpose, m_{cutoff} is fixed to 1.1 GeV, and the fraction of $K^*(892)$ is taken from Table 2.2. This signal model is also used when optimizing event selection criteria. This is the *only* place in this analysis where an assumed normalization (inclusive branching fraction) matters. But in this context, neither that assumption nor using a potentially "wrong" choice of model or model parameters can cause any bias; at worst they result in slightly non-optimal selection cuts.

A smearing of the photon energy (or m_{X_s}) distribution compared to Fig-

ure 2.9 arises from the Lorentz boost of the B meson from $\Upsilon(4S)$ decays in the overall center-of-mass frame (typically on the order of some 100 MeV). Figure 2.10 shows this prescription implemented for the signal model in the BABAR simulation of signal events. Here the photon energy plot is shown in terms of the energy in the overall center-of-mass frame, E^*_γ. Resolution effects in the detector are not taken into account in this figure, but they, of course, also yield an additional smearing.

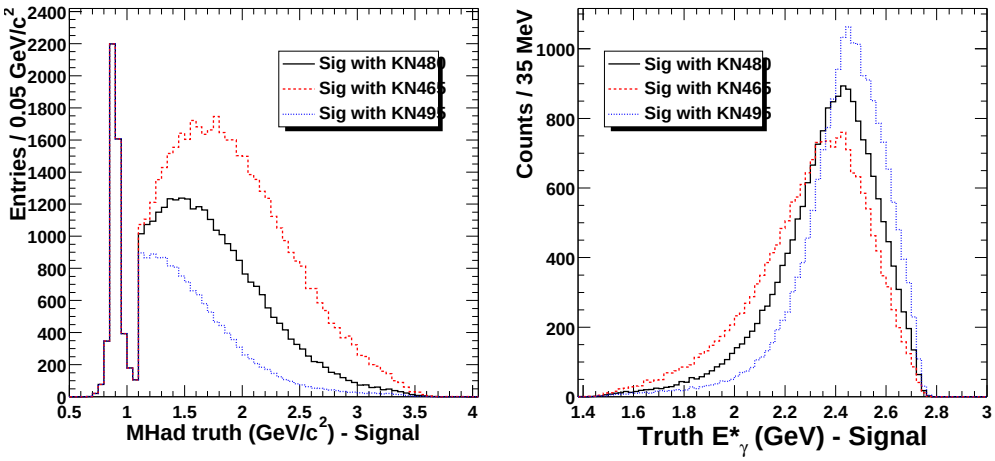

Figure 2.10: The m_{X_s} (left) and E^*_γ (right) signal model used for Monte Carlo simulation. E^*_γ is the energy of the photon in the $\Upsilon(4S)$ center-of-mass system. The fraction of $K^*(892)$ is as given in Table 2.2 for a fixed $m_{\rm cutoff}$ of 1.1 GeV. It is included in both plots. The m_{X_s} spectra for different m_b are normalized to the same $\mathcal{B}(B \to K^*\gamma)$, based on the BABAR measurement [18]. The E^*_γ spectra are all normalized to the same (arbitrary) integrals to allow easier visualization of the effect of a low cut-off in E^*_γ. The generated distributions are shown. They do not include reconstruction/resolution effects, only the Lorentz boost from the B rest frame into the overall CMS for the E^*_γ plot.

It is clear from the preceding discussion that the E_γ (m_{X_s}) spectrum is quite model-dependent, while the total $\mathcal{B}(B \to X_s\gamma)$ is much better understood theoretically. The total $\mathcal{B}(B \to X_s\gamma)$ is the integral of the $E_\gamma(m_{X_s})$ spectrum over the

entire kinematic range. This implies that, in order to avoid incurring a large model-dependent systematic uncertainty, an ideal analysis would have a flat 100% efficiency across the entire E_γ (m_{X_s}) spectrum — this could only be reached with no cuts on E_γ (m_{X_s}) and perfect reconstruction efficiency, clearly not a realistic assumption. For example, suppose that experimentally only photons above a threshold $E_\gamma^{\rm min}$ are selected. The measured branching ratio will then be the integral of the spectrum above the threshold. This is shown in the upper plots of Figure 2.8 as a function of minimum photon energy cut-off (or minimum bound of integration). It can be seen that as the threshold is raised the variation due to the model dependence is increased.

Thus, in order to relate an experimental measurement of $\mathcal{B}(B \to X_s\gamma)$ precisely to the theoretical predictions, a realistic analysis should be as inclusive as possible - *i.e.*, as few cuts on the E_γ or X_s system and as uniform selection efficiency as possible across the spectrum.

Chapter 3

The BaBar Experiment

The BaBar [35] experiment is run as a large collaboration of physicists and engineers from 75 institutions in 10 countries. It has been taking data at the PEP-II asymmetric B Factory [34] since 1999 and has recorded over 130 million $B\overline{B}$ pairs to date, a performance only rivaled by the Kek B Factory experiment Belle [36] in Japan which started operations almost concurrently with BaBar. The primary goal of these experiments is to investigate minute differences between the decay characteristics of the produced B mesons and their antiparticles (called B-bar mesons). This is part of the effort to understand the origin of the matter-antimatter asymmetry in nature (the universe contains a large amount of matter but very little anti-matter), *i.e.* to explore the mechanism of CP violation.

The abundant sample of $B\overline{B}$ mesons (pronounced "B B-bar") which give the experiment its name also offers a wealth of other analysis possibilities: precision

measurements of properties and decays of B and D mesons (especially rare decays), determination of CKM matrix elements, and many more.

The results presented in thesis were obtained from an event sample of about 90 million $B\overline{B}$ decays recorded with the BABAR detector [35] in the years 2000 to 2002. To demonstrate how these events were generated and collected, this chapter first gives a brief description of the accelerator and an overview of the different components of the BABAR detector. It then briefly summarizes the Monte-Carlo simulation of physics events and detector response and finally describes the data sets (real events taken with the BABAR detector as well as sets of simulated data) used in the $b \to s\gamma$ analysis.

3.1 The Asymmetric B-Factory PEP-II

The electron-positron collider PEP-II (**P**ositron-**E**lectron **P**roject II) is part of the Stanford Linear Accelerator Center (SLAC) in Northern California. SLAC's characteristic two mile long linear accelerator, the LINAC, accelerates electrons and positrons for collisions in BABAR and in fixed target experiments, as well as for the production of synchrotron radiation used for a variety of purposes by the Stanford Synchrotron Radiation Laboratory, SSRL (for example in the Sub-Picosecond Photon Source, SPPS). For use in BABAR, electrons and positrons reach their final energies of 9 GeV (electrons) and 3.1 GeV (positrons) in the LINAC, before they are injected in opposite directions into the PEP-II ring. There they are stored in two separate beam pipes on top of each other and brought into head-on collisions in one interaction region.

The BABAR detector wraps around this area and records the subsequent reactions.

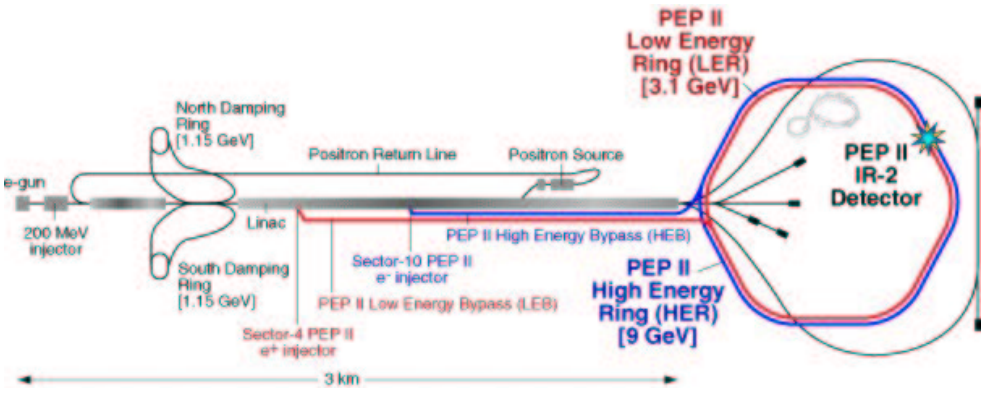

Figure 3.1: Detailed map of the BABAR accelerator system, showing the linac, the PEP-II rings, and the interaction region.

The majority of the data are taken at a center-of-mass (CMS) energy of 10.58 GeV, exactly on the peak of the $\Upsilon(4S)$ resonance (*on-resonance* data). This gives a maximum cross section for production of $B^0\overline{B}^0$ and B^+B^- pairs, since the $\Upsilon(4S)$ decays exclusively into pairs of charged and neutral B mesons. Other particles are also produced in these collisions: of most interest are initial pairs of c-quarks and τ-leptons, but pairs of lighter quarks and leptons are created in copious amounts also. See Table 3.1 for the detailed cross sections to all possible final states at on-resonance collision energies. Approximately 10% of all collisions are produced with a CMS energy about 40 MeV below the $\Upsilon(4S)$ peak (*off-resonance*), where the production of B mesons is kinematically suppressed, but all other particles are produced as before. This yields a control sample of continuum u-, d-, s-, and c-quark pairs, as well as τ-lepton pairs.

$e^+e^- \rightarrow$	Cross-Section (nb)
$b\overline{b}$	1.05
$c\overline{c}$	1.30
$u\overline{u}$, $d\overline{d}$, $s\overline{s}$	2.09
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 3.1: Production cross-section at $\sqrt{s} = M(\Upsilon(4S))$. The numbers quoted are effective cross-sections, as expected within statistics. The numbers are taken from reference [43].

At design luminosities of $3 \times 10^{33} cm^{-2} s^{-1}$, about 100 million b and c quark pairs are produced each year; regular operations exceed this benchmark since 2000[1], making PEP-II a true B Factory. It is called asymmetric because of the unequal energies of the electron and positron beams. The center-of-mass of the system resulting from the collision is then not at rest in the lab frame, but boosted in the direction of the higher-energy electron beam, with a Lorentz boost factor of $\beta\gamma = 0.56$. All decay particles, of course, will experience this boost also, resulting in a spatial separation of vertices from decays that happen at different times, an important condition for measuring time-dependent asymmetries in the decays of B and $\overline{B}$ mesons.

The event sample used for the present thesis corresponds to an integrated luminosity of 81.5 fb^{-1} collected on-resonance and 9.6 fb^{-1} collected off-resonance (see also Section 3.4).

[1]By May 2003 a peak luminosity of $6.1 \times 10^{33} cm^{-2} s^{-1}$ has been reached.

3.2 The BaBar Detector

In order to successfully investigate the broad range of physics questions reachable with a data sample obtained as described above, the BABAR detector must be a general-purpose instrument with great detection power for particles in a very wide dynamic range from 20 MeV (for slow π^0 and γ s) to 4-5 GeV (from asymmetric two-body B decays) and in the maximum possible angular acceptance region. Like every particle physics detector, it must have excellent abilities in vertex reconstruction and particle tracking, momentum and energy resolution, and in the identification of particle types ($e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, p, and also neutral hadrons: n, K_L^0). Additionally, there are requirements due to the forward boost of the decay particles in the lab frame. This puts the solid angle in the forward direction at a premium. Although the boost is rather small, the optimal detector acceptance is reached with an asymmetric detector. The boost also leads to most initial particles traveling close to the detector's z-axis (*i.e.* in the direction of the electron beam), this stresses the z-component of vertex reconstruction and the importance of minimizing multiple scattering within the detector.

The BABAR detector was designed to provide all the above features. Figure 3.2 shows a schematic of the detector.[2] Note the BABAR coordinate system as marked on the figure: the z-axis runs in the direction of the incoming e^- beam, the

[2]This and all following Figures in this chapter (unless marked otherwise) have been taken from the Nuclear Instruments and Methods (NIM) publication detailing the BABAR detector [35].

y-axis runs in vertical directions and the x-axis runs perpendicular to the z-axis in the horizontal plane. The polar angle, θ, is measured from the z- to the y-axis, and the azimuth angle, ϕ, is measured around the z-axis.

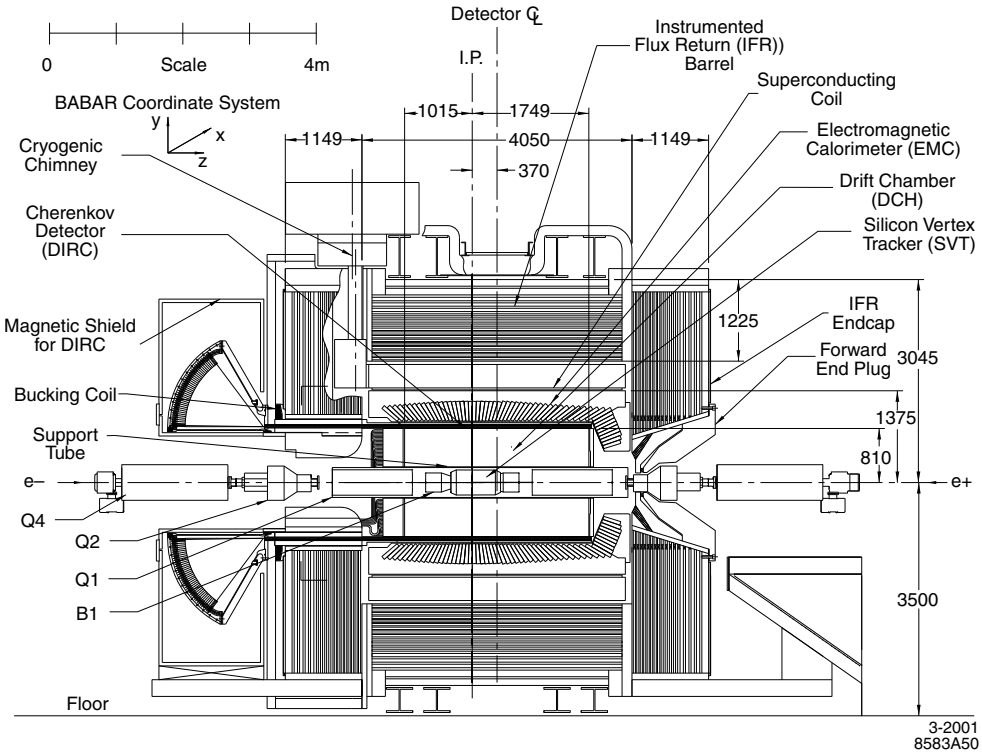

Figure 3.2: Schematic of the BABAR detector, showing - from the inside out- the SVT, DCH, DIRC, EMC, and IFR subcomponents. All dimensions are given in mm.

The following are its major components and their primary tasks, in order from closest to the beam-pipe to farthest from it:

- The Silicon Vertex Tracker (SVT)

  Precise position information on charged tracks and tracking of very low-energy

particles.

- The Drift Chamber (DCH)

  Main momentum measurement for charged particles and particle identification from measured specific energy loss, $\mathrm{d}E/\mathrm{d}x$.

- The Detector of Internally Reflected Cherenkov light (DIRC)

  Identification of charged hadrons: kaons, pions, protons.

- The Electromagnetic Calorimeter (EMC)

  Position and energy measurement for photons, identification of electrons from $\mathrm{d}E/\mathrm{d}x$.

- The 1.5 Tesla superconducting Magnet

  Solenoidal magnetic field, bending of charged tracks for momentum and charge determination.

- The Instrumented Flux Return (IFR)

  Hadron/muon discrimination and energy measurement.

All components, except the IFR, are encased in the magnet coil and its flux return. The measured curvature of the charged tracks in the magnetic field then allows the determination of their momentum and the sign of their charge.

The present analysis relies most heavily on the EMC, since the main characteristic of the signal is the high-energy photon. All other detector components are

used also, in the reconstruction of all other particles for the calculation of the shape of the event and for the identification of high-momentum electrons and muons (see Chapter 5). The detector subsystems are described in the following.

3.2.1 Silicon Vertex Tracker

The Silicon Vertex Tracker (SVT) is designed to provide very high precision trajectory measurements of charged particles in the immediate vicinity of the interaction region, and vertex information on the decay of short-lived particles. It is the only tracking device for particles with insufficient transverse momentum to reach the DCH. The charge collected in the silicon strips and the associated pulse shape yield information on the specific ionization energy loss.

The SVT is the innermost part of the BABAR detector. It consists of five concentric layers of double-sided silicon strip detectors, subdivided into modules in ϕ. The inner three layers are located as close as possible to the beam pipe, at radii of 32 to 54 mm (the beam-pipe is at 27.8 mm). Each of these layers has six modules assembled in a straight line. The outer two layers are farther away from the beam pipe (127 and 144 mm) with 16 and 18 modules each. They are arched to provide maximum coverage and smallest possible angle of incidence. The total active silicon area is 0.96 m^2 with about 150,000 strip readout channels. This layout is shown in Figure 3.3. The forward-backward or polar-angle coverage (+20.1° to -29.8°) is limited by magnets in the interaction region.

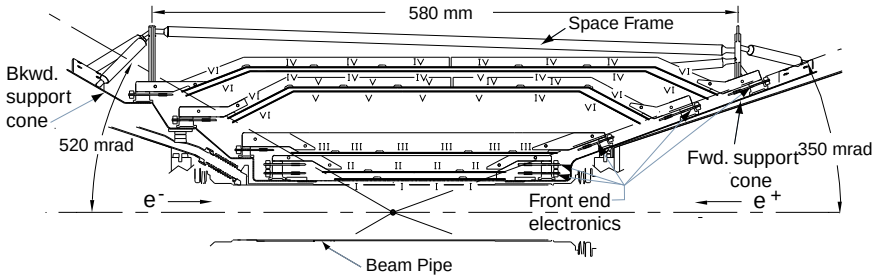

Figure 3.3: Layout of the Silicon Vertex Tracker (longitudinal cross section).

The strips on either side of the modules are rotated by 90° with respect to each other, with the strips closer to the interaction point running perpendicular to the beam (z-information), and the outer ones parallel to the beam (ϕ-information). Together they generate a stereo resolution on the hits they record. The pitch of the ϕ-strips varies between 50 μm in the innermost two layers and 100 μm in the outer layers. That of the z-strips is chosen similarly between 100 and 200 μm. The strips closer to the interaction region offer better resolution than the ones farther away from it, because of the smaller amount of material traversed before. This yields the highest resolution in the z-direction for the measurement of the distance between B and $\overline{B}$ decay vertices. Finally, the hits recorded in the SVT provide (entrance) angle information to the DCH to match up the trajectories reconstructed in the two subsystems.

To reconstruct space points from signals in adjacent silicon strips on both

sides of the sensors, the following procedure is used. First, channels are discarded that are not consistent with the event time as determined by the DCH. All good hits are then passed to the cluster-finding algorithm that groups adjacent strips with consistent times and calculates their charge pulse height. This is necessary, because on average one track will deposit charge in two or more strips due to track crossing angles as well as charge diffusion and induction. The center of the cluster is determined by averaging the positions of the first and last contributing strip, weighted by their collected charge. The clusters are then passed on to the pattern recognition algorithm that reconstructs tracks from all sub-detector information.

For optimal tracking reconstruction, a precise alignment of the SVT is necessary: the modules of the SVT need to be aligned locally, relative to each other, but also the SVT as a whole needs to be in alignment with the global detector system, as defined by the DCH. The local alignment is fairly stable over time. It only needs updating after major disruptions, like detector access or a magnet quench (motion due to heat expansion), it uses tracks from $e^+e^- \rightarrow \mu^+\mu^-$ events and cosmic rays. The global alignment is known to change continually over time, e.g. diurnally due to temperature changes. Thus it is usually updated on a run-by-run basis[3], every two to three hours. This way track reconstruction is always performed with up-to-date parameters.

The SVT is showing excellent performance in data taking. Its combined

[3]A run is a set of data taken in one session with the same machine and detector conditions.

hardware and software efficiency for single tracks has been measured to be 97%. It reaches a single hit resolution of 10 - 35 μm in both z- and ϕ-strips depending on the incident angle of the measured track. From measurements of the pulse shape for each hit, the ionization $\mathrm{d}E/\mathrm{d}x$ can be determined for each track. The resolution achieved on the best determination is approximately 14%. This energy loss is specific for each particle and yields an identification of the particle type with a separation of 2 σ between kaons and pions with momenta up to 500 MeV and between kaons and protons beyond 1 GeV. The SVT has thus fulfilled or surpassed its design goals. It is expected to hold up for another couple of years, even in the severe radiation environment brought about by increased luminosities.

3.2.2 Drift Chamber

Together, the SVT and the Drift Chamber comprise the tracking system for charged particles in the BABAR detector. The DCH provides measurements of drift times (and distances) for ionization events in the chamber caused by traversing charged particles. In this way it enables particle detection and extraction of momentum and angle information. DCH information provides reconstruction of trajectories as well as vertices of longer-lived particles decaying outside of the SVT volume. Its high precision enables the reconstruction of exclusive B- and D-meson decays with minimal background. In addition, the DCH provides ionization loss ($\mathrm{d}E/\mathrm{d}x$) measurements used in the identification of charged particles with low momenta.

The DCH is the next sensitive detector part outwards from the SVT. A structural support tube between these two tracking devices limits its inner radius at 236 mm, the DIRC the outer one at 809 mm. The DCH is cylindric, almost 3 m long, and offset from the interaction point with its center by 370 mm in forward direction. This yields a better forward coverage, with particles emitted at 17.2° in forward direction still traversing half of the chamber before exiting through the front plate. In the backward direction, this is true only up to particle angles of -27.4°. This layout is shown in Figure 3.4.

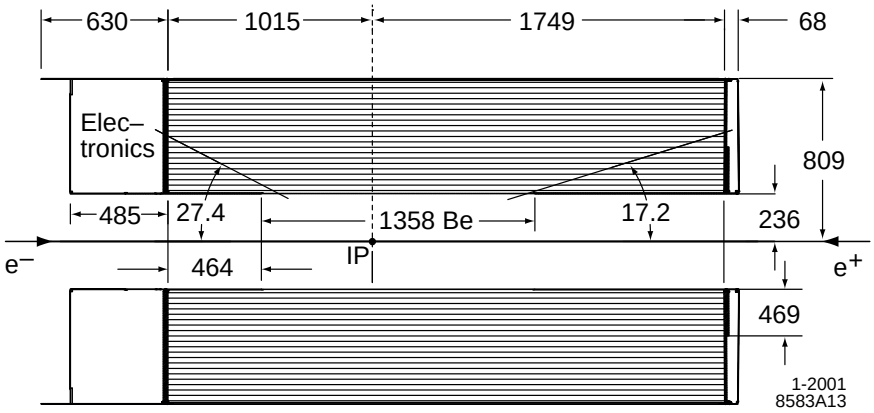

Figure 3.4: Layout of the Drift Chamber (longitudinal cross section). All dimensions are given in mm.

The chamber contains 40 radial layers of small hexagonal drift cells, 7,104 total, grouped by layers of four into 10 superlayers. Sequential layers are staggered by half a cell. Sense wires are placed inside each cell with various slight tilts towards the z-axis (same wire orientation within superlayers) to obtain longitudinal position

information. This set-up also enables local segment-finding and left-right ambiguity resolution within a superlayer. The stereo angles alternate between axial (A) and stereo (U,V) pairs from one superlayer to the next. Starting at lowest radii their order is AUVAUVAUVA, as shown in Figure 3.5. The tilt of the stereo wires increases with the radius of their layer (from ± 45 *mrad* to ± 76 *mrad*), with U (V) layers tilted to positive (negative) angles.

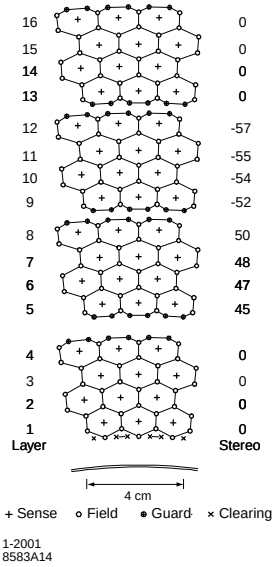

Figure 3.5: Schematic layout of the drift cells for the four innermost superlayers of the DCH (axial view).

The whole chamber is filled, at an overpressure of 4 mbar, with a 80:20 helium:isobutane gas mixture that minimizes multiple scattering, which is the dominant limitation on the resolution of tracks with momenta of less than 1 GeV (typical in B-decays). Up to 40 spatial and $\mathrm{d}E/\mathrm{d}x$ measurement points are obtained in the layers

for charged particles with transverse momenta p_T greater than 180 MeV/c.

The total amount of material in the DCH system (and in the support tube) is kept small to minimize the interaction impact on traversing particles. It also facilitates the matching of SVT and DCH tracks, improves the track resolution for high-momentum tracks, and avoids the degradation of performance in the DIRC and EMC.

For track reconstruction it is necessary to relate the drift time (measured) to the drift distance (which gives the position information) for each hit. For the extraction of dE/dx information, the relationship between gas gain and ionization loss is needed. The drift time-to-distance relationship is determined for each layer from Bhabha and $e^+e^- \to \mu^+\mu^-$ events from the measured drift time and the computed point of closest approach between the track (excepting the hit recorded on the wire in question) and the wire. The information is averaged over the hits within the same layer, keeping separate numbers for particles passing on the right and left side of the wires. A correction is also made for tracks with different entrance angles into the DCH. The overall gas gain is derived from measurements of the total charge deposited per drift cell after the measured number is corrected for geometric effects and variations in (gas) conditions and electronics response.

The performance of the DCH in data taking operations has been very stable and satisfies its design goals. The position resolution reached varies between 100 μm at medium drift distances (about 5 mm) and 250 μm for tracks either very close to

the wire or very far away (close to the cell boundary). The determination of $\mathrm{d}E/\mathrm{d}x$ has achieved a resolution of 7.5% to date.

3.2.3 Charged Particle Reconstruction

Trajectories of charged particles are reconstructed with the SVT and DCH information introduced above. Single hits are combined into track segments by pattern recognition algorithms if their event times match. For DCH track segments more hits (or whole segments) are added if this improves the track fit. Different track finding algorithms are employed for stand-alone tracks in each subdetector and for tracks not originating from the interaction region or not traversing the entire DCH. Each of them defines charged tracks by five parameters measured *at the point of closest approach to the z-axis* (POCA): d_0 (distance of closest approach (DOCA) in the x-y plane), z_0 (z-coordinate of the POCA), ϕ_0 (azimuth angle of the track at the POCA), λ (the dip angle relative to the transverse plane), and $\omega = 1/p_T$ (the curvature). Tracks found by the above procedure are then refit taking into account local variations in material and magnetic field (Kalman filter algorithm [37]). Then an attempt is made to extrapolate the track into the SVT, adding SVT track segments if they are found to be consistent by the track matching.

92% of the total solid angle in the center-of-mass system is covered by the tracking system. The DCH efficiency for tracks with transverse momenta $p_T > 200$ MeV is 98% as determined from the number of SVT tracks pointing into the

DCH volume compared to the number actually reconstructed there. The SVT efficiency for stand-alone tracks is greater than 80% down to $p_T \approx 70$ MeV/c. The resolution on the track parameters defined above can be calculated from Bhabha and μ-pair events. The p_T resolution is determined by the DCH only, using cosmic ray muons. The values obtained are:

$$\sigma_{d_0} = 23\mu\text{m} \qquad \sigma_{\phi_0} = 0.43 \text{ mrad}$$

$$\sigma_{z_0} = 29\mu\text{m} \quad \sigma_{\tan\lambda} = 0.53 \times 10^{-3}$$

$$\sigma_{p_T}/p_T = \quad (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\%$$

For the measurement of CP violation it is important that the resolution in the vertex displacement (along the z-axis), $\sigma_{\Delta z}$, between two neutral B mesons is about 25μm for transverse momenta above about 2 GeV.

3.2.4 Detector for Internally Reflected Cherenkov Radiation

The Detector for Internally Reflected Cherenkov Radiation (DIRC) is located between the DCH and the Electromagnetic Calorimeter at a radial distance of 837 mm. Its main purpose is the identification of and distinction between different types of hadronic particles, especially kaons and pions. This is particularly important for the full reconstruction of exclusive (rare) B decays and the determination of the B's flavor for CP violation measurements.

The DIRC is a new kind of ring-imaging Cherenkov light detector. It consists

of 144 long, straight bars of optically flat, synthetic fused silica quartz that run parallel to the beam direction, arranged as a 12-sided polygonal barrel. When a charged particle passes through the quartz, it produces Cherenkov radiation if its velocity v exceeds that of light in the medium c/n. The angle of emission with respect to the direction of the particle, θ_c, is related to $\beta = v/c$: $\cos\theta_c = 1/(\beta n)$, where $n = 1.473$ is the refractive index of the silica. Since one particle typically makes many Cherenkov photons at different azimuth angles around the photon direction, a light cone is produced. From its opening angle θ_c, a particle's mass and momentum can be deduced and thus its type is identifiable. The emitted light experiences total internal reflection inside the quartz bars and is thus transported to the ends of the optically flat bars with the angle of emission preserved. All light arriving at the forward end is reflected to the backward side with a mirror so that only the backward side of the bars needs to be instrumented. There it is allowed to expand into a 6 m^3 toroidal water reservoir. (Water, $n = 1.346$, is used because it is inexpensive and matches closely the index of refraction in the quartz thus minimizing internal reflection back into the quartz bars at the quartz–water interface.) The resulting (partial) Cherenkov rings on the back wall of the water chamber are detected by an array of 10,752 densely packed photo-multiplier tubes (PMTs). Both the detector layout and a schematic view of a typical photon path through the quartz and the water reservoir into the PMTs are shown in Figure 3.6.

The DIRC quartz bars (4.9 m long) cover the entire length of the DCH

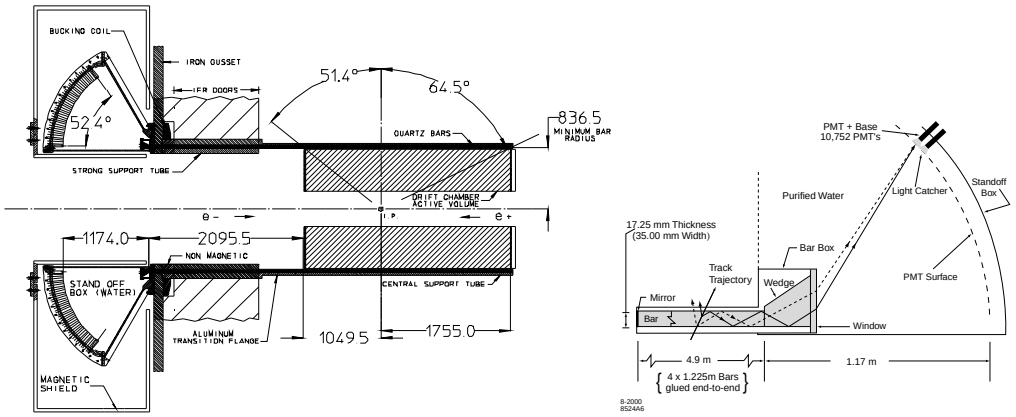

Figure 3.6: The DIRC system. On the left an elevated view of the overall DIRC geometry, on the right a detailed view of the quartz to water to PMT transition with example photon paths marked. Dimensions are given in mm.

giving a forward-backward coverage of 25.2° to -38.6°, equivalent to 83% of the polar angle in the CMS. The hexagonal arrays cover 94% of the total azimuth angle. The small radial thickness of 80 mm for the entire system (bars and support structures) and the small 17% of one radiation length this corresponds to keep the impact on the subsequent EMC measurement at a minimum, while allowing between 20 and 65 Cherenkov photons to be produced, depending on the polar angle.

The DIRC imaging uses both 3-dimensional position information and the arrival time of the signals. The ring image can then be associated with the DCH trajectory of a crossing particle. The resolution on the Cherenkov angle is found to be about 10.2 mrad, the time is measured to a resolution of about 1.7 ns.

The particle identification algorithm then maximizes the likelihood on the entire event by using individual track likelihoods for each of the following five particle

species hypotheses: electrons, muons, pions, kaons, and protons. The procedure combines information from the SVT, DCH, and the DIRC. For particles with momenta below about 700 MeV/c, mainly the dE/dx information from SVT and DCH is used for particle identification. DIRC information becomes dominant for momenta above that value, up to about 4.2 GeV/c (from two-body B decays). For momenta between 500 MeV/c and 3 GeV/c a kaon selection efficiency of $96.2 \pm 0.2\%$ is reached from a combinbation of all three systems with only $2.1 \pm 0.1\%$ pion misidentification. The separation between kaons and pions at 3 GeV/c is about 4.2 σ.

3.2.5 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (EMC) is designed to measure electromagnetic showers with very high precision and efficiency. It provides precise photon energy and angular position measurements over a wide kinematic range, from about 20 MeV to the kinematic limit of 9 GeV. This allows the measurement of photons from radiative and electromagnetic processes as well as π^0 and η decays and is the central point of importance for measuring the $b \to s\gamma$ decay with the fully-inclusive method presented in this thesis. It is based around the reconstruction of the high-energy (about $2 - 3$ GeV) signal photons and extracts the photon energy spectrum, so both high efficiency in the reconstruction and very good precision in the resolution of the energy measurement are needed, see Section 8.1. The upper kinematic limit for photon reconstruction in the EMC is reached in highly asymmetric π^0 decays from the $B \to \pi^0\pi^0$

decay mode. The EMC also measures the energy deposits from charged particles, which is the main ingredient for the identification of electrons and contributes to a variety of (semi-)leptonic physics analyses.

The EMC is composed of a cylindrical barrel and a conical forward endcap consisting of a total of 6,580 Thallium-doped CsI crystals. 5,760 crystals are located in the barrel, arranged in 48 identical rings; the remainder makes up eight rings in the endcap. This arrangement is presented in Figure 3.7. The system has full azimuthal

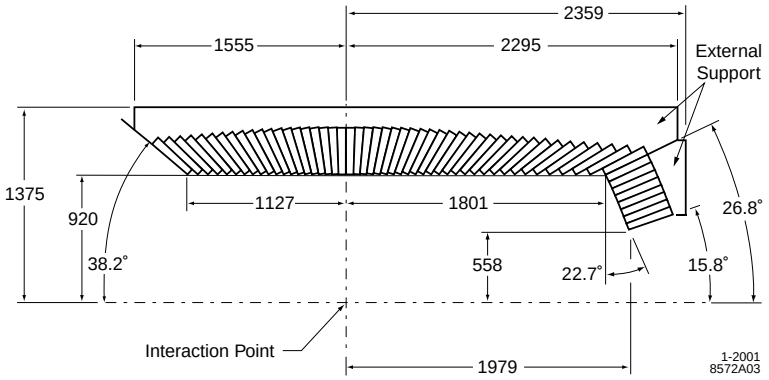

Figure 3.7: A longitudinal cross section of the EMC, showing the arrangement of the 56 crystal rings. The detector is axially symmetric around the z-axis, but has a forward offset from the interaction region. All dimensions given in mm.

coverage and extends in polar angle from $15.8 - 141.8°$, yielding a solid angle coverage of 90% total in the CMS.

The crystals have tapered trapezoidal longitudinal cross sections and increasing length in the forward direction to be able to avoid shower leakage of the expected higher energy particles there. At the front side of the EMC, the "pixellation" is given

by the frontal cross section of the crystals, 4.7×4.7 cm^2. The choice of CsI (Tl) for the crystal material offers advantages in energy resolution, stopping power, and radiation hardness. Two silicon pin diodes that are matched to the spectrum of the scintillation light are mounted on the rear face of each crystal. They read out the signal and pass it on to low-noise preamplifiers (see Fig. 3.8) that have a variable gain and can read out either of the two channels or average them.

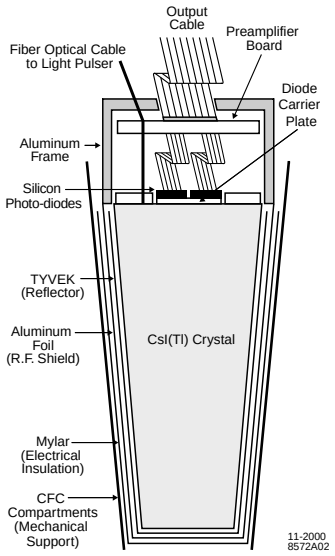

Figure 3.8: Schematic of a wrapped CsI (Tl) crystal and electronics.

The reconstruction of a particle's energy and angular position in the EMC has to take into account that electromagnetic showers from incident particles typically spread over many neighboring crystals, forming a *cluster*. Particles hitting the EMC in close vicinity of each other can result in merged clusters with several local maxima (*bumps*). A pattern recognition algorithm is then needed to efficiently identify single

or multiple bump clusters. Since charged particles leave energy deposits in the EMC as well, it is important to separate those clusters whose angular distribution can be matched with a charged track from the DCH/DIRC from the rest, which is then defined to be neutral clusters.

The relatively fine segmentation into crystals yields excellent angular resolution (between 3 and 12 mrad), parametrized as:

$$\begin{aligned} \sigma_\theta &= \sigma\phi \\ &= \left(\frac{3.87 \pm 0.07}{\sqrt{E(\,\mathrm{GeV})}} + 0.0 \pm 0.04 \right) \mathrm{mrad}. \end{aligned}$$

In order to realize the intrinsic performance expected of a CsI calorimeter, it is essential to set its energy scale with high precision and to determine any short- or long-term variations in its response. A multiple-step calibration system is therefore instated for the EMC. It first calibrates the energy measured in each individual crystal and then corrects for losses between crystals that occur because clusters spread over several crystals. (See left side of Figure 3.9 for an example of measured over expected energies.) The calibration proceeds as follows:

- A charge injection system linearizes the response of the front-end electronics to better than 1-2%.

- The individual crystals need to be calibrated at different energies because the light yield varies significantly from crystal to crystal and is generally non-uniform with energy and time.

- At low energies an absolute source calibration is performed by flushing aluminum tubes in front of the crystals with a neutron activated fluorine liquid radioactive source emitting 6.1 MeV test photons. This calibration is performed weekly to monitor the degradation of light yield due to radiation damage. It is statistically precise to better than 0.1%.
- At higher energies of $3 - 9$ GeV, the relation between polar angle and $e^{\pm}$ energy from Bhabha events is exploited for crystal calibration by comparing the measured values to predictions from Monte-Carlo simulations (MC). Statistical precision for this method is about 0.35%.

- To deduce the energy of the incident particle, it is necessary to correct the cluster energy reconstruction for energy losses due to shower leakage between crystals. This has to be done as a function of cluster energy and polar angle.
  - For cluster energies below 800 MeV a correction of about 6±1% is found by expressing the energy of photons from π^0 decays as a product of deposited energy and a correction function. The parameters of the latter are then determined iteratively.
  - At higher energy ($0.8 < \mathrm{E} < 9$ GeV) the correction is estimated from single-photon Monte-Carlo simulations or alternatively from radiative Bhabha events where the e^+ and e^- momenta and the photon direction allow a prediction of the photon energy.

The correction results in a very good agreement between the expected and the measured, corrected, peaks in cluster energy.

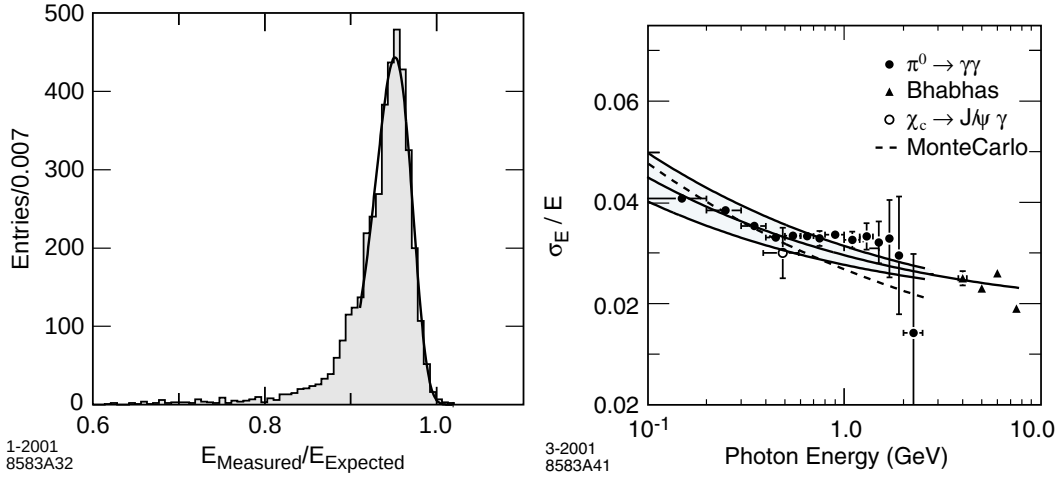

Figure 3.9: Left: The energy losses (between crystals) in the EMC, expressed as the ratio of measured over expected energy from Bhabha electrons of 7.5 GeV. Right: The energy resolution for the ECM as a function of cluster energy, measured for photons and electrons from various processes. The solid central curve is a fit to the $\frac{\sigma_E}{E}$ function and the area between the outer solid curves denotes the error of the fit.

The EMC has a typical light yield of about 50,000 electrons per MeV. The derived resolution as a function of cluster energy can be seen in the right plot in Figure 3.9 which also shows the π^0 and Bhabha sources of the calibration at different energies. At very low energies (6.13 MeV), the energy resolution reached in measurements with the radioactive source is $\sigma_E/E = 5.0 \pm 0.8\%$. At the other end of the spectrum (at 7.5 GeV), the Bhabha measurements find $\sigma_E/E = 1.9 \pm 0.07\%$. At intermediate energies, it is used that the mass resolution of reconstructed π^0 and η candidates is directly related to the daughter photon energy resolution. The deduced

photon energy resolution is parametrized as a smooth function of the cluster energy using a fit to the energy dependence:

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\%.$$

Isolated photons of roughly $1.5 - 3$ GeV, which are the centerpiece of this analysis, fall in the region that is extrapolated by the fit to the above function, as can be seen from Figure 3.9. An expected energy resolution of between 3% and 2.5% can be taken from the figure. A detection efficiency of nearly 100% is reached for these photons, provided they fall in the instrumented area (90% of the total solid angle in the CMS). Neutral hadrons can leave entries in the EMC that are unassociated with any charged track and will be a background to real photons, they can be reduced by using the shape of the radial and lateral distribution (parametrized in various moments) of the shower in the EMC, because hadronic interactions in the EMC leave different energy deposits than do electromagnetic interactions.

Electron identification in the EMC uses the shower energy, lateral moments of the shower, and the track momentum combined with DCH $\mathrm{d}E/\mathrm{d}x$ measurements and DIRC Cherenkov angles, but the main characteristic is the ratio of shower energy to track momentum E/p. The very tight selector used for the presented $b \to s\gamma$ analysis gives an electron efficiency of 88.1% for momenta between 0.5 and 2 GeV (which are of relevance here) while keeping the pion mis-identification rate around 0.15%.

3.2.6 Instrumented Flux Return

The Instrumented Flux Return (IFR) is designed to identify muons (complemented by the EMC) with high efficiency and purity, and to detect neutral hadrons, mainly K_L^0 and neutrons, over a wide range of momenta and angles. It is the outermost part of the detector and uses the segmented steel yoke of the super-conducting magnet as a muon filter and hadron absorber. For this task, it is instrumented in the segment gaps with resistive plate chamber (RPC) detectors with a two-coordinate readout. They consist of a thin gas filled core in which charged particles produce ionization which is detected on thin Aluminum strips on the top and bottom of the gas chamber. The strips are running in perpendicular directions to give two-dimensional position information of the ionization event. The iron plates with their high density compared to the other detector material and their high atomic number increase the chance that neutral long-lived particles will interact in the iron, producing charged secondaries that leave entries in the RPCs. The IFR is therefore able to detect these otherwise rarely interacting particles, but can only provide angle reconstruction of the incident particle. The calorimetric information from the secondaries cannot be related precisely enough to the original particle's energy to be useful. For charged particles, the IFR provides information on their penetration depth through the iron, thus enabling the identification of muons.

The hexagonal system consists of a central part (*Barrel*) and two plugs (*End Caps*). The Barrel is comprised of 19 active RPC layers interspersed with 18 layers of

steel plates. Each RPC plane is divided into three separate chambers that constitute the length of the barrel. The End Caps contain 17 layers of steel in between 18 active layers. They are split vertically into halves, each consisting of 6 chambers. They are movable, enabling a door-like access to the inner parts of the detector. This layout is shown in Figure 3.10.

