Measurement of CP Asymmetry in $b \to s\gamma$ Decays Using a Sum of Exclusive Final States

by

Minliang Zhao

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

Author .....
Department of Physics
July, 2008

Certified by .....
Gabriella Sciolla
Cecil and Ida Green Career Development Assistant Professor
Thesis Supervisor

Accepted by .....
THOMAS J. GREYTAK
Associate Head for Education, Physics

Measurement of CP Asymmetry in $b \to s\gamma$ Decays Using a Sum of Exclusive Final States

by

Minliang Zhao

Submitted to the Department of Physics
on July, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

In this thesis, I describe a study of the CP asymmetry in $b \to s\gamma$ decays. This CP asymmetry is expected to be less than 0.01 in the Standard Model, but could be enhanced up to the order of 0.10 by non-Standard Model contributions. We use a sample of 383×10^6 $B\overline{B}$ events collected by the *BABAR* detector at the PEP-II asymmetric B factory and reconstruct sixteen flavor-specific B decay modes containing a high energy photon and a hadronic X_s system. We measure the CP asymmetry to be $-0.011 \pm 0.030_{stat} \pm 0.014_{syst}$ for a photon energy threshold at 1.6 GeV and the hadronic system mass between 0.6 and 2.8 GeV/c^2.

Thesis Supervisor: Gabriella Sciolla
Title: Cecil and Ida Green Career Development Assistant Professor

Acknowledgments

I dedicate this thesis to my parents, for their love and support throughout my life.

Throughout the years of study at MIT, I have constantly received enlightenment and inspiration from many faculty members from the Department of Physics and I express my gratitude to all of them and particularly to Prof. Gabriella Sciolla for giving me the opportunity to undertake my research as part of the *BABAR* Collaboration and guiding me through. It was pleasant to work with Steve Sekula, who consistently helps me on my research. This work cannot be done on without his generous contribution. Denis Dujmic and Ray Cowan are always patient whenever I have a question and seek helps from them. I owe my special thanks to Prof. Peter Fisher for his advice and encouragement on my teaching assistant's work. I slowly cast away my fear for teaching and stand up to speak. I am also grateful for his and Richard Yamatomo's endless effort that is put into organizing the lepton quark study group, trying to create an ideal learning environment for all the students.

I also want to take the chance thank my undergrad mentors at University of Hong Kong for introducing me into the exciting field of particle physics. It was Prof. C.D. Beling who taught me the very first particle physics class, showed me the excitement of this field, and encouraged me to keep exploring. I was fortunate to have Prof. Jason Pun, Alan Poon, and K.S. Cheng to guide me on my final year project. The enjoyment through this very first research experience helped me make up the mind of undertaking more research through a PhD program.

My acknowledgements would not be nearly complete without mentioning the review committee members for their valuable comments to improve this analysis. I owe my special thanks to Prof. Masahiro Morii at the Harvard University. It is his criticism and suggestions that bring this project to its successful completion.

Finally, I want to thank my family and friend, nearby or far apart, for their love and faith in me throughout these many years, for sharing my happiness and sadness whenever I need them. This thesis would not have been possible without their accompany.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The universe is composed of matter particles (protons, neutrons and electrons) rather than their antimatter partners (antiprotons, antineutrons and positrons). The Big Bang should have created equal amounts of matter and antimatter. Why is there now so much of one and so little of the other? *CP* violation is one of the essential elements in answering this question.

CP violation was first discovered in 1964 in the decays of neutral kaons. Its discoverers James Cronin and Val Fitch were awarded the Nobel Prize in Physics in 1980, consequently. In the beginning of this century, it was revealed in the B meson system as well. To date this topic remains a vibrant area of theoretical and experimental work.

CP violation can be explained and calculated through the Standard Model of particle physics or physics beyond the Standard Model. The Standard Model successfully incorporates all known forces (except gravity) and particles into a common framework, and correctly predicates the outcome of the particle interactions, which have been confirmed by a large variety of experiments. Today, it is a well established theory applicable over a wide range of conditions. This section gives an overview of the Standard Model and its extensions.

1.1 The Standard Model and Beyond

1.1.1 The Standard Model

The Standard Model is a theory describing the building blocks of the universe and the fundamental interactions between them. It combines the unified electromagnetic(EM) and weak ("electroweak") interaction and the strong interaction into a theory represented by the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The interactions are mediated by three types of spin-1 gauge bosons (Table 1.1): the strong interaction is carried by the gluons, the weak by the W and Z bosons, and the electromagnetic force by the photons.

Interaction	Acts on	Particles experiencing the force	Mediating boson masses (GeV/c^2)
Weak	Flavor	Quarks, Leptons	$W^{\pm}$: 80.4; Z: 91.2
Electromagnetic	Electric Charge	Electric Charge	γ: 0
Strong	Color Charge	Quarks, Gluons	Gluons: 0

Table 1.1: Properties of the Standard Model interactions and force mediating bosons. Boson masses shown are theoretical values.

Generation 1	Generation 2	Generation 3	Electric Charge	Interactions
Quarks				
Up u	Charm c	Top t	2/3	EM, weak, strong
Down d	Strange s	Bottom b	-1/3	EM, weak, strong