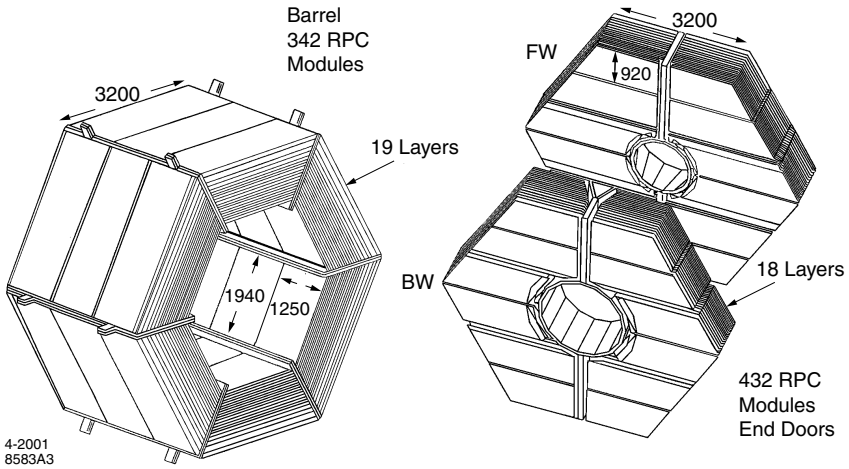

Figure 3.10: Overview of the IFR: Barrel sectors and forward (FW) and backward (BW) end doors. The shape of the RPC modules and their dimensions (in mm) are indicated.

Additionally, inside the barrel and directly outside the EMC is a novel feature of the BABAR detector: a cylindrical RPC in two layers. It measures charged particles that are leaving the EMC. Another novelty is the IFR's segmentation in the barrel: the thickness of the 18 iron layers increases (from 2 to 10 cm) with increasing radial position. The motivation for this rests with the K_L^0 detection efficiency, which improves with finer segmentation. Since most K_L^0 s interact within the first RPC layers, in this

innermost region the iron plates should be kept thin. At larger radial positions, K_L^0 detection becomes less important and a more cost-effective coarser segmentation can be chosen.

The IFR reconstruction algorithm starts by joining adjacent hits in one layer (one read-out coordinate) into a cluster. Then two-dimensional clusters are formed by combining one-dimensional clusters of the same coordinate in different layers. Finally, two-dimensional clusters in different coordinates are combined into the final three-dimensional IFR clusters. A second algorithm extrapolates charged tracks from the DCH to IFR clusters. Unmatched IFR clusters identify neutral hadrons; the algorithm tries to match them with neutral EMC clusters based on production angle. To distinguish muons from charged hadrons, the reconstruction compares the total number of interactions lengths worth of material traversed before the last IFR entry with the number expected for the same particle from simulations. It also uses the shape of the signal distribution over adjacent RPC strips and the quality of the fit to the cluster shape.

For muons with momenta in the range $1.5 < p < 3.0$ GeV a detection efficiency of almost 90% has been determined from final states, such as $\mu\mu\gamma$. Fake rates for pions in the same range are below 8%. For neutral hadrons the detection efficiency increases roughly linearly from 20 to 50% in the range $1 < p < 4$ GeV. The angular resolution is determined to be about 30 to 60 mrad, depending on whether an EMC match could be found.

Of the active RPC modules, 75% reach efficiencies above 90%. The RPCs are extremely sensitive to upward fluctuations in temperature which cause a decrease in their performance with time. In data-taking a correction has to be applied for this non-static behavior.

3.3 Detector and Event Simulation

As already mentioned above, simulated events are necessary to compare expected behavior of particles and decays with actually measured events. A Monte-Carlo (MC) simulation technique [38] is used for this task. It involves two steps:

- First, an event generator for physics processes of interest is needed. It must start from the initial e^+e^- pair and evolve through intermediate states (*e.g.* the $\Upsilon(4S)$ resonance) to the final state observable particles. This stage must include correctly both hadronization effects (through phenomenological fragmentation models) and subsequent final state particle decays (mostly taken from measured rates) to arrive at the stable particles that would be measured in real data. For event generation in BABAR a special package was developed: EvtGen [39]. It focuses in particular on the generation of B-decays. Their decay parameters are specified in large decay tables and taken from various signal models (for $b \to s\gamma$ see Section 2.4). Pre-existing generators are interfaced for other tasks. JetSet [40], for example, is used for simulation of continuum $q\bar{q}$ events.

- Second, if these events are to be analyzed in the same way as real data, then they have to also be passed through a simulation of the detector. This procedure has to model the detector geometry, material and fields that a particle traversing the detector would encounter. This includes the following processes that the particle might be subjected to: energy loss in the detector, (multiple) scattering on detector material, (hadronic) interactions with the detector, and decays of long-lived particles inside the detector volume. For this task the GEANT package [41, 42] is used in BABAR.

The generated events are then processed through the regular event reconstruction algorithms so that two sets of information are available for these events: the originally generated values including information about the entire decay chain (*MC truth* values) and the reconstructed ones (*reconstructed MC* values). The latter include effects of detector performance, like angular acceptance and resolution in the single detector components. This allows the study of data acquisition and reconstruction effects which may have to be corrected in imprecisely measured values or incorrect particle-type assignments. It also enables the study of specific analysis requirements by making the MC truth values of variables accessible. That means that it is possible in this way to break up the expected data sample into signal and background events and to optimize and evaluate signal efficiencies and background expectations with these events.

The operation of the BABAR detector during data taking continuously provides a better understanding of the detector response and the reconstruction requirements. Also the understanding of certain physics processes that are modeled in the generators improves with new measurements on BABAR and other data. Thus both the event generator and the detector simulation get updated regularly, typically after every run. New MC data then need to be produced with the new parameters.

3.4 Data Sets

Table 3.2 lists all the Monte-Carlo and BABAR data sets used in this analysis. BABAR's data taking periods are divided into so-called *Runs* with roughly similar running conditions between major shutdowns (typically this happens yearly). Here all data from Run 1 and Run 2, collected between March 2000 and June 2002, are used. This data set contains integrated on-resonance luminosities of 20.1 fb^{-1} in Run 1 and 61.4 fb^{-1} in Run 2 (81.5 fb^{-1} total). The complete set corresponds to a total of $88.5 \pm 1.0 \times 10^6$ $B\overline{B}$ pairs. The Monte-Carlo events used in the present analysis all stem from the fourth round of the BABAR **S**imulation **P**roduction and are identified as *SP4 Monte-Carlo*. The cross sections used for the $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, $c\overline{c}$, and $\tau^+\tau^-$ processes are the same as given in Table 3.1 and are estimated from the calculations in the event generators (no error is available on these numbers). The luminosity listed for each Monte Carlo sample is simply a consequence of the number of generated events and the assumed cross section (or branching fraction). It is easily rescaled for any change

in that assumed value.

The algorithms used in the reconstruction of events change with better knowledge gained about the detector during data-taking, in roughly the same way as the MC simulations. Different versions of the reconstruction package are here named after the numbered software release they are implemented into. The latest version is always used for the production of new MC data sets. BABAR data taken in previous runs need to be reprocessed when new reconstruction improvements are adopted. (This is a task that increases in complexity as the collected amount of data increases.) For this thesis, BABAR data and MC events have been reconstructed with "release-10" software.

The Kagan-Neubert (KN) signal model used for simulation is described in Section 2.4. In order to estimate the dependence of the analysis efficiency on the signal model, the KN signal is generated with several different sets of input parameters: the b quark mass (m_b), which labels the sample, and an appropriately prescribed value of the b-quark momentum inside the B meson (the μ_π or λ_1 parameter). Appendix B and Table B.1 therein provide more details of the KN model, and specify how the second parameter was set for the generator.

To be able to compare the on-resonance BABAR data set to all MC subsets and the off-resonant sample, their sizes must correspond to that of the on-resonance data. The yields obtained in all these sets after the analysis cuts must therefore be scaled appropriately. For $B\overline{B}$ and signal Monte Carlo, one uses *not* the luminosity

Data Set	Source	Events	Cross Section or $\mathcal{B}$	Luminosity
uds	SP4	124230000	2.09 nb	$59.4\,\text{fb}^{-1}$
$c\overline{c}$	SP4	61417500	1.30 nb	$47.2\,\text{fb}^{-1}$
$\tau^+\tau^-$	SP4	66264000	0.94 nb	$70.5\,\text{fb}^{-1}$
$B^0\overline{B}^0$	SP4	99293277	0.543 nb	$182.9\,\text{fb}^{-1}$
B^+B^-	SP4	99273616	0.543 nb	$182.9\,\text{fb}^{-1}$
$B^0 \to K^{*0}(892)\gamma$	SP4	111000	$4.03 \pm 0.43 \times 10^{-5}$	$2526\,\text{fb}^{-1}$
$B^\pm \to K^{*\pm}(892)\gamma$	SP4	117000	$4.03 \pm 0.43 \times 10^{-5}$	$2663\,\text{fb}^{-1}$
$B^0 \to X_s^0\gamma$ (KN465)	SP4	70805	$3.51 \pm 0.37 \times 10^{-4}$	$184.3\,\text{fb}^{-1}$
$B^\pm \to X_s^\pm\gamma$ (KN465)	SP4	65601	$3.51 \pm 0.37 \times 10^{-4}$	$170.7\,\text{fb}^{-1}$
$B^0 \to X_s^0\gamma$ (KN480)	SP4	83992	$2.10 \pm 0.22 \times 10^{-4}$	$366.4\,\text{fb}^{-1}$
$B^\pm \to X_s^\pm\gamma$ (KN480)	SP4	80889	$2.10 \pm 0.22 \times 10^{-4}$	$352.6\,\text{fb}^{-1}$
$B^0 \to X_s^0\gamma$ (KN495)	SP4	56325	$1.07 \pm 0.11 \times 10^{-4}$	$482.0\,\text{fb}^{-1}$
$B^\pm \to X_s^\pm\gamma$ (KN495)	SP4	56464	$1.07 \pm 0.11 \times 10^{-4}$	$483.2\,\text{fb}^{-1}$
$B^0 \to X_s^0\gamma$ (KN flat)	SP4	475666	n/a	n/a
$B^\pm \to X_s^\pm\gamma$ (KN flat)	SP4	569644	n/a	n/a
On-Resonance	Run 1+2			$81.5 \pm 1.2\,\text{fb}^{-1}$
Off-Resonance	Run 1+2			$9.59 \pm 0.15\,\text{fb}^{-1}$

Table 3.2: Monte Carlo and real data sets. KNxxx = Kagan & Neubert model [26] with "xxx" the value of the b quark mass and "flat" as explained in the text. Each $\mathcal{B}(B \to K^*\gamma)$ is set to the weighted average of the BABAR measurements for the two charge states [18]. Each $B \to X_s\gamma$ sample excludes $B \to K^*\gamma$ events; all branching ratios have been normalized to that for $B \to X_s\gamma$ using the K^* (892) fractions from the constant-m_{cutoff} column of Table 2.2.

values of the samples, but rather the number of $B\overline{B}$-pairs in the samples (also called the B-count). The number for on-resonance data is estimated from its relation to the number of μ-pairs and multihadronic events in the on- and off-resonance samples [56]. A 1.1% uncertainty on the method has been determined [57], which will be taken into account as a systematic error for this analysis (Section 8). Normalization of the continuum samples is done in terms of luminosity. For the Monte Carlo sets the luminosity is determined from the cross-sections and number of generated events as mentioned

above (uncertainties on the cross sections are not taken into account). The luminosity for the data samples is determined from the number of measured μ-pairs in each sample and the very well determined cross section for μ-pair production, with an error of roughly 1.5% (mostly from reconstruction effects). This makes use of the experimentally very clean handle on μ-pair reconstruction. Finally, in normalizing off-resonance to on-resonance data, the *relative* luminosity is used (with an uncertainty of less than 0.5% [57] because many effects cancel out in the ratio). All normalization uncertainties turn out to be negligible compared to other uncertainties (see Section 8.4).

It is worth noting that, while the MC samples were produced to correspond to the range of data-taking conditions, the background samples in particular are not uniformly spread in proportion to actual data. For example, a breakdown into 2000 *vs.* 2001 *vs.* 2002 conditions is shown in Table 3.3. A study of the possible effect of this is presented in Section 8.4.

To study particular sources of background (π^0/ η) as described in Section 6.3, a different preselection of events is used than for the regular analysis. It makes use of the total available real data sample, but a statistically slightly different set of $B\overline{B}$ Monte Carlo events. A summary of these data sets is given in Table 3.4.

Data Set	Luminosity in fb^{-1} (fraction)		
	2000	2001	2002
uds	17.1 (0.29)	28.7 (0.48)	13.6 (0.23)
$c\bar{c}$	15.9 (0.34)	17.2 (0.36)	14.1 (0.30)
$\tau^+\tau^-$	17.2 (0.24)	38.6 (0.55)	14.7 (0.21)
$B^0\overline{B}^0$			
B^+B^-			
On-Resonance	20.34 (0.25)	35.61 (0.44)	25.55 (0.31)
Off-Resonance	2.61 (0.27)	3.79 (0.40)	3.19 (0.33)

Table 3.3: Background Monte Carlo and real data sets broken down by year. For the Monte Carlo samples, this indicates the conditions used for simulation and reconstruction. The "fractions" in parentheses are relative to the total luminosity for each sample.

Data Set	Source	Events	Cross Section	Luminosity(fb^{-1})
$B^0\overline{B}^0$	SP4	98993277	0.543 nb	$182.4\,\text{fb}^{-1}$
B^+B^-	SP4	99283611	0.543 nb	$182.9\,\text{fb}^{-1}$
On-Resonance	Run 1+2			$81.5 \pm 1.2\,\text{fb}^{-1}$
Off-Resonance	Run 1+2			$9.59 \pm 0.15\,\text{fb}^{-1}$

Table 3.4: Monte Carlo and real data sets used for $\pi^0(\eta)$ studies ("π^0 Skim").

Chapter 4

Analysis Overview

Various considerations contribute to the mapping out of an analysis strategy. Amongst them are the sources of backgrounds to the desired signal and their relative contributions to the measurement, but also the different possible techniques for selecting signal over background. These are described for the $b \to s\gamma$ analysis in the first two parts of this chapter. Then an overview is given of the specific analysis method used in this thesis. The chapter finishes with a synopsis of previous experimental measurements of the $b \to s\gamma$ decay and a comparison of their techniques.

4.1 Backgrounds

The cross sections and branching ratios listed in Table 3.2 demonstrate very clearly the relative amounts of different hadronic events that are produced in e^+e^- collisions at the $\Upsilon(4S)$ energy: for every $B\overline{B}$ pair created in the collisions in BABAR

more than four events are produced that include light-quark or tau pairs[1]. Furthermore, fewer than one in 1000 $B\overline{B}$ events yields a signal $B \to X_s\gamma$ decay. Therefore, the $b \to s\gamma$ analysis must be carried out in the presence of large amounts of other events, backgrounds, which need to be suppressed to arrive at a sample that consists only of signal events.

Fortunately, the energy of the photons from $B \to X_s\gamma$ decay is high enough to provide a good experimental handle for the selection. The $B \to X_s\gamma$ decay is essentially a 2-body decay of a B meson, modulo the Lorentz-boost from the B rest frame into the CMS and modulo QCD mass effects. Therefore, the photon energy (in the CMS frame) peaks around 2.4 GeV (roughly half the B mass). The spectrum then falls steeply towards higher energies and has essentially vanished by about 2.8 GeV, but has a long tail to lower energies, extending below 1.6 GeV, as can be seen in the right half of Figure 2.10. There are no other hadronic B decays expected to yield a *direct* photon in this energy range.

To single out the signal, it is useful to start with a hadronic event selection, which suppresses lepton-pair events and $e^+e^- \to \gamma\gamma$ processes as well as beam backgrounds.[2] To this, one adds the requirement of a high-energy photon (see Figure 2.10). But after this simplest possible fully-inclusive selection, the signal is still overwhelmed

[1] There are also more than 40 events with electron or muon pairs expected per $B\overline{B}$ event, but these are easily distinguished from hadronic events and are not of concern.

[2] Beam background events stem from interactions of beam-particles with the magnets, with the beam pipes, or with particles from the other beam, all of which can leave signals in the detector volume faking an event. How many of these events end up in the data sample depends on the control of the beams' positions, dispersions, *etc.* (beam "tune") during data-taking.

by hadronic backgrounds, as can be seen in Figure 4.1.

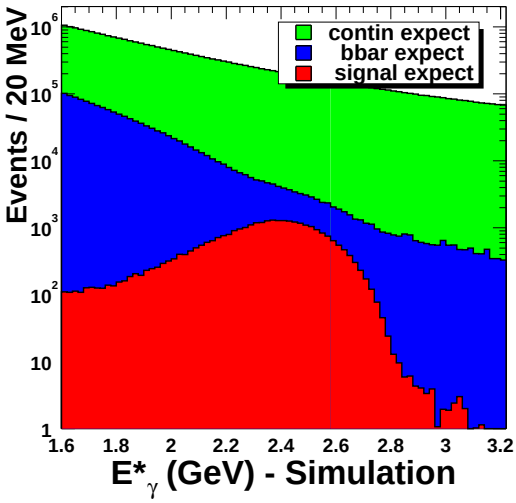

Figure 4.1: The $SP4$ Monte Carlo expected E^*_γ distributions for signal and background photon candidates, normalized to $81.5\,\mathrm{fb}^{-1}$. These distributions are after the B hadronic skim (Section 5.1) but before the "photon quality" cuts that are designed to reduce backgrounds from π^0s, ηs and hadrons.

These backgrounds consist mainly of:

- Non-radiative $B\overline{B}$ events:
  - Decays of inclusively produced mesons. The high-energy photon comes most often from a π^0 or η parent produced in a B or $\overline{B}$ decay. The $\pi^0(\eta)$ can decay asymmetrically into two photons, of which one can have quite high energy. The spatial distribution of the measured particles in the event (the event shape) is the same as that for the $b \to s\gamma$ signal, *i.e.*, isotropic.
  - Hadronic particles. The high-energy photon is faked by the hadronic interaction in the calorimeter of (*e.g.*) a pion, kaon, proton, or neutron produced

in the B decay. The event shape is the same as that for the $b \to s\gamma$ signal, *i.e.*, isotropic.

- Continuum events ($u\overline{u}$, $d\overline{d}$, $s\overline{s}$ and $c\overline{c}$):
  - Decays of inclusively produced mesons. The high-energy photon comes most often from the asymmetric decay of a π^0 or η parent into two photons. The event has a "dijet" shape, with the photon in one of the jets.
  - Initial State Radiation (ISR). A high-energy photon is radiated (usually at a small angle) from one of the colliding $e^{\pm}$ before the collision. The ensuing $q\overline{q}$ from the e^+e^- collision recoils against the high-energy photon in a "three-jet" event shape.
  - Hadronic particles. Analogous to the corresponding $B\overline{B}$ entry, except that the event has a dijet shape.

Both continuum and non-radiative $B\overline{B}$ backgrounds also have additional small contributions, mostly from radiative decays of other particles. Finally, there is expected to be a small contribution (of order 5% as explained in Section 2.3) from $b \to d\gamma$ decays, which is not separable in a fully-inclusive analysis. ($B \to X_c\gamma$ processes are not penguin decays and yield photons at substantially lower energies, so very little background of this kind is expected in the signal E^*_γ region.)

The experimental challenge is to reduce the backgrounds to such a level that the statistical and systematic errors on the background subtraction are substantially

smaller than the signal, while avoiding the large model-dependent systematics that arise when limitations on the γ or X_s system are imposed. For the BABAR analysis presented here, a compromise between both requirements is found by using a minimum cut on E^*_γ of 2.0 GeV. Section 9.1 presents the basis for this choice.

4.2 Methods

For an inclusive measurement of the $b \to s\gamma$ branching fraction and spectrum two approaches are possible. One uses a semi-inclusive method that fully reconstructs several exclusive X_s states and combines them with the selected high-energy photon. The second approach is solely centered around the high-energy photon and does not use any reconstruction of or requirements on the X_s system. This makes this method fully-inclusive. The two approaches are complementary and may be combined into one measurement. Such combination has clear statistical advantages, but also potential drawbacks (systematic precision, model dependence). Both methods can, of course, also be pursued separately and used to cross-check and confirm each other. This is the current situation at BABAR. The fully inclusive analysis is the topic of this thesis, but in the following a brief summary of both methods (in the way they are implemented in BABAR analyses) is given, with a discussion of advantages and disadvantages of each.

4.2.1 Semi-Inclusive Method

The semi-inclusive analysis is, perhaps, the more intuitive approach to measuring $b \to s\gamma$: The X_s system is predicted to consist of the various resonances and non-resonant states listed in Table 2.1. Reconstructing various final states of kaons and pions and combining them with high-energy photons to fully reconstruct one of the B mesons in the signal event is then a straightforward expansion of exclusive measurements of, for example, the mode $B^0 \to K^{*0}(892)\gamma$, $K^{*0} \to K^+\pi^-$. Currently final states with one kaon and up to four pions, at most one of them neutral (12 final states), are considered and added together for this analysis. This semi-inclusive method represents only on the order of one half of all possible final states. The inclusion of more final states is presently under study within the BABAR collaboration.

The analysis can be outlined as follows. Quality cuts are placed on the high-energy photon, and event shape variables are used to suppress the continuum backgrounds. $B\overline{B}$ backgrounds have to be estimated from Monte Carlo simulations and subtracted. Events are required to be reconstructed in one of the exclusive modes where quality cuts on the daughter candidates reduce backgrounds. The full reconstruction of the signal B decay also yields kinematic constraints on the full B decay system. Since the X_s system is reconstructed, the recoil hadronic mass is an accessible variable and the spectrum can be measured in bins of m_{X_s}. It can, of course, easily be converted into the corresponding photon energy spectrum using the kinematic relation between the two variables. In terms of the hadronic mass, the region with dominating

$B\overline{B}$ backgrounds is now the high end of the spectrum. Due to limitations in the reconstruction of final states and the large uncertainties of the background subtraction, only a limited range of the full spectrum is considered for this analysis with an upper cut-off of $m_{X_s} < 1.094$ GeV (equivalent to a minimum E^*_γ cut-off of about 2.1 GeV). The background subtracted data-yield in each bin must be corrected for the signal efficiency, which is estimated from signal MC. However, the yields measured directly in this way only represent the 12 reconstructed modes and not the full number of final states that contribute to the X_s system. The number in each bin therefore must be corrected for the fraction of unmeasured states to give the inclusive spectrum. It can then be fit to extract the theoretical model parameters and extrapolated to the full range to yield the inclusive branching fraction.

The advantages of this method lie in the availability of information about the hadronic system and the reconstructed B meson. Because of the kinematic constraints on the decay products from the full reconstruction of the signal B, this is a relatively tight selection and there is not much non-B background to contend with. Yet, an additional background contribution is introduced that comes from wrongly reconstructed signal events. Furthermore, the hadronic mass resolution can be up to an order of magnitude better than the resolution of the smeared photon energy spectrum. The resolution uncertainty (smearing) of the photon energy spectrum (measured in the $\Upsilon(4S)$ rest frame) stems from the fact that the two B mesons are not created at rest in the $\Upsilon(4S)$ frame. The resulting Doppler smearing of the photon energy is enhanced

by the additional resolution uncertainty from the calorimeter. The total smearing is on the order of 150 MeV. The energy spectrum can be corrected for this effect, but the use of the m_{X_s} spectrum may allow for a more precise measurement of the spectral shape.

The disadvantages of the semi-inclusive method are mainly connected to the large uncertainty in modeling the way in which the X_s system is composed of exclusive final states. It arises from the hadronization of the inclusive spectrum. Since not all final states that contribute to the $b \to s\gamma$ rate can be reconstructed, the correction for the unmeasured fraction must be taken from Monte-Carlo simulations. The theoretical description of hadronization effects is given by QCD, but the large value of α_s at the involved distances forbids a perturbative calculation. Simulations therefore rely on phenomenological models whose parameters must be tuned to give results that reproduce measured distributions. But the higher resonance states and non-resonant contributions at high hadronic mass are not well determined experimentally, and so there is no good way to verify that the simulation results model the data in this region well. This approach therefore shows a strong dependence on the signal model resulting in a large uncertainty on the deduced branching ratio and spectrum. In addition, background contributions from wrongly reconstructed signal events have to be identified and corrected, which complicates the method and increases the systematic effects.

Preliminary results obtained at BABAR with this method on an on-resonance

data sample of 20 fb^{-1} were presented at the 2002 International Conference on High Energy Physics (ICHEP). The result obtained is:

$\mathcal{B}(B \to X_s\gamma)^{preliminary}_{semi\text{-}inclusive} = (4.3 \pm 0.5(\text{stat}) \pm 0.8(\text{syst}) \pm 1.3(\text{th})) \times 10^{-4}$

4.2.2 Fully-Inclusive Method

The second method for measuring the $b \to s\gamma$ decay is centered around the reconstruction of only the high-energy photon. Direct photons in the energy range of interest have few other direct sources, making this approach possible. No requirements are made on the X_s system, which allows contributions from all possible hadronic final states to be measured, making this truly a fully inclusive analysis.

This method works with roughly the same quality cuts on the high-energy photon as the semi-inclusive analysis. Continuum backgrounds in particular are rejected by an analysis of the event shape and the requirement of a high-momentum lepton which in signal decays comes from the second B in the event. This exploits the fact that high-momentum leptons are produced much more often in $B\overline{B}$ decays than in continuum events. This approach eliminates almost all of the continuum background in the sample. Any remainder is subtracted using off-resonance data scaled to the on-resonance luminosity. Unfortunately, no similarly powerful handle exists for backgrounds from $B\overline{B}$ decays. They must be estimated from tuned MC simulations and subtracted. In the lowest photon energy region the $B\overline{B}$ backgrounds dominate and introduce large systematic uncertainties. Their influence is reduced by placing a

minimum cut on E_γ^* of 2.0 GeV. The subtracted data-yield is treated in bins of E_γ^* (m_{X_s} is, of course, not accessible here) and corrected for efficiency. This final spectrum can then be fit and extrapolated to the full range (the considered full range starts at 1.6 GeV as discussed in Section 2.3.1) to obtain the inclusive branching fraction. From the spectrum's moments the different theoretical model parameters can be deduced.

The obvious disadvantage of this method is the large amount of $B\overline{B}$ and continuum background that must be dealt with. While the combination of shape analysis and lepton requirements is able to remove almost all of the continuum background, it is not very efficient and greatly reduces the statistics in the signal.[3] There is no equivalent technique available to suppress the $B\overline{B}$ backgrounds which dominate at low E_γ (the full reconstruction of the signal B in the semi-inclusive analysis is an advantage in this respect). The subtraction using Monte Carlo data brings about the additional task of verifying (better: tuning) the production of the main background processes in the simulation and the associated contribution to the systematic error. Another limitation of this method is the resolution smearing of the photon energy by about 150 MeV, as mentioned above. To first order this *does not* affect the determination of the branching fraction. It *does* influence the extraction of any model parameters, though. For this purpose the measured E_γ^* spectrum can be parametrized as a convolution of a function describing the resolution effects with the true spectrum in the B rest frame. This will give a calculable additive correction to the moments once they

[3]Even with these drawbacks, this *is* the optimal solution for maximizing the overall statistical power, which depends on both signal statistics and statistical effects from the background subtraction.

are extracted from the unblinded data.

The main strength of this approach is its complete inclusiveness to all hadronic final states, which is reached with this simple method. No additional uncertainties due to hadronization effects are introduced. This helps to keep the total measurement error low. The E^*_γ cut-off does introduce dependence on the signal model, but this has to be weighed against the systematic error from the use of the simulation to subtract the $B\overline{B}$ backgrounds. With a data sample of higher statistics and an even better understanding of the $B\overline{B}$ background in data and simulation, the cut on E^*_γ may be reduced to include more of the original spectrum and to reduce the model dependence.

At the ICHEP 2002 conference, also preliminary results from this second method were presented. At that point the analysis was performed on a data sample of 54 fb^{-1} and yielded the following result:

$\mathcal{B}(B \to X_s\gamma)^{preliminary}_{fully\text{-}inclusive} = (3.88 \pm 0.36(\text{stat}) \pm 0.37(\text{syst}) \pm^{0.44}_{0.23}(\text{th})) \times 10^{-4}$

Any combination of this method with a semi-inclusive approach that reconstructs only specific final states will take away from this advantage by introducing many more dependencies on theoretical assumptions. Still, a combination may be desirable for the gain in efficiency. Eventually, a complete analysis of all errors that enter into the combination will be needed to decide on whether a combination is overall beneficial. The investigation of the combined approach is planned as the next major step for this measurement at the BABAR experiment.

4.3 Analysis Outline

The analysis presented here is designed to be fully inclusive with minimal requirements on the high-energy photon and X_s system. The general approach is to make a series of cuts which discriminate against the backgrounds (most effectively against continuum backgrounds) without introducing any significant model-dependence for the signal.

This selection is described in detail in Chapter 5. Events are pre-selected with a very loose selection that starts from the reconstructed data format. To pass the pre-selection stage, events must exhibit a hadronic signature and at least one photon candidate above 1.5 GeV. They are then passed through a dedicated reconstruction process for this analysis that computes relevant quantities and stores them in object-oriented n-tuples that were specifically designed for this measurement. The subsequent C++ analysis first selects candidates for good high-energy isolated photons using a set of quality requirements. Photons that are consistent with coming from calorimeter interactions of hadronic particles or from π^0 and η decays are rejected to suppress background from both continuum and $B\overline{B}$ decays. The vetoed events give one possibility for a $B\overline{B}$ background control sample to verify the Monte Carlo modeling. A second more inclusive sample is described below. Next, a combination of shape variables into a linear Fisher discriminant is used to reduce (especially) continuum background. The requirement of the presence of a lepton, which in signal events comes from the recoiling non-signal B, is also particularly effective in this regard, de-

spite a considerable cost in signal efficiency. B events are much more likely to have high-momentum leptons than are continuum events, and the lepton direction relative to the photon candidate is also a valuable signature.

Since the selection criteria involve only global event quantities and the recoiling B meson (for the lepton requirement), it is possible to reduce the continuum background without imposing stringent requirements on E_γ (m_{X_s}). The large $B\overline{B}$ background at low energies forces the use of a cut ($E^*_\gamma > 2.0\,\mathrm{GeV}$) at the moment. It may be viable to lower this cut-off at a later point, when these backgrounds are better understood and will not introduce as much of a systematic uncertainty.

After applying the cuts, the remaining continuum background is estimated using the data taken at CMS energies below the $\Upsilon(4S)$, where only continuum events are produced. All off-resonance data surviving the analysis cuts are scaled to the on-resonance integrated luminosity and subtracted from the on-resonance sample. The off-resonance CMS energy lies 40 MeV below the peak of the $\Upsilon(4S)$ resonance. Since the resonance is very narrow (the full width is only 14 MeV [19]) this is far enough away to completely avoid $B\overline{B}$ production from the tails. Yet, it is close enough that the production cross sections for the continuum are practically the same as at the $\Upsilon(4S)$ energy. It is advantageous to use the real off-resonance data for this subtraction task instead of Monte-Carlo simulation continuum background events, because the Monte Carlo is not regarded as reproducing the real data reliably enough. Using off-resonance data ensured that no potential dependence on the parameter tune of the Monte Carlo

is introduced. But since only about 12% of the total BABAR data sample is taken off-resonance, this method is statistically limited.[4]

On the other hand, the remaining $B\overline{B}$ background *is* taken from the simulation. For the primary components that are contributing to the backgrounds in this analysis, independent measurements are used to correct or at least check that simulation. These checks are important, because there are portions of this background (especially hadronic fakes) which are not expected to be well-simulated. The dominant $B\overline{B}$ backgrounds arise from π^0 and η decays to two photons. For these cases, the inclusive B-decay fractions are measured *vs.* parent $\pi^0(\eta)$ energy using independently-selected BABAR data samples of π^0 and η decays and compare the results to the Monte Carlo expectations. The results of these comparisons are used to re-weight the generic $B\overline{B}$ Monte Carlo events which are used in the subtraction. As a cross check for this method the sample of events that fail the π^0 and η vetoes in the signal event selection are used in a similar manner confirming the Monte Carlo tune. An analogous test sample was also defined for the hadronic component of the $B\overline{B}$ background. The uncertainty in the $B\overline{B}$ background is expected to be primarily systematic.

The E^*_γ distribution in 100 MeV bins yields the signal spectrum. All subtractions are performed bin-by-bin for the full range of $1.7 < E^*_\gamma < 3.5\,\mathrm{GeV}$. To extract a branching fraction, only the range $2.0 - 2.7$ GeV is used. The areas outside this region ($1.7 < E^*_\gamma < 2.0\,\mathrm{GeV}$ and $2.7 < E^*_\gamma < 3.5\,\mathrm{GeV}$) are dominated by $B\overline{B}$ and continuum

[4]When the off-resonance data-yield is scaled to the on-resonance luminosity, its statistical error is scaled in the same way, making this the dominant contribution to the total statistical error.

background, respectively, and can be used to check subtraction methods and other simulation issues.

The analysis is done blind in the region $1.9 < E^*_\gamma < 2.9\,\mathrm{GeV}$. This means that the whole analysis technique, including the determination of related systematic errors, is fixed using solely simulated MC events before the BABAR data in this region can be looked at (or un-blinded).

4.4 Other Measurements

This section lays out the current status of experimental measurements of the $b \to s\gamma$ decay.

The first pure penguin process was observed experimentally in 1993 by the CLEO collaboration [2] long after this class of decays was postulated by theorists in 1977 [1]. This first measurement of the rate for the exclusive state $B \to K^*\gamma$ was evidence for the existence of penguin mediated processes and of the $b \to s\gamma$ decay in particular. However, due to the limitations in the modeling of the $b \to s\gamma$ fragmentation to exclusive final states, the inclusive branching fraction (which has implications for extensions of the Standard Model) cannot be deduced from this measurement. Therefore, focus was shifted soon to the inclusive measurement of $b \to s\gamma$.

CLEO, again, was first to publish a branching fraction for the inclusive rate in 1995 [4]. This measurement was done on only 2.2 million $B\overline{B}$ pairs. It shows the

expected large (statistical) errors. This result was confirmed by the ALEPH collaboration in 1998 [6] using 0.8 million $B\overline{B}$ pairs identified in Z^0 decays. Two further measurements were published in 2001 and 2002. The first is by the Japanese B-factory BELLE [7] and uses their first data set of 6 million $B\overline{B}$ decays. The second is a refined result by the CLEO experiment [5]. It uses their upgraded detector CLEO II and the final data set of 10 million $B\overline{B}$ decays; it supersedes their original measurement. Two as yet unpublished but public measurements (a semi-inclusive and the antecessor of the present fully-inclusive analysis) were presented by the BABAR collaboration at the 2002 ICHEP conference, as discussed briefly in Section 4.2. Since their results are already widely cited by theorists, they are included in this summary. All these experimental results and the theoretical prediction by Gambino and Misiak [16] are summarized in Table 4.1 and Figure 4.2. The current world average including all experimental numbers is found to be [10]: $\mathcal{B}(B \to X_s\gamma)_{worldAvg} = (3.34 \pm 0.38) \times 10^{-4}$.

Experiment	Reference	Min. E^*_γ	$\mathcal{B}(B \to X_s\gamma) \times 10^{-4}$
CLEO 95	[4]	2.20 GeV	$2.32 \pm 0.57(stat) \pm 0.35(sys)$
ALEPH 98	[6]	2.20 GeV	$3.11 \pm 0.80(stat) \pm 0.72(sys)$
BELLE 01	[7]	$\sim$ 2.24 GeV	$3.36 \pm 0.53(stat) \pm 0.42(sys)^{+0.50}_{-0.54}(th)$
CLEO 01	[5]	2.00 GeV	$3.21 \pm 0.43(stat) \pm 0.27(sys)^{+0.18}_{-0.10}(th)$
BABAR 02 s	[8]	2.1 GeV	$4.30 \pm 0.50(stat) \pm 0.80(sys) \pm 1.3(th)$
BABAR 02 f	[9]	2.1 GeV	$3.88 \pm 0.36(stat) \pm 0.37(sys) \pm^{0.43}_{0.23}(th)$

Table 4.1: Previously-measured values of $\mathcal{B}(B \to X_s\gamma)$. The BELLE measurement is semi-exclusive and cuts on $m_{X_s} < 2.05\,\text{GeV}$ rather than on E^*_γ; on average, that cut corresponds to an E^*_γ of 2.24 GeV. BABAR's previous preliminary measurements are also included since they are already widely used by theorists (s for semi-, f for fully-inclusive).

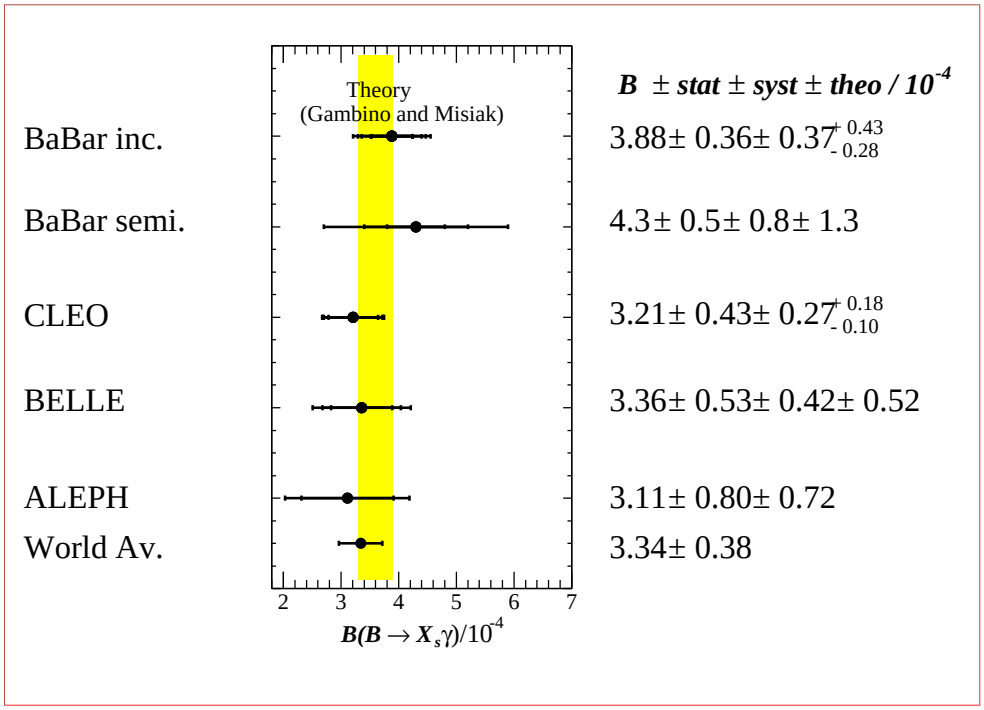

Figure 4.2: Previous experimental measurements of the branching ratio $\mathcal{B}(B \to X_s\gamma)$ compared to the NLO theoretical expectations for the Standard Model.

To facilitate a later comparison with the analysis presented here, the different analysis methods and parameters leading to the above experimental numbers shall be briefly discussed here.

CLEO's original analysis used a combination of a fully inclusive and a semi-inclusive analysis. The two approaches are complementary with little correlation. The former is based on the presence of a high-energy photon and the particular expected shape of signal events versus different kinds of backgrounds. As discussed above, the semi-inclusive approach uses pseudo-reconstruction to combine a kaon and several

pions with the high-energy photon into one object consistent with a B meson. Not all the possible hadronic final states can be reconstructed in this way, so the measured partial rate needs to be extrapolated back to the fully inclusive space. This leads to a strong model dependence because of large QCD uncertainties in the fragmentation to specific final states. The modeling of the signal composition and spectrum relies on work by Ali and Greub [30] which by today has been shown to rely on untenable assumptions [16, 26]. The minimum photon energy cut-off used in this analysis is 2.2 GeV.

The approach taken by the ALEPH collaboration is necessarily different from the other experiments, since this experiment is the only one not running at a CMS energy of the $\Upsilon(4s)$ resonance but instead on the Z^0 peak. The signal sample is tagged for b-quark content, which results in far less problems with background from continuum decays. A semi-inclusive analysis method is then applied to this sample. The signal region extends down to photon energies of 2.2 GeV. The extrapolation and signal modeling are done according to [30].

The number from the BELLE experiment is from a purely semi-inclusive analysis approach. It uses the signal model following the prescription in [26]. Due to their analysis method of reconstructing final states of a kaon and pions, m_{X_s} is their natural variable. BELLE uses a mass cut-off of $m_{X_s} < 2.05$ GeV to avoid large backgrounds from non-signal B decays. The corresponding photon energy cut-off is about 2.24 GeV.

The recent, second CLEO measurement is the most sophisticated to date. It

is also the first to use the obtained photon energy spectrum to extract the HQET parameter $\bar{\Lambda}$. This analysis is, again, a hybrid analysis. It combines a shape variable method with pseudo-reconstruction and with the requirement of a high-momentum lepton. This analysis considers events down to photon energies of 2.0 GeV. Signal modeling is again done with the prescription described in [30] but the correction for the unmeasured region in E_γ^* uses the model from [26].

All experimental results quoted so far have been extrapolated down to E_γ^* =250 MeV. Gambino and Misiak prescribe that this should be raised to 1.6 GeV. To be able to compare new branching ratio measurements (with the new range of extrapolation) to the previous results, the latter have to be scaled by a factor of 3.57/3.73=0.957 (from their calculations for the two ranges).

Chapter 5

Event Selection

The fully inclusive study of $b \to s\gamma$ requires hadronic events which include a high-energy photon originating directly from a B meson decay. To provide such an event sample, selection criteria are employed which are designed to yield a maximum reduction of background events while maintaining a high signal efficiency.

Several successive stages of event selection are necessary in the process. These are described in the following sections.

- The first selection step is a fast (skim) preselection that restricts the sample to hadronic events with a high-energy photon present. It includes rough cuts against Bhabha events ($e^+e^- \to e^+e^-$) and the most obvious continuum events. This step is described in Section 5.1.

- The reconstructed high-energy photon candidates are subjected to a set of crite-

ria to ensure the quality of their reconstruction and their identity as real photons. See Section 5.2 for details.

- Vetoes are applied that are designed to reject a high-energy photon if it is consistent with being the decay product of a $\pi^0 \to \gamma\gamma$ or $\eta \to \gamma\gamma$ decay. A full explanation can be found in Section 5.3.

- The energy range of the high-energy photon is limited to reduce systematic errors associated with large backgrounds at the low end of the photon energy spectrum. More details are given in Section 5.4.

- Signal and continuum background exhibit different event topologies. Several event shape variables are used for background reduction, most of them combined into a Fisher discriminant. See Section 5.5 for a discussion of the event variables used and the benefits of using them in a linear discriminant method.

- Events with multiple high-energy photon candidates are removed from the event sample; see Section 5.6.

- Most of the continuum reduction is achieved by the requiring the presence of a lepton in the event. This is presented in Section 5.7.

The order of cuts is relevant both to the histograms shown for each quantity before a cut is made on it and to the numerical effects of each cut, but it does not affect the overall result. This is due to the fact that most of the cut values have been selected

using an optimization procedure described in Section 5.8, which studies the effects of each cut with *all* other cuts applied. Monte Carlo simulation expectations for this selection are then presented in Section 5.9. Finally the deduction of an overall signal efficiency is described in Section 5.10, which concludes this chapter.

Tables 5.1 and 5.2 summarize the step-by-step efficiencies for the different signal models described in Section 2.4 and the X_s component. The KN computations are for an m_{X_s} cutoff of $m_{\mathrm{cutoff}} = 1.1\,\mathrm{GeV}$. Tables 5.3 through 5.6 show the $SP4$ efficiencies for each category of background. The cuts are made sequentially and each cut efficiency is given relative to the previous cut. Cumulative efficiencies are given for sets of associated cuts. The first column of each table provides a reference to the section below where the cuts are described in detail. Each section contains a description of the cuts and plots of the relevant distributions for signal and background. The background plots overlay the different components described in Section 4.1 and their sum. In several cases the cuts were chosen to be identical to those used in a previous published analysis, so the original work is referenced and only a brief description given here.

5.1 Skim Cuts

The selection starts with a preselection. This process is conceptually similar to skimming cream off milk: it rejects the majority of background events (about 97%), keeping only events that appear roughly similar to the signal events that are sought.

Sec.	Description	Neutral X_s			Charged X_s		
		KN465	KN480	KN495	KN465	KN480	KN495
5.1	BGFMultiHadron	0.97	0.96	0.96	0.97	0.97	0.96
5.1	nTrk > 2	1.00	1.00	1.00	1.00	1.00	1.00
5.1	$1.5\,\text{GeV} <$ e1Mag $< 3.5\,\text{GeV}$	0.76	0.77	0.77	0.77	0.77	0.77
5.1	$R_2^* < 0.9$	1.00	1.00	1.00	1.00	1.00	1.00
5.1	p1Mag $< 2.5\,\text{GeV}$	0.98	0.98	0.98	0.98	0.98	0.98
5.1	nHEgamma > 0	1.00	1.00	1.00	1.00	1.00	1.00
	Cumulative Skim	**0.72**	**0.73**	**0.72**	**0.73**	**0.73**	**0.73**
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.97	0.97	0.97	0.97	0.97	0.97
5.2	No problematic channels	0.97	0.98	0.97	0.97	0.98	0.97
5.2	nCrystals > 4	1.00	1.00	1.00	1.00	1.00	1.00
5.2	Second Moment < 0.002	0.98	0.98	0.98	0.98	0.98	0.98
5.2	Lateral Moment < 0.45	1.00	1.00	1.00	1.00	1.00	1.00
5.2	Bump Isolation > 25 cm	0.91	0.91	0.91	0.90	0.90	0.90
5.3	π^0 veto	0.93	0.94	0.95	0.93	0.94	0.94
5.3	η veto	0.93	0.94	0.95	0.92	0.94	0.95
	Cumul. Photon Quality	**0.72**	**0.74**	**0.75**	**0.71**	**0.73**	**0.74**
5.4	$\mathbf{2.0\,\text{GeV} < E_\gamma^* < 2.7\,\text{GeV}}$	**0.85**	**0.91**	**0.94**	**0.85**	**0.90**	**0.94**
5.5	$R_2^* < 0.55$	0.97	0.97	0.97	0.97	0.97	0.97
5.5	Fisher Discriminant > 0.575	0.59	0.61	0.64	0.56	0.59	0.62
	Cumulative Event Shape	**0.57**	**0.59**	**0.62**	**0.55**	**0.57**	**0.60**
5.6	**Single Gamma Cut**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	Cumulative before Tag	**0.25**	**0.29**	**0.32**	**0.24**	**0.28**	**0.31**
5.7	Electron Tag $p_e^* > 1.25\,\text{GeV}$	0.038	0.039	0.040	0.039	0.041	0.043
5.7	$\cos\theta_{\gamma e} > -0.7$	0.880	0.877	0.903	0.882	0.871	0.860
5.7	Muon Tag $p_\mu^* > 1.5\,\text{GeV}$	0.022	0.022	0.022	0.025	0.022	0.026
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.765	0.809	0.794	0.812	0.798	0.787
5.7	Emiss$> 0.8\,\text{GeV}$	0.978	0.977	0.974	0.978	0.967	0.965
	Cumulative Lepton Tag	**0.049**	**0.050**	**0.052**	**0.053**	**0.052**	**0.055**
	TOTAL Efficiency	**0.0123**	**0.0144**	**0.0165**	**0.0130**	**0.0144**	**0.0169**
	Expected Events, $81.5\,\text{fb}^{-1}$	387	269	157	407	269	161

Table 5.1: Selection efficiencies and expected number of events for the X_s component of the signal in $81.5\,\text{fb}^{-1}$, from $SP4$ Monte Carlo samples for different values of the Kagan Neubert parameter m_b. This is denoted as KNxxx, where xxx is $100^* m_b$.

		K^{*0}	K^{*+}
5.1	BGFMultiHadron	0.93	0.94
5.1	nTrk > 2	1.00	1.00
5.1	$1.5\,\text{GeV} <$ e1Mag $< 3.5\,\text{GeV}$	0.79	0.79
5.1	$R_2^* < 0.9$	1.00	1.00
5.1	p1Mag $< 2.5\,\text{GeV}$	0.98	0.98
5.1	nHEgamma > 0	1.00	1.00
	Cumulative Skim	**0.72**	**0.73**
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.97	0.97
5.2	No problematic channels	0.97	0.97
5.2	nCrystals > 4	1.00	1.00
5.2	Second Moment < 0.002	0.98	0.98
5.2	Lateral Moment < 0.45	1.00	1.00
5.2	Bump Isolation > 25 cm	0.91	0.91
5.3	π^0 veto	0.97	0.97
5.3	η veto	0.98	0.98
	Cumul. Photon Quality	**0.80**	**0.79**
5.4	$\mathbf{2.0\,GeV < E_\gamma^* < 2.7\,GeV}$	**0.91**	**0.91**
5.5	$R_2^* < 0.55$	0.92	0.93
5.5	Fisher Discriminant > 0.575	0.80	0.75
	Cumulative Event Shape	**0.74**	**0.70**
5.6	**Single Gamma Cut**	**1.00**	**1.00**
	Cumulative before Tag	**0.39**	**0.37**
5.7	Electron Tag $p_e^* > 1.25\,\text{GeV}$	0.043	0.044
5.7	$\cos\theta_{\gamma e} > -0.7$	0.897	0.910
5.7	Muon Tag $p_\mu^* > 1.5\,\text{GeV}$	0.028	0.026
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.816	0.812
5.7	Emiss$> 0.8\,\text{GeV}$	0.945	0.944
	Cumulative Lepton Tag	**0.058**	**0.058**
	TOTAL Efficiency	**0.0224**	**0.0213**
	Expected Events, $81.5\,\text{fb}^{-1}$	80	76

Table 5.2: Selection efficiencies and expected number of events for the K^* component of the signal in $81.5\,\text{fb}^{-1}$, from $SP4$ Monte Carlo samples.

Sec.	Description	$u\overline{u}, d\overline{d}, s\overline{s}$		
		π^0/η	ISR	other
5.1	BGFMultiHadron	0.900		
5.1	nTrk > 3	0.999		
5.1	1.5 GeV $<$ e1Mag $<$ 3.5 GeV	0.162		
5.1	$R_2^* < 0.9$	0.994		
5.1	p1Mag $<$ 2.5 GeV	0.849		
5.1	nHEgamma > 0	0.938		
	Cumulative Skim	**0.115**		
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.96	0.91	0.94
5.2	No problematic channels	0.97	0.98	0.97
5.2	nCrystals > 4	1.00	1.00	1.00
5.2	Second Moment < 0.002	0.95	0.97	0.44
5.2	Lateral Moment < 0.45	0.99	1.00	0.93
5.2	Bump Isolation > 25 cm	0.44	0.93	0.60
5.3	π^0 veto	0.53	0.95	0.88
5.3	η veto	0.63	0.94	0.85
	Cumulative Photon Quality	**0.13**	**0.71**	**0.16**
5.4	**2.0 GeV $< E_\gamma^* <$ 2.7 GeV**	**0.33**	**0.36**	**0.31**
5.5	$R_2^* < 0.55$	0.74	0.96	0.76
5.5	Fisher Discriminant < 0.575	0.25	0.13	0.25
	Cumulative Event Shape	**0.18**	**0.12**	**0.19**
5.6	**Single Gamma Cut**	**1.00**	**1.00**	**1.00**
5.7	Electron Tag $p_e^* > 1.25$ GeV	0.0035	0.0021	0.0032
5.7	$\cos\theta_{\gamma e} > -0.7$	0.0741	0.1892	0.1294
5.7	Muon Tag $p_\mu^* > 1.5$ GeV	0.0048	0.0034	0.0046
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.0294	0.1111	0.0583
5.7	Emiss> 0.8 GeV	0.7027	0.6667	0.7222
	Cumulative Lepton Tag	**0.00028**	**0.00052**	**0.00050**
	TOTAL Efficiency	**2.58e-07**	**1.89e-06**	**5.50e-07**
		4.59e-07		
	Expected Events, 81.5 fb^{-1}	36	25	18
		78		

Table 5.3: Expectations for $u\overline{u}$, $d\overline{d}$, $s\overline{s}$ background in 81.5 fb^{-1}, using $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays, photons from initial-state radiation, and other).

Sec.	Description	$c\bar{c}$		
		π^0/η	ISR	other
5.1	BGFMultiHadron	0.960		
5.1	nTrk > 3	0.999		
5.1	$1.5\,\text{GeV} <$ e1Mag $< 3.5\,\text{GeV}$	0.099		
5.1	$R_2^* < 0.9$	0.998		
5.1	p1Mag $< 2.5\,\text{GeV}$	0.927		
5.1	nHEgamma > 0	0.827		
	Cumulative Skim	**0.072**		
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.96	0.91	0.95
5.2	No problematic channels	0.98	0.98	0.97
5.2	nCrystals > 4	1.00	1.00	1.00
5.2	Second Moment < 0.002	0.92	0.98	0.42
5.2	Lateral Moment < 0.45	0.99	1.00	0.92
5.2	Bump Isolation > 25 cm	0.49	0.93	0.56
5.3	π^0 veto	0.49	0.94	0.84
5.3	η veto	0.66	0.94	0.84
	Cumulative Photon Quality	**0.13**	**0.71**	**0.14**
5.4	$\mathbf{2.0\,GeV < E_\gamma^* < 2.7\,GeV}$	**0.28**	**0.36**	**0.25**
5.5	$R_2^* < 0.55$	0.79	0.99	0.79
5.5	Fisher Discriminant < 0.575	0.23	0.18	0.22
	Cumulative Event Shape	**0.18**	**0.18**	**0.17**
5.6	**Single Gamma Cut**	**1.00**	**1.00**	**0.99**
5.7	Electron Tag $p_e^* > 1.25\,\text{GeV}$	0.0112	0.0075	0.0111
5.7	$\cos\theta_{\gamma e} > -0.7$	0.0315	0.2269	0.0652
5.7	Muon Tag $p_\mu^* > 1.5\,\text{GeV}$	0.0076	0.0046	0.0099
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.0405	0.1136	0.0488
5.7	Emiss$> 0.8\,\text{GeV}$	0.8667	0.9531	0.8000
	Cumulative Lepton Tag	**0.00057**	**0.00213**	**0.00096**
	TOTAL Efficiency	**2.83e-07**	**7.09e-06**	**4.22e-07**
		1.26e-06		

Table 5.4: Expectations for $c\bar{c}$ background in $81.5\,\text{fb}^{-1}$, using $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays, photons from initial-state radiation, and other).

Sec.	Description	$\tau\bar{\tau}$		
		π^0/η	ISR	other
5.1	BGFMultiHadron	0.272		
5.1	nTrk > 3	0.995		
5.1	$1.5\,\text{GeV} <$ e1Mag $< 3.5\,\text{GeV}$	0.131		
5.1	$R_2^* < 0.9$	0.899		
5.1	p1Mag $< 2.5\,\text{GeV}$	0.808		
5.1	nHEgamma > 0	0.993		
	Cumulative Skim	**0.026**		
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.96	0.91	0.87
5.2	No problematic channels	0.97	0.98	0.96
5.2	nCrystals > 4	1.00	1.00	1.00
5.2	Second Moment < 0.002	0.97	0.99	0.65
5.2	Lateral Moment < 0.45	1.00	1.00	0.98
5.2	Bump Isolation > 25 cm	0.32	0.95	0.53
5.3	π^0 veto	0.47	0.97	0.96
5.3	η veto	0.95	0.98	0.97
	Cumul. Photon Quality	**0.13**	**0.80**	**0.26**
5.4	$\mathbf{2.0\,GeV < E_\gamma^* < 2.7\,GeV}$	**0.34**	**0.35**	**0.31**
5.5	$R_2^* < 0.55$	0.05	0.81	0.07
5.5	Fisher Discriminant < 0.575	0.21	0.05	0.26
	Cumulative Event Shape	**0.01**	**0.04**	**0.02**
5.6	**Single Gamma Cut**	**1.00**	**1.00**	**0.94**
5.7	Electron Tag $p_e^* > 1.25\,\text{GeV}$	0.011	0.021	0.023
5.7	$\cos\theta_{\gamma e} > -0.7$	0.143	0.163	0.333
5.7	Muon Tag $p_\mu^* > 1.5\,\text{GeV}$	0.002	0.013	0.011
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.000	0.111	0.333
5.7	Emiss$> 0.8\,\text{GeV}$	1.000	0.800	0.667
	Cumulative Lepton Tag	**0.0016**	**0.0039**	**0.0076**
	TOTAL Efficiency	**1.80e-08**	**1.08e-06**	**2.61e-07**
		1.66e-07		
	Expected Events, $81.5\,\text{fb}^{-1}$	1	9	2
		13		

Table 5.5: Expectations for tau background in $81.5\,\text{fb}^{-1}$, using $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays, photons from initial-state radiation, and other).

Sec.	Description	B^0B^0		B^+B^-	
		π^0/η	other	π^0/η	other
5.1	BGFMultiHadron	0.991		0.992	
5.1	nTrk > 2	1.000		1.000	
5.1	$1.5\,\text{GeV} <$ e1Mag $< 3.5\,\text{GeV}$	0.025		0.024	
5.1	$R_2^* < 0.9$	0.999		0.999	
5.1	p1Mag $< 2.5\,\text{GeV}$	0.982		0.980	
5.1	nHEgamma > 0	0.960		0.972	
	Cumulative Skim	**0.023**		**0.023**	
5.2	$-0.74 < \cos\theta_\gamma < 0.93$	0.97	0.95	0.97	0.95
5.2	No problematic channels	0.98	0.98	0.98	0.98
5.2	nCrystals > 4	1.00	0.99	1.00	0.99
5.2	Second Moment < 0.002	0.92	0.44	0.94	0.50
5.2	Lateral Moment < 0.45	0.99	0.93	0.99	0.94
5.2	Bump Isolation > 25 cm	0.65	0.68	0.64	0.69
5.3	π^0 veto	0.42	0.82	0.40	0.82
5.3	η veto	0.69	0.82	0.71	0.82
	Cumul. Photon Quality	**0.17**	**0.17**	**0.16**	**0.20**
5.4	$\mathbf{2.0\,GeV < E_\gamma^* < 2.7\,GeV}$	**0.116**	**0.122**	**0.110**	**0.109**
5.5	$R_2^* < 0.55$	0.98	0.99	0.98	0.99
5.5	Fisher Discriminant < 0.575	0.45	0.27	0.46	0.29
	Cumulative Event Shape	**0.44**	**0.27**	**0.45**	**0.29**
5.6	**Single Gamma Cut**	**0.0713**	**0.0573**	**0.0778**	**0.0713**
5.7	Electron Tag $p_e^* > 1.3\,\text{GeV}$	0.047	0.038	0.053	0.051
5.7	$\cos\theta_{\gamma e} > -0.75$	0.64	0.69	0.59	0.60
5.7	Muon Tag $p_\mu^* > 1.55\,\text{GeV}$	0.025	0.020	0.025	0.020
5.7	$\cos\theta_{\gamma\mu} > -0.7$	0.56	0.67	0.58	0.62
5.7	Emiss$> 1.2\,\text{GeV}$	0.98	0.81	0.98	0.83
	Cumulative Lepton Tag	**0.043**	**0.032**	**0.045**	**0.036**
	TOTAL Efficiency	**3.58e-07**	**4.10e-06**	**8.35e-06**	**5.16e-06**
		2.30e-05		**7.78e-06**	
	Expected Events, $81.5\,\text{fb}^{-1}$	280	52	297	49
		332		346	

Table 5.6: Expectations for $B\overline{B}$ background in $81.5\,\text{fb}^{-1}$, according to $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays and other; the latter includes hadronic interactions faking photons).

This is a very fast selection because it uses quantities in a tag database that are (pre-) determined for each event at the time of event reconstruction and can be accessed quickly.

First a basic hadronic sample is created by requiring the event to be "flagged" by a multihadronic tag-bit (*BGFMultihadron*), which requires

- $R_2^{ch} < 0.98$, where R_2^{ch} is a second Fox–Wolfram moment [47][1] computed in the overall CMS using only charged tracks. (This cut mainly rejects residual Bhabhas.)

- at least three charged tracks.

The requirements on the charged tracks are subsequently tightened, requiring that at least three tracks satisfy the criteria in Table 5.7.

The main characteristic of $b \to s\gamma$ signal events being the high-energy photon, the primary background-rejection of the skim selection is reached with the requirement of a high-energy neutral cluster (a calorimeter cluster not associated with a charged track) in the EMC. A cut is made on the quantity "e1Mag" which is the highest CMS energy of all neutral EMC clusters: $1.5 < \text{e1Mag} < 3.5$ GeV. Additional loose cuts are applied on the overall Fox–Wolfram moment R_2^* ($R_2^* < 0.9$) and on the CMS momentum of the highest-momentum charged track ("p1Mag" < 2.5 GeV). Where

[1] R_2 is really the ratio of the second to the zeroth Fox–Wolfram moment: $R_2 \equiv H_2/H_0$ with $H_n = \sum_{i,j} \frac{|p_i| \cdot |p_j|}{E_{vis}} \cdot P_n(\cos\theta_{ij})$, where $P_n(x)$ are the Legendre polynomials, $p_{i,j}$ are the particle momenta, θ_{ij} is the opening angle between particles i and j, and E_{vis} is the total visible energy of the event. R2 will be largest for events in which all particle momenta are aligned tightly along one axis.

Variable	Cut
Minimum Transverse momentum	0.1 GeV
Maximum momentum	10 GeV
Minimum number of DCH hits	12
Minimum fit χ^2 probability	0
Maximum DOCA in XY plane	1.5 cm
Minimum Z DOCA	-10 cm
Maximum Z DOCA	10 cm

Table 5.7: Quality cuts used for good track selection. Transverse momentum is measured with respect to the beam direction (Z-axis in BaBar), DOCA stands for the distance of closest approach to the beam spot, *i.e.* the interaction point.

are R_2^* is defined analogous to R_2^{ch} but using all charged tracks *and* all photons. The star will be used throughout this thesis to denote that a quantity is calculated in the CMS frame. Finally, it is required that at least one high-energy photon candidate can be reconstructed in each event. Reconstruction of these candidates starts with neutral bumps in the EMC that cannot be matched with any charged track and that show the characteristics described in Table 5.8. These candidates are called *GoodPhotonsLoose*. They are then required to satisfy the high-energy requirement of 1.5 GeV $< E_\gamma^* <$ 3.5 GeV (in the CMS frame).

5.2 Photon Quality Cuts

Initial high-energy photon candidates are defined according to Table 5.8 with the additional energy requirement of 1.5 GeV $< E_\gamma^* <$ 3.5 GeV in the CMS frame. Photon quality cuts are applied to ensure that these photons were measured correctly

Variable	Cut
Minimum photon CMS energy	30 MeV
Maximum lateral moment	0.8

Table 5.8: Quality cuts used for good photon selection. The lateral moment [48] of a cluster is a measure for how evenly the energy is deposited radially around the cluster center. It is defined as the ratio of the sum of the energies of all but the two most energetic crystals, weighted by the square of their distance to the cluster center, and the sum over all crystal energies including the two most energetic ones, weighted in the same manner.

and are not likely to be bremsstrahlung, fakes from hadronic particles interacting in the EMC, or split-offs of charged particles (*i.e.* are not radiated off them):

- the photon must be measured within the fiducial volume of the EMC, resulting in a cut on angular acceptance of $-0.74 < \cos\theta_\gamma < 0.93$;
- the photon's EMC cluster must be of good quality, *i.e.* it must not contain dead or noisy crystals;
- the EMC cluster must contain at least 4 crystals;
- the second moment[2] of the cluster must be below 0.002.

The first three cuts are designed to remove incompletely measured or poorly reconstructed photons, *e.g.* with clusters containing dead or noisy channels. The fourth one is used to suppress background from $\pi^0 \to \gamma\gamma$ events in which the two photons are

[2]The second moment of a cluster is defined as: $S = \Sigma E_i(\Delta\Theta_i)^2/\Sigma E_i$ where $\Delta\Theta_i = \begin{pmatrix}\theta_{clust} - \theta_i \\ \phi_{clust} - \phi_i\end{pmatrix}$ and $\Theta_{clust} = \Sigma E_i\Theta_i/\Sigma E_i$ with a similar expression for ϕ_{clust}. All sums are over the crystals in the cluster.

merged to form a single cluster. This merged π^0 background to high-energy photons has been studied extensively in a BABAR analysis that searched for the rare decay $B^0 \to \gamma\gamma$ [59]. A maximum second moment cut is particularly effective at removing the latter kind of background, but also helps significantly with backgrounds from hadrons that are incorrectly reconstructed as photons in the EMC (fake photons). This is due to the fact that hadronic and electromagnetic showers have a different shape in the EMC. Electromagnetic showers are typically very narrow, whereas hadronics are much more spread out. For the cuts on the quantities introduced above, the same cut values as in reference [18] are used.

In this analysis, the second moment cut is reinforced with a new, tighter cut on the lateral moment of the EMC cluster, lateral moment < 0.45; this yields further suppression of hadronic backgrounds (especially important for $B\overline{B}$ backgrounds) while affecting the signal only minimally. (See Tables 5.3 through 5.6, where the hadronic component is reflected in the column called "other".) The lateral moment makes use of another shape difference between electromagnetic and hadronic showers, the latter leave much more irregular energy deposits that are fairly broadly spread giving rise to large lateral moments, whereas electromagnetic showers leave well focused symmetric deposits in the EMC. This was first used by the Argus collaboration [48]. The cut value has been chosen based on properties of simulated hadrons *vs.* photons identified from a sample of virtual Compton scattering events (see Section 6.4.1).

Finally, a cut on the minimal distance to the nearest EMC bump is helpful against all the backgrounds except (obviously) those due to the isolated photons from ISR. The cut value of 25 cm has been found from the optimization procedure described in Section 5.8.

Distributions for signal and background Monte Carlo events are shown in Figure 5.1 for the lateral moment with all prior cuts applied. Here only the distribution for hadronic fakes in the $B\overline{B}$ background is contrasted with the signal distribution, other parts of the background are not shown to facilitate the comparison. Figure 5.2 then shows the complete background distributions for continuum and $B\overline{B}$ Monte Carlo background events for the distance to the nearest bump. In this figure as for many of the following ones, "Continuum MC" refers to the sum of ISR and non-ISR contributions, "non-ISR" includes photons from π^0 or *eta* decays and hadronic fakes. Again, here and in the following all prior cuts are applied, with the cut order the same as in Tables 5.1 through 5.6.

For events with more than one photon of high energy, all of these cuts are applied to *each* high-energy photon candidate.

5.3 π^0 and η Vetoes

The decays $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ are a principle source of background high-energy photons. This background can be reduced by eliminating any photon candidate which is part of a reconstructed photon pair with invariant mass close to the π^0 or

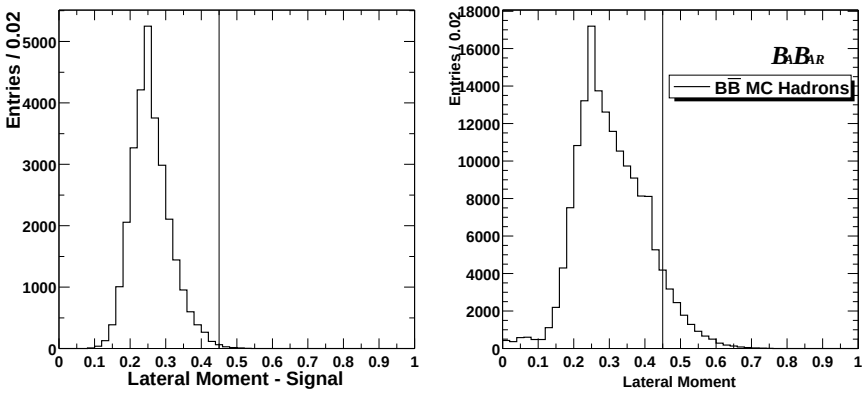

Figure 5.1: Distribution of the lateral moment of the EMC cluster associated with a high-energy photon candidate, from $SP4$ simulation. The complete signal but only the hadronic fake part of the $B\overline{B}$ background is shown. The vertical lines show the maximum value of the lateral moment accepted for a candidate.

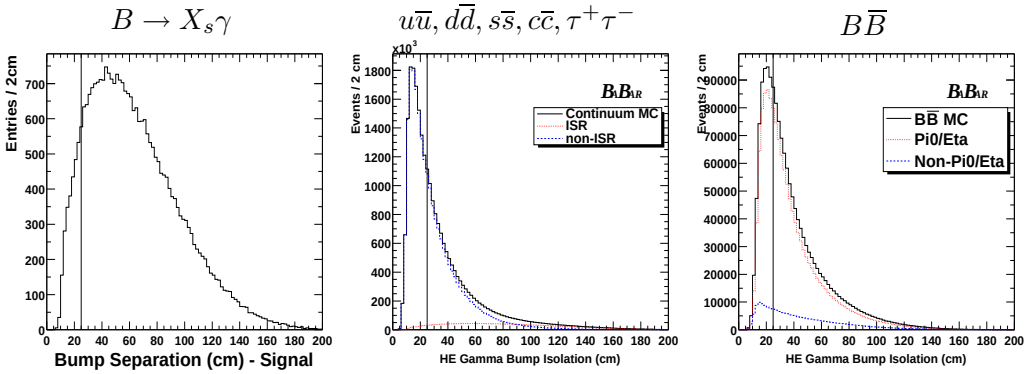

Figure 5.2: Distribution of the distance (cm) at the EMC between a high-energy photon candidate and the nearest other EMC bump, from $SP4$ simulation. The vertical lines show the minimum distance required for a candidate.

η mass. To test the π^0 or η hypothesis, high-energy photons are combined with any other photon in the event that fulfills a minimum energy requirement. The photon is vetoed as coming from $\pi^0(\eta)$ decays if the invariant mass of the combination lies within

the $\pi^0(\eta)$ mass window $115(508) < M_{\gamma\gamma} < 155(588)\,\mathrm{MeV}$. Figure 5.3 shows, for both signal and background, the energy spectrum of all photons in the event other than the high-energy candidate. Figures 5.4 and 5.5 show the invariant mass distributions for all possible combinations of such photons with the high-energy photon.

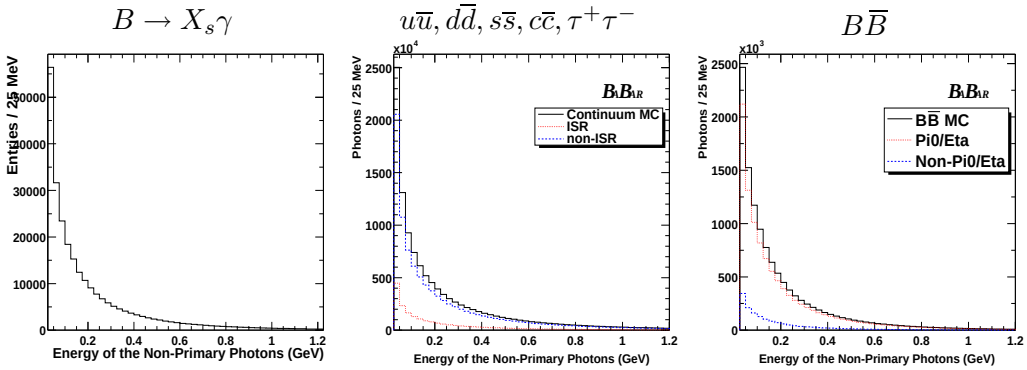

Figure 5.3: Lab-frame energy distribution for all photons in an event except the high-energy candidate, $SP4$ simulation. All of these are combined with the high-energy photon to check their consistency of coming from a π^0 or *eta* decay. The vertical lines show the two sets of minimum lab energies required of such photons to veto their combination with the high-energy photon as a π^0 or an η. The cut values are shown in detail in Table 5.9.

Most of the background stems from this source. Most problematic is the contribution from non-radiative $B\overline{B}$ events which dominates at lower photon energies but is less significant in the middle and high range of the considered spectrum. To best reduce this particular background, a lower cut is chosen on the minimum lab energy $E_{\gamma 2}$ of the second $\pi^0(\eta)$ daughter photon for $E^*_\gamma < 2.3$ GeV than above that value. The cut values on $E_{\gamma 2}$ for the different vetoes in both regions are then determined by the optimization procedure laid out in Section 5.8. Table 5.9 summarizes the E^*_γ-

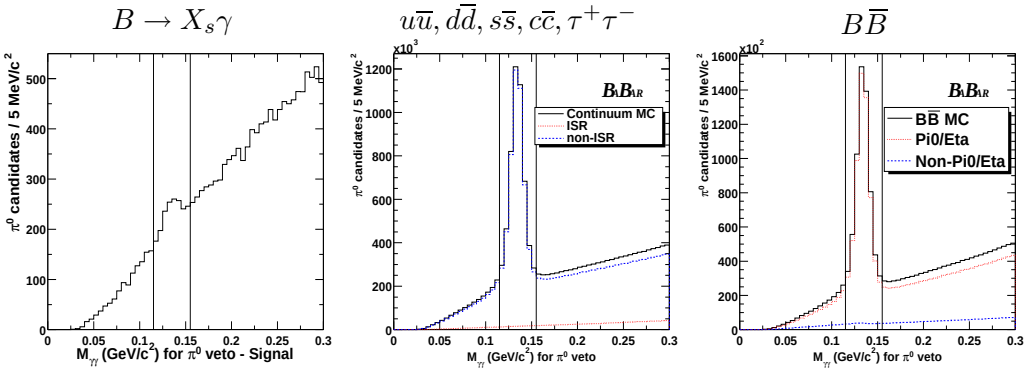

Figure 5.4: The $M_{\gamma\gamma}$ distribution for combinations of the high-energy photon with all other photons meeting the minimum energy requirement, $SP4$ simulation, π^0 region. The vertical lines show the mass cuts employed in the veto.

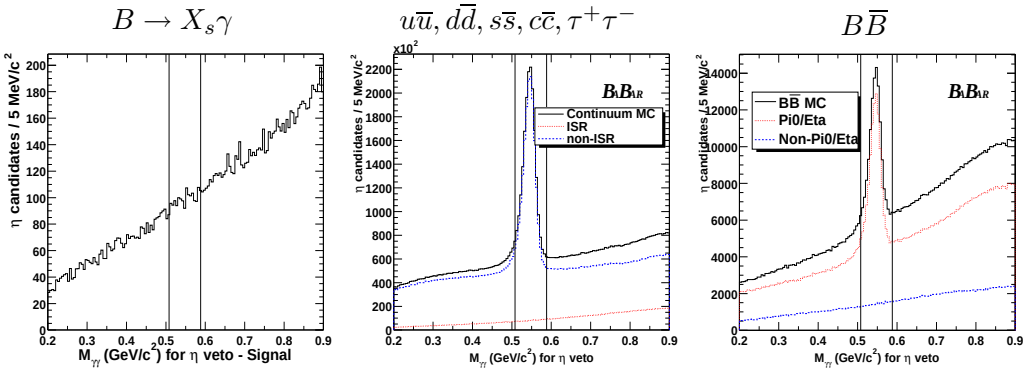

Figure 5.5: The $M_{\gamma\gamma}$ distribution for all combinations of the high-energy photon with all other photons meeting the minimum energy requirement, $SP4$ simulation, η region. The vertical lines show the mass cuts employed in the veto.

dependent cuts. This method improves the signal-over-background ratio quality factor $\mathrm{S}^2/(\mathrm{S}+\mathrm{B}')$ by about 9% over an E^*_γ-independent treatment. Here S stands for the number of signal events and B' is the statistically weighted number of background events.

	$< 2.3\,\mathrm{GeV}$	≥ 2.3 GeV
π^0 veto	$E_{\gamma 2} \geq 40$ MeV	$E_{\gamma 2} \geq 80$ MeV
η veto	$E_{\gamma 2} \geq 175$ MeV	$E_{\gamma 2} \geq 275$ MeV

Table 5.9: The minimum $E_{\gamma 2}$ requirements for reconstructing a π^0 or η candidate with the high-energy photon in the different regions of E^*_γ.

5.4 High-Energy Photon Energy Cut

Figure 5.6 shows the E^*_γ distribution for all high-energy photon candidates after the photon quality cuts and the $\pi^0(\eta)$ veto. The kinematic limit of the signal distribution is at $E^*_\gamma \approx 2.8$ GeV; the region above this, dominated by continuum background, can be used as a control sample. The $B\overline{B}$ background rises sharply below $E^*_\gamma \approx 2.2$ GeV.

The choice of the minimum E^*_γ cut is a trade-off between the systematic uncertainties arising from subtraction of the modeled $B\overline{B}$ background (which increases rapidly as E^*_γ decreases) and the introduction of model dependence in extracting the signal branching ratio (as described in Section 2.4). A cut of $E^*_\gamma > 2.0$ GeV is imposed; the large $B\overline{B}$ backgrounds at lower values and the associated increased systematic errors, as discussed in Section 6, do not allow any lower cutoff. The details and motivations of this choice will be provided in Section 9.1.

A cut is also made at the upper end, $E^*_\gamma < 2.7$ GeV, because of the very small amount of signal expected this close to the kinematic limit. The effect of only this upper cut on the signal efficiency at this stage of the selection is about 92% for

the $B \to K^*\gamma$ modes, but 99.7% for the KN480 samples. As already discussed in Section 2.4 in more detail, the K^* resonance makes up between roughly 10 and 27% of the total signal spectrum, depending on the model parameters chosen (see Table 2.2).

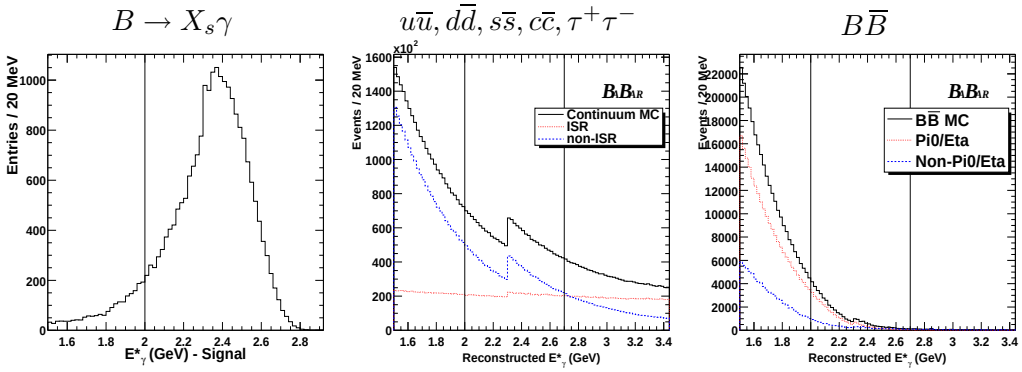

Figure 5.6: The E^*_γ distribution of the high-energy photon, $SP4$ simulation. The vertical lines show the cuts used at this stage of the event selection.

5.5 Event-Shape Cuts

Figure 5.7 shows the different shape characteristics for signal and continuum background events in the center-of-mass (CMS) frame, *i.e.*, the e^+e^- collision rest frame. The signal and $B\overline{B}$ backgrounds are isotropic in this frame. The ISR and inclusive $\pi^0(\eta)$ background have different shapes: they show the hadronization products from the original quark–antiquark pair bundled tightly together in two "jets" moving in opposite directions. In ISR events the jets are only back-to-back in the center-of-mass of the system without the photon; in the overall CMS, the photon can be seen as the third jet in a three-jet system. Photons from inclusive $\pi^0(\eta)$ decays are then

found within one of the jets. In ISR events the photon recoils against the $q\overline{q}$ system.

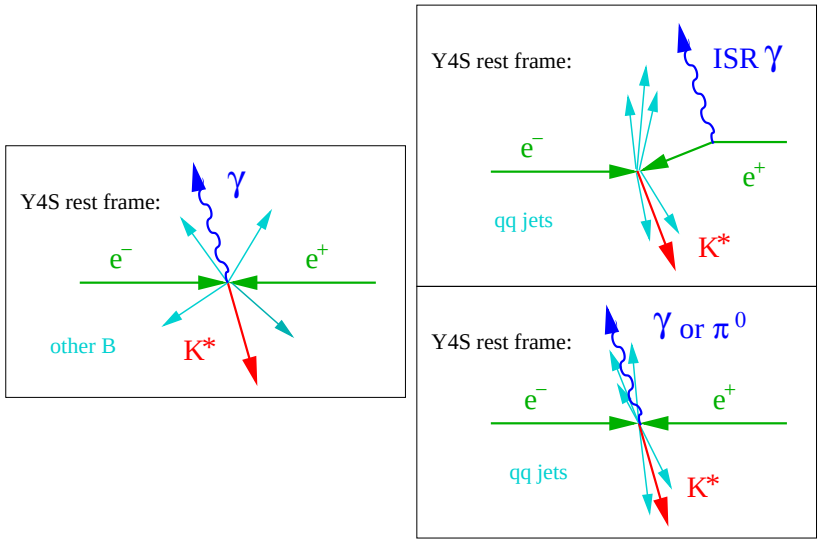

Figure 5.7: The Topology of Signal events (left) and the two types of continuum backgrounds

Use is made of a set of variables that are designed to exploit these different characteristics. It is useful to introduce a "recoil" (or "prime") frame, defined from the beam momenta and the reconstructed high-energy photon candidate by $\vec{p}(e^+) + \vec{p}(e^-) - \vec{p}(\gamma) = 0$. This is therefore the $q\overline{q}$ rest frame for an ISR event, in which the ISR background has a "two-jet" structure. The variables are computed from all reconstructed charged tracks fulfilling the criteria listed in Table 5.7 and all reconstructed photons satisfying a minimum energy requirement of 30 MeV and a maximum lateral moment cut of 0.8.

R_2^* The second Fox–Wolfram moment, computed in the overall center-of-mass frame using *all* reconstructed particles.

E^*_{f1}- E^*_{f9} Nine 10° cones are defined around the high-energy photon in the overall center-of-mass frame, from 0° to 90°. E^*_{fi} is the energy flow through the ith cone, *i.e.* the energy sum of all reconstructed particles (charged and neutral) whose momentum vectors lie within the ith cone. The sum excludes the high-energy photon.

E^*_{b1}- E^*_{b9} The same as for E^*_f except that the cones are in the backward direction, 180 to 90° from the high-energy photon vector.

$R_2^{'}/R_2^*$ Here $R_2^{'}$ is the second Fox–Wolfram moment computed in the recoil frame, which for ISR events is the $q\overline{q}$ rest frame; in this case the computation is done *omitting* the photon candidate. R_2^* is as defined above.

In the present analysis all of these shape variables, except for R_2^*, enter via a linear combination called a Fisher discriminant. This multivariate technique will be presented in Section 5.5.2. However, to illustrate the discriminating power of each type of variable separately, the results of a study using single-variable cuts will be shown first, in Section 5.5.1

5.5.1 Single-Variable Cut Method

This section describes a single-variable cut method which was studied in detail before the current multivariate tool was employed.[3] The requirement of the presence of a lepton in the event was used already in this study. All cut values were

[3]This analysis was presented at the ICHEP-2002 conference [9].

optimized according to the procedure outlined in Section 5.8. The results from this method are later (at the end of Section 5.5.2) used as a basis of comparison for any improvements from the multivariate technique.

For this initial approach, two broader energy cones were defined.

E_f^* The energy sum in the overall center-of-mass frame of all reconstructed particles (charged and neutral) whose momentum vectors lie within a 30° cone of the high-energy photon vector. The sum excludes the high-energy photon itself.

E_b^* The same as for E_f^* except that the cone is 140 to 180° from the high-energy photon vector.

Cuts are made in sequence on R_2^* (Fig. 5.8), E_f^* (Fig. 5.9), E_b^* (Fig. 5.10), and $R_2^{'}/R_2^*$ (Fig. 5.11). Note that the plots are for the initial version of the analysis; they do not reflect changes to photon quality cuts that were introduced later and are used in the current version of the analysis.

		$u\overline{u}, d\overline{d}, s\overline{s}$			$c\overline{c}$		
Sec.	Description	π^0/η	ISR	other	π^0/η	ISR	other
5.5.1	R2 < 0.45	0.59	0.92	0.60	0.60	0.95	0.61
5.5.1	$E_f < 1.1\,\mathrm{GeV}$	0.87	0.99	0.88	0.92	0.99	0.91
5.5.1	$1.6\,\mathrm{GeV} < E_b < 3.5\,\mathrm{GeV}$	0.54	0.30	0.49	0.58	0.32	0.48
5.5.1	$R2'/R2 < 1$	0.91	0.52	0.90	0.92	0.63	0.90
	Cumulative Event Shape	**0.25**	**0.14**	**0.23**	**0.29**	**0.19**	**0.24**

Table 5.10: Event shape efficiencies for continuum background in the ICHEP-2002 analysis, using $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays, photons from initial-state radiation, and other).

Sec.	Description	$B^0\bar{B}^0$		B^+B^-	
		π^0/η	other	π^0/η	other
5.5.1	R2 < 0.45	0.96	0.98	0.96	0.97
5.5.1	$E_f < 1.1\,\mathrm{GeV}$	0.98	0.95	0.97	0.94
5.5.1	$1.6\,\mathrm{GeV} < E_b < 3.5\,\mathrm{GeV}$	0.61	0.32	0.65	0.38
5.5.1	$R2'/R2 < 1$	0.91	0.90	0.90	0.88
	Cumulative Event Shape	**0.52**	**0.27**	**0.54**	**0.31**

Table 5.11: Event shape efficiencies for $B\overline{B}$ background in the ICHEP-2002 analysis, according to $SP4$ Monte Carlo. The background is divided into its principal components ($\pi^0/\eta \to \gamma\gamma$ decays and other; the latter includes hadronic interactions faking photons).

Sec.	Description	Neutral X_s			Charged X_s		
		KN465	KN480	KN495	KN465	KN480	KN495
5.5.1	R2 < 0.45	0.92	0.91	0.90	0.93	0.92	0.91
5.5.1	$E_f < 1.1\,\mathrm{GeV}$	0.98	0.98	0.98	0.98	0.98	0.98
5.5.1	$1.6\,\mathrm{GeV} < E_b < 3.5\,\mathrm{GeV}$	0.72	0.73	0.75	0.71	0.74	0.75
5.5.1	$R2'/R2 < 1$	0.93	0.93	0.95	0.92	0.92	0.93
	Cumulative Event Shape	**0.60**	**0.61**	**0.62**	**0.59**	**0.61**	**0.62**

Table 5.12: Event shape efficiencies for expected signal in the ICHEP-2002 analysis, according to $SP4$ Monte Carlo. For all Kagan-Neubert models (labelled by 100 times m_b in GeV), a minimum m_{X_s} cutoff of 1.1 GeV is used.

Sec.	Description	K^{*0}	K^{*+}
5.5.1	R2 < 0.45	0.80	0.83
5.5.1	$E_f < 1.1\,\mathrm{GeV}$	0.99	0.99
5.5.1	$1.6\,\mathrm{GeV} < E_b < 3.5\,\mathrm{GeV}$	0.76	0.69
5.5.1	$R2'/R2 < 1$	0.97	0.96
	Cumulative Event Shape	**0.59**	**0.55**

Table 5.13: Event shape efficiencies for expected K^* signal component in the ICHEP-2002 analysis, according to $SP4$ Monte Carlo.

Table 5.10 confirms the expectation (from event shape considerations, see Fig. 5.7) that the R_2^* cut is particularly effective against non-ISR components of the continuum background, while the R_2'/R_2^* cut is primarily effective against ISR. The E_b^* cut helps with *all* types of background, including (unlike the other shape cuts) $B\overline{B}$ (Table 5.11). As seen in Figure 5.10, this occurs mostly for the hadronic fake component of the $B\overline{B}$ background, but even the $\pi^0(\eta)$ component tends to have slightly lower E_b^* than the signal distribution. Tables 5.12 and 5.13 show the corresponding signal efficiencies.

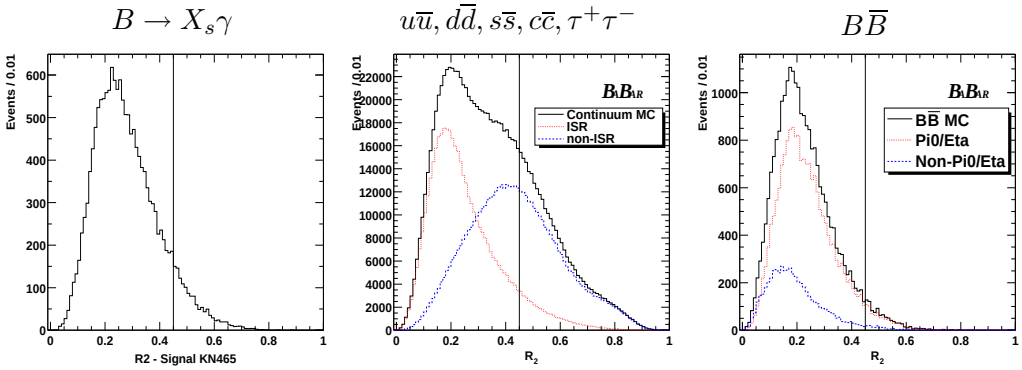

Figure 5.8: The ICHEP-2002 R_2^* distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted in this single-variable cut study.

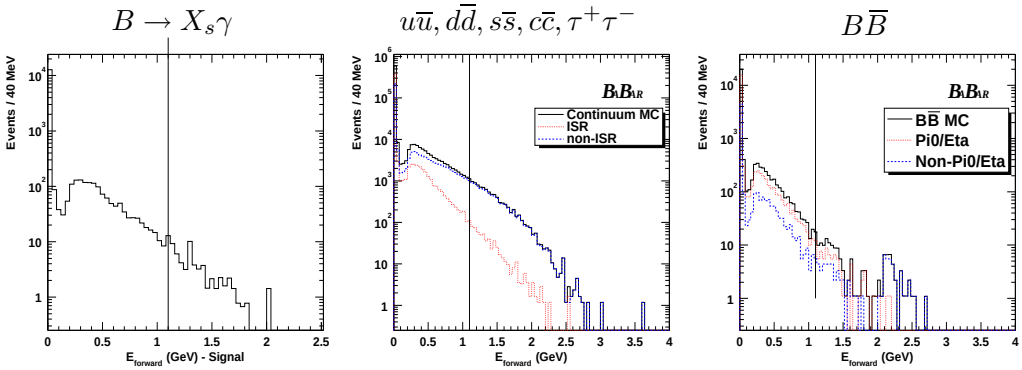

Figure 5.9: The ICHEP-2002 E_f^* distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted by the E_f^* cut in this single-variable cut study.

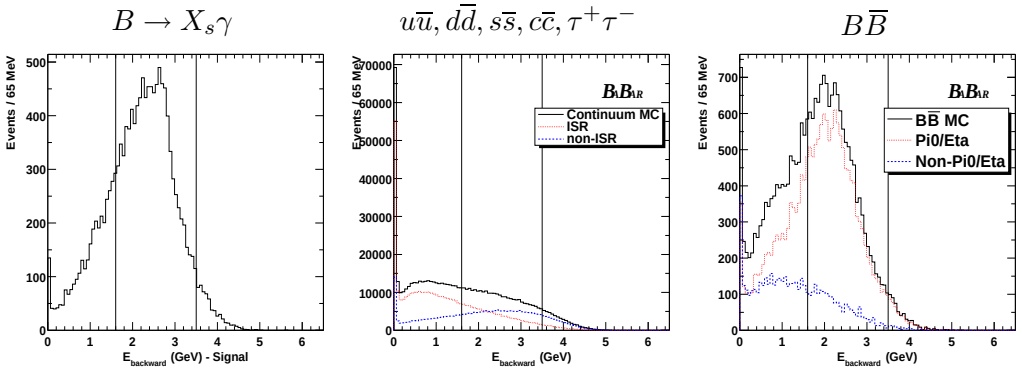

Figure 5.10: The ICHEP-2002 E_b^* distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted by the E_b^* cut in this single-variable cut study.

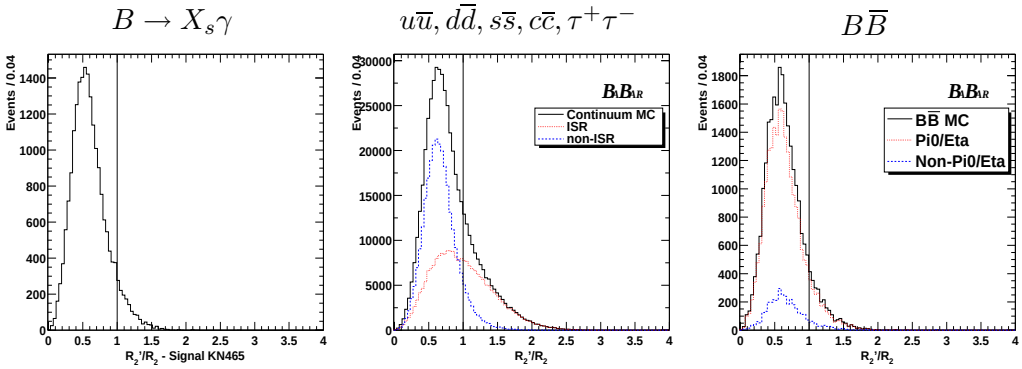

Figure 5.11: The ICHEP-2002 R_2'/R_2^* distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted in this single-variable cut study.

5.5.2 Fisher Discriminant

Of the described event shape variables none has a particularly strong discrimination power, but all have some, especially between signal and continuum background events. Hence it is advantageous to combine them (or subsets of them) into a single discriminant variable according to an algorithm proposed by Fisher in 1936 [49, 50].[4]

Linear discriminant methods define a discriminating function for signal and background to be a linear combination of the input variables chosen to characterize an event. The Fisher discriminant algorithm determines the coefficients of the combination in a way that optimizes the discrimination between signal and background, or, in other words, the classification of events into either category.

Consider p input variables. It is easiest to visualize the process by thinking about the p-dimensional space that these variables span. Each event occupies one point in the p-dimensional variable space. If the input variables are chosen appropriately, signal and backgrounds will cluster in separate regions in this space. Projecting an event onto any line in this space then yields a linear combination of the p variables:

$$FD = -O + \sum_{i=1}^{p} C_i \cdot V_i.$$

With V_i the variables, C_i their coefficients and O an offset that is discussed below. This linear combination can be considered as a new variable, the Fisher discriminant

[4]Sir Roland A. Fisher (1890-1962) not only had a substantial part in laying the foundation for what today is recognized as the academic discipline of statistics but also developed a large part of the theoretical framework for population and quantitative genetics (called eugenics at the time). His influence in bringing the mathematical and quantifiable approach to biology was great.

variable, FD. The challenge is then to find the line that provides maximum separation between signal and background events.

Fisher's algorithm also takes into account the scatter within the signal and background category by maximizing

$$\lambda = \frac{(\bar{u}_{signal} - \bar{u}_{background})^2}{\sigma^2_{signal} + \sigma^2_{background}}$$

where $\bar{u}$ denotes the mean of the distributions and σ their widths.

Various combinations of the introduced shape variables have been studied as input to the Fisher linear discriminant method, single variable cuts then being used for the shape variables *not* included in the Fisher discriminant. The best performance is found for a Fisher discriminant incorporating $R_2^{'}/R_2^*$ and all 18 E_f^* and E_b^* cone variables (and combined with a single variable cut on R_2^*).[5] Since the shape variables primarily discriminate between signal and continuum background, only samples for those two categories are employed in the training; the algorithm requires these samples to be of equal size. Samples of about 3300 signal and continuum $SP4$ events which already satisfy the photon quality cuts, the E_γ^* cut, and the requirement of a high-momentum lepton in the event, and which contain only one high-energy photon candidate are passed to the training algorithm. This comprises half the available signal MC statistics. The other half may be used in an independent evaluation sample. The signal sample mixes KN480 and $B \to K^*\gamma$ Monte Carlo events, with a $B \to K^*\gamma$

[5]The optimal coefficients of the linear combination are computed using the BABAR *Cornelius++* package [51] in a ROOT implementation [52] (Fisher "training").

fraction of 0.161 at the generator level (see Table 2.2). The background sample mixes uds, $c\bar{c}$, and $\tau^+\tau^-$ Monte Carlo events with equal luminosities. The training algorithm determines the linear coefficients for the shape variable combination; it also yields an offset that results in the final Fisher Discriminant variable being centered at zero for the full input data, with negative values for more background-like and positive ones for more signal-like events. Therefore the signal can be selected from the background by requiring a minimum value for the Fisher discriminant. All input variables, their Fisher parameters, and the offset are summarized in Table 5.14.

To judge the improvement that this linear discriminant method yields over the single variable cut method, a comparison was made as follows. The baseline is the version of this analysis that is based on single variable cuts (as presented in Section 5.5.1) with the addition of the lateral moment cut and the energy-dependent π^0 and η vetoes, and with the E^*_γ range expanded to 2.0 to 2.7 GeV. For the multivariate method the three cuts on E^*_f, E^*_b and $R_2^{'}/R_2^*$ were replaced by one cut on the Fisher discriminant. The value of this cut is determined by a preliminary optimization of the Fisher variable (before full optimization of the other cuts). The new method shows increases of the signal over background ratio by $\approx 14\%$ and of the quality factor $\mathrm{S}^2/(\mathrm{S}+\mathrm{B}')$ by $\approx 18\%$.

The full optimization of all cut variables determines the final cut values on the R_2^* ($R_2^* < 0.55$) and Fisher discriminant distribution (Fisher discriminant > 0.575) as shown in Figures 5.12 and 5.13.

Input variables	Coefficient	Coeff$*\Delta$(mean)
R_2'/R_2^*	-0.8627	0.731
E_{f1}^* (0-10°)	-0.8715	0.097
E_{f2}^* (10-20°)	-0.7637	0.094
E_{f3}^* (20-30°)	-0.4282	0.017
E_{f4}^* (30-40°)	-0.2881	-0.013
E_{f5}^* (40-50°)	-0.1552	-0.012
E_{f6}^* (50-60°)	-0.1226	-0.014
E_{f7}^* (60-70°)	-0.0567	-0.005
E_{f8}^* (70-80°)	-0.0016	-0.000
E_{f9}^* (80-90°)	0.0343	0.001
E_{b1}^* (170-180°)	0.1380	0.035
E_{b2}^* (160-170°)	0.0756	0.030
E_{b3}^* (150-160°)	-0.0664	-0.010
E_{b4}^* (140-150°)	-0.1387	0.003
E_{b5}^* (130-140°)	-0.2529	0.058
E_{b6}^* (120-130°)	-0.1587	0.040
E_{b7}^* (110-120°)	-0.0512	0.010
E_{b8}^* (100-110°)	0.0319	-0.005
E_{b9}^* (90-100°)	0.0344	-0.003
Offset	-1.5884	

Table 5.14: The Fisher Discriminant variables and their coefficients. The numbers are obtained from Fisher training using $SP4$ samples of signal and continuum (defined in the text) The "Offset" is subtracted from the resulting linear combination so as to center at 0 the distribution for the full input data set. The last column crudely indicates the separation "power" of the input variables in the combination. It is the product of the coefficient of a variable and the difference of the mean of the variable for signal events and the mean for background events. This does not account for correlations or the shapes of the distributions.

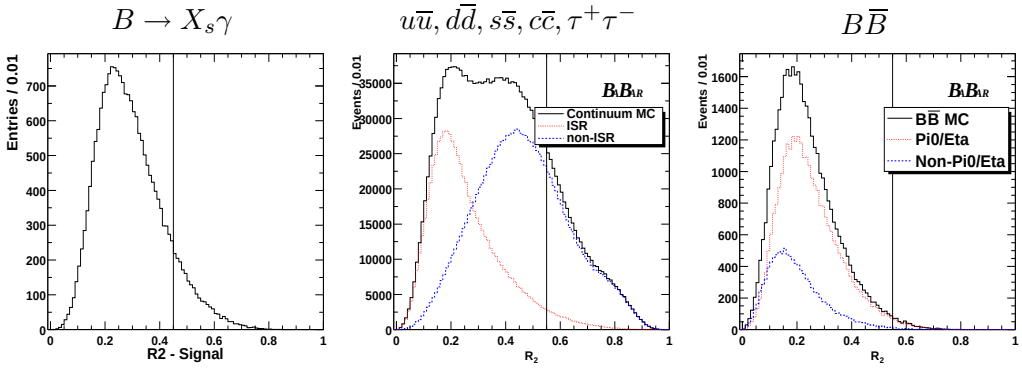

Figure 5.12: The current R_2^* distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted.

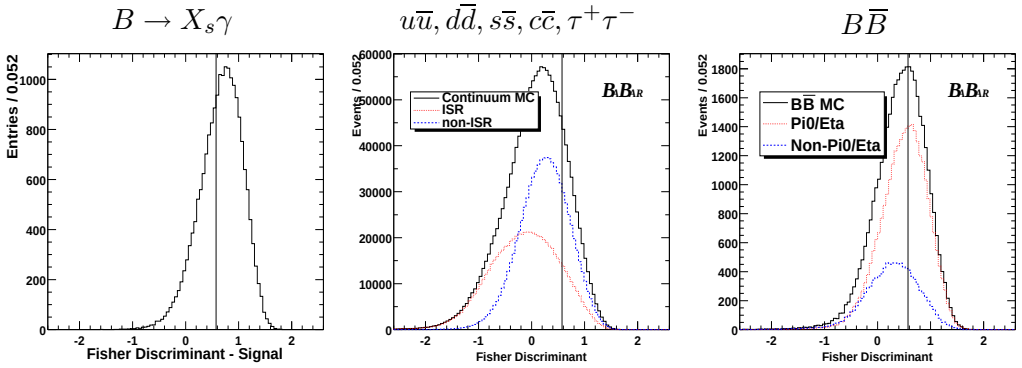

Figure 5.13: The Fisher discriminant distribution from $SP4$ simulation after previous cuts. The vertical line shows the maximum value accepted.

5.6 Removal of Multiple-Candidate Events

At this stage of the analysis, there are few (less than 0.1%) signal Monte Carlo events remaining with multiple photon candidates. Those events are simply discarded. Table 5.15 tracks how the fraction of such events changes through the cut categories. It can be seen that most of the apparent ambiguities that pass the

E^*_γ 1.5 GeV skim cut are cleared up by the photon quality cuts, leaving only 0.5% actual ambiguities. The minimum E^*_γ requirement of 2.0 GeV (Section 5.4) amounts to selecting the higher-energy photon in most of those 0.5% of events. Monte Carlo statistics are inadequate to tell if event shape cuts have any further effect.

Stage in succesion of cuts	Events	Multi-candidate fraction
After skim cuts	32442	0.025
After photon quality cuts	24193	0.005
After 2.0 GeV $< E^*_\gamma <$ 2.7 GeV	21914	< 0.001
After event shape cuts	13239	$\ll 0.001$

Table 5.15: Fractions of signal events with multiple high-energy photon candidates, at various stages of selection cuts, from $SP4$ Monte Carlo. A sum of $K^*\gamma$ and $X_s\gamma$ is used, the latter generated from the KN480 model.

5.7 Lepton Tags

Approximately 20% of B meson decays involve a lepton (electron or muon) in the final state ($\mathcal{B}(B \to e + anything) = 0.1041 \pm 0.0029, \mathcal{B}(B \to \mu + anything) = 0.103 \pm 0.005$ [19]). Leptons do also arise from the fragmentation of $u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}$, but are significantly less abundant at momenta greater than 1 GeV. Therefore a high-momentum lepton from the recoiling B meson can be required in $B \to X_s\gamma$ events to suppress continuum background. Since this selection makes no requirement on the $B \to X_s\gamma$ decay products, it adds essentially no model-dependence to the selection.

5.7.1 Lepton Identification Efficiency and Purity

The leptons are selected from tracks satisfying the good tracks requirements of table 5.7, but are required to satisfy tighter cuts on the "distance of closest approach" to the interaction point: DOCA(xy) $< 0.4\,$cm, and DOCA(z) $< 4.0\,$cm. In this sample electrons and muons are identified as follows:

The electron selector [60] applies the following very tight criteria:

- a cut on $540 < dE/dx < 860$ as measured in the EMC;
- at least three crystals in the EMC cluster;
- energy over momentum ratio of $0.89 < E/p < 1.2$ for the track;
- lateral moment below 0.6;
- Zernike[6] moment 42 [53] of the EMC cluster of -10 $< A_{42} <$ 0.11;
- A transverse momentum ($p_\perp$) dependent cut on the azimuthal position of the EMC cluster centroid relative to the azimuth angle of the track's entrance point into the EMC, $\Delta\phi$[7]: $\Delta\phi > -0.02$ and $\Delta\phi < (0.03 + 0.1 \cdot e^{-1.9 \cdot p_\perp})$.

[6]The Zernike moment A_{42} is a measure for how asymmetric the energy of an EMC cluster is distributed about its maximum. It is one of the terms of the expansion of the lateral shower shape in terms of Zernike moments:
$A_{nm} = \sum_{r_i \le R_0}^{n} \frac{E_i}{E} \cdot f_{nm}(\frac{r_i}{R_0}) \cdot e^{-im\phi_i}$, $R_0 = 15\,$cm This is defined for all crystals i that contribute to the cluster and are less than 15cm from the cluster centroid, with E_i being the energy of the single crystals, E the total energy of the cluster, and
$f_{nm}(\rho_i \equiv \frac{r_i}{R_0}) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho_i^{n-2s}}{s!((n+m)/2-s)!((n-m)/2-s)!}$, where $n, m \ge 0$ integers, $n-m$ even and $m \le n$.

[7]$\Delta\phi = q \times (\phi_{EMC} - \phi_{cluster})$. Where q is the charge of the track, $\phi_{cluster}$ is the azimuth angle of the cluster centroid, and ϕ_{EMC} is the azimuth angle of the track's entrance point into the EMC surface. This cut exploits longitudinal differences in shower shape: electromagnetic clusters shower earlier than do hadronic ones. Because the EMC lies within a magnetic field, a longitudinal difference results in an azimuthal difference as well.

- the Cherenkov angle measured for this track in the DIRC is consistent (within 3σ) with the electron hypothesis. This requirement is applied only if the there are at least ten DIRC photons measured for this track.

The Muon Selector [61] requires a track to satisfy the following tight muon criteria:

- the energy deposit in the EMC is in the range 50 MeV $< E_{cal} <$ 400 MeV;
- a minimum of three layers with hit strips in the IFR;
- a maximum difference on the number of interaction lengths traversed in the entire detector as theoretically expected using a muon hypothesis for the track (λ_{expec})– and as obtained from the actual measurement (λ): $\Delta\lambda = \lambda_{expec} - \lambda < 1$;
- a minimum measured number of interaction lengths for the track of $\lambda > 2.2$;
- the χ^2 per degree of freedom of the IFR hit strips in the track extrapolation $\chi^2_{trk} < 5$;
- the χ^2 per degree of freedom of the IFR hit strips in a fit of the cluster $\chi^2_{fit} < 3$;
- $T_c > 0.3$ where T_c is the ratio of the number of layers with hit strips over the difference of the number of the last and the number of the first layer with hits, modified to take into account that the innermost layer is numbered -1;
- maximum average multiplicity of hit strips per layer $\bar{m} < 8$;
- maximum standard deviation of the latter: $\sigma_m < 4$.

On these tight muon candidates the selector then imposes a veto against kaons. If a candidate passes the loose kaon requirements [62] listed in the following, it is rejected as a muon candidate (this is effectively a tight veto against kaons):

- the maximum likelihood for a kaon is greater than that for a pion;
- the likelihood for a kaon is greater than that for a proton.

These likelihoods are determined using information from the DIRC only, or from the DIRC, SVT, and DCH combined. For all track momenta of interest in this analysis, the DIRC likelihood information is used (if there is at least one expected photon for an electron with the measured momentum). For track momenta greater 1.5 GeV SVT and DCH likelihood information is added (if at least 3 SVT dE/dx sample hits and at least 10 DCH dE/dx hits are available for the track). The total likelihood is the product of the subsystem likelihoods.

Figures 5.14 and 5.15 show the efficiency and purity for electrons and muons using these selectors. Electrons can be selected with high efficiency (90%) and purity with a momentum above approximately 0.75 GeV. Muons can be identified with reasonable efficiency (60%) above a momentum of approximately 1 GeV. However the pion fake rate is an order of magnitude higher for muons than for electrons. The kaon fake rate is negligible in both cases for momenta above 1.0 GeV.

A procedural problem in including lepton tagging is that actual particle-ID criteria are not modeled well enough in the Monte Carlo simulation. As a remedy

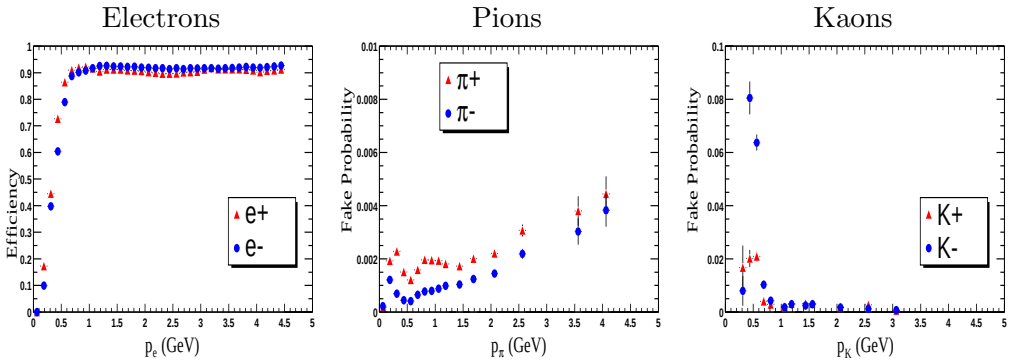

Figure 5.14: The electron efficiency and fake probability for pions and kaons.

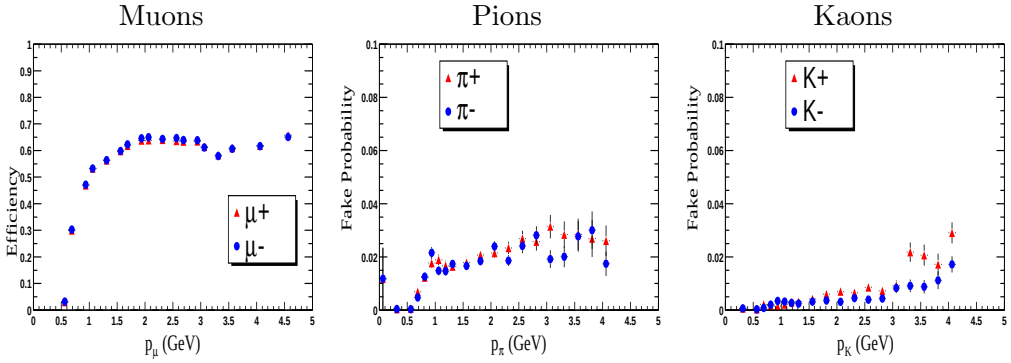

Figure 5.15: The muon efficiency and fake probability for pions and kaons

the Monte Carlo efficiency and purity distributions are replaced with those measured in data, shown in figures 5.14 and 5.15. The efficiency and purity of the selector is measured from a number of independent data control samples and used to compile a look-up table (LUT) binned in particle species, momentum, and polar angle. For example, the probability that an electron will pass the electron selector has been measured from a sample of radiative Bhabhas while the probability of a kaon passing

the electron selector and thus faking an electron has been measured from a sample of kaons from $D \to K\pi$ decays. Similarly the probability of a pion passing the electron selector has been measured using pions from a sample of 1-3 tau decays.

In the Monte-Carlo all reconstructed tracks are matched to the underlying, generated (truth) particle (*truth-matching*) and subsequently for each track the electron selector LUT probability for that particle species is compared with a random number uniformly distributed between 0 and 1. If the random number is less than the probability in the LUT, the track is identified as an electron.

The validity of this procedure is checked by comparison with a different analysis that also uses a lepton tag, the determination of sin2β [63]. The sin2β lepton tag requires leptons to pass the same selectors as above, with kinematic requirements $p_e^* > 1.0\,\mathrm{GeV}$, $p_\mu^* > 1.1\,\mathrm{GeV}$. In the sin2$\beta$ analysis the tag efficiency is determined directly from the $B\overline{B}$ data sample. This fully reconstructs one B meson (which determines its flavor), the flavor of the second B is then deduced from a lepton in the rest of the event. This uses the fact that primary leptons from B decays will have the same charge as the original b-quark, thereby determining whether the meson is a B or a $\overline{B}$. The efficiency of the lepton identification and tag is then determined from the number of events with correct, incorrect and no flavor tags from leptons. When this directly measured efficiency is compared to the uncorrected Monte Carlo simulation (no LUT applied) it is found to differ by a factor of 0.95 ± 0.02. For the present analysis and its signal sample it was verified that, when using the sin2β lepton cuts, the tag efficiency

measured in the uncorrected Monte Carlo simulation is the same as found in the sin2β analysis. The application of the LUT corrections should then result in a correction factor equal to 0.95 ± 0.02. The correction factor found is 0.913 ± 0.012, in agreement with this expectation.

Lepton Tag Cuts and Distributions

Figure 5.16 shows the CMS momentum distributions for the highest-energy electron (if any) satisfying the particle-ID requirements noted above, while Figure 5.17 shows the corresponding distribution for muons. Cuts on the lepton momentum are optimized using the procedure described in section 5.8 resulting in the requirements: $p_e^* > 1.25\,\mathrm{GeV}$ and $p_\mu^* > 1.5\,\mathrm{GeV}$. If both an electron and a muon satisfy their respective cuts, the electron is given preference, because its identification is more reliable. Figure 5.17 suggests that there might be some further continuum discrimination from a maximum p_μ^* cut at around 2.2 GeV; yet after a cut on $\cos\theta_{\gamma\mu}$ (see below) and a cut on missing-energy in the event (see section 5.7.2), no events survive in that region.

In addition to the lepton tag requirement (on lepton presence and momentum), the tagging lepton candidate can be used to construct a new shape variable, the angle between the electron (muon) and the photon in the center-of-mass frame. Figures 5.18 and 5.19 show that the cosine of this angle gives additional discrimination. To further suppress continuum backgrounds, events are required to satisfy $\cos\theta_{\gamma e} > -0.7$ and $\cos\theta_{\gamma\mu} > -0.7$ (chosen according to the optimization procedure

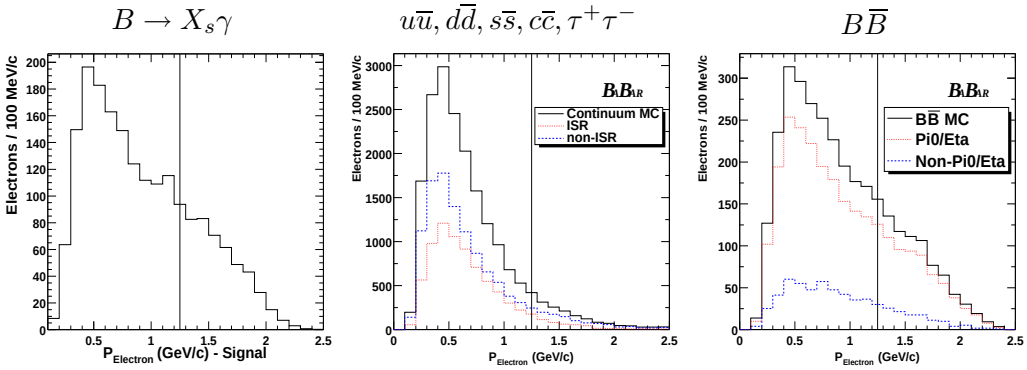

Figure 5.16: The p_e^* distribution from $SP4$ simulation, for the highest-momentum particle satisfying the Very Tight Electron PID. The vertical line shows the minimum tagging value.

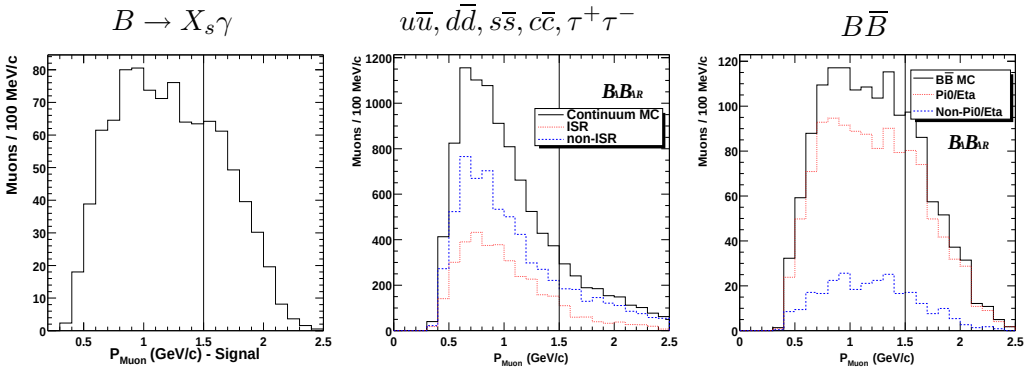

Figure 5.17: The p_μ^* distribution from $SP4$ simulation, for the highest-momentum particle satisfying the Tight Muon PID. The vertical line shows the minimum tagging value.

discussed below).

Looking at these distributions for *signal* events, the photon and any real lepton derive from different B's, so the distribution of $\cos\theta_{\gamma e}$ ($\cos\theta_{\gamma\mu}$) is expected to be isotropic. Especially for muons, there is also a surplus at ($\cos\theta_{\gamma\mu} = -1$) due to pions from the X_s decay faking a lepton. This expected distribution may be seen in

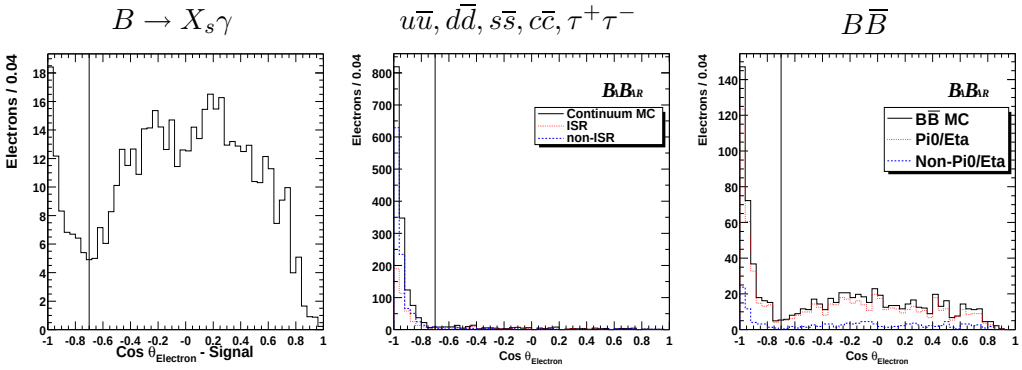

Figure 5.18: The $\cos\theta_{\gamma e}$ distribution from $SP4$ simulation for electron tags. The vertical line shows the minimum accepted value.

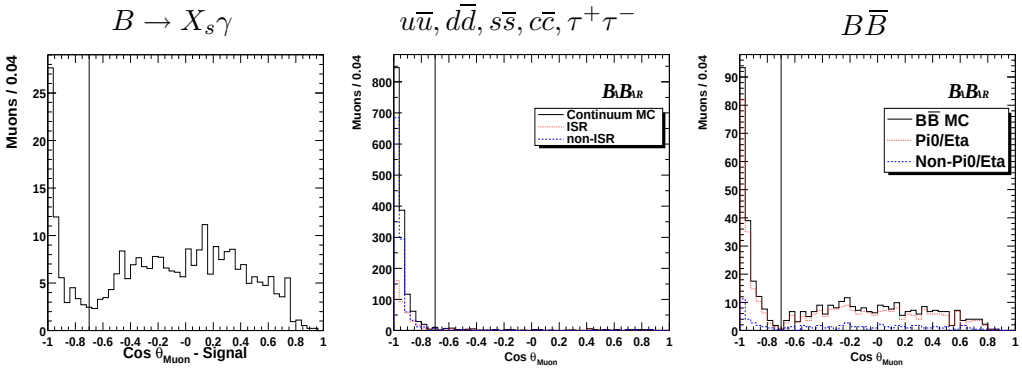

Figure 5.19: The $\cos\theta_{\gamma\mu}$ distribution from $SP4$ simulation for muon tags. The vertical line shows the minimum accepted value.

the black histogram of figure 5.20, for which the bump-isolation and event-shape cuts have been removed. The subsequent isolation cut on the high-energy photon depletes the region of $\cos\theta_{\gamma e}$ ($\cos\theta_{\gamma\mu}$) near +1, as it removes events with the tagging lepton close to the photon. Furthermore, the shape cuts tend to remove events for which the tag direction is relatively close to a line along or opposite to the photon direction.

This effect is more pronounced for the lower multiplicity final states such as $B \to K^*\gamma$ which tend to be less isotropic, so that an aligned tag can make them look particularly "jet-like".

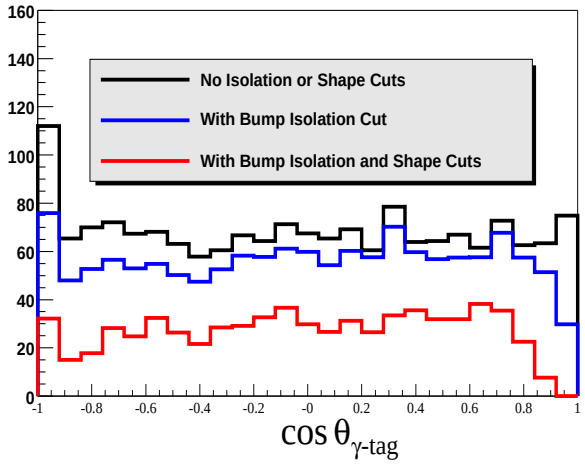

Figure 5.20: Signal $\cos\theta_{\gamma-tag}$ before and after the bump isolation and shape cuts (R_2^*,E_f^*,E_b^*, and $R_2^{'}/R_2^*$) according to $SP4$ simulation. Electron and muon tags have been combined.

The $\cos\theta_{\gamma e}$ and $\cos\theta_{\gamma\mu}$ distributions for $B\overline{B}$ *background* show two components. A largely-flat component is expected when the tagging lepton and the photon candidate come from *different* (hence largely uncorrelated) B's; while the backward peak arises when they come from the *same* B, given that the photon takes up close to half of the available energy.

Finally, the *continuum background* is peaking strongly toward $\cos\theta_{\gamma e} = -1$, $\cos\theta_{\gamma\mu} = -1$. In the CMS frame this would be expected to peak toward $\cos\theta_{\gamma e} = \pm 1$, $\cos\theta_{\gamma\mu} = \pm 1$ as a consequence of the "jet-like" topology. Phase space

considerations imply that the probability of two high-momentum particles occurring in the same jet ($\cos\theta_{\gamma e} = +1,\ \cos\theta_{\gamma\mu} = +1$) is less than the probability of them occurring in different jets ($\cos\theta_{\gamma e} = -1,\ (\cos\theta_{\gamma\mu}) = -1$). In addition, the photon isolation (bump-distance) cut strongly suppresses the forward peak for electrons (less so for muons since they are minimum ionizing and produce low-energy bumps).

5.7.2 Real *vs.* Fake Tags

Matching a reconstructed (and identified) particle to its true identity from the generator (or truth) information can be used to break down Monte Carlo distributions for tag candidates according to whether the tag has been properly identified ("real" lepton) or not (hadronic "fake"). Figures 5.21 and 5.22 show the tag-candidate composition before tagging cuts on momentum and the angle to the high-energy photon, while figures 5.23 and 5.24 show the tag composition after those tagging cuts. Even in the selected higher-momentum region, it is clear that the cuts on $\cos\theta_{\gamma e}$ and $\cos\theta_{\gamma\mu}$ reduce the fraction of fakes, at least for signal and $B\overline{B}$. The electrons are considerably purer than the muons, as expected from Figs. 5.14 and 5.15. The principle contamination is from pions faking muons.

Missing Energy Cut

The source of real lepton tags in signal events is the semi-leptonic decay $B \to Xl\nu$ while the continuum background contains a significant fraction of fake lep-

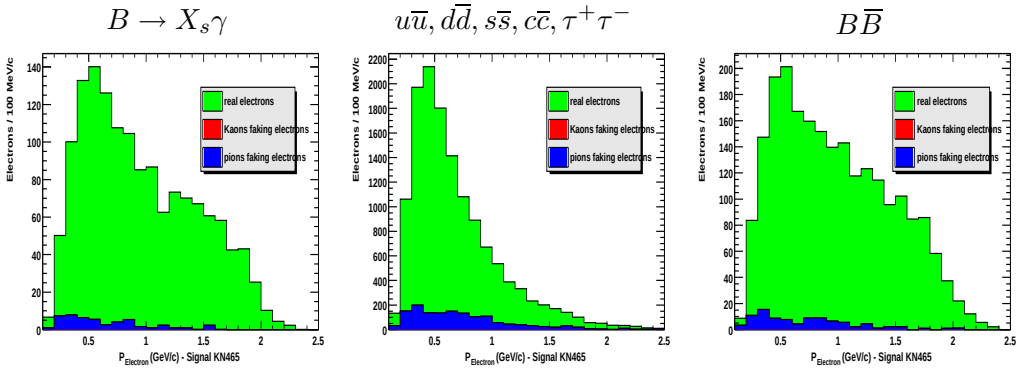

Figure 5.21: The tag composition for candidate electron tags from $SP4$ simulation *vs.* electron momentum (p_e^*) before tagging cuts.

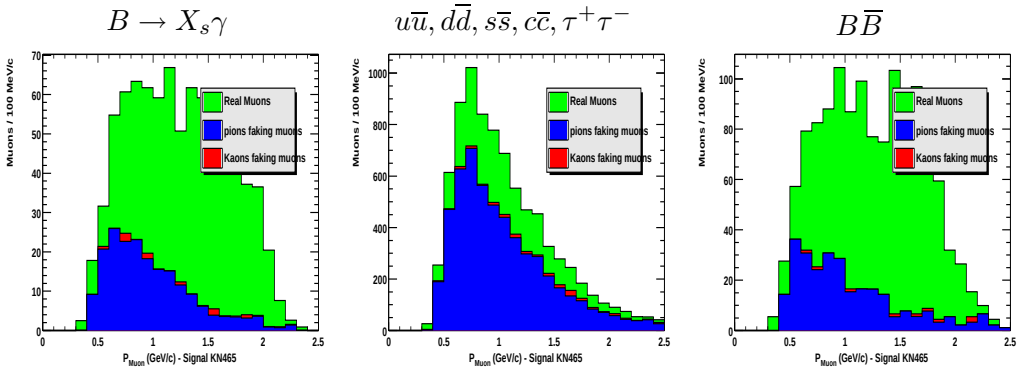

Figure 5.22: The tag composition for candidate muon tags from $SP4$ simulation *vs.* muon momentum (p_μ^*) before tagging cuts.

tons. Thus the signal tags will be associated with missing energy from the undetected neutrino which will be absent in the continuum fakes. Figure 5.25 shows the missing energy distribution after all prior cuts. Additional signal–continuum separation can be gained with a cut of $E_{miss} > 0.8$ GeV (again from optimization).

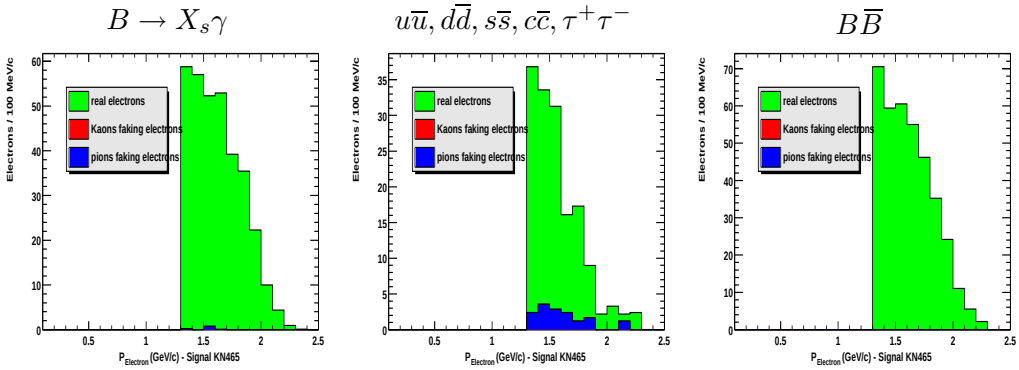

Figure 5.23: The tag composition for electron tags from $SP4$ simulation *vs.* electron momentum (p_e^*) after tagging p_e^* and $\cos\theta_{\gamma e}$ cuts.

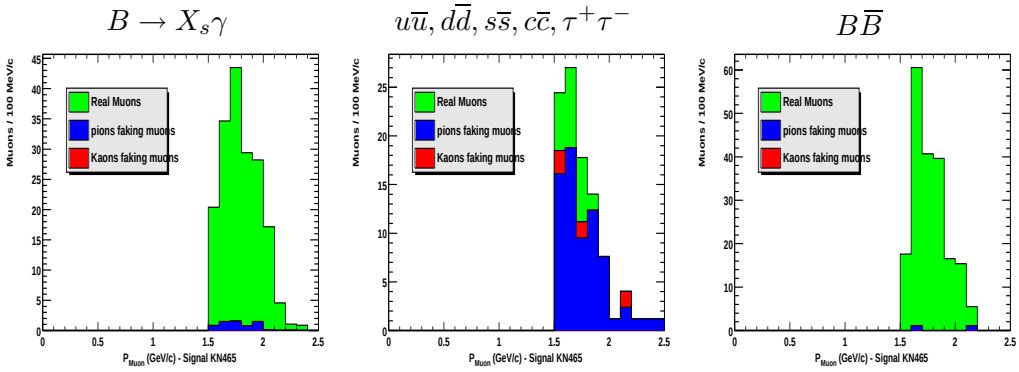

Figure 5.24: The tag composition for muon tags from $SP4$ simulation *vs.* muon momentum (p_μ^*) after tagging p_μ^* and $\cos\theta_{\gamma\mu}$ cuts.

5.8 Cut Optimization

The cut values used in this analysis were largely selected by an optimization procedure. Both the variables to be cut on and the numerical values of these cuts are selected using Monte Carlo samples for the signal and *both* continuum and $B\overline{B}$ background.

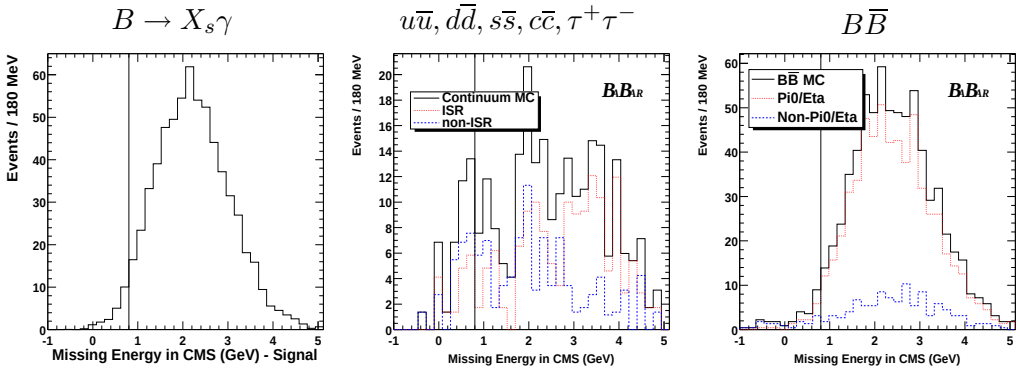

Figure 5.25: The center-of-mass frame missing energy distribution (in GeV) from $SP4$ simulation, for signal and background events passing electron or muon tagging cuts. The vertical lines show the minimum accepted value.

Traditionally the cut values are optimized by maximizing the resulting value of $S^2/(S+B)$ for the final selection. However, while using $S^2/(S+B)$ as a criterion is relevant for analyses in which backgrounds can be separated from signal by a fit of signal and background spectra (*e.g.* , reconstructed B mass spectra in the exclusive and semi-exclusive Radiative Penguin analyses),this is not correct for an analysis which has no such quantity to fit, and which instead relies on subtracting yields obtained in a smaller data set (*i.e.*, the off-resonance data). The real goal is of course to maximize $S/\Delta S$. Appendix C addresses this point (for statistical uncertainties only). It shows how the statistical precision depends upon the fraction f of on-resonance running. In particular, equation C.5 shows that B in $S^2/(S+B)$ needs to be replaced by

$$B' \equiv B + C/(1-f) \quad , \tag{5.1}$$

where B is now just the $B\overline{B}$ background, and C is the continuum background.

The goal is then to choose each cut so as to maximize $S^2/(S+B')$, where S and B' are the signal and effective background (eq. 5.1) expectations for $81.5\,fb^{-1}$ from $SP4$ Monte Carlo simulation. The optimization needs to be iterative, since optimal values of some cuts may depend on values of others. It begins with informed "guesses" for cut values. During the procedure, at each step, *all* cuts are applied to the Monte Carlo signal ($K^*\gamma$ plus $X_s\gamma$ using KN480) and background ($B\overline{B}$ plus continuum) samples, except for the particular cut under study. That cut only is relaxed, and $S^2/(S+B')$ is determined as a function of the cut value. When there is a clear maximum, the cut is set to that value and kept there for the subsequent optimization of the next variable until all variables have been optimized. If the maximum is broad, a cut value at the higher signal efficiency end of the $S^2/(S+B')$ plateau is chosen.

For this analysis, each "round" of optimization starts by redetermining optimal Fisher coefficients (Section 5.5.2) using signal and continuum Monte Carlo events selected with the latest values of all other relevant cuts. This cycle is repeated until no further changes to any cuts are found. Convergence is reached in two or three full rounds. A final set of $S^2/(S+B')$ distributions for each variable (with no further cut changes) is then produced.

The optimization is done using an E^*_γ range from 2.0 to 2.7 GeV. The set of variables subject to optimization is:

- Minimum bump distance;
- Minimum energies for second photons used in the π^0 and η vetoes, separately

for each parent and each E^*_γ range;

- The two shape cuts described in Section 5.5: one on R^*_2, the other on the Fisher discriminant;

- The minimum cuts on tagging variables p^*_e, p^*_μ, $\cos\theta_{\gamma e}$, and $\cos\theta_{\gamma\mu}$ (separately for muons and electrons);

- The missing energy cut after tagging;

Figures 5.26 and 5.27 illustrate optimization verification distributions for several cut variables. Note that the optimization procedure is done on a statistically somewhat smaller sample than the final analysis. Therefore, the statistical sensitivity implied by the maxima on these plots differs slightly from that quoted in Section 5.9.

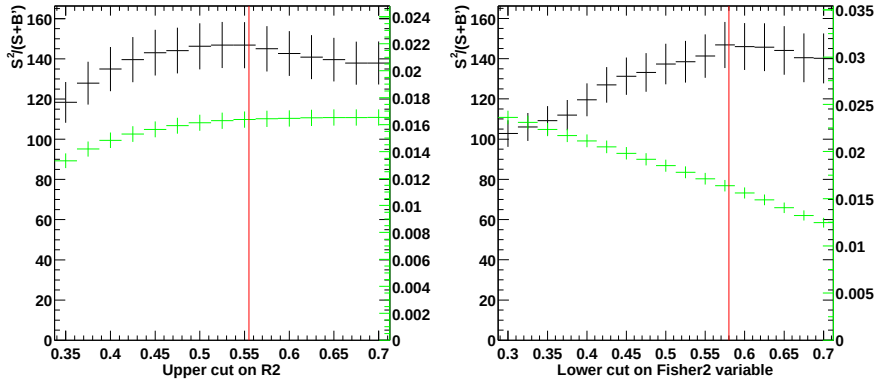

Figure 5.26: Sample cut optimization plots, for the maximum accepted value of R^*_2 (left) and the minimum accepted value of the Fisher discriminant (right). Each plot shows how $S^2/(S + B')$ (black points, left scale) and the signal efficiency (green points, right scale) vary with the cut value, using $SP4$ Monte Carlo simulation samples. All other cuts are held at their nominal values.

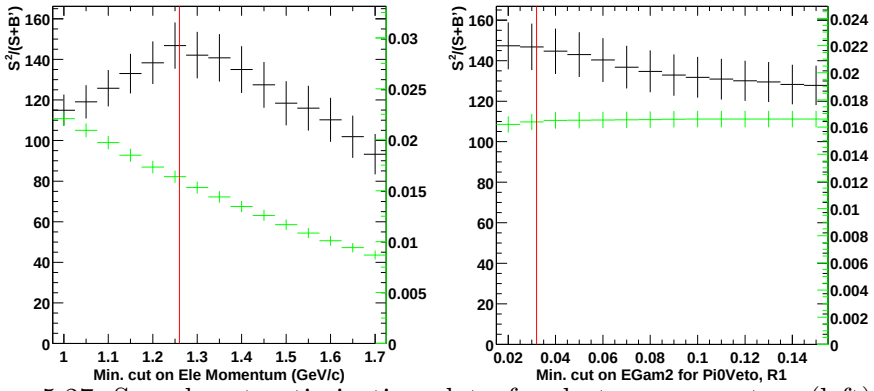

Figure 5.27: Sample cut optimization plots, for electron momentum (left) and minimum energy of the second photon in the π^0 veto in the range $2.0 < E^*_\gamma < 2.3$ (right). Each plot shows how $\mathrm{S}^2/(\mathrm{S}+\mathrm{B}')$ (black points, left scale) and the signal efficiency (green points, right scale) vary with the cut value, using $SP4$ Monte Carlo simulation samples. For each, all other cuts are held at their nominal values.

5.9 Expectations

If the Monte Carlo predictions for the signal, $B\overline{B}$ background, and continuum background were all correct, then – applying the prescriptions of appendix C to the bottom-line numbers of Tables 5.1 – 5.6 and with on-resonance and off-resonance data sets of $81.5\,\mathrm{fb}^{-1}$ and $9.59\,\mathrm{fb}^{-1}$, respectively – it is expected to extract the signal with a statistical precision of $\approx 7.8\%$.

Table 5.16 gives the expected number of signal and background events in 100 MeV energy bins after all cuts except the E_γ^* cut, according to $SP4$ Monte Carlo simulation. These expectations are derived using the cross-sections and luminosities given in Table 3.2. Fig. 5.28 shows these expectations in visual form.

E^*(GeV)	$Cont(MC)$	$B\overline{B}$	$BkgTotal$	$Signal$	$S^2/(S+B')$
$1.6-1.7$	21.2 ± 5.7	589.9 ± 16.2	611.1 ± 17.2	1.9 ± 0.6	0.0 ± 0.0
$1.7-1.8$	25.1 ± 6.2	494.6 ± 14.8	519.7 ± 16.1	3.1 ± 0.8	0.0 ± 0.0
$1.8-1.9$	12.8 ± 4.3	441.6 ± 14.0	454.4 ± 14.7	6.5 ± 1.2	0.1 ± 0.0
$1.9-2.0$	17.6 ± 5.1	318.1 ± 11.9	335.7 ± 13.0	13.8 ± 1.7	0.4 ± 0.1
$2.0-2.1$	16.3 ± 4.8	265.6 ± 10.9	281.9 ± 11.9	28.6 ± 2.5	1.8 ± 0.4
$2.1-2.2$	16.1 ± 4.9	172.4 ± 8.8	188.5 ± 10.0	58.3 ± 3.6	8.9 ± 1.5
$2.2-2.3$	14.9 ± 4.8	87.3 ± 6.2	102.2 ± 7.8	97.8 ± 4.6	29.3 ± 4.7
$2.3-2.4$	20.3 ± 5.7	78.4 ± 5.9	98.7 ± 8.2	153.7 ± 5.6	55.6 ± 7.8
$2.4-2.5$	19.1 ± 5.4	34.8 ± 3.9	53.8 ± 6.7	170.0 ± 5.5	74.9 ± 10.6
$2.5-2.6$	27.6 ± 6.6	23.2 ± 3.2	50.7 ± 7.3	118.7 ± 4.3	34.9 ± 5.8
$2.6-2.7$	43.8 ± 8.3	7.6 ± 1.8	51.3 ± 8.5	62.3 ± 2.7	8.0 ± 1.5
$2.7-2.8$	25.6 ± 6.3	6.7 ± 1.7	32.3 ± 6.5	11.9 ± 0.7	0.5 ± 0.1
$2.8-2.9$	40.1 ± 8.1	4.5 ± 1.4	44.5 ± 8.2	0.5 ± 0.2	0.0 ± 0.0
$2.9-3.0$	37.6 ± 7.6	2.7 ± 1.1	40.3 ± 7.7	0.0 ± 0.0	0.0 ± 0.0
$3.0-3.1$	41.8 ± 8.0	2.2 ± 1.0	44.0 ± 8.1	0.0 ± 0.0	0.0 ± 0.0
$3.1-3.2$	38.9 ± 7.9	2.2 ± 1.0	41.2 ± 7.9	0.0 ± 0.0	0.0 ± 0.0
$3.2-3.3$	48.3 ± 8.8	0.9 ± 0.6	49.2 ± 8.8	0.0 ± 0.0	0.0 ± 0.0
$3.3-3.4$	37.5 ± 7.4	0.9 ± 0.6	38.4 ± 7.5	0.0 ± 0.0	0.0 ± 0.0
$3.4-3.5$	24.7 ± 6.1	0.9 ± 0.6	25.6 ± 6.1	0.0 ± 0.0	0.0 ± 0.0
$Total$	529.2 ± 28.6	2534.5 ± 33.6	3063.6 ± 3063.7	727.1 ± 11.5	63.8 ± 2.9

Table 5.16: Monte Carlo expectations ($SP4$) for signal and background events in 81.5 fb^{-1} in bins of E^*_γ after all cuts except that on E^*_γ. $BkgTotal$ is the sum of $Cont(MC)$ and $B\overline{B}$ contributions. $Signal$ is the expected signal using the KN480 model plus 16.1% $B\to K^*\gamma$ contribution (see Table 2.2). $S^2/(S+B')$ is the figure of merit used in the optimization of the statistical significance. All errors are statistical.

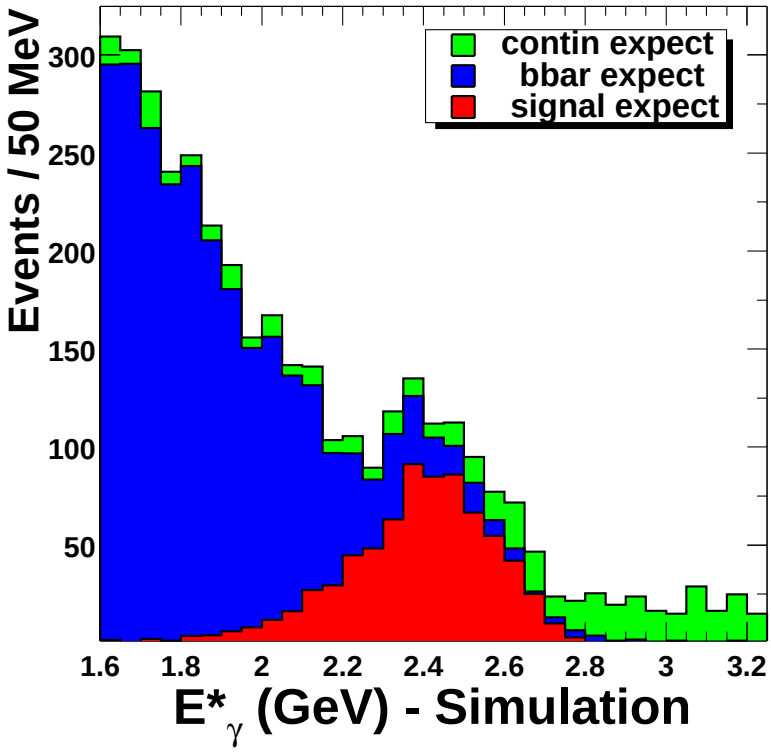

Figure 5.28: The E^*_γ distribution as expected in $81.5\,\mathrm{fb}^{-1}$ of data after all cuts (except that on E^*_γ itself), broken up into $B \to X_s\gamma$, continuum ($u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}, \tau^+\tau^-$) and $B\overline{B}$ contributions, from $SP4$ Monte Carlo simulation.

5.10 Overall Efficiency

To compute an overall signal efficiency with the defined signal model for a particular m_b, the KN and K^* contributions must be weighted according to a chosen K^* fraction. For the central result, $m_b = 4.80\,\mathrm{GeV}$ and the Kagan-Neubert prescription as described in Section 2.4 will be used. For this model a K^* fraction of 0.161 is given in Table 2.2, which assumes $m_{\mathrm{cutoff}} = 1.1\,\mathrm{GeV}$. The resulting bottom-line efficiency for $2.0\,\mathrm{GeV} < E_\gamma^* < 2.7\,\mathrm{GeV}$ is $\epsilon_{2.0-2.7\ \mathrm{GeV}} = (1.56 \pm 0.04)\%$ (a weighted average over the K^* and KN480 efficiencies), with a relative uncertainty of 1.6% from Monte Carlo statistics.

If a $\mathcal{B}(B \to K^*\gamma) = 4.03 \times 10^{-5}$ (see Table 3.2) is assumed for the K^* part of the signal along with the K^* fraction of 0.161 (this puts the total branching fraction to $\mathcal{B}(B \to X_s\gamma) = 2.50 \cdot 10^{-4}$), then the corresponding number of neutral-B (charged-B) signal events in $81.5\,\mathrm{fb}^{-1}$ would be 80(76) K^* plus 269(269) KN480, for a total of 694. *Regardless* of any normalizing assumption, for this K^* plus KN480 model 26.4% of all $B \to X_s\gamma$ events passing the selection cuts are from $B \to K^*\gamma$.

The above overall efficiencies are only interesting as an impression of the general losses the presented event selection incurs and for the optimization of selection criteria. For the extraction of a total branching fraction for the $b \to s\gamma$ transition and the determination of the E_γ^* spectral shape a measurement in bins of photon energy is performed. Thus an energy-dependent efficiency correction (bin-by-bin) is needed. To obtain it, Monte Carlo simulations are again used. The number of passing signal

events in a particular bin is divided by the number of events generated in that bin. This is again done separately for the K^* and the X_s parts of the signal. For the X_s component, the generated spectrum shows a varying statistical precision over the full E_γ^* range. To counteract this effect and to be independent of the particular shape of any specific choice of parameters in the Kagan and Neubert model, a dedicated, large set of Monte Carlo X_s events has been generated that artificially produced the same number of events in every energy bin. It still uses the Kagan and Neubert prescription for the particle composition of the hadronic final state, but any Kagan and Neubert shape prediction may at most influence the distribution *within* each bin. The obtained efficiency spectra (*vs.* E_γ^*) are shown in Figure 5.29.

To arrive at a final efficiency correction factor for data, the spectra have to be combined. The expected fraction of X_s/K^* must be determined individually for each bin and the two contributions weighted accordingly. This derived correction is then applied to the final data E_γ^* spectrum. It yields two alternative ways of deriving the total branching fraction (both of which will be tried and then compared): the fitted integral over the spectrum in the E_γ^* signal region may be taken or the corrected event counts in the contributing bins may be simply added.

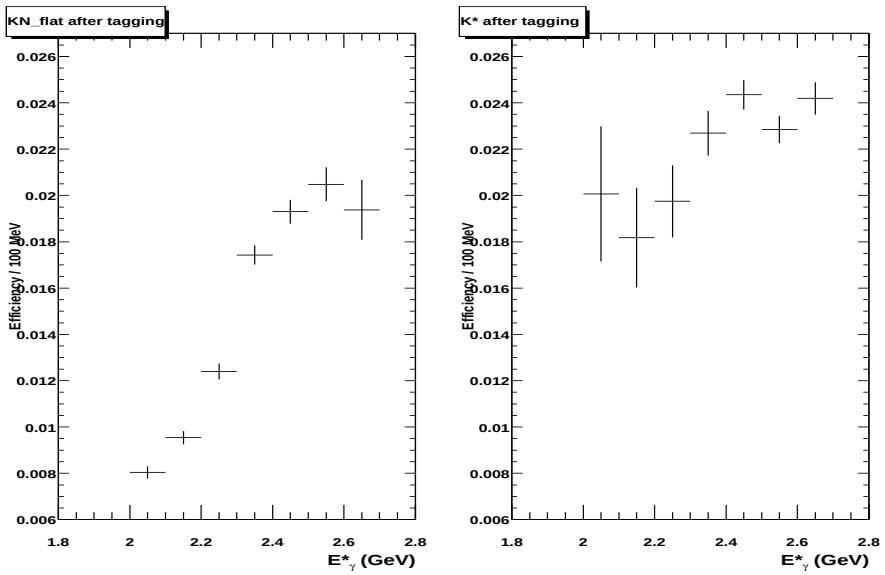

Figure 5.29: Signal efficiency as a function of E^*_γ for the X_s (on the left) and the K^* (on the right) component or the signal. The X_s efficiency is obtained with a Monte Carlo sample that has the same number of events generated in each energy bin to arrive at a largely uniform statistical uncertainty on the efficiency over the full energy range.

Chapter 6

BB Background

This chapter discusses in detail the $B\overline{B}$ background that is expected in this analysis. Since this background must be subtracted directly from Monte-Carlo simulations, predictions for different components must be checked against appropriate data samples and, if necessary, corrected. First, the expectations from the simulation are outlined and the general methods of checking and correcting the predictions are discussed. The central part of the chapter focuses individually on the most important components to the $B\overline{B}$ background and describes how appropriate, dedicated control samples are defined for checks between data and simulation — and the corrections derived from them. Finally, the method of implementing all corrections and the fully corrected $B\overline{B}$ E^*_γ spectrum are presented.

6.1 Monte Carlo Expectations

Fig. 5.28 shows the E_γ^* distribution for signal, continuum, and $B\overline{B}$ background after all cuts. The $B\overline{B}$ background is dominant at low E_γ^*. As described in Section 4.1 the main contributions to this background are photons from $\pi^0(\eta)$ decays which make up the single largest contribution, but also "photons" faked by hadronic interactions in the EMC. Table 5.6 gives the $SP4$ Monte Carlo expectations for the efficiency and yield for $\pi^0(\eta)$ and for "other" sources of $B\overline{B}$ background. The latter category includes the hadronic fakes. The figures throughout Section 5 also show the distributions broken down in this way.

A more complete summary of the breakdown of $B\overline{B}$ events by the "true" identity of the candidates is shown in Table 6.1. It shows both the true identity of the candidate photons and their true parent particles, as generated in the Monte-Carlo simulation. The fractions in the table are produced before lepton-tagging cuts are applied, to provide more meaningful statistics. These fractions are not expected to change when the lepton requirements are applied, because the probability of one B meson decaying into both a high-energy photon and a high-momentum lepton is very low. Thus, lepton requirements effectively cut on the "other" B in the event and will not have much influence on distributions on the high-energy photon side.

When determining the true particle identities and the parentage from the generated (truth) information in the simulation, the method for mapping reconstructed to generated particles (truth-matching) becomes important. Any method will make

Truth-match	Parentage	Fraction
Photon	π^0	0.642
Photon	η	0.175
Photon	ω	0.024
Photon	η'	0.011
Photon	B	0.010
Photon	J/ψ	0.008
Photon	Other	0.002
Photon	Total	0.871
π^0 (merged)	Any	0.022
None	Hadronic	0.107

Table 6.1: Breakdown of SP4 $B\overline{B}$ backgrounds by high-energy "photon" origin. The reconstructed photon candidate is matched to the generated true particle ("Truth-match" column). The type of parent as obtained from "truth" information classifies the source of these background photons. χ^2-matching is used to obtain these numbers. The fractions are for $B^0\overline{B}^0$ and B^+B^- combined, in the region $2.0 < E^*_\gamma < 2.7\,\mathrm{GeV}$, and before lepton-tagging cuts. Row 5 shows direct photons from (radiative semileptonic) B decays. The Hadronic category includes any candidates without a match to the generator level information. They are overwhelmingly hadronic. The sample used to derive this breakdown includes about 12,000 events.

use of one list of particles that contains all reconstructed objects (charged tracks or neutral clusters) and a second list that consists of the generated particles and their decay products. The second list also keeps track of the relation between each decaying particle and its and daughters. Different algorithms vary in how they map one list to the other and how products from, for example, interactions with the detector material are treated. Two methods are currently in use in the BABAR collaboration and both of them have been used for the present analysis.

- "χ^2-matching" is the current default in the BABAR experiment. It is used throughout most of this thesis. This method matches objects from both lists

using a comparative fit to their spatial positions in the detector. Pairs of objects that yield the lowest χ^2 are matched to each other. The χ^2-based truth matching can match a neutral candidate only to a photon or a (merged) π^0. Unfortunately, hadronic fakes cannot be matched, since products from hadronic interactions are not put into the truth lists. In this situation, any non-truth-matched "photon" is identified in this study as a hadronic fake. This is correct in almost all cases, since it is improbable that an actual high-energy photon would fail to be truth-matched.

- "GHIT-based matching" matches objects in both lists, if they share a sufficient number of single hits in the detector (GHITs). This is a more reliable approach to finding a map between the two lists, because it is less prone to errors due to, for example, overlapping particles. For the truth list, his method considers EMC clusters from hadronic interactions as split-offs, and therefore really parts of the original hadron. It therefore matches hadronic fakes to the originating hadron (mostly neutrons), a clear advantage over the χ^2-based method. On the other hand, unfortunately, all π^0, even ones where both photons are merged into one EMC entry, are considered as decaying and both the π^0 and its daughter photons are entered separately into the truth list. Identification of merged π^0 is then not possible with this method because they cannot be separated from regular π^0 decays. When comparing both matching algorithms, furthermore, photon candidates that are identified as merged π^0 with χ^2-matching, here are

identified as true photons from π^0 decays, as products from hadronic fakes, or even as radiated off of electrons. At this time, this method is only used for the $\pi^0(\eta)$ control samples discussed below. But because of its more stable and intuitive matching, future extensions of this analysis will most likely switch completely to the GHIT-based method.

It is clear from this discussion that any differences between the matching algorithms only affect the hadronic and merged π^0 components of the backgrounds. According to the simulation and the χ^2-matching, this is only slightly more than 10% of the total expected as can be seen from Table 6.1. It also shows that of the non-$\pi^0(\eta)$ category about 60% stems from hadronics and merged π^0s; the rest consists of a variety of other radiative meson decays. Figure 6.1 shows how the untagged $B\overline{B}$ background breaks down as a function of E^*_γ. Note that the hadronic component dominates at the highest E^*_γ (where the $B\overline{B}$ yield becomes low); the merged-π^0 component is combined with it, but its fraction decreases strongly as E^*_γ increases. These estimates are, of course, purely Monte-Carlo predictions. I must be examined whether or not they reflect the actual composition of $B\overline{B}$ backgrounds. This makes the extensive checks necessary that are discussed in the remainder of this chapter.

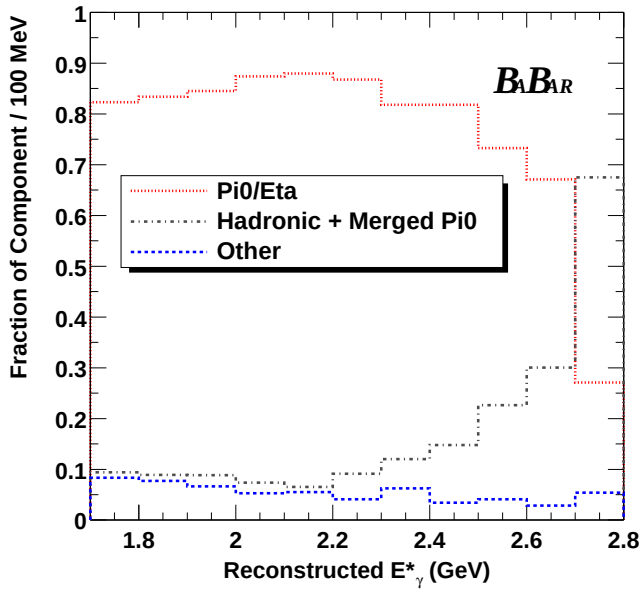

Figure 6.1: Breakdown of $SP4$ simulated $B\overline{B}$ backgrounds *vs.* E^*_γ, before tagging. These fractions are for the combined $B^0\overline{B}^0$ and B^+B^- samples, with $1.7 < E^*_\gamma < 2.8\,\mathrm{GeV}$. The "other" category includes the miscellaneous photon parents listed in Table 6.1.

6.2 Checking and Correcting the Monte Carlo Predictions

Since this analysis relies upon Monte Carlo simulation to estimate and subtract the $B\overline{B}$ background, it must be established how well the simulation models the real data. For this task the total $B\overline{B}$ sample is divided into parts according to the origin of the background high-energy photon.

The dominant component of the $B\overline{B}$ background comes from decays of $\pi^0(\eta)$ mesons into two photons. To study it in more detail in the simulation and to compare

it to data, a dedicated selection of inclusive π^0s and ηs (*i.e.* from any source in $B\overline{B}$ background decays) yields a control sample that is independent of the signal sample. Within this set of events, π^0s and ηs are treated separately, and corrections are computed in terms of $\pi^0(\eta)$ energy. While defining a generic $B\overline{B}$ background Monte Carlo sample is not a problem, a trick must be used to arrive at a sample of generic $B\overline{B}$ background events in real data: To get rid of the continuum events present, the appropriately scaled off-resonance sample must be subtracted, which leaves a continuum-free sample of $B\overline{B}$ decays. It still contains signal events along with the $\pi^0(\eta)$ background. But since the high-energy photon in signal events never is part of a real π^0 or *eta*, it can only contribute to combinatoric $\pi^0(\eta)$ candidates and will be subtracted out with the method presented below. The correction factors for the simulated $B\overline{B}$ background are then calculated in bins of $\pi^0(\eta)$ energy as the ratios of the off-resonance subtracted data yield over the generated $B\overline{B}$ yield. The corrections can be applied to the $B\overline{B}$ Monte-Carlo photon spectrum by using Monte-Carlo truth information to identify the energy of the parent $\pi^0(\eta)$ event-by-event. The treatment of the $\pi^0(\eta)$ component is described in more detail in Section 6.3.

The next-largest, hadronic, component is particularly problematic for two reasons. First, much of it is expected to be due to neutrons[1] and it is known that the Monte-Carlo generator is not very reliable in estimating baryon yields, especially those from $B\overline{B}$ decays. Second, the remaining background after cuts is in a "tail" region

[1]A charged hadron is much less probable to yield a high-energy *neutral* cluster in the EMC, so mainly n, $\overline{n}$ and K_L^0 sources of this component must be considered.

which is challenging to simulate correctly. On the other hand, the Monte Carlo-*predicted* hadronic background is not very large. In order to keep the influence of this component as small as possible, additional emphasis was put on rejecting clusters consistent with hadronic interactions or merged π^0 during the photon quality selection (Sec. 5.2). Hadronic interactions produce photon clusters with different lateral profiles, giving rise to a cut on the lateral moment to decrease this background. Because of the relatively small fraction of hadronic (plus merged-π^0) background, it is sufficient to use an "anti"-veto data and simulation event sample here, instead of a dedicated selection. Events that are *rejected* from the normal selection by the (anti-hadron) cuts on lateral moment or second moment are now used to study how well the simulation predicts the amount of hadron or merged-π^0 production resulting in a high-energy photon signature. This, along with systematic checks, is described in Section 6.4.

Finally, checks and corrections for the remaining meson-decay components from Table 6.1 are described in Section 6.5.

One handle on the *total* $B\overline{B}$ contribution is to compare data and Monte Carlo predictions in an E_γ^* "sideband" below the range where a signal is expected. As will be shown in Section 7.1, the agreement between on-resonance data and the Monte Carlo estimate for the $B\overline{B}$-dominated control region $1.7 < E_\gamma^* < 1.9\,\mathrm{GeV}$ is reasonable.

6.3 Studies of the $\pi^0(\eta)$ Component

6.3.1 Selecting Samples of Inclusive π^0 and η Events

To compare the modeling of the $\pi^0(\eta)$ component of the $B\overline{B}$ Monte-Carlo simulation to that in data, an inclusive sample of π^0s and ηs is needed that should be as similar to the background events in the $b \to s\gamma$ measurement. Therefore, the $\pi^0(\eta)$ decay products have to include a photon of high energy.

The preselection of events for this sample uses a dedicated $\pi^0(\eta)$-skim selection that roughly corresponds to the skim selection in the $b \to s\gamma$ signal selection (Sec 5.1). Events must show the multihadronic characteristics required by the BGF-Multihadron tag-bit and contain at least three charged tracks of good quality. The minimum requirement on the highest CMS energy of all neutral EMC clusters is loosened here to $1.0 <$ e1Mag < 3.5 GeV. This effectively lowers the minimum energy requirement on the high-energy photon to allow for more symmetric $\pi^0(\eta) \to \gamma\gamma$ decays in the sample. The following loose cuts on the overall Fox–Wolfram moment R_2^* and the CMS momentum of the highest-momentum charged track are exactly the same as in the main selection. For this special selection, a reconstructed high-energy photon in the broader energy range $1.0 < E_\gamma^* < 3.5$ GeV (same range as that for e1Mag) is required, but, in addition, here it is also demanded that at least one π^0 or η candidate can be reconstructed from that photon in combination with any other photon in the event.

The subsequent selection proceeds again as described in Chapter 5 for the signal $b \to s\gamma$ analysis. There are only two exceptions: First, obviously no π^0 or η vetoes are applied to the high-energy photon, since these are the particles to be studied here. Second, the momentum requirements for the lepton-tags are loosened to 1.0 GeV for electrons and 1.1 GeV for muons to gain statistics. Again, this is not expected to alter the final distributions, because the high-energy photon and high-momentum lepton are improbable to be decay particles from the same B so that the signal side should essentially not be affected by this cut.

In the event sample defined in this way, π^0 and η candidates are reconstructed from the high-energy photon candidate and any other photon in the event. To be used in the reconstruction of a $\pi^0(\eta)$ candidate, second photons must satisfy a minimum energy requirement of 50(150) MeV and some of the same quality cuts placed on the high-energy photon (see Sec. 5.2). In particular, they must:

- be reconstructed within the same angular acceptance of $-0.74 < \cos\theta_{\gamma 2} < 0.93$;
- satisfy the same maximum second moment requirement of 0.002; and
- show a lateral moment of the EMC cluster of less than 0.45 (also the same as for the high-energy photon).

In addition, the energy of the reconstructed meson must be in the region $1.0 < E_{\pi^0(\eta)} < 3.5$ GeV and its invariant mass must lie within a broad window of 50-250 MeV and 400-700 MeV for π^0s and ηs, respectively.

Of course, this way of reconstructing $\pi^0(\eta)$ decays yields a number of random combinations that underlie the true candidate distribution as shown in Figure 6.2 for π^0s. Choosing the invariant mass windows as wide as mentioned above (more than 10σ) then enables a fit to the signal $\pi^0(\eta)$ and combinatoric background distributions. Thus, the combinatoric background can be subtracted off and the inclusive yield of true $\pi^0(\eta)$ candidates can be obtained. See the next section, 6.3.2, for details about the fits and yields.

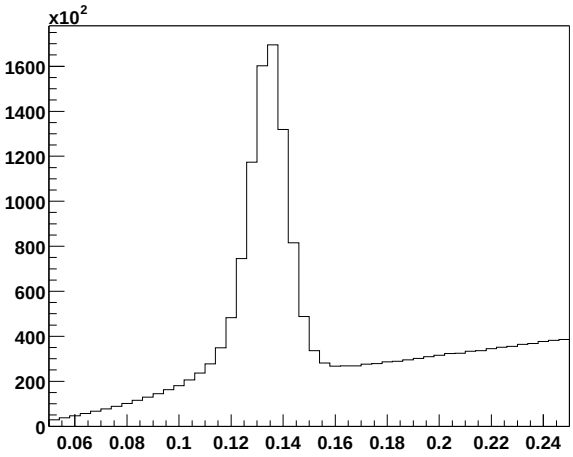

Figure 6.2: Invariant mass distribution for all π^0 candidates from the dedicated $\pi^0(\eta)$ selection in on-resonance data, before lepton-tagging cuts are applied. Shown in the histogram is the number of events versus the π^0 ($\gamma\gamma$) invariant mass.

6.3.2 Corrections to the $\pi^0(\eta)$ Monte Carlo Expectation

To obtain the most precise corrections to the $\pi^0(\eta)$ component of the $B\overline{B}$ Monte-Carlo simulations, the correction must be a function of the $\pi^0(\eta)$ reconstructed

energy in the CMS, $E^*_{\pi^0(\eta)}$. To obtain this function, the full sample of π^0s and ηs is treated in bins of $\pi^0(\eta)$ energy, with signal and background fits and their subtraction performed in each bin. This reduces the statistics for each fit and may introduce convergence problems. However, if the bin size is adjusted in such a way as to always keep a sufficient and roughly equal amounts of events in each bin, this problem can be largely avoided. Note that from here on, π^0s and ηs are treated separately, because of the much larger fraction — and therefore sample size — of π^0s.

For the latter, it was found that the $E^*_{\pi^0}$ distribution is best treated in a total of eight bins, with seven equidistant bins of 200 MeV width between $E^*_{\pi^0} = 1.0$ GeV and 2.4 GeV and one larger bin at $2.4 < E^*_{\pi^0} < 3.0$ GeV in the area with lower $B\overline{B}$ statistics in general. The combinatoric background in the invariant mass spectra is fit with a polynomial of order one, two or three, and a Crystal-Ball-like function is used to parametrize the "signal" π^0 part. Crystal-Ball shapes are well-suited to describe mass peaks with low-end radiative tails. Their general definition combines a central Gaussian distribution with a power-law tail at the lower end of the spectrum. For the fit here, the central, single Gaussian is replaced by a double-Gaussian distribution to allow a broader width versus height variation, and the lower energy tail is kept in the definition:

$$x > x_0 : \; GG(x) = 2A_{gg} \cdot \{f \cdot \mathrm{erfc}(\tfrac{x_0-\mu_1}{\sqrt{2}\sigma_1})^{-1} \cdot G_1(x) + (1-f) \cdot \mathrm{erfc}(\tfrac{x_0-\mu_2}{\sqrt{2}\sigma_2})^{-1} \cdot G_2(x)\}$$

$$x \leq x_0 : \; CB(x) = GG(x_0) * \{\tfrac{p\sigma_1}{\lambda} \cdot (x_0 - x + \tfrac{p\sigma_1}{\lambda})\}^{-p}$$

Here $G_i(x) = \frac{1}{\sqrt{2\pi}\sigma_i} \cdot \exp(-\frac{(x-\mu_i)^2}{2\sigma_i^2}), i = 1, 2$ are the two single Gaussian distributions

that are combined into the central double Gaussian $GG(x)$. μ_i and σ_i are their peak positions and widths. A_{gg} is the area of the double Gaussian from x_0 to infinity, f is the ratio of the first Gaussian's area (from x_0) over A_{gg}. The complementary error functions erfc$(y/(\sqrt{2}\sigma))$ are defined as the integrals from $-y$ to $+y$ over the corresponding normalized Gaussians, overall they serve normalization purposes. x_0 is the point at which the Gaussian changes into the Crystal-Ball power function $CB(x)$,with p the exponent of the polynomial and λ the second parameter of the function that determines how far the tail reaches. $CB(x)$ is normalized by the value of $GG(x)$ at the transition point x_0 to ensure that the entire function is continuous across x_0.

There is a multitude of parameters that have to be determined for each sample in the fit: A_{gg}, f, μ_1, σ_1, μ_2, $\frac{\sigma_1}{\sigma_2}$, λ, p and x_0. To minimize the statistical uncertainty on the values of these parameters, the largest possible sample should be used for the fit that determines them. Various approaches were tried to this end. The following procedure with three fit stages was found to give the smallest deviation between the fitted yield and the true number of π^0s in the generic $B\overline{B}$ sample: First, the mass distributions of the truth-matched π^0s in the generic $B\overline{B}$ sample are fit with the above function. The parameters μ_2, $\frac{\sigma_1}{\sigma_2}$ are given as initial conditions and kept fixed. This works because the second Gaussian is only a small correction to the first. It adds an extra "tail" on the high end of the spectrum. The pure signal sample that this fit works on is especially suited to study the signal shape alone as expressed in the Crystal-Ball function above. The *decay* of π^0s is modeled very well in the

simulation, so no large differences between the data and simulation are expected for the shape of the signal function. It is the *production* of π^0s in B decays and with that the overall normalization of the π^0 mass spectrum that are known to be problematic in the simulation. Therefore, the parameters f, λ and x_0 are not expected to be influenced by any simulation–data differences (they describe the π^0 decay), and they are now also fixed at the values found from this $B\overline{B}$ Monte Carlo truth-match fit for the subsequent fit stages. This is a good position to start including the background parts in the fits, i.e. to return to the complete reconstructed event sets. In each sample (on-resonance, off-resonance, and generic non-truth-matched $B\overline{B}$ Monte Carlo) the full fit of the Crystal-Ball plus background polynomial is performed on the "untagged" π^0 mass distributions, i.e. on the distributions obtained before the lepton tagging cuts are applied. This is the second stage of the fitting. Values for all fit parameters are determined for each event sample separately. They are all afterwards kept fixed to these values, expect for A_{gg} and the polynomial background parameters. Finally, in the third stage of fitting, the event samples — *with* all lepton-tag cuts required — are fit again with only A_{gg} floating. A_{gg} as well as the total yield are determined. Considering the latter in all $E^*_{\pi^0}$ bins that these fits were performed in then gives the final $E^*_{\pi^0}$ spectra for data and Monte Carlo. From these the corrections factors are to be derived.

To describe the background shape in the last two fit stages, a polynomial of third order is used in the untagged samples and a polynomial of first, second or

third order in the tagged samples. In the last case, the exponent depends on the available statistics. For the π^0 component, all $E^*_{\pi^0}$ bins in all samples (on-resonance, off-resonance, and generic $B\overline{B}$ Monte Carlo) yield a satisfactory fit performance. For an example of the fit performance in the presence of ample statistics, see Figure 6.3 which presents the fits to the tagged on-resonance data sample (*i.e.* stage three). The two lines show the performance of the background plus Crystal-Ball fit to the entire sample and the result of only the Crystal-Ball (using the determined parameters) to the background subtracted π^0 yield.

The resulting π^0 energy spectra for the subtracted on-resonance data and the $B\overline{B}$ simulation can be seen in the upper row in Figure 6.4 both before and after lepton-tagging cuts are applied to the samples. It is clear that some correction is needed for the simulated $B\overline{B}$ π^0 yield. The correction factors are obtained in bins of $E^*_{\pi^0}$ from the ratios of subtracted data over $SP4$ $B\overline{B}$ Monte Carlo yields:

$$\pi^0 \text{ Selection Ratio}_{(i)} = \frac{\text{OnResonance data}_{\text{(i)}} - \text{OffResonance data}_{\text{(i)}}}{B\overline{B} \text{ expectation}_{(i)}} \quad . \qquad (6.1)$$

Different numbers are obtained from the tagged and untagged samples. They are tabulated along with their errors in Table 6.2 and shown graphically in the bottom plots in Figure 6.4. The errors quoted include statistical effects from all samples and uncertainties from the fits. They have to be taken into account as systematic errors when the corrected $B\overline{B}$ E^*_γ spectrum is subtracted from the one in data. Note that the untagged corrections show slightly smaller errors due to the larger statistics of the samples. However, the tagged numbers reflect more closely the π^0 background events

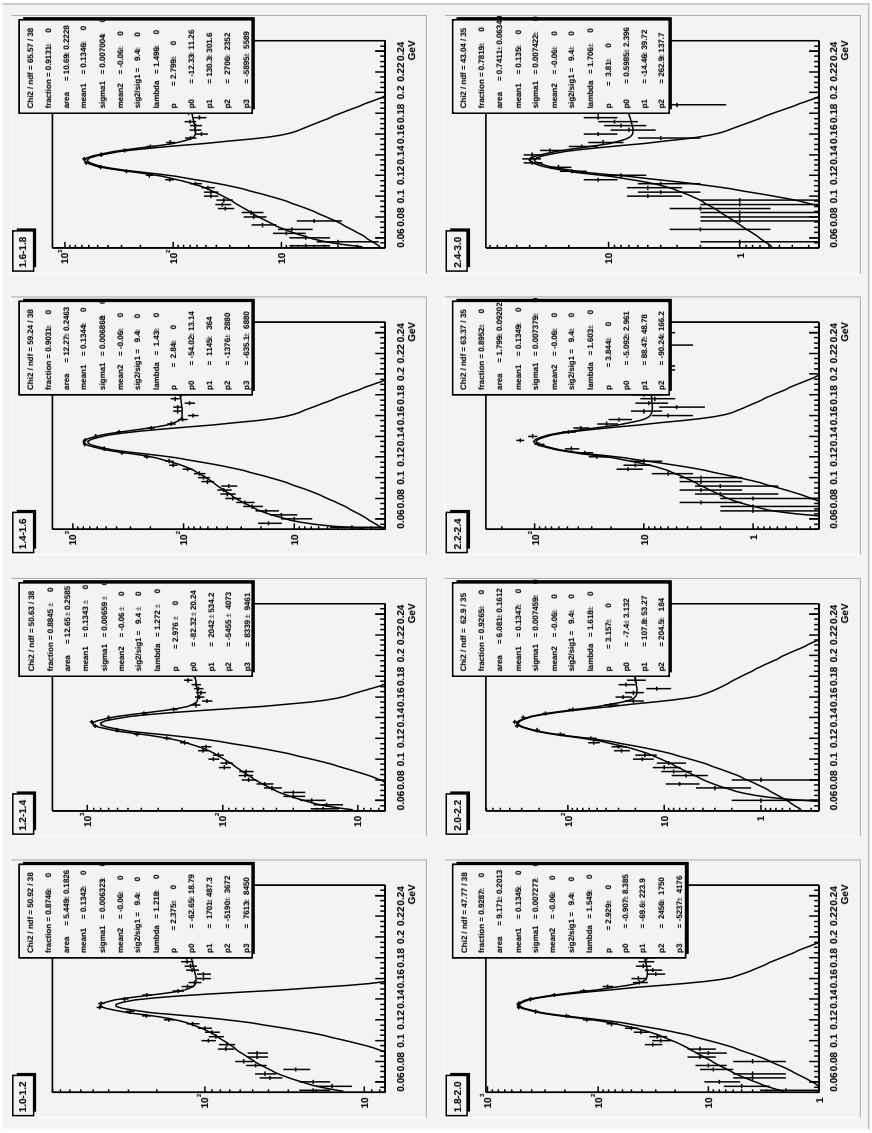

Figure 6.3: Examples of π^0 mass fits in the dedicated $\pi^0(\eta)$ sample from on-resonance data, after lepton-tagging cuts are applied. The $E^*_{\pi^0}$ range is indicated in the upper left hand corner of each plot. The two lines show the performance of the background plus Crystal-Ball fit to the entire sample and the result of only the Crystal-Ball (using the determined parameters) to the background-subtracted signal π^0 yield.

left over in the $b \to s\gamma$ analysis and so are used for the correction.

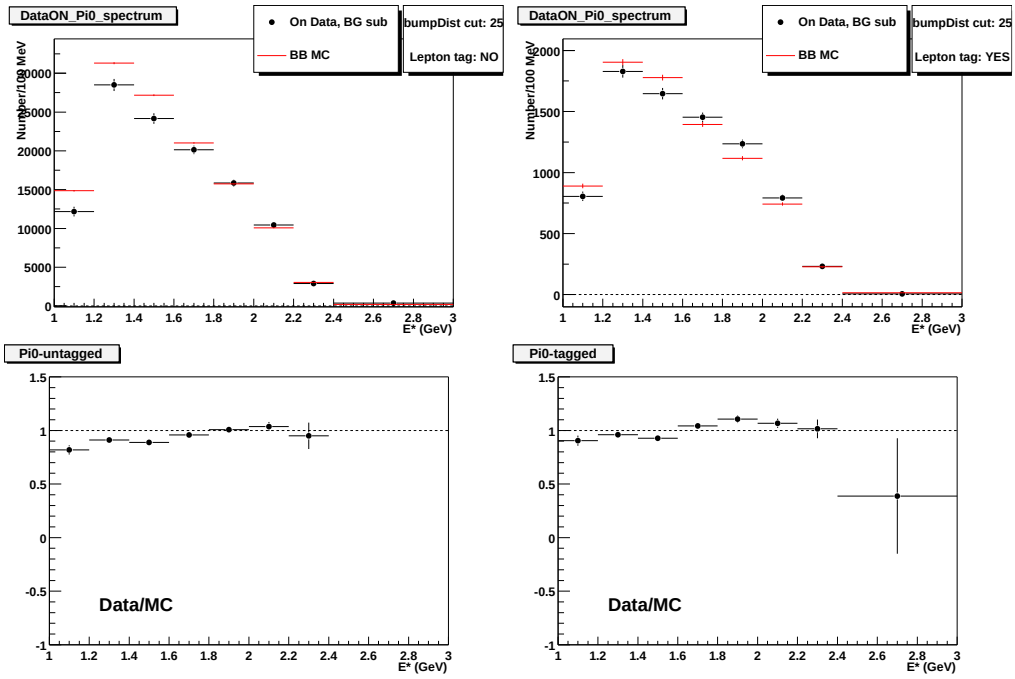

Figure 6.4: The π^0 spectra in $B\overline{B}$ simulated events and off-resonance subtracted data before (top left) and after (top right) lepton-tagging cuts are applied; and the derived correction factors (data/simulation ratios) for the π^0 component as a function of π^0 energy before (bottom left) and after (bottom right) tagging cuts are applied.

$E^*_{\pi^0}-$ Range	π^0 Correction factors	
	Tagged	Untagged
$1.0 - 1.2$	0.9046 ± 0.0472	0.8187 ± 0.0428
$1.2 - 1.4$	0.9603 ± 0.0297	0.9104 ± 0.0254
$1.4 - 1.6$	0.9265 ± 0.0290	0.8893 ± 0.0250
$1.6 - 1.8$	1.0415 ± 0.0319	0.9576 ± 0.0270
$1.8 - 2.0$	1.1055 ± 0.0357	1.0080 ± 0.0315
$2.0 - 2.2$	1.0666 ± 0.0436	1.0366 ± 0.0422
$2.2 - 2.4$	1.0152 ± 0.0877	0.9498 ± 0.1232
$2.4 - 3.0$	0.3880 ± 0.5394	1.7087 ± 0.8551

Table 6.2: Correction factors for the π^0 component of the generic $B\overline{B}$ background simulation as a function of π^0 energy, before and after lepton-tagging requirements are applied. Note the varying bin-widths.

Fitting the η distributions poses more challenges, because of the significantly lower statistics in the samples. Especially for off-resonance data which only comprises roughly 10% of the total collected luminosity, it is non-trivial to reach fits of good quality. The region from 1.0 to 2.6 GeV is considered for the fits, it is divided into five bins of varying width (between 200 and 500 MeV) as can be seen in Table 6.3. For the total, measured η mass spectra, the background is again parametrized as a polynomial of order one (or two or three for the lepton-tagged stage), as in the π^0 case. However, the general signal shape is different for ηs, so that a different parametrization has to be found here. Again, a variation on a Crystal-Ball function yields the best results. For the η signal, one central Gaussian peak with two independent power-law tails, one on each side, is used (the regular Crystal-Ball has only one tail at the low side of the spectrum):

$$f(x) = \begin{cases} g_1 \cdot \{\frac{p_1\sigma}{\lambda_1} \cdot (x_1 - x + \frac{p_1\sigma}{\lambda_1})\}^{-p_1} & x < x_1 \\ A \cdot \frac{1}{\sqrt{2\pi}\sigma_i} \cdot \exp(-\frac{(x-\mu)^2}{2\sigma^2}) & x_1 < x < x_2 \\ g_2 \cdot \{\frac{p_2\sigma}{\lambda_2} \cdot (x_2 - x + \frac{p_2\sigma}{\lambda_2})\}^{-p_2} & x > x_2 \end{cases}$$

Here $x_i = \mu - \lambda_i * \sigma$ are the connection points between the Gaussian and the power functions (they are not independent parameters anymore), μ and σ are the Gaussian's peak position and width, A is its area. λ_i and p_i are the widths and exponents of the power functions and g_1, g_2 are determined by requiring that the entire function be

continuous across the connection points, x_1 and x_2.

The free parameters in the fit are the area, mean and σ of the Gaussian and the λ and p parameters for the two tails. Due to the small sample size for the η component, the fixing of parameter values from fits to higher-statistics samples becomes even more important here than in the π^0 case. As before, the fitting is done in three stages. However, because a double-tailed Gaussian is easier to fit than the double Gaussian before, there are some differences as to which parameters are fixed at which stage. In the first stage, again, the signal portion of the fitting function is fit to the truth-matched η candidates in the simulated $B\overline{B}$ events. For the ηs it is not necessary to fix any parameter to initial conditions prior to this step. However, the values obtained from this fit are again used to fix some parameters before the background is included in the subsequent stages. Due to the large combinatoric background for η candidates, it is necessary to fix all tail parameters at this stage; that is λ_1, p_1, λ_2, and p_2. Next, the untagged distributions are fit in all samples with the sum of the signal Crystal-Ball and the background polynomial functions. As in the π^0 case, all parameters, except the area under the Gaussian are then fixed separately in each sample from these fits. In the last stage, they are used to fit the tagged distributions. The integral of the signal Gaussian and the overall yield is obtained in each bin of E^*_η and the final E^*_η spectra for data and Monte Carlo are obtained.

The challenges involved with fitting the η distributions can be seen in Figure 6.5 which shows the performance of the fits for the sample with the lowest statistics

— off-resonance data after all (relaxed) lepton-tagging cuts are applied. The line for the background polynomial plus double-tail Crystal-Ball fit to the entire spectrum is included, as is the fit of only the Crystal-Ball (with the determined parameters) to the background-subtracted signal-η spectrum. The larger contribution of the combinatoric background can also be seen from these plots.

The η energy spectra resulting from these fits are shown in the upper plots of Figure 6.6 for the off-resonance subtracted data and for $B\overline{B}$ Monte Carlo before and after lepton-tagging cuts are applied. The discrepancy between the spectra gives rise as before to energy-dependent correction factors in the form of the ratios of data over Monte Carlo yields in bins of E^*_η bins, in completely analogy to equation 6.1. The ratios are listed in Table 6.3 for the lepton-tagged and -untagged samples. The corrections predicted from both samples are in agreement. The numbers from the tagged sample are used for the analysis, because this sample corresponds closer to the backgrounds found in the actual $b \to s\gamma$ analysis after all cuts. The energy-dependent distribution of the correction factors is again also shown in graphic form, see the lower plots in Figure 6.6.

6.3.3 Systematic Studies and Checks of the $\pi^0(\eta)$ Component

The corrections that were found for the π^0 and η component of the simulated $B\overline{B}$ background and the way they are applied back to the Monte Carlo $B\overline{B}$ sample must now be checked. Of primary interest is, of course, whether the corrected $B\overline{B}$

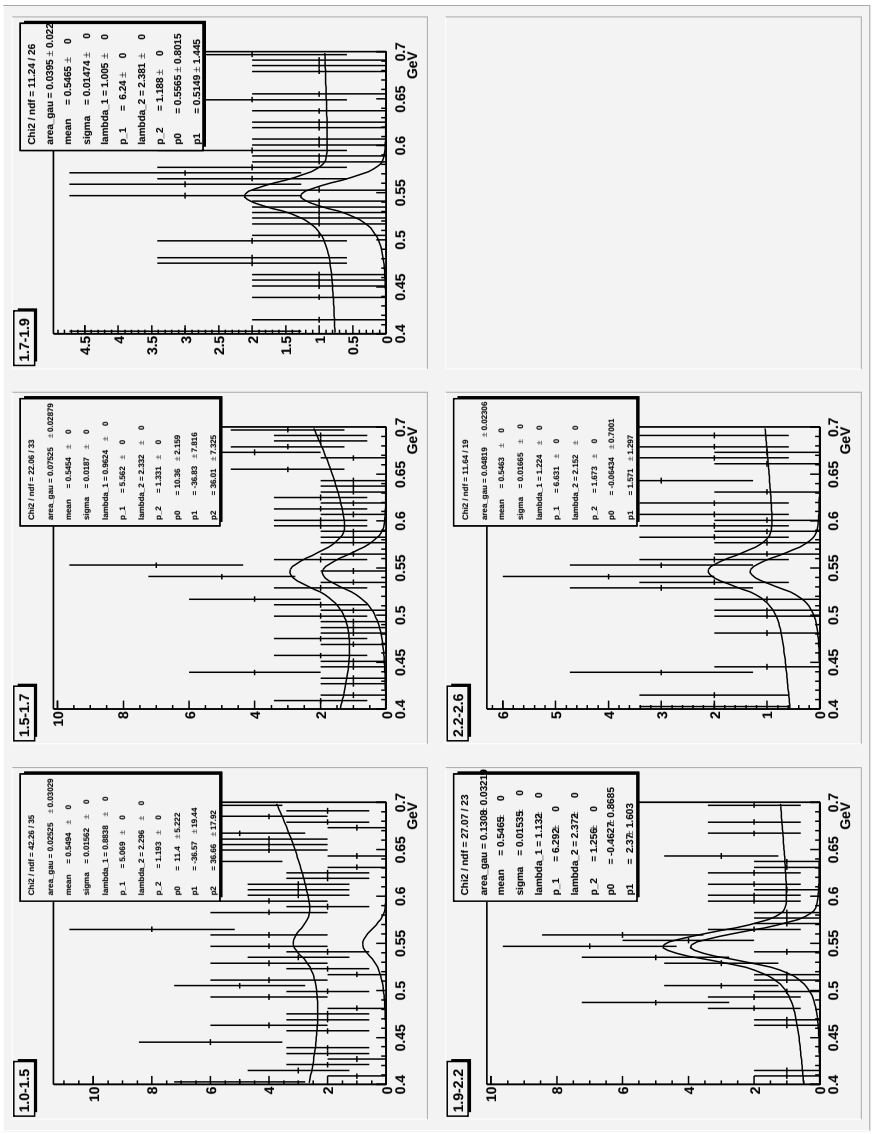

Figure 6.5: Examples of η mass fits in the dedicated $\pi^0(\eta)$ sample from off-resonance data, after lepton-tagging cuts are applied. The E^*_η range is indicated in the upper left hand corner of each plot. The two lines show the performance of the background polynomial plus double-tail Crystal-Ball fit to the entire spectrum and the result of only the Crystal-Ball (using the determined parameters) to the background-subtracted signal η yield.

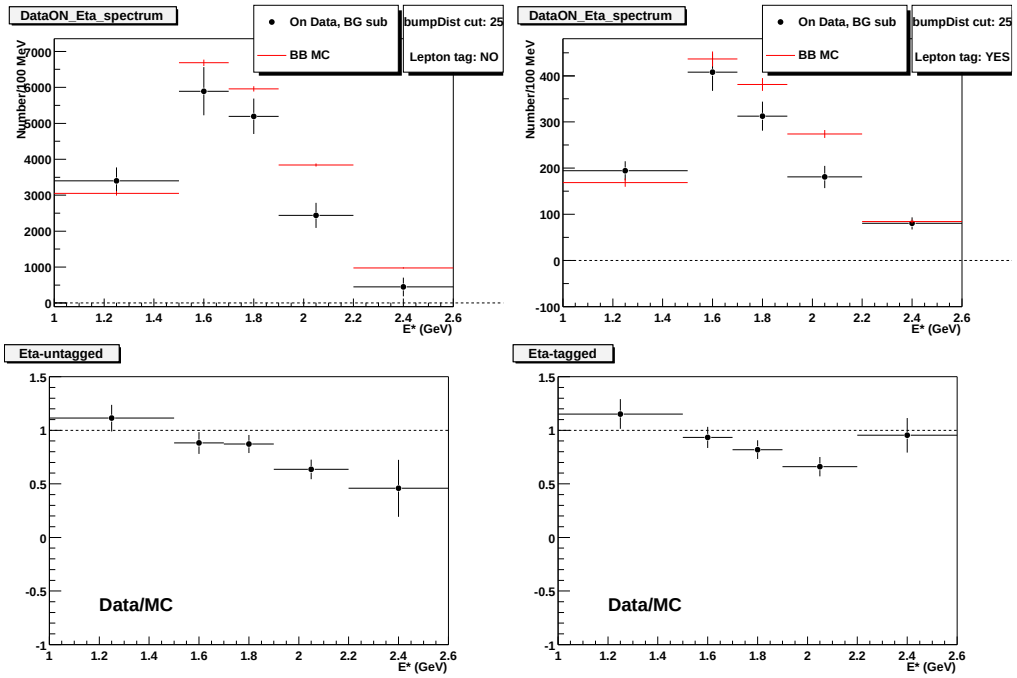

Figure 6.6: The η spectra in $B\overline{B}$ simulated events and off-resonance subtracted data before (top left) and after (top right) lepton-tagging cuts are applied; and the derived correction factors (data/simulation ratios) for the η component as a function of η energy before (bottom left) and after (bottom right) tagging cuts are applied.

Monte Carlo $\pi^0(\eta)$ samples now agree better with the samples in data. Also, the uncertainty connected to the methodology of obtaining and applying the corrections must be evaluated. This section presents the checks and studies that are done in order to get an understanding of these issues.

The π^0 and η corrections are defined depending on the reconstructed energy of the candidate in the CMS frame. When working them back into the simulation, however, they are applied to the high-energy photons, if they are identified as daughter particles of a $\pi^0(\eta)$ via the mapping to the true generated objects. That means that

$E^*_\eta-$ Range	η Correction Factors	
	Tagged	Untagged
$1.0-1.5$	1.1520 ± 0.1385	1.1138 ± 0.1243
$1.5-1.7$	0.9337 ± 0.0988	0.8816 ± 0.1011
$1.7-1.9$	0.8195 ± 0.0875	0.8719 ± 0.0837
$1.9-2.2$	0.6603 ± 0.0906	0.6354 ± 0.0912
$2.2-2.6$	0.9533 ± 0.1617	0.4593 ± 0.2659

Table 6.3: Correction factors for the η component of the generic $B\overline{B}$ background simulation as a function of η energy, before and after lepton-tagging requirements are applied. Note the varying bin-width.

here the parent $\pi^0(\eta)$ is not reconstructed and no reconstructed energy is available for it. Instead, the truth (generated) energy of the parent is used to identify the weight (or correction) to be assigned to the photon (and the event). The effect of this must be studied for introducing a bias, because reconstructed values for any variable include resolution (smearing) and reconstruction (mis-measurement) effects so that a $\pi^0(\eta)$ with a true energy in a certain energy bin may end up in a different one in the reconstructed distribution. This effect is expected to be most pronounced for many narrow bins in the distribution, so the rather large energy bins (200-500 MeV) used here for the corrections might ameliorate the consequences.

To study this effect, and to confirm that the method is working, the following "sanity check" is executed. In the dedicated $\pi^0(\eta)$ sample used for the study of these corrections, a modified $b \to s\gamma$ analysis is carried out. The allowed energy range for the high-energy photons is broadened to $1.0-3.5$ GeV and the π^0 and η vetoes are not only removed, but essentially inverted: events are *selected* if the high-energy photon

can be combined with another photon in the event to make a $\pi^0(\eta)$ in the mass range $115(508) < m_{\gamma\gamma} < 155(588)$ MeV. (The second photon is required to have a minimum energy of 50 (150) MeV.) This negated-veto or "antiveto" selection effectively yields a sample of π^0s and ηs that can be defined in simulated and data event sets because it does not rely on any Monte Carlo truth information. After all cuts are applied to these control samples, the E^*_γ spectrum obtained in simulated $B\overline{B}$ antiveto events can be compared to the one obtained from the on-resonance sample after the off-resonance data subtraction. However, in this case the distribution obtained in $b \to s\gamma$ signal Monte Carlo must be subtracted from the data, because the antiveto sample contains a non-negligible component of $b \to s\gamma$ signal events. This is due to the fact that combinatoric $\pi^0(\eta)$ candidates are not rejected in the antiveto selection. The sample is thus enriched with random combinations of signal photons and lower energy gammas unless this component is subtracted out. In the following, the sample obtained from applying an inverted π^0 veto (π^0-antiveto sample) is treated separately from the one obtained from the complement of the η veto (η-antiveto sample). The left plots in Figures 6.7 and 6.8 show the antiveto E^*_γ distributions after all cuts for π^0s and ηs in the properly subtracted data sample compared to the still uncorrected prediction from $B\overline{B}$ background Monte Carlo. Here as in the dedicated $\pi^0(\eta)$ sample, the curves do not match perfectly. Some correction is needed to get a better agreement. This can also be seen from the plots in the lower left of these Figures, where the ratio of the above curves is shown.

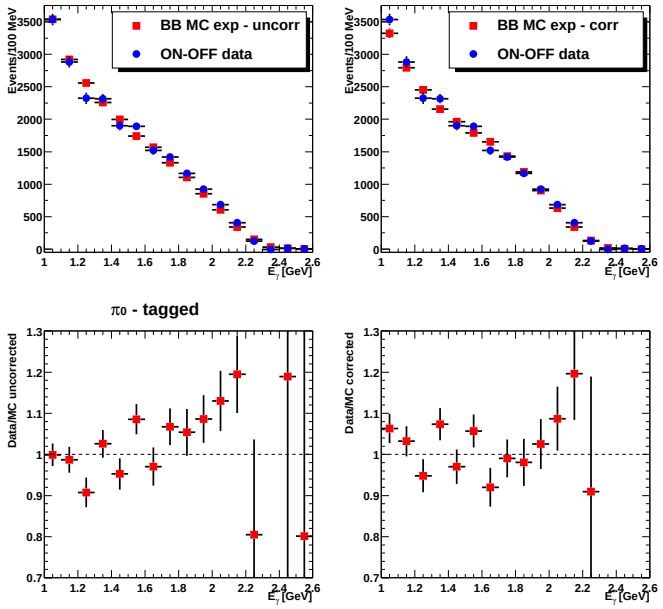

Figure 6.7: Comparison of E_γ^* spectrum after all analysis cuts for off-resonance subtracted data and simulated $B\overline{B}$ background in the π^0 antiveto sample. The left (right) side shows the spectra before (after) the π^0 corrections are applied: "Sanity Check".

The $B\overline{B}$ simulation sample is then corrected with the appropriate correction factors obtained in the last section (Tables 6.2 and 6.3), according to the photon parent particle and its energy. The resulting corrected E_γ^* spectrum from the simulation is compared to the (subtracted) data distribution in the right plots of Figures 6.7 and and 6.8 for π^0s and ηs, respectively. The plots on the lower right show the ratio of the curves after correction to ease the comparison. The corrected case shows better, although not perfect, agreement between data and Monte Carlo, especially in the region of interest, $2.0 - 2.7$ GeV. This validates the correction factors obtained earlier and the technique used to derive them.

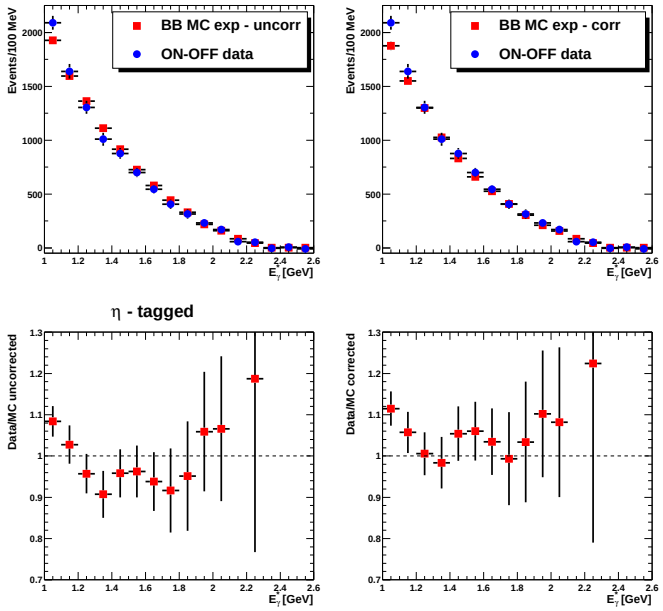

Figure 6.8: Comparison of E^*_γ spectrum after all analysis cuts for off-resonance subtracted data and simulated $B\overline{B}$ background in the η antiveto sample. The left (right) side shows the spectra before (after) the η corrections are applied: "Sanity Check".

Furthermore, no shift in the spectra is observed. This is important, because there is an inconsistency in applying corrections obtained from *reconstructed* energy spectra to spectra of the *generated* truth energy of the parent. This may cause a shift in the corrected spectra. These effects are seen to be negligible.

The uncertainty on the correction factors, although dominated by statistics (especially due to the small fraction of off-resonance luminosity), has to be taken as a systematic uncertainty on the corrected Monte Carlo $B\overline{B}$ distribution.

A second check can be performed. It uses the entirety of the $\pi^0(\eta)$ event sample that was the starting point for the determination of the corrections. Again,

the E^*_γ distributions can be compared before and after corrections are applied. After the correction, any differences should decrease considerably. Since this check uses the entire, regular $b \to s\gamma$ signal selection, without making antiveto cuts or reconstructing π^0s and ηs in any other way, this is essentially a regular $b \to s\gamma$ study on a different initial sample. Therefore, the signal region in the data sample must be kept blind, in order not to incur any bias. This method then yields a large control region from $1.0 - 1.9$ GeV below the signal region, in which data and (corrected) simulation can be compared. Figure 6.9 shows this comparison. The agreement between data and simulation improves when the corrections are applied. However, it is not complete yet, because there are still about 20% of the $B\overline{B}$ background uncorrected in the simulation. These other components are discussed in the following sections.

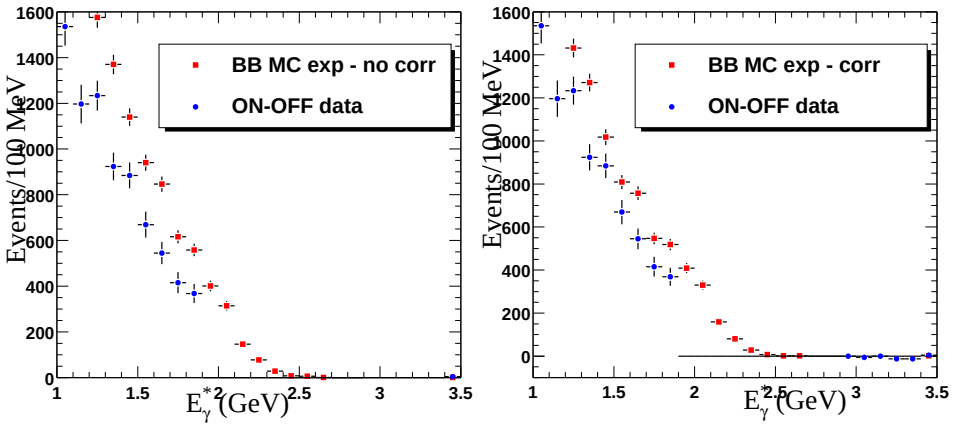

Figure 6.9: E^*_γ comparison of data and simulation in the $\pi^0(\eta)$ skim sample, before and after $\pi^0(\eta)$ corrections are applied to the simulated $B\overline{B}$ sample.

6.4 Studies of the Hadronic Component

6.4.1 Defining a Hadronic Control Sample

In order to study how well the simulated $B\overline{B}$ background represents hadronic particles that are faking a high-energy photon in the EMC, a hadronic control sample is defined by considering events that have been rejected by those photon quality cuts in Section 5.2 that show good rejection power for hadrons. Instead of rejecting such events, the opposite is done, and a sample of photons that do *not* pass those cuts is selected. This is effectively a hadronic "antiveto" sample. The primary cut that discriminates against faked photons is the requirement that the second moment of the EMC cluster's energy distribution be less than 0.002. To avoid overlap with the standard data sample, the hadronic control sample should at a minimum negate that cut. In fact, the second moment cut is also effective against merged π^0s, and negating that cut also, will leave these candidates in the sample. This is beneficial because merged π^0s are another category of faked photons. The sample could therefore also be called the "fake photon control sample". As already explained in Section 5.2, also the lateral moment of a cluster is sensitive to differences between electromagnetic and hadronic showers in the EMC. This was first observed by the Argus experiment in the context of electron identification [48]. The lateral moment is small when most of the shower's energy is in just two crystals, and becomes larger as the energy is more spread out. The lateral moment also is more uniformly defined across the EMC than is the

second moment[2]. This property of the lateral moment is exploited in the additional cut on the lateral moment that was added to the photon selection (Sec. 5.2). In the context of creating a hadronic control sample, the lateral moment's separation of true photons and hadrons is useful. An inverted lateral moment cut is added to the inverted second moment cut to define this control sample and to keep it as purely hadronic as possible. Where the antiveto cuts are to be set is determined from the following considerations. Figure 6.10 shows how photons from virtual Compton scattering ($e^+e^- \to e^+e^-\gamma$) that are precisely identifiable as photons in data, are distributed in second moment *vs.* lateral moment. Note that some real photons fail the standard second moment and lateral moment cuts (a point to be addressed in Section 8.1) these will end up in the control sample. According to χ^2 truth-matching information, about about 52.8% of the sample consists of true hadrons, 19.6% are merged π^0s, and the rest, 27.7%, consists of true photons. Once the sample is defined, the analysis is continued following the standard selection. This can again yield two distributions for the data–simulation comparison: one tagged sample with all analysis cuts applied and a second untagged sample that does not have the lepton-tagging cuts applied. Figure 6.11 shows the predictions for Monte Carlo truth-matched hadrons and merged π^0s which satisfy all event-selection cuts before tagging.

Based on these distributions, the hadronic (and merged-π^0) control sample

[2]The second moment is defined in a two-dimensional space of $\Delta\theta$ and $\Delta\phi$ relative to the cluster centroid, rather than using the distance from each crystal to the centroid. This implies a varying relationship of second moment to cluster shape across the EMC, and also a distorted relationship in the two dimensions (except near $\theta_{\rm lab} = 90\,{\rm deg}$). The lateral moment, on the other hand, *is* symmetrically defined.

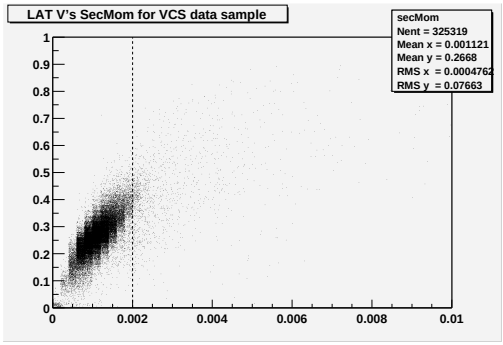

Figure 6.10: Lateral moment *vs.* second moment of EMC clusters for real photons from virtual Compton scattering. The vertical line shows the location of the second moment cut in the current analysis.

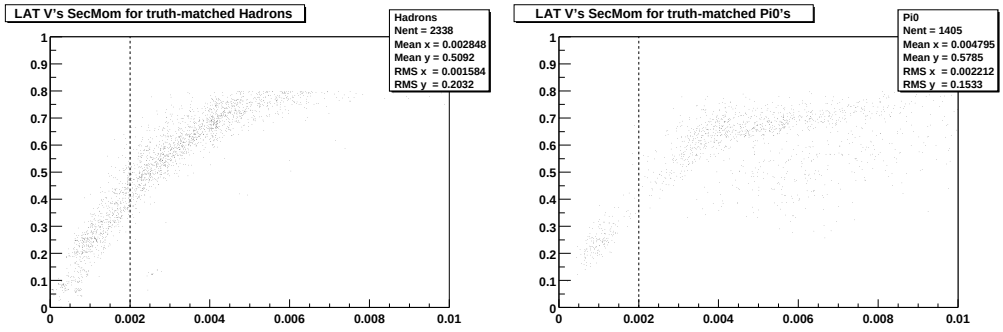

Figure 6.11: Lateral moment *vs.* second moment of EMC clusters for Monte Carlo truth-identified hadrons (left) and merged π^0s (right) in the $SP4$ $B\overline{B}$ sample. All standard selection cuts prior to tagging, *except* those on second moment and lateral moment, have been made. The vertical lines show the second moment cut. The upper cutoff on the lateral moment is an effect of the reconstruction criteria for a good photon as listed in Table 5.8.

requires a second moment > 0.002 and a lateral moment > 0.55. Figure 6.12 compares distributions of both quantities in such an antiveto *data* sample with expectations from the $SP4$ Monte Carlo simulation. For the latter, a normalized sum of all continuum plus $B\overline{B}$ simulations is used. Generally good agreement is found in the shape of these distributions. Although the absolute normalization is not in perfect agreement, this supports the validity of the hadronic antiveto method, because the same fraction of events will be found in the antiveto sample for both data and simulation. This is important, because valid corrections can only be obtained from the control sample if the results gained there can be correctly extrapolated into the space of the regular $b \to s\gamma$ selection.

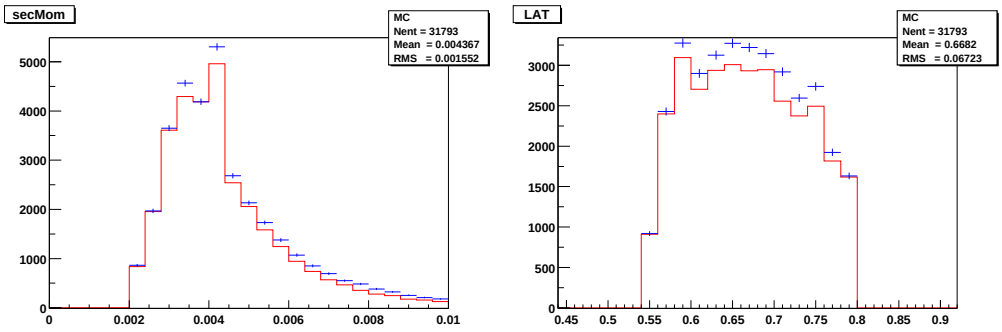

Figure 6.12: Comparison of EMC cluster shape quantities between $SP4$ simulation (red histograms) and on-resonance data (blue points with statistical errors) for the hadronic antiveto selections prior to tagging cuts. Shown are the second moment (left) and lateral moment variable (right).

E^*_γ(GeV))	$Cont(Off)$	$B\overline{B}$	$BkgTotal$	$Signal$	On
$1.7-1.8$	8.5 ± 8.5	72.7 ± 8.9	81.2 ± 12.4	0.0 ± 0.0	105.0 ± 10.2
$1.8-1.9$	34.1 ± 17.1	59.4 ± 8.1	93.6 ± 18.9	0.0 ± 0.0	85.0 ± 9.2
$1.9-2.0$	0.0 ± 0.0	47.3 ± 7.2	47.3 ± 7.2	0.7 ± 0.7	78.0 ± 8.8
$2.0-2.1$	0.0 ± 0.0	46.3 ± 7.1	46.3 ± 7.1	0.0 ± 0.0	58.0 ± 7.6
$2.1-2.2$	8.5 ± 8.5	33.0 ± 6.0	41.6 ± 10.4	0.0 ± 0.0	46.0 ± 6.8
$2.2-2.3$	17.1 ± 12.1	24.2 ± 5.2	41.2 ± 13.1	0.0 ± 0.0	32.0 ± 5.7
$2.3-2.4$	8.5 ± 8.5	8.8 ± 3.1	17.3 ± 9.1	0.7 ± 0.7	17.0 ± 4.1
$2.4-2.5$	0.0 ± 0.0	11.0 ± 3.5	11.0 ± 3.5	2.1 ± 1.2	6.0 ± 2.4
$2.5-2.6$	0.0 ± 0.0	4.4 ± 2.2	4.4 ± 2.2	0.0 ± 0.0	7.0 ± 2.6
$2.6-2.7$	0.0 ± 0.0	6.6 ± 2.7	6.6 ± 2.7	0.1 ± 0.0	6.0 ± 2.4
$2.7-2.8$	0.0 ± 0.0	3.3 ± 1.9	3.3 ± 1.9	0.7 ± 0.7	4.0 ± 2.0
$2.8-2.9$	0.0 ± 0.0	1.1 ± 1.1	1.1 ± 1.1	0.0 ± 0.0	1.0 ± 1.0
$2.9-3.0$	0.0 ± 0.0	2.2 ± 1.6	2.2 ± 1.6	0.0 ± 0.0	3.0 ± 1.7
$3.0-3.1$	0.0 ± 0.0	1.1 ± 1.1	1.1 ± 1.1	0.0 ± 0.0	1.0 ± 1.0
$3.1-3.2$	0.0 ± 0.0	1.1 ± 1.1	1.1 ± 1.1	0.0 ± 0.0	3.0 ± 1.7
$3.2-3.3$	8.5 ± 8.5	1.1 ± 1.1	9.6 ± 8.6	0.0 ± 0.0	2.0 ± 1.4
$3.3-3.4$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 1.0
$3.4-3.5$	0.0 ± 0.0	1.1 ± 1.1	1.1 ± 1.1	0.0 ± 0.0	1.0 ± 1.0
$Total$	85.3 ± 27.0	330.2 ± 19.1	415.5 ± 33.0	4.3 ± 1.7	459.0 ± 21.4

Table 6.4: Expectations from $SP4$ $B\overline{B}$ simulation and observation in data for the hadronic/merged-π^0 antiveto sample described in the text. This study was done on $54.6\,\mathrm{fb}^{-1}$ of on-resonance data. The comparison is valid as is, but to extract event yields for the current case, all entries have to be scaled by the ratio of luminosities: 81.5/54.6. $Cont(Off)$ is the continuum background from off-resonance data. $B\overline{B}$ is the $B\overline{B}$ background from $SP4$ MC. $BkgTotal = Cont(Off) + B\overline{B}$. $Signal$ is the expected signal using $\mathcal{B}(B \to X_s\gamma) = 3.45 \times 10^{-4}$. On is on-resonance data. All errors are statistical.

6.4.2 Data *vs.* Monte Carlo Expectations for the "Hadronic" Sample

Table 6.4 shows the E^*_γ distributions after tagging for hadronic antiveto selections on the various Monte Carlo and data samples. In particular, it shows the comparison between on-resonance data and the expectation from $SP4$ Monte Carlo (for $B\overline{B}$ plus a small signal contribution) and off-resonance data (for the continuum

component). Contributions from the $B \to X_s\gamma$ signal that feed into this control sample are not significant (this is different for the antiveto sample that is used for the systematic check of the $\pi^0(\eta)$ corrections obtained with the dedicated selection, see Section 6.3.3). Figure 6.13 compares subtracted hadronic antiveto data (on-resonance data minus off-resonance data minus simulated signal) after tagging cuts are applied to $SP4$ Monte Carlo $B\overline{B}$ predictions. The simulation shows a much softer spectrum than that measured in data. Therefore, any corrections will depend strongly on E^*_γ.

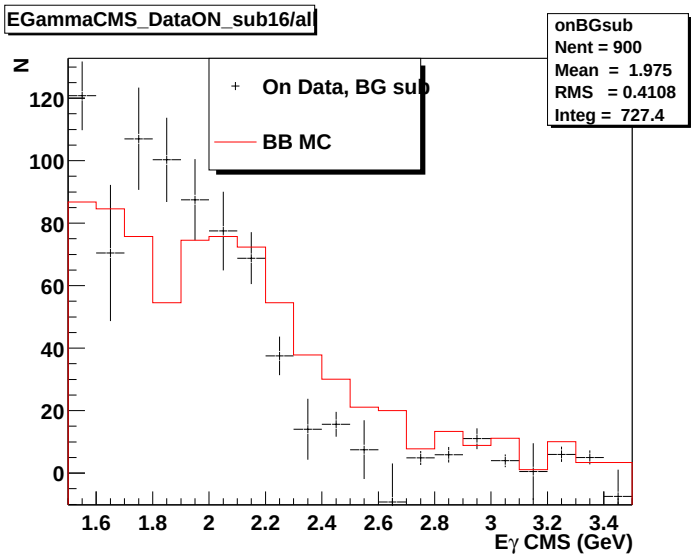

Figure 6.13: Comparison of subtracted data to $SP4$ $B\overline{B}$ simulation for tagged hadronic control samples. The data points show on-resonance data that pass the antiveto cuts, minus the corresponding off-resonance data scaled to the same luminosity, minus a small expectation from $B \to X_s\gamma$ signal simulation. The histogram shows the sum of $B^0\overline{B}^0$ and B^+B^- simulations, normalized to the data's luminosity (This study was done on $54.6\,\mathrm{fb}^{-1}$ of on-resonance data which enters the overall normalization. The comparison and derivation of corrections is valid, but to extract event yields all entries have to be scaled by the ratio of luminosities: 81.5/54.6.)

To extract effective correction factors for the hadronic component of the $B\overline{B}$ background, a ratio analogous to equation 6.1 is defined, although here this is taken in bins of E^*_γ (the corrections can then be directly applied to the hadronic part of the $B\overline{B}$ photon spectrum after all $b \to s\gamma$ analysis cuts have been applied). The obtained ratios are listed in the second column of Table 6.5, which shows the ratio of the event counts of subtracted data over $B\overline{B}$ Monte Carlo in each bin, after the lepton-tags have been applied. There is a clear trend of the values falling with increasing E^*_γ.

E^*_γ Range	Hadron Ratios
2.0-2.1 GeV	1.024 ± 0.208
2.1-2.2 GeV	0.951 ± 0.165
2.2-2.3 GeV	0.688 ± 0.150
2.3-2.4 GeV	0.371 ± 0.266
2.4-2.5 GeV	0.521 ± 0.167
2.5-2.6 GeV	0.355 ± 0.451
2.6-2.7 GeV	-0.450 ± 0.638

Table 6.5: Ratios of subtracted "hadronic" antiveto tagged data to $SP4$ $B\overline{B}$ simulation, determined for E^*_γ between 2.1 and 2.7 GeV. The "Hadron Ratio" is defined as in eq. 6.1, and in this case totals $(80.2 \pm 20.3)/(90.8 \pm 10.0)$. Quoted uncertainties are statistical only.

6.5 Other $B\overline{B}$ Contributions

With the $\pi^0(\eta)$ and the fake photon components discussed in the last sections, the large majority of the total $B\overline{B}$ background (as modeled in the simulation) already has corrections for differences between data and simulation. The rest of the $B\overline{B}$ background is expected to be around 5-6% (see Table 6.1) with the largest con-

tributions from ω and η' decays. Since any mismodeling in $B\overline{B}$ components of such small fractions can have only minimal effects on the overall $B\overline{B}$ uncertainty, no dedicated selections are studied here. It is useful, however, to look at these components and verify to which degree the simulation is trustworthy. If exact correction factors cannot be determined, at least a systematic uncertainty due to these backgrounds will be assigned.

First, inclusive ω decays are considered. They have been studied by an analysis group in BABAR [64]. The yield as a function of the ω momentum in the CMS frame, p^*_ω, is investigated in both on- and off-resonance data and in the $B\overline{B}$ simulation. The data samples are subtracted and the ratio of the data over $B\overline{B}$ Monte Carlo yield is measured in 10 bins of p^*_ω, covering 0.0-3.0 GeV. All bins are 250 MeV wide, except for the last (lowest statistics) bin, which has triple this size. The ω corrections are listed in Table 6.6 and also shown in Figure 6.14. For the present analysis, only the high end of this spectrum will be relevant due to the high photon momenta of interest here. Unfortunately, in that region the ratios show the largest uncertainties. Applying them will therefore be similar to assigning a large systematic error on the omega fraction in the $B\overline{B}$ simulation.

The last source of background photons in $B\overline{B}$ decays to be considered individually is from direct photons in η' decays. Again, use is made of an existing BABAR study of inclusive η' decays [65]. The η' yield is studied in bins of the reduced η' momentum in the CMS frame, $x' = p^*_{\eta'}/\sqrt{E^2_{beam} - m^2_{\eta'}}$. This variable is used because

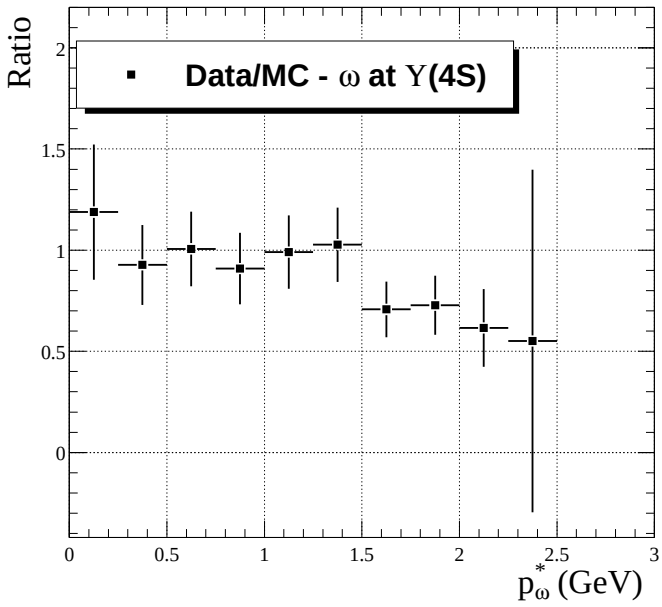

Figure 6.14: Ratio of inclusive ω yields in data and simulation *vs.* CMS ω momentum, p^*_ω for an inclusive sample of ω decays as measured in [64], the plot is taken from that document. Note that the bin with the last entry at the upper end of the spectrum really should span the entire range from 2.25 to 3 GeV.

$p^*_\omega-$ Range	ω Correction Factors
0.00−0.25	1.1883 ± 0.3340
0.25−0.50	0.9275 ± 0.1974
0.50−0.75	1.0067 ± 0.1845
0.75−1.00	0.9089 ± 0.1765
1.00−1.25	0.9909 ± 0.1815
1.25−1.50	1.0274 ± 0.1837
1.50−1.75	0.7075 ± 0.1371
1.75−2.00	0.7284 ± 0.1462
2.00−2.25	0.6155 ± 0.1926
2.25−3.00	0.5513 ± 0.8460

Table 6.6: Correction factors for the ω component of the generic $B\overline{B}$ background simulation as a function of ω CMS frame momentum. Note the one much larger bin at the high momentum end of the spectrum.

it is not sensitive to the differences in beam-energy between on- and off-resonance data taking. The inclusive branching fraction of B mesons into η' states is determined in terms of x' by subtracting the measured yields in on- and off-resonance data, but no differences with respect to the simulation are studied. Because the fraction of η' on the total $B\overline{B}$ background is so small, it is sufficient here to determine the simulated branching fraction of B mesons into η' states (in bins of x') directly from the generator level information in a Monte Carlo sample of 6 million generic B^+B^- and $B^0\overline{B}^0$ events. Correction factors can then be determined in each bin as the ratio of the branching fraction in data as measured in [65] and the branching fraction in the simulation as determined from the generator level information. The corrections must then be applied to the $B\overline{B}$ background simulation for true η' parent particles depending on their x' values. Only two bins are used in the data measurement. They can

be seen in Table 6.7 which lists the ratios and their error. The latter is dominated by the uncertainty of the data measurement. Again, applying these corrections will effectively mean to include a large systematic error on the η' fraction of the $B\overline{B}$ simulation, but this will only be a small effect in the overall correction of the modeling of $B\overline{B}$ background. It is therefore fully sufficient for the present study.

x'-Range	η' Correction Factors
0.10−0.39	0.488 ± 0.147
0.39−0.52	1.771 ± 0.579

Table 6.7: Correction factors for simulated inclusive η' simulation in bins of the reduced η' momentum. Corrections are deduced from the branching fractions for $B \to \eta' X$ in data and simulation (on generator level). The data values are taken from [65].

Sources of high-energy background photons from $B\overline{B}$ decays other than the ones discussed individually above comprise less than 2% of the total expected. No correction factors are applied to them, but large errors (of ±100%) are assigned to each such photon in the simulation in the finally selected sample.

6.6 Corrected Overall $B\overline{B}$ Background

The derived correction factors (or systematic errors) are applied as weights to all components of the $B\overline{B}$ simulation depending on the photon parent and the value of its characteristic kinematic variable (as determined from generator level information). This is effectively a tune of the high-energy photon production in the generic $B\overline{B}$

simulation to the levels observed in complementary data sets.

The corrections are applied to the simulated $B\overline{B}$ background sample that is left after all cuts of the regular $b \to s\gamma$ event selection are applied. The E^*_γ spectrum for the such tuned $B\overline{B}$ background events in the the regular $b \to s\gamma$ analysis is shown in the left plot in Figure 6.15 in comparison with the subtracted data distribution. . To get a sense for the magnitude of the corrections applied, this spectrum can be compared to the original uncorrected spectrum. The ratio between the two, *i.e.* the net correction in each E^*_γ bin (which has contributions from all components discussed above) is shown in the right plot of Figure 6.15. From the total event counts in both samples also an overall correction factor of 0.924 ± 0.097 can be obtained.

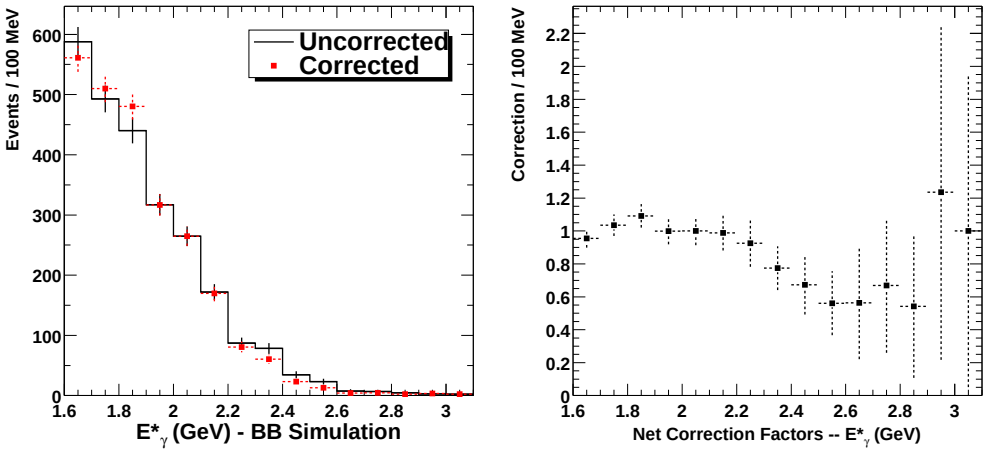

Figure 6.15: Corrected and uncorrected $B\overline{B}$ background spectra after all $b \to s\gamma$ analysis cuts on the regular $b \to s\gamma$ event selection sample (left plot) and net correction factors in bins of E^*_γ (right plot) obtained as ratios of corrected over uncorrected spectra.

The performance of the corrections is then evaluated by comparing the cor-

rected $B\overline{B}$ simulation to the real off-resonance subtracted data yield in the control region of 1.7-1.9 GeV, below the signal region. This comparison is shown in Figure 6.16. Reasonable agreement is seen.

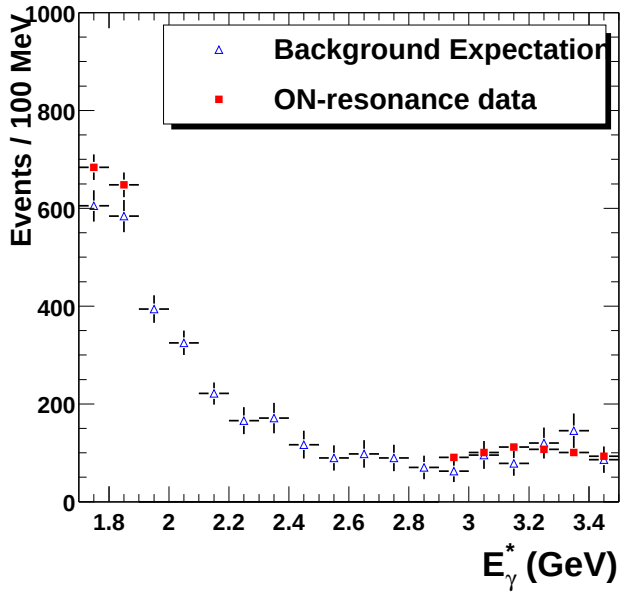

Figure 6.16: E^*_γ distribution for the fully corrected $B\overline{B}$ simulation in comparison to on-resonance data. The expected signal region is kept blind, leaving only the low energy control region for the comparison.

The tuned $B\overline{B}$ background spectrum from the simulation is now ready to estimate the data $B\overline{B}$ background component and to be subtracted from the (continuum subtracted) data spectrum to yield the raw $b \to s\gamma$ signal spectrum result.

Chapter 7

Comparisons of Data to Expectations

7.1 Summary of Data and Monte Carlo Expectations

Table 7.1 repeats the $SP4$ Monte Carlo expectations for the number of signal and background events in 100 MeV E^*_γ bins, that were already presented in Table 5.16. However, here also the continuum expectation as obtained from the scaled off-resonance data is included, and the observed number of on-resonance data events (still blinded in the region where most of the signal is expected). All expectations are derived using the cross-sections and luminosities given in Table 3.2. The $B\overline{B}$ background numbers presented here include the corrections derived in the previous chapter, their errors, which now include the uncertainties from the correction, will be treated

as a systematic error in the next chapter. The quoted background total is the sum of the corrected $B\overline{B}$ Monte Carlo plus off-resonance continuum, in line with the strategy for deriving the signal, as outlined in Section 4.1.

7.2 Off-resonance Data *vs.* Continuum Monte Carlo Predictions

Even though the continuum Monte Carlo sample is not used for the final subtraction of the continuum background, it *is* used for event selection studies and cut optimization. A comparison with off-resonance data is of interest not only for this reason, but because it can clarify aspects of the simulation which might otherwise be difficult to assess directly. Table 7.1 shows that the data are systematically higher than the Monte Carlo predictions by roughly a factor of two to three.

Note that if the actual number of off-resonance data events is used in the computation of the statistical significance described in Section 5.9, the expected statistical uncertainty on the extracted signal increases from $\approx 6.3\%$ to $\approx 8.5\%$ for the chosen E^*_γ range of 2.0 to 2.7 GeV. This assumes the value of the signal branching fraction that is predicted by the theory (see Section 2.3.1). The two other energy ranges that will be considered in this thesis (see Section 9.1) are 1.9 to 2.7 GeV and 2.1 to 2.7 GeV. Using them would result in estimated statistical uncertainties of 9.0% and 8.3%, respectively, with the off-resonance numbers used in the computation. The

$E^*_\gamma(GeV)$	$Cont(MC)$	$Cont(Off)$	$B\overline{B}$	$BkgTotal$	$signal$	On
$1.6-1.7$	21.2 ± 5.7	76.5 ± 25.5	563.2 ± 45.3	639.7 ± 30.0	1.9 ± 0.6	756.0 ± 27.5
$1.7-1.8$	25.1 ± 6.2	93.5 ± 28.2	511.7 ± 45.5	605.1 ± 32.0	3.1 ± 0.8	684.0 ± 26.2
$1.8-1.9$	12.8 ± 4.3	102.0 ± 29.4	482.1 ± 49.8	584.1 ± 32.9	6.5 ± 1.2	648.0 ± 25.5
$1.9-2.0$	17.6 ± 5.1	76.5 ± 25.5	317.7 ± 36.9	394.2 ± 28.1	13.8 ± 1.7	*blind*
$2.0-2.1$	16.3 ± 4.8	59.5 ± 22.5	265.6 ± 40.1	325.1 ± 25.0	28.6 ± 2.5	*blind*
$2.1-2.2$	16.1 ± 4.9	51.0 ± 20.8	170.3 ± 36.6	221.3 ± 22.6	58.3 ± 3.6	*blind*
$2.2-2.3$	14.9 ± 4.8	85.0 ± 26.9	80.9 ± 31.5	165.9 ± 27.5	97.8 ± 4.6	*blind*
$2.3-2.4$	20.3 ± 5.7	110.5 ± 30.6	60.7 ± 8.9	171.2 ± 31.1	153.7 ± 5.6	*blind*
$2.4-2.5$	19.1 ± 5.4	93.5 ± 28.2	23.4 ± 3.2	116.9 ± 28.4	170.0 ± 5.5	*blind*
$2.5-2.6$	27.6 ± 6.6	76.5 ± 25.5	13.0 ± 2.4	89.5 ± 25.6	118.7 ± 4.3	*blind*
$2.6-2.7$	43.8 ± 8.3	93.5 ± 28.2	4.3 ± 1.4	97.8 ± 28.2	62.3 ± 2.7	*blind*
$2.7-2.8$	25.6 ± 6.3	85.0 ± 26.9	4.5 ± 1.4	89.5 ± 26.9	11.9 ± 0.7	*blind*
$2.8-2.9$	40.1 ± 8.1	68.0 ± 24.0	2.4 ± 1.0	70.4 ± 24.1	0.5 ± 0.2	*blind*
$2.9-3.0$	37.6 ± 7.6	59.5 ± 22.5	3.3 ± 1.2	62.8 ± 22.5	0.0 ± 0.0	91.0 ± 9.5
$3.0-3.1$	41.8 ± 8.0	93.5 ± 28.2	2.2 ± 1.0	95.7 ± 28.2	0.0 ± 0.0	101.0 ± 10.0
$3.1-3.2$	38.9 ± 7.9	76.5 ± 25.5	2.2 ± 1.0	78.7 ± 25.5	0.0 ± 0.0	112.0 ± 10.6
$3.2-3.3$	48.3 ± 8.8	119.0 ± 31.8	0.9 ± 0.6	119.9 ± 31.8	0.0 ± 0.0	107.0 ± 10.3
$3.3-3.4$	37.5 ± 7.4	144.5 ± 35.0	0.9 ± 0.6	145.4 ± 35.0	0.0 ± 0.0	101.0 ± 10.0
$3.4-3.5$	24.7 ± 6.1	85.0 ± 26.9	0.9 ± 0.6	85.9 ± 26.9	0.0 ± 0.0	93.0 ± 9.6
$Total$	529.2 ± 28.6	1648.9 ± 118.4	2510.2 ± 109.6	4159.0 ± 123.0	727.1 ± 11.5	5649.0 ± 75.2

Table 7.1: Expectations for $81.5\,\mathrm{fb}^{-1}$ using $SP4$ Monte Carlo and off-resonance data, after all cuts except that on E^*_γ, compared to on-resonance data. $Cont(MC)$ is the continuum background estimated from $SP4$ MC. $Cont(off)$ is the continuum background from off-resonance data. $B\overline{B}$ is the corrected $B\overline{B}$ background from $SP4$ MC. $BkgTotal$ is the sum of expected $B\overline{B}$ plus off-resonance continuum data. (Note: in table 5.16, this is purely MC.) *Signal* is the expected signal using the KN480 model plus the prescribed 16.1% $B\to K^*\gamma$ contribution. On is on-resonance data. All errors are statistical.

predictions of the continuum Monte Carlo yield 6.0% and 5.5% in the same ranges. Note that none of these results allow for uncertainties due to the limited $B\overline{B}$ Monte Carlo statistics. The latter is treated as a systematic uncertainty. Using the numbers from the $B\overline{B}$ Monte Carlo simulation without applying the corrections discussed in the previous chapter results in an increase of less than 0.1% in the significance estimates.

7.2.1 Studies of Off-resonance Events *vs.* Predictions

To look into the off-resonance data *vs.* continuum Monte Carlo comparison more closely, the region between 2.0 and 2.7 GeV is studied. For this task, the tagging momentum cuts can either be kept at their values used in the analysis, to correspond most closely to the final analysis sample, or they can be relaxed to $p_e^* > 1.0\,\mathrm{GeV}$, $p_\mu^* > 1.1\,\mathrm{GeV}$, to provide better statistics. Table 7.2 summarizes the results.

	Off-Resonance Data / Simulation		
	All Tags	Electrons	Muons
Before event shape cuts	1.140 ± 0.002		
After event shape cuts	1.216 ± 0.006		
After p_e^*/ p_μ^* cut	1.59 ± 0.07	1.74 ± 0.11	1.44 ± 0.07
	(1.54 ± 0.05)	(1.68 ± 0.08)	(1.43 ± 0.07)
After $\cos\theta_{\gamma e}$/ $\cos\theta_{\gamma\mu}$ cut	2.42 ± 0.33	2.47 ± 0.41	2.33 ± 0.55
	(1.96 ± 0.14)	(2.20 ± 0.21)	(1.72 ± 0.18)
After missing-E cut	2.52 ± 0.37	2.45 ± 0.44	2.70 ± 0.70
	(1.95 ± 0.14)	(2.12 ± 0.21)	(1.77 ± 0.19)

Table 7.2: Ratios of off-resonance data over predictions of the $SP4$ continuum simulation, showing the effect of the event shape and lepton tagging cuts. The simulation is normalized to the off-resonance luminosity of $9.59\,\mathrm{fb}^{-1}$. For the numbers in parentheses, the tagging momentum cuts are relaxed to $p_e^* > 1.0$ GeV, $p_\mu^* > 1.1$ GeV. All errors are statistical only.

It stands out that there is a discrepancy (about 14%) already *before* any event shape or lepton-tagging cuts. It is increased by the event shape cuts to about 22%. When applying the lepton-tags, the agreement deteriorates noticeably with each cut (momentum, angle and missing energy), and the simulation underestimates the data by a factor of 2.5 after all cuts. Two separate effects are believed to cause this behavior.

First, much of the initial discrepancy is believed to be due to inadequacies in that part of the continuum cross section which results in a high-energy photon. To be more specific, the majority of the continuum events in the sample (see Tables 5.3 and 5.5) are expected to be from ISR. For such events, the the predictions the simulation depends upon the implementation of the cross sections and form factors at e^+e^- center-of-mass energies near 8 GeV (instead of the values at the $\Upsilon(4S)$ energy) since a high-energy photon of about 2 GeV has already been radiated off (this is not implemented correctly and usually does not influence analyses). The discrepancy increases slightly between E^*_γ of 2.0 and 2.7 GeV, but the variation is contained within $\approx \pm 3\%$. The event shape cuts (especially the requirement on the Fisher discriminant) then increase these differences because they enrich the ISR component in the samples.

The second effect is related to the lepton tags. For events in which a lepton candidate is found, a possible dependence on the lepton's momentum in the CMS frame, p^*, is studied. Figure 7.1 compares off-resonance data to the continuum simulation just before the tagging momentum cuts, and shows the momentum distributions

at the top, their ratios at the bottom. While the ratio for muon tags shows no clear pattern, the ratio for electron tags increases with p_e^*. Thus, in the standard selection (with a cut of $p_e^* > 1.25\,\mathrm{GeV}$) the discrepancy between data and simulation is larger than in the relaxed cut sample (see the numbers in parentheses in Table 7.2).

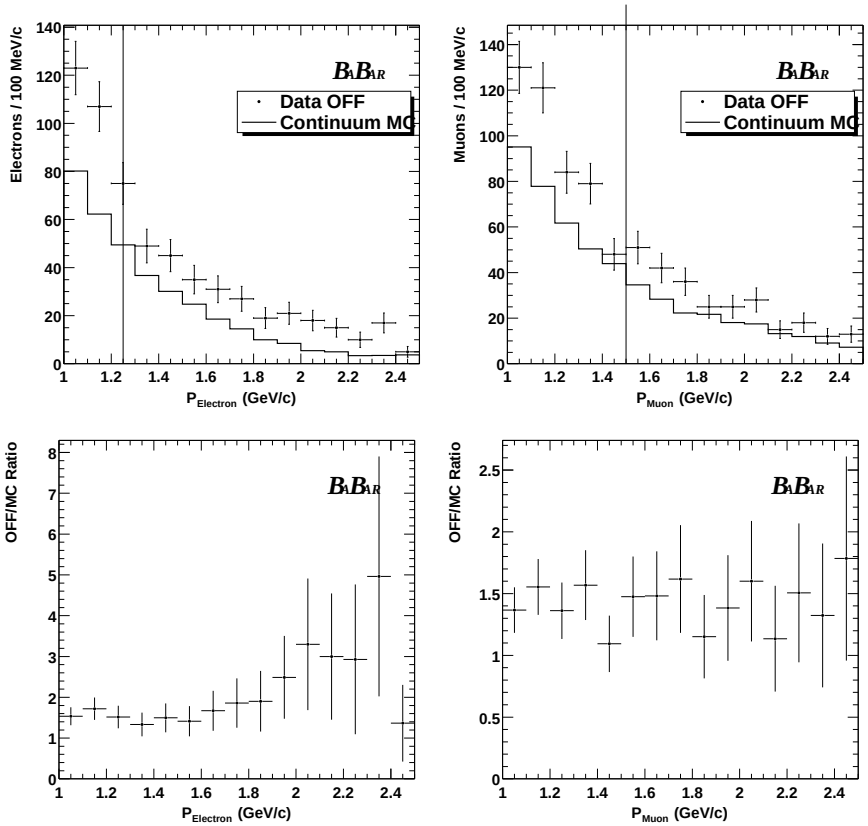

Figure 7.1: Comparison of off-resonance data to $SP4$ continuum simulation (normalized to the off-resonance data luminosity) as a function of lepton momentum, just before lepton-tagging momentum cuts: electrons (left), muons (right). The corresponding data/simulation ratios are shown below the spectra.

It was suggested [66] that for a basic selection like the one used here, which is based on simple multihadronic requirements plus the presence of a high-energy

photon, pure-QED processes can survive subsequent cuts and show up enriched in a tagged electron sample. Since these processes are not included in the simulation of quark–anti-quark pairs, this can contribute to the excess of events in the data sample. The Argus experiment found a "magic" number (an effective multiplicity), $N_{\rm Magic} = N_{\rm tracks} + N_{\rm photons}/2$, to yield good discrimination of these events. Figure 7.2 compares distributions of this quantity just after an electron-momentum cut at 1.0 GeV. The small peak at the very lowest values is evidence for such a pure QED contribution, which is of course not allowed for in the simulation. When a minimum "Nmagic" value of 6 is required, the data/MC ratio for *all* tags after all cuts decreases from about 2.52 ± 0.37 to 1.78 ± 0.33. However, such a cut is *not* adopted into the regular $b \to s\gamma$ analysis because it could introduce a multiplicity-dependence (and thus a dependence on the fragmentation-model) into the signal efficiency estimation.

While this discrepancy might result in non-optimal cut values from the optimization procedure (which relies on the Monte Carlo simulation of the continuum), it does *not* introduce any bias on the final result, since off-resonance data is used for the subtraction of continuum background.

7.2.2 Other Possible Studies

To cross-check the comparisons of off-resonance data to $SP4$ Monte Carlo continuum predictions, use could also be made of the fully orthogonal samples obtained by using the $\pi^0(\eta)$ antiveto selection. This ansatz will lessen problems introduced by

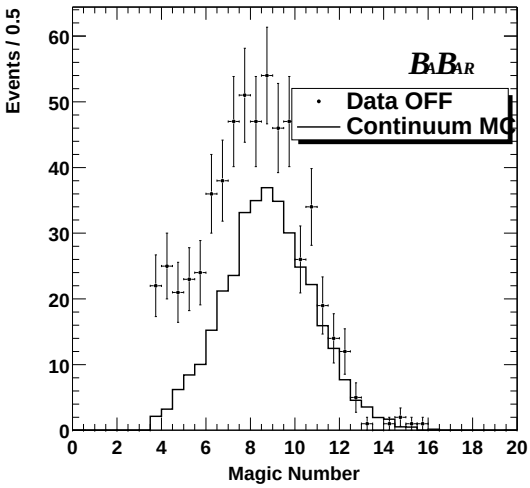

Figure 7.2: "Magic number" distribution for off-resonance data (points) *vs.* $SP4$ continuum simulation (histogram, normalized to the same luminosity), after an electron-momentum cut of p_e^* above 1.0 GeV.

insufficiently well modeled ISR events, because the sample explicitly reconstructs π^0s and ηs (although they may be combinatoric and contain an ISR photon that way). The study should be done before any lepton-tagging cuts are applied to circumvent the accumulation of the pure QED events mentioned in the previous section.

This was studied on an event sample of 54.6 fb^{-1}: The ratio of data to simulation integrated over 2.0 GeV $< E_\gamma^* <$ 2.7 GeV came to 1.02 ± 0.01, with negligible E_γ^*-dependence. This compares to 1.140 ± 0.002 for the standard photon selection. (The second sample may still contain some QED events, but before lepton-tagging this is not expected to be a significant effect.) This result confirms the ansatz for understanding the differences between off-resonance data and simulation. Since in the $\pi^0(\eta)$ sample data and simulation are essentially in agreement, the main differences

in the regular sample before tagging *do* seem to stem from the way the production of high-energy photons is modeled in the simulation, especially that of ISR events (for which energy effects also play a role).

7.3 On-resonance Comparisons Outside the Signal Region

Table 7.1 also presents the actual on-resonance data along with the expectations from Monte Carlo in the same bins, leaving blind the region where most of the signal is expected. The low energy bins $1.7 < E^*_\gamma < 1.9$ GeV are dominated by $B\overline{B}$ background and so provide a control region for evaluating the accuracy of the prediction. The high-energy region $E^*_\gamma > 2.9$ GeV is exclusively populated by continuum background and so provides another control region. Fig 7.3 also shows this comparison of on-resonance data to the expected background. The errors are dominated by statistics, due to limited off-resonance data. It can be seen that within errors there is agreement in both these regions, between the on-resonance data and the $B\overline{B}$ Monte Carlo plus off-resonance data expectation. In particular, the check yields

$$\text{Data} - \text{expectation} = 142 \pm 86 \text{ events} \quad (1.7\,\text{GeV} < E^*_\gamma < 1.9\,\text{GeV}) \; . \tag{7.1}$$

This includes the small contribution from signal, which is expected to consist of no more than 10 events for the assumed branching fraction in the $SP4$ simulation. If this outcome is assigned to the understanding of the total $B\overline{B}$ contribution, it translates

to a $14.3 \pm 8.5\%$ underestimate by the corrected $B\overline{B}$ $SP4$ simulation *in this region.* Within errors this agrees with the 7% error on the $B\overline{B}$ corrections that will be discussed in Section 8.5. The deviation from the latter is not entirely unexpected, because the derived $B\overline{B}$ corrections were calculated primarily for an energy range of the photon of about $2.0 - 2.7$ GeV.

On the high end of the spectrum (2.7-2.9 GeV) the continuum yield in off-resonance data can be compared to the same in on-resonance data. No background from $B\overline{B}$ events is expected in this region. Any shift between the two distributions would point toward a systematic error associated with the assumption that the $u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}$ background at the $\Upsilon(4S)$ energy can be estimated from continuum events taken in a lower CMS energy system. Looking at Table 7.1 and Figure 7.3, the measured and expected contributions are found in agreement. However, the limitations of this comparison are clear: the region has very limited statistics which complicates any conclusion of significance.

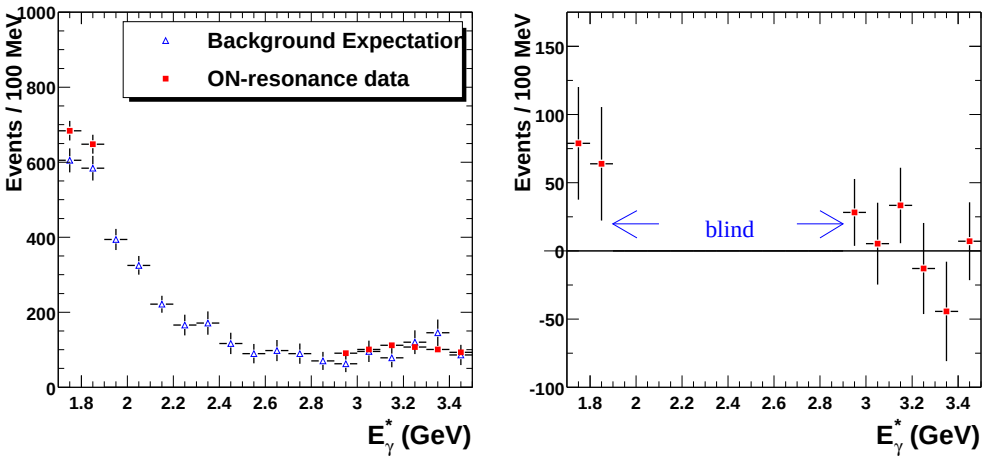

Figure 7.3: The E^*_γ distribution in on-resonance data compared to expectations as presented in Table 7.1. The expected signal region is kept blind. Note that the signal expectation ($SP4$ simulation) from the table has not been added in. An excess in the on-resonance spectrum is therefore expected in the signal. The left plot shows the two spectra, while the right plot shows the data minus the expectation (the measured signal).

Chapter 8

Systematic Errors and Checks

A summary of the considered systematic uncertainties is given in Table 8.1. The first column references the section where the details of the individual systematics are discussed. All errors are expressed as percentages of $\mathcal{B}(B \to X_s\gamma)$.

Sub-Section	Systematic	$\%\mathcal{B}(B \to X_s\gamma)$
8.1	Photon Selection	3.6
8.2	Event shape cuts	0.0
8.3	Lepton tags	2.0
8.4	Normalization	1.1
8.5	BB background estimation	5.1
8.6	Signal efficiency	4.3
	Total	7.9

Table 8.1: The systematic error expressed as a percentage of $\mathcal{B}(B \to X_s\gamma)$ for $2.0 < E^*_\gamma < 2.7\,\mathrm{GeV}$. The left column gives the subsection in which the estimation of this error is discussed in detail. The total is the quadratic sum.

This table does not contain the uncertainty that arises from the dependence of

the results on the choice of the theoretical model; this is discussed in Section 8.7. This uncertainty is not counted with the systematic errors due to the analysis technique, but quoted as a separate (theoretical) error on the extracted branching fraction.

8.1 Photon Selection Systematics

The photon selection implemented in this analysis is identical to that used in the published $B \to K^*\gamma$ measurement [18], with the exception of the energy-dependent π^0 and η vetoes and the addition of the lateral moment cut. Therefore the numbers found in that study are largely taken over for this analysis. The estimation of the systematic uncertainties associated with the selection is discussed below. The uncertainties are given as a percentage of $\mathcal{B}(B \to X_s\gamma)$ and are summarized in Table 8.2.

Uncertainty	$\%Br(B \to X_s\gamma)$
Photon efficiency	± 2.5
Second moment cut	± 0.8
Lateral moment cut	± 0.9
Energy scale	± 0.5
Energy resolution	± 0.5
Photon isolation cut	± 2.0
π^0/η veto	± 1.0
Total	± 3.6

Table 8.2: The systematic uncertainty in $\mathcal{B}(B \to X_s\gamma)$ associated with the photon selection.

The photon efficiency is checked by considering a sample of $\tau^+\tau^-$ events in which one tau decays to a charged hadrons and one or two π^0s [45]. The ratio of the

observed number of decays with one and two π^0 in data is compared to that measured in the Monte Carlo simulation (equation 8.1) and is measured as a function of the π^0 energy.

$$R = \frac{N^{data}(\tau \to h^{\pm}\pi^0\pi^0)/N^{data}(\tau \to h^{\pm}\pi^0)}{N^{MC}(\tau \to h^{\pm}\pi^0\pi^0)/N^{MC}(\tau \to h^{\pm}\pi^0)} \tag{8.1}$$

The photon efficiency influences all these event counts since the π^0s are reconstructed from their decay to two photons. If all effects that can change the efficiency are not modeled well in the simulation, the above ratio will be different from one. Both energy resolution and reconstruction differences between data and Monte Carlo can influence the simulated efficiency. To reach better agreement, the simulated efficiency is then adjusted by "smearing" the photon energy and "killing" a certain percentage of photons in the simulation. The systematic uncertainty of the correction procedure (due to resolution effects, relative branching fraction differences, the modeling of the detector response and occupancy differences between τ and $B\overline{B}$ events) gives a 2.5% (systematic) uncertainty on the efficiency to reconstruct a photon.

The second moment cut is checked by comparing data and Monte Carlo samples of photons from virtual Compton scattering (VCS): $e^+e^- \to e^+e^-\gamma$. This sample was already introduced in Section 5.2 and 6.4.1. The efficiencies for a cut at 0.002 are 0.988 and 0.996, in data and Monte Carlo, respectively. No correction is applied, but the difference (0.8%) is assigned as systematic uncertainty.

The same is done for the maximum lateral moment cut of 0.45. The VCS data sample yields an efficiency of 0.986, whereas the VCS simulation finds 0.995.

Again the difference (0.9%) is assigned as the error.

For the photon energies of interest, the calibration of the calorimeter is performed as a function of the energy and polar angle ($\cos\theta$) in the lab frame [35]. It uses the electrons from Bhabha decays as described in Section 3.2.5. The resulting energy scale of the calibrated EMC has been checked with symmetric $\eta \to \gamma\gamma$ decays and also a sample of photons from virtual Compton scattering (VCS) [46]. In the η sample, both photons are required to pass the photon selection criteria of Section 5.2 and to be in the same 500 MeV energy bin in the range $0.0 < E_{lab} < 3.0$ GeV. The invariant mass is compared to the nominal η mass [19]. Any deviation in mass is directly proportional to a deviation in energy scale. An uncertainty of 0.5% has been found. The dominant effect of a small shift in the energy scale is to change the signal efficiency of the energy cut (see Fig. 5.6 in Section 5.4). The left plot in Figure 8.1 shows how the efficiency depends upon the energy scale. An uncertainty of 0.5% in energy scale results in about 0.5% uncertainty in efficiency. The $B\overline{B}$ background in this analysis is significantly more sensitive to the energy scale because it depends so steeply on the measured energy (a roughly exponential increase with falling energy). A shift in energy scale of 0.5% results in a 5% uncertainty in the $B\overline{B}$ background estimate. However, no separate systematic error is added for this effect, since the overall uncertainty on the $B\overline{B}$ simulation (including the corrections) is derived from the comparison of data with Monte Carlo and is already sensitive to any shift in energy scale that might be present. An uncertainty of 0.5% on the photon efficiency is assigned in Table 8.2.

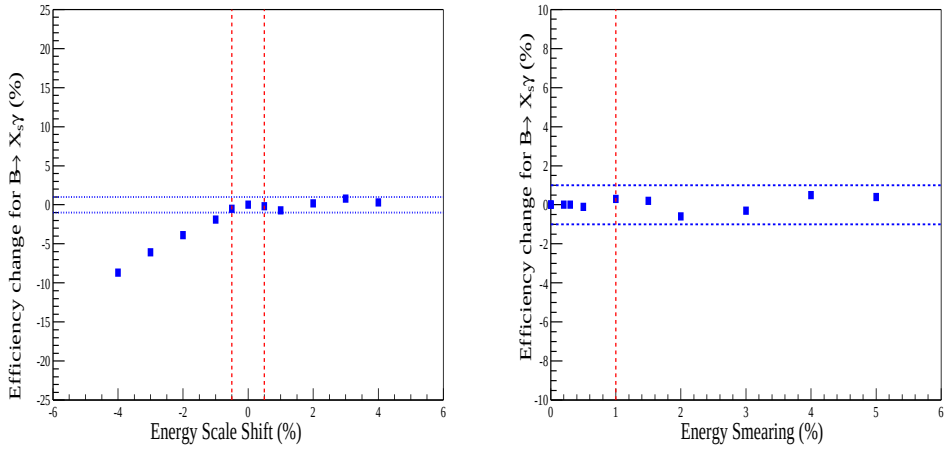

Figure 8.1: The change in efficiency for signal $B \to X_s\gamma$ (X_s plus K^*) $SP4$ Monte Carlo versus energy scale uncertainty (on the left) and additional energy resolution smearing (on the right). The measured 0.5% uncertainty in energy scale results in a 0.5% uncertainty in signal efficiency (left). A $1-3\%$ smearing in energy resolution results in an uncertainty of 0.5%.

The resolution of the calorimeter can be checked with a variety of processes [35]. For example, the width of the π^0 mass peak can be compared with Monte Carlo expectations (see also Section 3.2.5). In the energy range of interest ($1.5\,\mathrm{GeV} < E_{lab} < 3.5\,\mathrm{GeV}$) there is agreement between data and Monte Carlo within statistics. The VCS photon sample also provides a precise test of the resolution for high-energy photons and confirms again that the simulation agrees well with data. The right plot in Figure 8.1 shows the change in signal efficiency if the photon energy is smeared. It indicates that the analysis is rather insensitive to any mis-modeling of the resolution. A systematic uncertainty of 0.5% is ascribed that covers a smearing of $1-3\%$.

This analysis requires photons to be isolated by at least 25 cm from any other charged or neutral cluster. To estimate the systematic uncertainty associated with this requirement, radiative Bhabha photons from data have been embedded in both generic $B\overline{B}$ data and $B\overline{B}$ MC events which are then passed through the regular reconstruction chain. The efficiency of the isolation cut is compared in both cases. A deviation of 2% has been observed and is assigned as the uncertainty.

The systematic uncertainty associated with the veto requirements against π^0s and ηs is estimated also via an embedding method. An photon is embedded in both off-resonance data and continuum Monte Carlo, completely analogously to the case above. Here, the π^0 mass distributions are compared in both samples, for all combinations of the high-energy photon with any second photon in the event (if that satisfies the minimum energy requirements of Section 5.3). This yields a discrepancy of 1% which is assigned as an uncertainty. However, the use of energy-dependent vetoes might introduce an energy or multiplicity dependence for these cuts. These effects will be discussed in Section 8.6.

8.2 Systematics from Event Shape Requirements

No systematic was associated with the shape cuts used in the $B \to K^*\gamma$ analysis since the data and Monte Carlo are in very good agreement [18]. It is assumed that this systematic is negligible in this analysis also. This assumption has been checked by comparing efficiencies in continuum simulation and off-resonance data, since the

event shape requirements are mainly effective against this type of background.[1]

8.3 Lepton Selection Systematics

The standard particle identification (PID) efficiency look-up-tables (introduced in Section 5.7.1) are used to ensure that the efficiency of the tag requirement is modeled correctly in the simulation. The uncertainty on this correction is computed by changing all PID efficiencies coherently by 1σ of their total errors. These are given by the statistical errors of the control samples that yielded the PID tables. Figure 8.2 shows that a 1σ smearing results in a 2% systematic uncertainty.

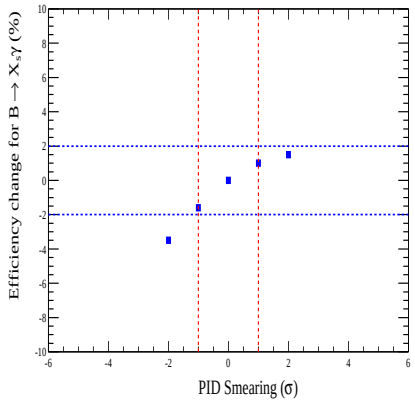

Figure 8.2: The change in efficiency for signal $B \to X_s\gamma$ (X_s plus K^*) $SP4$ Monte Carlo versus the change in the particle identification efficiency (in units of PID standard deviations). The PID efficiencies are changed in phase to obtain the points plotted.

[1]Since efficiencies only rely on shapes of distributions, the large normalization differences between data and Monte Carlo do *not* influence this comparison.

8.4 Data Sample Systematics

There are some small systematic effects associated with the sizes of the data sets listed in Table 3.2. The dominant error is from $B\overline{B}$-"counting". The number of $B\overline{B}$ events in the data sample is derived from its relation to the number of μ-pairs and multihadronic events in the on- and off-resonance samples [56]. Several effects introduce systematic uncertainties: the cuts used in the selection of the hadronic and muon samples, and most notably the variation of the machine conditions (beam backgrounds, machine tunes, *etc.*) over time. A total $B\overline{B}$-counting uncertainty of 1.1% is determined [57] which translates directly into a branching fraction uncertainty. For the off-resonance data only the *ratio* of off- to on-resonance luminosities is important, and any errors essentially cancel out.

Any potential effect from the unevenly distributed background Monte Carlo samples noted in Table 3.3 is found to be negligible since the efficiencies agree in all sub-samples with different running conditions.

8.5 $B\overline{B}$ Subtraction Systematics

The $B\overline{B}$ background is estimated by Monte Carlo simulation. It was shown in Chapter 6 that this simulation needs tuning in order to agree better with data on the amount of background high-energy photons that will be present in this analysis. There is a systematic uncertainty associated with the tune and correction of this background.

In the following the derivation of this uncertainty is explained first and the number obtained is then checked for consistency.

It has been shown that the correcting tune for the $B\overline{B}$ background simulation has been derived from the comparative study in data and simulation of control samples for all separate sources of background that are listed in Table 6.1. The single errors on the corrections are obtained from the statistical uncertainty in the control samples and, to a much lesser extent, from any systematic effects in fit convergences. Since the data has to be off-resonance subtracted, the limited off-resonance luminosity has the dominant statistical effect. Note that most correction factors (and their errors) were derived in bins of kinematic quantities other than the photon energy, E^*_γ. To arrive at correction factors for each bin in E^*_γ, the corrections were applied in the simulation according to the photon candidates true identity and true value of the used kinematic variable. To determine the error in each E^*_γ bin (as shown in Table 7.1) the uncertainties on all corrections contributing to the total in that bin and the Monte Carlo statistical uncertainty are summed in quadrature. (An average uncertainty of 7% is found on the corrections only.) This is the total systematic error on the number of $B\overline{B}$ events to be subtracted in each bin. It varies from bin to bin, due to the different statistical significance of the bins.

Overall, the error on the corrections can be checked by using the sum of all contributing bins. A net data/Monte Carlo correction factor of 0.924 ± 0.0797 for the region $2.0 - 2.7$ GeV in E^*_γ was derived from the comparison of the corrected and

uncorrected spectra in Section 6.6. When the statistical part of this uncertainty is deconvoluted from the error on this correction, a fractional error of 6.6% is found, confirming the number above.

Extracting the corrected number of $B\overline{B}$ background events from the corrected spectra in Chapter 6 (or from Table 7.1 which lists the separate bins) yields a total of 616 ± 50 $B\overline{B}$ events in the relevant energy range of $2.0 - 2.7$ GeV, which includes the total uncertainty. To turn this into an error on the final branching ratio, the number of (expected) signal events after all cuts is needed. The absolute 1σ number of $B\overline{B}$ events to be subtracted over the signal expectation then yields the fractional error. The number of expected signal events assuming the Kagan-Neubert model with m_b =4.8 GeV can be taken from Table 7.1. The specific construction of the Kagan-Neubert signal model prescribes 16.1% of the total branching fraction to be from K^* for this value of m_b. When combined with the $B \to K^*\gamma$ branching ratio that is measured from BABAR, however, the total branching fraction prediction in this case is only $\mathcal{B}(B \to X_s\gamma)_{KN480} = 2.50 \times 10^{-4}$, because Kagan and Neubert fix only the shape but not the area of the spectrum. This number is significantly below the theoretical branching ratio for the total signal of $B \to X_s\gamma$, $\mathcal{B}(B \to X_s\gamma) = 3.57 \times 10^{-4}$ [16] which describes the actually measured branching fraction much more closely. Thus, to arrive at a realistic number of expected signal events in the considered energy range, the number from Table 7.1 is scaled to the latter branching fraction number and yields 983 signal events. Altogether, this results in an anticipated systematic uncertainty on

the signal of 6.7% due to the $B\overline{B}$ subtraction. This number has to be recomputed once the actual number of measured signal events is extracted.

As a consistency check for the estimated error on the corrections, the method described in Section 7.3 is used. It compares the expectation in the $1.7 < E^*_\gamma < 1.9\,\mathrm{GeV}$ region with the measured on-resonance yield. The low energy region is dominated by $B\overline{B}$ background and contains $\pi^0(\eta)$, hadronic and all other components. In this *low-energy* region, the corrected $B\overline{B}$ simulation underestimated the yield in data by about $14.3 \pm 8.5\%$. Within the error this agrees with the 7% correction uncertainty stated above for the *signal* region.

8.6 Systematic Effects from the Simulation

When considering how the general production of simulated events can influence the correctness of the signal efficiency, two effects must be considered: uncertainties due to the limited available statistics in the generated samples and, more fundamentally, errors in the fragmentation mechanism that is assumed in the generator.

First, the systematic uncertainty is estimated that arises in determining the signal efficiency due to the limited Monte Carlo statistics. The central value of the signal efficiency used is based on the Kagan-Neubert signal model with $m_b = 4.80\,\mathrm{GeV}$ (the error from using this model is estimated separately in the next section). Using the KN480 and K^* columns from Table 5.1, with a K^* fraction of 0.161, and taking

into account the actual Monte Carlo statistics from Table 3.2, a 1.6% uncertainty is found on the efficiency.

Potentially more important is any sensitivity to the details of the JetSet fragmentation model used for the inclusive X_s system in the Kagan-Neubert Monte Carlo samples. (This is not expected to be an issue in the modeling of the decay of the K^* system, because as an exclusive state it is very well determined from isospin conservation.) The BABAR semi-inclusive $b \to s\gamma$ analysis [8, 44] found evidence that the distributions of both total multiplicity in the X_s final state and the ratio of the number of neutral to charged pions differed significantly between data and the simulation. The presented fully-inclusive analysis, however, is sensitive to this *only* to the extent that the detection efficiency varies with these quantities. The efficiencies of the photon selection cuts or of the lepton-tag momentum and angle requirements are not expected to show sensitivity to details of the X_s system. On the other hand, event-shape and missing energy cuts might. Hence it is relevant to check, for example, the effect of introducing the Fisher linear discriminant method into the analysis.

The possible efficiency variations were investigated for the case of the signal simulation with the KN480 model.

Table 8.3 shows efficiency *vs.* multiplicity, while Table 8.4 shows efficiency *vs.* the π^0/π ratio. The event counts determine the precision for the numbers. For this study the order of cuts has been changed somewhat: the photon selection and energy cut (of $2.0 < E^*_\gamma < 2.7\,\mathrm{GeV}$) are made first ("pre-tag" columns). Directly afterwards,

Mult	Number of MC events			Efficiency relative to pre-tag	
	Pre-tag	Post-tag	All cuts	Post-tag	All cuts
2	12857	862	374	6.71%	2.91%
3	30710	2125	918	6.92%	2.99%
4	24587	1763	637	7.17%	2.59%
5	17236	1190	372	6.90%	2.16%
≥ 6	17776	1267	347	7.13%	1.95%
Total	103166	7207	2648	6.99%	2.57%

Table 8.3: Dependence of KN480 efficiency on X_s multiplicity for $SP4$ Monte Carlo simulation, using the sum of both charge states. "Multiplicity" is defined as the total number of charged and neutral kaons and pions (excluding pions from K^0_S decay) plus etas that, according to Monte Carlo truth, decay to $\gamma\gamma$. The cuts applied in the pre- and post-tag columns are described in the text. Additional Monte Carlo events are included here that were not available for Table 8.4.

$r = \pi^0/\pi^\pm$	Number of MC events			Efficiency relative to pre-tag	
	Pre-tag	Post-tag	All cuts	Post-tag	All cuts
$r = 0$	27711	1922	894	6.94%	3.23%
$0.0 < r \leq 0.5$	16931	1140	460	6.73%	2.72%
$0.5 < r \leq 1.0$	19667	1301	584	6.61%	2.97%
$1.0 < r < \infty$	5613	391	155	6.97%	2.76%
$r = \infty$	11548	716	329	6.20%	2.85%
Total	81470	5470	2422	6.71%	2.97%

Table 8.4: Dependence of KN480 efficiency on X_s $\pi^0/\pi^\pm$ multiplicity ratio for $SP4$ Monte Carlo simulation, using the sum of both charge states. Pions from K^0_S decay are not counted. The cuts applied in the pre- and post-tag columns are described in the text.

the lepton-tag momentum and angle cuts are made at the values shown in Table 5.1 ("post-tag" columns) and not till the end are the event shape and E_{miss} cuts applied.

Looking first at the effects in the π^0/π ratio, no clear trend can be seen in Table 8.4. However, it appears that the efficiency for states *without* π^0s is a factor of about 1.13 larger than for states *with* π^0s. The semi-inclusive analysis in BABAR[44]

finds, on average, that the data have relatively more events with π^0s than predicted by the JetSet model; in fact, their numbers suggest that the actual fraction *without* a π^0 could be as low as 0.5 times the prediction. To estimate the effect of such a mis-modeling, half the events in the category without π^0s in Table 8.4 ($r = 0$) are moved to other categories. This results in a fractional reduction of 2.2% in the overall KN480 signal efficiency. This is taken as the systematic uncertainty for charged–neutral fragmentation effects in the simulation.

Turning now to multiplicity effects, Table 8.3 shows a decrease for multiplicities of 5 and larger. This may be fully explained (within statistics) as a reflection of the E^*_γ dependence on efficiency: the average multiplicity in the fragmentation model decreases slowly with E^*_γ, while the efficiency increases with E^*_γ. However, it is of interest to ask what the impact would be if the multiplicity dependence in Table 8.3 were real. If the mean multiplicity in the JetSet model were low by 0.5 (as suggested by the semi-inclusive analysis [44]), and this was represented by moving half the events from each multiplicity bin to the next higher bin, the overall efficiency would change by -3% of itself.

However, as mentioned, the observed effect may just be a consequence of the efficiency's energy-dependence. This is investigated in the following. The statistical precision of this check is 5 to 10% of the efficiency if the regularly produced signal Monte Carlo is used. This is mostly due to the Kagan-Neubert spectrum falling off relatively sharply at both ends, so that the available statistics, especially in the high-E^*_γ

range (lowest multiplicity) and the low-E_γ^* range (highest multiplicity), is insufficient for these studies. Thus, an additional large sample of signal events was generated that follows the Kagan-Neubert prescription for the modeling of the inclusive X_s system at $m_b = 4.80$ GeV, but does not follow their overall spectrum prescription. Instead, in each bin of E_B^* (the photon energy in the rest-frame of the signal B meson) the same number of events are created, yielding a flat spectrum and more balanced statistical distribution over all photon energies for this study. This sample is used solely for efficiency studies. A second advantage of this new sample lies in the fact that it is entirely independent of the original sample that was used for the "training" of the Fisher discriminant code and the optimization of the analysis cuts.

To decouple any energy dependence effects from the multiplicity effects, the photon energy range is divided into bins and the multiplicity study is repeated in these bins. This is a simple approach, but it requires a lot of statistics to yield significant answers. Three bins in E_γ^* are used, splitting the total range of $2.0 < E_\gamma^* < 2.7$ GeV at 2.2 and 2.4 GeV. A summary of the findings is recoded in Table 8.5. The most significant difference between low and high multiplicities is observed for the highest energy bin ($2.4 - 2.7$ GeV). Using this bin, event yields with multiplicities of 3 and 4 can be compared to those with multiplicities of 5 and higher. If again the assumption is made that the JetSet predictions are low by half a unit in multiplicity, the event counts before and after this shift can be compared. A decrease of 3.3% is found.

To cross-check this result, a second method of decoupling any energy de-

E^*_γ Range	Efficiency, Multiplicity = 2, 3	Efficiency, Multiplicity ≥ 5
$2.0 - 2.2.$ GeV	$(1.85 \pm 0.09)\%$	$(1.73 \pm 0.06)\%$
$2.2 - 2.4.$ GeV	$(2.77 \pm 0.08)\%$	$(2.62 \pm 0.10)\%$
$2.4 - 2.7.$ GeV	$(3.35 \pm 0.09)\%$	$(2.76 \pm 0.21)\%$

Table 8.5: Dependence of KN480 efficiency on X_s multiplicity in bins of E^*_γ for $SP4$ Monte Carlo simulation. The sum of both charge states is used and the Monte Carlo simulation with a "flat" distribution of events over the E^*_γ range. Three tables like the one in Table 8.3 are combined, summarized and compared here. Note the largest difference in the highest bin.

pendence effects is used. It fits a function (a straight line) to the energy-dependent efficiency distribution for the event shape and lepton-tagging cuts, similar to the one introduced in Section 5.10. The fit is then used to estimate an extra weight for events with lower energy (where the efficiency is lower). The multiplicity dependent efficiency of the such weighted events is compared. Some deviation is found between high and low efficiencies, but due to the larger errors associated with the fit and the rescaling the deviation is not significant. Therefore this method is only used as a check and confirms the finding of the first study that a small dependence on multiplicity is observed.

Thus, any numbers are taken from the first study. The systematic uncertainty due to multiplicity effects in the efficiency is given there as 3.3%. This is taken as the fractional systematic uncertainty on the overall efficiency due to multiplicity effects.

For the combination of the two fragmentation effects, a fractional systematic uncertainty of 4.0% is assigned. In the end, this has to be combined in quadrature with the overall statistical uncertainty on the Monte Carlo signal samples which yields

a total systematic uncertainty on the signal efficiency of 4.3%.

8.7 Model-Dependence Systematics

Characteristics of the signal model that introduce a systematic uncertainty are the shape of the model-spectrum, the cut-off point in the hadronic mass where the non-resonant component of the spectrum is cut out and replaced with the K^* peak, and the fraction of $B \to K^*\gamma$ on the full branching fraction. All of these effects must be evaluated for their influence on changing the extracted branching fraction or spectrum. This is because one particular model will be used to extrapolate to the full spectrum after a cut is made on the photon energy. The three effects are treated separately in the following subsections.

8.7.1 Dependence on Spectral Shape (m_b)

Part of the theoretical uncertainty related to the use of a signal model in the extrapolation to the full spectrum is related to the different spectral shapes that are possible in the model. For the model of Kagan and Neubert [26], the spectrum is parametrized in the variables m_b and μ_π (or λ_1), with μ_π fixed for each value of m_b as explained in Section 2.4 and Appendix B. The authors prescribe that to evaluate the theoretical uncertainty, the b-quark mass m_b must be varied between 4.65 and 4.95 GeV. With Monte Carlo samples available for this analysis with m_b values of 4.65, 4.80, and 4.95 GeV (see Table 5.1), $m_b = 4.80\,\text{GeV}$ has been used as the central

value throughout this analysis. The theoretical uncertainty is then obtained as the difference in efficiency when using the other two signal samples. Efficiency here refers only to the effect of restricting the E_γ^* range of the photons, not to the losses of other analysis cuts because the bin-by-bin efficiency correction will take care of these effects. A fixed $m_{\rm cutoff} = 1.1\,{\rm GeV}$ is used for all models, and the appropriate $K^*(892)$ fractions for each value of m_b are taken from Table 2.2. On generator level in the signal simulation the separately generated spectra for the K^* and X_s components are then combined as described above. For each model the fraction contained within the three E_γ^* regions of interest, $1.9 - 2.7$, $2.0 - 2.7$ and $2.1 - 2.7$ GeV is determined. Apart from resolution effects, this is a good approximation for the m_b dependence of the energy-range efficiency. Table 8.6 shows the results, with asymmetric errors for the various E_γ^* ranges as the bottom line.

	Efficiencies for E_γ^* range in GeV		
Model	1.9 to 2.7	2.0 to 2.7	2.1 to 2.7
KN465	0.934	0.888	0.817
KN480	0.961	0.935	0.890
KN495	0.967	0.955	0.935
% variation	+0.6, -2.7	+2.2, -5.0	+5.2, -8.1

Table 8.6: Dependence of the fraction of events inside different "signal regions" on the choice of m_b in the Kagan-Neubert model; μ_π is set as suggested by KN (described in Appendix B). The given fraction of events is determined on generator level and represents the efficiency of a cut on the energy range. It leads to the model-dependence errors given. All numbers include the K^* contributions. The % variation is relative to (KN480 + K^*).

As a check, the model-dependence can be estimated using a set of models with a fixed $K^*(892)$ fraction and $m_{\rm cutoff}$ adjusted to the values given in the last

column of Table 2.2. The result is found to be identical within 0.5%. Therefore, the above numbers are taken as the estimates of the m_b-dependence in the efficiency.

8.7.2 Sensitivity to Boundary Between $K^*\gamma$ and X_s

The signal model, described in Section 1, is composed of a mixture of $B \to K^*\gamma$ and $B \to X_s\gamma$ with a sharp boundary at 1.10 GeV. In reality this boundary is more of a smooth transition. To take this into account, an uncertainty is computed by varying the boundary between 1.0 and 1.2 GeV *without* varying either m_b or the K^* fraction. This study has been carried out with the same generator level method as described above. For this study, m_b is kept fixed at 4.80 GeV. Again fractions of the full spectrum that lie within the considered relevant energy ranges are determined for all three values of the K^*–X_s boundary. Results are shown in Table 8.7. Because the m_{X_s} spectrum is steeply falling near 1.1 GeV, a larger effect on the X_s component is expected when moving to a higher cutoff than to a lower. That is verified from the table. The asymmetric errors from the bottom line of the table are used.

The K^* component of the spectrum stays constant from one cut-off to another since the K^* spectrum is used in full. From the significantly smaller uncertainties observed in this study it is seen that the sensitivity to this cut-off is much smaller than to m_b. This reflects the fact that the signal model is rather stable to the details of the inclusion of the K^*.

	Events for E^*_γ range in GeV		
Cutoff	1.9 to 2.7	2.0 to 2.7	2.1 to 2.7
1.00 GeV	0.962	0.938	0.895
1.10 GeV	0.961	0.935	0.890
1.20 GeV	0.958	0.931	0.882
% variation	+0.1, -0.2	+0.3, -0.5	+0.6, -0.9

Table 8.7: Dependence of expected energy-range efficiency on the choice of the hadronic mass cutoff between K^* and X_s components. The latter is modeled using the Kagan-Neubert model with $m_b = 4.80\,\text{GeV}$ (KN480). The % variation is relative to the cutoff of 1.10 GeV.

8.7.3 Sensitivity to Assumed $B \to K^*\gamma$ Branching Fraction

Another question is the dependence of the signal efficiency on the fraction of $B \to K^*\gamma$ in the signal model. This effect is investigated using the K^* plus KN480 model. The expected ratio of $\mathcal{B}(B \to K^*\gamma)$ to $\mathcal{B}(B \to X_s\gamma)$ is 0.120 ± 0.021 (assuming the measured world average for $\mathcal{B}(B \to X_s\gamma)$ and the BABAR measurement of $\mathcal{B}(B \to K^*\gamma)$ as given in Section 4.4). To see the effect of varying the $B \to K^*\gamma$ fraction by 1σ, the following approach is taken: The expected $B \to K^*\gamma$ event fraction of 0.161 for a K^* plus KN480 model (see Table 2.2) is varied by the same *fractional* error as the one quoted above, *i.e.* 0.161 ± 0.037 and the effect on the efficiency is computed. A change in the efficiency of 0.7% is observed, which is assigned as the systematic uncertainty. The reason for the low sensitivity to such a change is the fact the the selection efficiency is not too much different for $B \to K^*\gamma$ than for the non-resonant part of $B \to X_s\gamma$.

8.7.4 Overall Model-Dependence Uncertainty

Adding in quadrature the numbers in the previous three subsections, the following overall model-dependence uncertainties on the efficiency are obtained :

- 1.9 to 2.7 GeV: +0.9%, -2.8%
- 2.0 to 2.7 GeV: +2.3%, -5.1%
- 2.1 to 2.7 GeV: +5.2%, -8.2%

Chapter 9

Summary and Conclusions

9.1 Selection of Photon Energy Range

The choice of the E_γ^* range used for the extraction of a branching fraction involves balancing the statistical, systematic and model dependence errors. To reduce model dependence it is advantageous to cut at low E_γ^*. However, the sharply rising $B\overline{B}$ background contribution (as E_γ^* decreases) offsets the reduction in model dependence of such a cut with an increased systematic uncertainty in the $B\overline{B}$ background estimation. The choice of region has to be made just before unblinding the data, using percentage statistical and systematic uncertainties relative to an assumed branching fraction.

The Kagan and Neubert signal model with $m_b = 4.8$ GeV (mix of KN480 plus $K^*\gamma$) is used for this task, because it was used to define the event selection

strategy. This model, as mentioned, keeps the branching fraction for $B \to K^*\gamma$ fixed at the value measured by BABAR [18]. The total branching fraction for $B \to X_s\gamma$ is then this number divided by the $K^*\gamma$ fraction for this model, resulting in $\mathcal{B}(B \to X_s\gamma)_{KN480+K^*} = 2.50 \times 10^{-4}$. This clearly underestimates the theoretical prediction that is given in Section 2.3.1 by about 30%, because Kagan and Neubert only describe the shape but not the overall normalization of the spectrum and do not account for fixing the absolute $\mathcal{B}(B \to K^*\gamma)$. Thus, for this step, the signal predictions are scaled to the more realistic theoretical expectation for the branching fraction, $\mathcal{B}(B \to X_s\gamma)_{theo} = 3.57 \times 10^{-4}$.

Table 9.1 assembles the percentage errors affecting this measurement for three different choices of the E^*_γ lower bound. It brings together the statistical significances that are computed in Section 7.2, the systematic error estimates as described in Section 8, and the model dependence uncertainties that are examined in Section 8.7.4.

The total error is the quadratic sum of the three components. These results lead to the choice of $2.0 < E^*_\gamma < 2.7\,\text{GeV}$ for the optimal region that should be considered for the extraction of the branching fraction.

To date, the data still have to be unblinded, so no measured branching fraction number or spectral shape is available yet. The presented full determination of all involved analysis errors is a large step towards the final results and can already be used for comparisons of the analysis performance as will be shown below.

E_γ^* (GeV)	Stat.	Sys.	Mod.	Total
1.9-2.7	9.0	9.5	2.1	13.3
2.0-2.7	8.5	7.9	4.0	12.3
2.1-2.7	8.3	7.0	6.9	12.9

Table 9.1: The statistical, systematic and model-dependence uncertainties expressed in % $\mathcal{B}(B \to X_s\gamma)$ for different E_γ cut choices. The numbers are obtained with the Kagan and Neubert signal model with $m_b = 4.8$ GeV and then scaled to the theoretical branching fraction number of $\mathcal{B}(B \to X_s\gamma)_{theo} = 3.57 \times 10^{-4}$. For purposes of this comparison, the average of the asymmetric plus and minus model-dependence errors is used. The total uncertainty is the sum in quadrature of the contributions.

9.2 Comparisons

With the determination of all involved analysis uncertainties finalized, the performance of this analysis can be already be studied in comparison to other measurements.

Looking back to Table 4.1 and comparing the total fractional errors of the measurements there, the second CLEO measurement was the most precise one listed, with an overall error of 16.5%. The analysis presented here shows a significant improvement in this regard with an expected total error of only 12.3%. If the total error is broken up into statistical, systematic, and model-dependent (or theoretical) uncertainties, the same trend is observed for each. It is noteworthy in particular, that as the only truly fully-inclusive analysis this measurement shows the smallest model-dependent uncertainty. This is a very good validation and justification of the analysis technique used.

The world average for the $B \to X_s\gamma$ branching fraction was determined in Section 4.4. It used all previously published experimental results, including the two preliminary BABAR numbers (from the semi-inclusive analysis and from the antecessor of the present fully-inclusive measurement). The number found was:

$$\mathcal{B}(B \to X_s\gamma)_{oldAvg} = (3.34 \pm 0.38) \times 10^{-4}.$$

It is interesting to see the effect that a new measurement with the presented errors will have on this number.

The contributing measurements have to undergo a weighted averaging. The χ^2 minimization technique described for this problem in [67] (and already used in [10]) is then applied for this task. All statistical and systematic errors are treated as uncorrelated, the model-dependence as a correlated error. To evaluate the effect of the new smaller measurement errors on the total average, the following considerations and assumptions are made: The full result of the previous BABAR fully-inclusive measurement [9] is already contained in the existing calculation and has to be replaced when the new result is included because the two are not independent. Therefore it makes sense to use the previously obtained central value (3.88×10^{-4}) with the percentage errors that are measured for the present analysis. Assuming the same central value as for the previous measurement should not alter the central value of the average too much (except for weighting effects) and therefore seems adequate for determining the change in error.

The weights that are found for the individual measurements are compared

in Table 9.2 for the previous situation and after the addition of the new measurement error presented herein. The improvement in technique is clear when comparing the weight for BABAR's previous fully inclusive measurement with the weight for the present one, it increased by a factor of 1.7. The new small errors found here shift the largest weight to this measurement for the first time, the influence on other measurements is negligible.

Experiment	Reference	Weight	
		previous	now
ALEPH 98	[6]	0.99	0.99
BELLE 01	[7]	0.93	0.98
CLEO 01	[5]	3.28	3.30
BABAR 02 s	[8]	-0.49	-0.43
BABAR 02 f	[9]	2.39	——
BABAR 03 f	this thesis	——	4.05

Table 9.2: Weights of the contibuting measurements in the world average calculation of the branching fraction $\mathcal{B}(B \to X_s\gamma)$. The values before and after the inclusion of the new measurement errors are shown. For the presented measurement the same central value as previously measured [9] is assumed. For BABAR s indicates semi-, f indicates fully-inclusive).

The χ^2-curve obtained from the weighted averaging is shown in Figure 9.1, the minimum is indicated by the arrow. The new world average is found to be:

$$\mathcal{B}(B \to X_s\gamma)_{newAvg} = (3.45 \pm 0.34) \times 10^{-4}.$$

No information can be obtained from the central value of the new average, but it is clear that with the new fully-inclusive BABAR measurement the error on the world average will be pushed below 10% for the first time (it was 11.4% previously

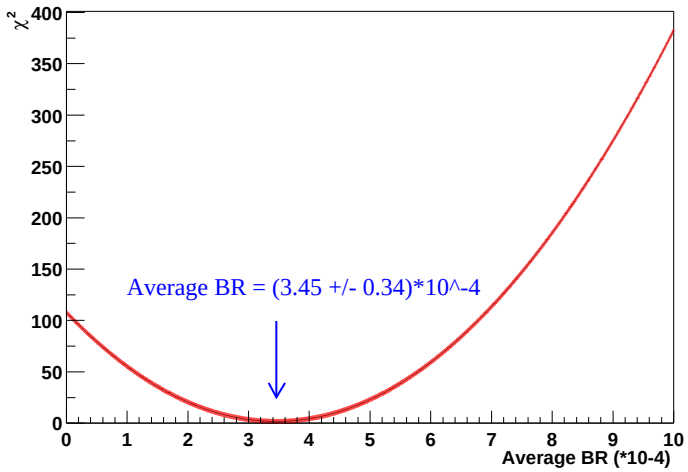

Figure 9.1: χ^2 distribution for the extraction of a new average branching fraction. The minimum is indicated by the arrow. For the new measurement the same central value is assumed as was determined with the previous version of this analysis but with the percentage errors that are measured for the present analysis.

and now is reduced to 9.7%).

Under the speculative assumption that the central value of the total average has stayed the same, the consequences of this number with the smaller error can be investigated in terms of the restrictive power for new physics scenarios. (Of course this may change when the actual measured branching fraction number for the present analysis becomes available.) Looking at Figure 2.6 in Section 2.3.1, a branching fraction of then $\mathcal{B}(B \to X_s\gamma) = (3.3 \pm 0.4) \times 10^{-4}$ yielded a constraint on the mass of the charged Higgs of $M_H > 350$ GeV at 99% confidence level. This is for a two Higgs-doublet model. It was predicted [16] that a measurement of the same central value with an uncertainty of only 10% would push this limit up to $M_H > 475$ GeV. With

some degree of optimism regarding the constancy of the central value when the new measurement is included, a constraint of that order seems possible with the current analysis.

9.3 Conclusions

An analysis of the radiative penguin process $b \to s\gamma$ has been presented. It is fully inclusive in that it does not make any requirements on the hadronic strange system in the decay and therefore allows all possible hadronic final states to contribute. The analysis was done blind on approximately 88 million $B\overline{B}$ decays ($81.5\,\mathrm{fb}^{-1}$) collected with the BABAR detector between 2000 and 2002. The method uses bins in E^*_γ so that both a branching fraction and a spectral shape can be obtained for this process.

The event selection is based solely around the high-energy photon expected from the signal decay and the typical signal event shape. A Fisher linear discriminant method has been used to combine several shape variables and their correlations into one "optimal" quantity. Clear advantages in the suppression of especially continuum background have been shown for this method as compared to using single-variable cuts. Requiring the presence of a high-momentum lepton in the event yields an significant additional reduction of continuum background. It is, however, also connected to a decided loss of signal efficiency.

Very little continuum background is left after all analysis cuts. It is sub-

tracted off with data taken below the $\Upsilon(4S)$ resonance scaled to the on-resonance data luminosity. Only continuum event production is expected for off-resonance collision energies. The limited statistics in the off-resonance data sample is responsible for most of the statistical error on the result.

Most of the background left after all analysis cuts stems from high-energy photons which arise largely from two-photon decays of $\pi^0(\eta)$ mesons in $B\overline{B}$ events. This $B\overline{B}$ background rises steeply towards lower E^*_γ. It must be subtracted off using Monte Carlo predictions. To correct for differences between data and simulation in the production of these mesons, the Monte Carlo must be tuned to resemble the data more closely. Dedicated samples of the background processes were obtained in data and simulation and the ratios of observed event yields are used as correction factors. This method is quite involved and its uncertainty introduces the largest contribution to the systematic error.

To reduce the systematic error due to the $B\overline{B}$ background, a cut is made on the E^*_γ range and only signal within a signal region is considered. Any losses from this cut must be corrected for in the end using a model that parametrizes the signal spectral shape. This correction then introduces a dependence of the final result on the theoretical model used. This effect is stronger the more of the spectrum is excluded from the signal region. It must therefore be balanced with the uncertainties from the $B\overline{B}$ simulation. The optimal range for this purpose was found to be $E^*_\gamma = 2.0{-}2.7$ GeV

.

All expected errors on the final branching fraction have been determined to a total of 12.3%. The data need yet to be unblinded. From the error estimate alone, however, it can already be deduced that the current measurement will yield the most precise result on $B \to X_s\gamma$ to date. It will lower the error on the world average for this measurement considerably, giving rise to tighter constraints on new physics models. Possible consequences for a two Higgs-doublet model have been demonstrated that would increase the lower limit on the charged Higgs mass to 475 GeV from the $b \to s\gamma$ branching fraction number alone.

The unblinding of the data and the determination of a central value for the branching fraction as well as the extraction of a spectrum are currently underway at the BABAR experiment. Final results are expected within the next months (summer 2003).

Appendix A

Illustration of the Penguin in the Diagram

Richard Feynman's formulation of relativistic quantum mechanics, which is expressed in terms of vertex factors and propagators, yields a wonderful way of representing particle physics processes graphically. The resulting Feynman diagrams are not only accurate contractions of the theoretical description, but are furthermore very descriptive illustrations of the underlying process. Particle physicists are generally used to "thinking" in Feynman diagrams, to the extent that for many the introduction of a decay or transition is only understandable and imaginable when the corresponding Feynman diagram is drawn.

The diagrams represent the propagators for different particle types by different line styles (solid lines for fermions, broken lines for bosons, sine-waves for photons

and helices for gluons). However, there is no prescription for the shape of the lines, except that kinks should only appear at interaction vertices. While it is common to use straight lines as much as possible, the validity of the description is not broken by distorting the line *shape*, as long as the lines don't cross over each other, or the distortion resembles a different line *style*.

Therefore, the penguin diagram for the $b \to s\gamma$ transition shown in Figure A.1 may be drawn in a different way that better expresses the "penguin-ness" of the process: see Figure A.2.

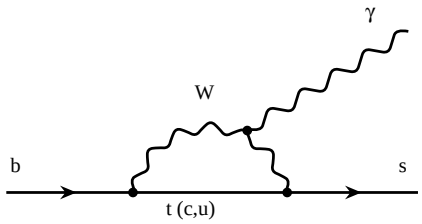

Figure A.1: Standard drawing of the leading order Feynman diagram for the $b \to s\gamma$ process. This is the minimalist approach to rendering a penguin diagram.

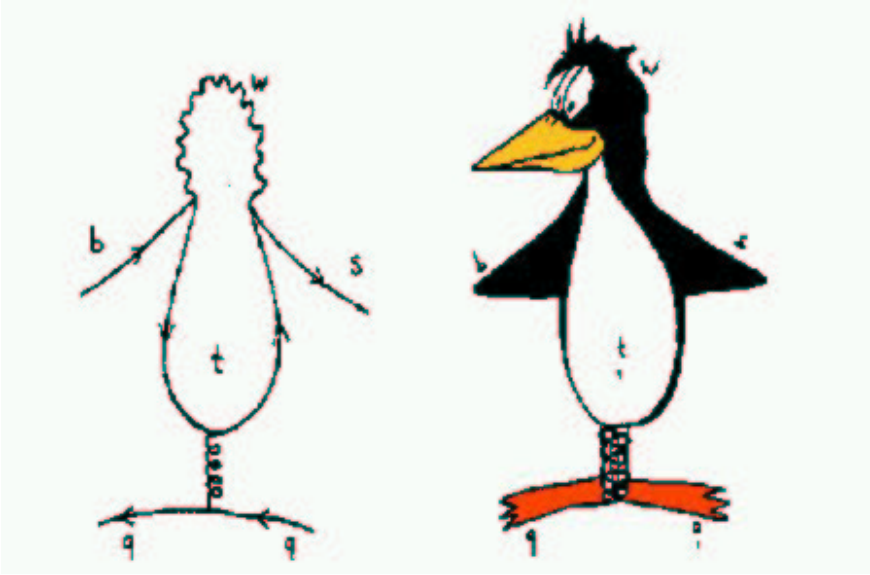

Figure A.2: Illustration of the "penguin-ness" of the leading order Feynman diagram for the $b \to s\gamma$ transition. The figure is reproduced from [68].

Appendix B

Monte Carlo Signal Generator

The model used for Monte-Carlo generation of the E_γ spectrum is that of Kagan and Neubert [26]. The calculation begins at the parton level where the NLO $b \to s\gamma$ matrix element is calculated. This results in a photon spectrum $P(y_p)$ where $0 \leq y_p (= 2E_\gamma/m_b) \leq 1$. The maximum energy of the photon is given by the kinematic limit $E_\gamma^{max} = m_b/2$. The effect of placing the parton in the B meson is modeled by convoluting the photon spectrum with a shape function $F(y^{'})$ that represents the effect of the motion of the b-quark inside the B meson. The transformation $y \to y^{'} = ym_B/m_b$ is affected because the new kinematic limit for the photon is $E_\gamma^{max} = \mathrm{m_B}/2$ by virtue of the smearing of the kinematic limit due to the b-quark's motion. The shape function is given by equation B.1.

$$F(y^{'}) = N(1 - y^{'}/\overline{\Lambda})^a \cdot e^{(1+a)y^{'}} \tag{B.1}$$

Signal	m_b	μ_π^2
KN460	4.65	0.52
KN475	4.75	0.37
KN480	4.80	0.30
KN495	4.95	0.14

Table B.1: The input parameters used in the Monte Carlo signal generation. The range of parameters follows the prescription given by Kagan and Neubert in reference [26].

Here N is the normalization, $\overline{\Lambda} \approx m_B - m_b$, and a is related to the kinetic energy of the b quark inside the B meson, μ_π^2, by equation B.2.

$$\mu_\pi^2/3 = \overline{\Lambda}^2/(1+a) \tag{B.2}$$

The two input parameters are m_b and μ_π. Figure 2.8 shows how the shape function varies for different choices of these parameters. The m_b plots are made by keeping $\mu_\pi/\overline{\Lambda}$ constant. It can be seen that the dependence on m_b is much greater than that on μ_π. Consequently Kagan and Neubert prescribe that the model dependence is "fairly represented" by varying m_b in the range $4.65 - 4.95\,\mathrm{GeV}$ keeping $\mu_\pi/\overline{\Lambda}$ constant. This constant is set by the reference point $m_b = 4.8\,\mathrm{GeV}$, $\mu_\pi^2 = 0.3\,\mathrm{GeV}^2$, $m_B = 5.2788\,\mathrm{GeV}$. The generator is implemented in the *EvtGen* package (see Section 3.3). The input parameters used for signal generation in the simulation are given in Table B.1. In addition there are other parameters that have been set constant throughout the generation. The minimum energy of the photon is given by $E_\gamma \geq \frac{1}{2}(1-\delta)m_b$ where the cutoff parameter δ is set to be 0.9. The renormalization scale μ_b is set at the b-quark mass m_b. The ratio of the charm and beauty quark masses $z = m_c/m_b$ is set to 0.0841.

Appendix C

Background-Subtraction Statistics and Off-Peak Running

The question addressed here is the uncertainty of signal yield due to backgrounds for a measurement in which there are two types of background present: one type which can be simulated or otherwise computed, and one for which subtraction of independently measured data is necessary. This relevant for the presented fully-inclusive analysis, because continuum backgrounds are directly subtracted using off-resonance data, which is independently measured and statistically limited. (Other measurements rely on a fitting technique in an appropriately chosen variable, in which a portion of the background is directly visible, and the remainder is constrained using sideband and/or a fit to the off-resonance data.) The main consequence of the limited statistics in the subtracted sample for an analysis using a straight subtraction is the

rethinking the standard "$S^2/(S+B)$" cut optimization formula.

C.1 Formulas

Define three components to the event yield in a measurement at some standard integrated luminosity:

S: Signal of interest,

B: Simulatable background,

C: Background requiring subtraction using off-resonance data.

For example, let B be the $B\overline{B}$ background surviving after all cuts. To compute it, one might rely either on pure simulation, or on deduction from measured events in which the high-energy γ is replaced by its presumed mesonic parent, or some combination. On the other hand, let C be the continuum background, for which one must rely on off-resonance data subtraction. The uncertainty due to B could then be treated as systematic, while the uncertainty due to C would be explicitly statistical. Any potential uncertainties from adjusting the off-resonance data for the different kinematic endpoint are not considered for this study.

The quantity to be maximized for separating background from signal is the quality factor

$$Q \equiv \frac{S}{\Delta S} \quad . \tag{C.1}$$

If a fraction f of the total luminosity sample is devoted to on-resonance running and $1-f$ to off-resonance, then the nominal yields are

$$N_{on} = \mathrm{S} + \mathrm{B} + \mathrm{C} \tag{C.2}$$

and

$$N_{off} = (1-f)\mathrm{C} \quad . \tag{C.3}$$

One can then solve for an estimate of the signal. (The notation used here does not distinguish this estimate from the "true" signal.)

$$\mathrm{S} = N_{on} - \mathrm{B} - \frac{1}{1-f}\mathrm{C} \quad , \tag{C.4}$$

and for its statistical uncertainty (expressed in terms of the contributing yields)

$$(\Delta \mathrm{S})^2 = \mathrm{S} + \mathrm{B} + \frac{\mathrm{C}}{1-f} \ . \tag{C.5}$$

Of course, for an *overall* error one should add to this in quadrature the systematic uncertainty in B.

The extra weight given here to the subtracted background ought to be taken into account in cut optimization (for f fixed at the standard BABAR value), by maximizing $\mathrm{S}^2/(\mathrm{S}+\mathrm{B}+\mathrm{C}/(1-f))$ rather than the usual $\mathrm{S}^2/(\mathrm{S}+\mathrm{B}+\mathrm{C})$. This is relevant only to analyses in which an actual subtraction is necessary; it would *not* be the right choice when fitting is used for background determination.

An interesting question is: given values of S, B and C (*i.e.* for a given analysis performance), what is the optimum value of f, the on-resonance luminosity fraction?

This may be defined as the value which minimizes $(\Delta \mathrm{S})^2$, *i.e.*, which maximizes Q (for the assumed and fixed performance of the analysis). Letting

$$r \equiv \frac{\mathrm{C}}{\mathrm{S}+\mathrm{B}} \quad , \tag{C.6}$$

the optimum f resulting from this minimization is

$$f_{\mathrm{opt}} = 1 + r - \sqrt{r(1+r)} \quad . \tag{C.7}$$

This has the expected limiting behavior: $f_{\mathrm{opt}} \to 1$ as $r \to 0$, and $f_{\mathrm{opt}} \to 1/2$ for $r \gg 1$. For $f = f_{\mathrm{opt}}$,

$$Q_{\mathrm{opt}} = \frac{\mathrm{S}(\sqrt{1+r} - \sqrt{r})}{\sqrt{\mathrm{S}+\mathrm{B}}} \quad . \tag{C.8}$$

This implies that Q_{opt} for large r is $1/(2\sqrt{r})$ times Q_{opt} for $r \approx 0$.

C.2 Numerical Examples

Table C.1 shows values of f_{opt} for a range of r. Also shown are the ratio of the optimum Q relative to its value with $\mathrm{C} = 0$ (no continuum background), the cost (in terms of Q) of using the standard BABAR $f \approx 0.90$) instead of the optimum value, and also what would happen if the on-resonance fraction were reduced to 0.80. Asymptotically, for $r \gg 1$, the statistical "quality ratio" Q/Q_{opt} approaches $2\sqrt{f(1-f)}$, as can be seen in the table.

Note that inclusion of a realistic uncertainty on B (the $B\overline{B}$ background in the case of fully inclusive $b \to s\gamma$) would improve these quality ratios, at least a bit. Carrying this computation farther requires a model for that extra component.

$r \equiv \mathrm{C}/(\mathrm{S}+\mathrm{B})$	f_{opt}	$Q_{\mathrm{opt}}/\sqrt{(\mathrm{S}+\mathrm{B})/\mathrm{S}^2}$	$Q_{(f=0.90)}/Q_{\mathrm{opt}}$	$Q_{(f=0.80)}/Q_{\mathrm{opt}}$
0.00	1.000	1.000	0.949	0.894
0.01	0.910	0.905	0.999	0.965
0.02	0.877	0.869	0.997	0.982
0.03	0.854	0.842	0.989	0.991
0.05	0.821	0.801	0.967	0.999
0.10	0.768	0.733	0.916	0.997
0.20	0.710	0.648	0.845	0.976
0.30	0.676	0.592	0.801	0.955
0.50	0.634	0.518	0.748	0.924
0.70	0.609	0.467	0.718	0.903
1.00	0.586	0.414	0.691	0.882
2.00	0.561	0.318	0.651	0.848
3.00	0.536	0.268	0.636	0.835
5.00	0.523	0.213	0.622	0.822
10	0.512	0.154	0.612	0.811
100	0.501	0.050	0.601	0.801
1000	0.500	0.016	0.600	0.800

Table C.1: Optimum on-resonance running fraction f, corresponding value of $Q \equiv \mathrm{S}/\Delta\mathrm{S}$ (relative to that for $r = 0$), and cost of using non-optimum f values.

For the presented analysis, 983 signal events (assuming the branching fraction predicted by theory), 616 corrected $B\overline{B}$ events and 569 off-resonance data events survive all cuts. This yields $r = 0.36$ which implies an optimal on-resonance running fraction of $f_{\mathrm{opt}} = 66\%$. Vice versa, from BABAR's 88% on-resonance luminosity fraction a degradation in performance of about 20% can be seen. Changing to an on-resonance fraction of 20% would improve the situation so that only 5% of precision would be lost. If this analysis was the only one of concern a change like this would clearly be advantageous. However, there are a multitude of other analyses done on BABAR data (like the CP violation measurement and any studies that fully reconstruct

B decays) which rely critically on the total on-resonance statistics and favor as high an on-resonance fraction as possible. The 12% currently in use represent the compromise found that does not impact either groups of measurements too much.

Bibliography

[1] J. R. Ellis *et al.*, Nucl. Phys. B **131**, 285 (1977).

[2] CLEO Collaboration, R. Ammar *et al.*, Phys. Rev. Lett. **71**, 674 (1993).

[3] J. Hewett & J. Wells, Phys. Rev. D **55**, 5549 (1997) [hep-ph/9610323].

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

[5] CLEO Collaboration, S. Chen *et al.*, Phys. Rev. Lett. **87**, 251807 (2001) [hep-ex/0108032].

[6] ALEPH Collaboration, R. Barate *et al.*, Phys. Lett. B **429**, 169 (1998).

[7] BELLE Collaboration, K. Abe *et al.*, Phys. Lett. B **511**, 151 (2001) [hep-ex/0103042].

[8] BaBar Collaboration, B. Aubert *et al.*, Contributed to the 31st International Conference on High Energy Physics (2002) [hep-ex/0207074].

[9] BaBar Collaboration, B. Aubert *et al.*, Contributed to the 31st International Conference on High Energy Physics (2002) [hep-ex/0207076].

[10] C. Jessop, "A World Average for $B \to X_s\gamma$", SLAC-PUB-9610 (2002).

[11] A. Seiden, "Particle Physics - A Comprehensive Introduction", to be published in Addison Wesley (2003).

[12] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963);
M. Kobayashi and K. Maskawa, Prog. Theo. Phys. **49**, 652 (1973).

[13] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, in: N. Svartholm (ed.), "Elementary Particle Theory", Almquist and Wiksell (1968).

[14] S. Weinberg, "The Quantum Theory of Fields, Volume III, Supersymmetry", Cambridge University Press (2000);
H.E. Haber and G.L. Kane, Phys. Rep. **117**, 75 (1985).

[15] A.J. Buras *et al.*, Nucl. Phys. B **424**, 374 (1994) [hep-ph/9311345];
J.R. Ellis, D.V. Nanopoulos and K.A. Olive, Phys. Lett. B **508**, 65 (2001) [hep-ph/0102331].

[16] P. Gambino and M. Misiak, Nucl. Phys. B **611**, 338 (2001) [hep-ph/0104034].

[17] Belle Collaboration, Y. Ushiroda *et al.*, "Ise-Shima B physics and CP violation", 71 (2001) [hep-ex/0104045].

[18] BaBar Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **88**, 101805 (2002) [hep-ex/0110065].

[19] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66**, 1 (2002).

[20] A.F. Falk, M. Luke and M. Savage, Phys. Rev. D **49**, 3367 (1994) [hep-ph/9308288].

[21] A. Buras and M. Misiak, submitted to Acta Phys. Pol. B (2002) [hep-ph/0207131].

[22] For a review, see: G. Buchalla, A.J. Buras and M.E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996) [hep-ph/9512380].

[23] S. Bertolini, F. Borzumati and A. Masiero, Phys. Rev. Lett. **59**, 180 (1987);
N.G. Deshpande *et al.*, Phys. Rev. Lett. **59**, 183 (1987).

[24] A.F. Falk and M. Neubert, Phys. Rev. D **47**, 2982 (1993) [hep-ph/9209269].

[25] K. Adel and Y. P. Yao, Phys. Rev. D **49**, 4945 (1994) [hep-ph/9308349];
C. Greub, T. Hurth and D. Wyler, Phys. Lett. B **361**, 146 (1995) [hep-ph/9506374];
C. Greub and T. Hurth, Phys. Rev. D **56**, 2934 (1997) [hep-ph/9703349];
A. J. Buras, A. Kwiatkowski and N. Pott, Phys. Lett. B **414**, 157 (1997) [hep-ph/9707482].

[26] A. Kagan and M. Neubert, Euro. Phys. Jour. **7**, 5 (1999) [hep-ph/9805303].

[27] K. Chetyrkin, M. Misiak, and M. Münz, Phys. Lett. B **400**, 296 (1997) [hep-ph/9612313].

[28] S. Baek, P. Ko, and W.Y. Song, Jour. High En. Phys. **0303**, 054 (2003) [hep-ph/0208112].

[29] Z. Ligeti *et al.*, Phys. Rev. D **60**, 34019 (1999) [hep-ph/9903305].

[30] A. Ali and C. Greub, Phys. Lett. B **361**, 146 (1995) [hep-ph/9506374].

[31] C. Bauer, Phys. Rev. D **57**, 5611 (1998) [hep-ph/9710513].

[32] Z. Ligeti, private communication.

[33] S. Veseli and M.G. Olsson, Phys. Lett. B **367**, 309 (1996) [hep-ph/9508255].

[34] SLAC-R-0418

[35] BaBar Collaboration, B. Aubert *et al.*, Nucl. Instrum. Methods **479**, 1 (2002) [hep-ex/0105044].

[36] Belle Collaboration, A.Abashian *et al.*, Nucl. Instrum. Methods A.**479**, 117 (2002).

[37] P. Billoir, Nucl. Instrum. Methods A.**225**, 335 (1984).

[38] For a summary see *e.g.* : G.S. Fishman, "Monte Carlo: Concepts, Algorithms, and Applications", Springer-Verlag (1996).

[39] A. Ryd *et al.*, BaBar Analysis Document 522 (2003) (BaBar internal), http://www.slac.stanford.edu/~lange/EvtGen (external access).

[40] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).

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

[42] GEANT4 Collaboration, S. Agostinelli *et al.*, SLAC-PUB-9350 (2002).

[43] BaBar Collaboration, H. Quinn *et al.* (eds.), SLAC-R-504 (1997).

[44] F. DiLodovico *et al.*, "Study of $b \to s\gamma$ using the sum of exclusive modes", BaBar Analysis Document 220 (2002) (BaBar internal).

[45] T. Colberg *et al.*, "Study of π^0 Reconstruction Efficiency in Run I Data Using $\tau\tau$ 1-on-1 Decays", BaBar Analysis Document 196 (2000) (BaBar internal).

[46] The BaBar Neutrals Analysis working group, F. DiLodovico *et al.*.

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

[48] A. Drescher *et al.*, Nucl. Instrum. Methods A.**237**, 464 (1985).

[49] C.J. Huberty ,"Applied Discriminant Analysis", Wiley-Interscience (1988).

[50] R.A. Fisher, Ann. Eugen. **7**, 179 (1936).

[51] BABAR note **461**, "B^0 AntiB^0 Tagging Studies", (2001) (BABAR internal).

[52] R. Brun, Nucl. Instrum. Methods A.**389**, 81 (1997);
BABAR Computing Group, T. Adye *et al.*, Comput. Phys. Commun. **150**, 197 (2003) [physics/0206067].

[53] R. Sinkus and T. Voss, Nucl. Instrum. Methods A.**391**, 360 (1997).

[54] BABAR Collaboration, B. Aubert *et al.*, "Search for $B \to \rho\gamma$", to be submitted to Phys. Rev. Lett (2003).

[55] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **86**, 2515 (2001) [hep-ex/0102030].

[56] C. Hearty, "Measurement of the Number of $\Upsilon(4S)$ Mesons Produced in Run 1 (B Counting)", BABAR Analysis Document 134 (2001) (BABAR internal only).

[57] C. Hearty (BABAR), private communication.

[58] A. Gaidot *et al.*, "Linear Discriminant Analysis and Neural Network Approach for BaBar Tagging (Part II)", BABARTagging Note # 7 (BABAR internal only).

[59] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 241803 (2001) [hep-ex/0107068].

[60] BABAR Electron Identification Analysis Working Group, U. Langenegger (ed.), "Cut-Based Electron Identification", BABAR Analysis Document 90 (2000). (BABAR internal)

[61] The BABAR Muon Identification Analysis Working Group, F. Fabozzi (ed.) "Muon Identification in the BABAR Experiment", BABAR Analysis Document 60 (2000). (BABAR internal)

[62] S. Spanier (ed.), "Kaon Selection at BaBar", BABAR Analysis Document 116 (2001). (BABAR internal)

[63] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **66**, 032003 (2002) [hep-ex/0201020].

[64] P.F. Giraud (ed.), "Measurement of inclusive eta and omega momentum spectra and multiplicities", BABAR Analysis Document 163 (2001). (BABAR internal)

[65] A. Hicheur (ed.), "Study of the inclusive and semi-exclusive η' spectrum", BABAR Analysis Document 179 (2001). (BABAR internal)

[66] U. Langenneger (BABAR), private communication.

[67] A.G. Frodesen, O. Skjeggestad, and H. Tøfte, "Probability and Statistics in Particle Physics", Universitetsforlaget (1979).

[68] T. Hurth, Invited Contribution to Rev. Mod. Phys. [hep-ph/0212304].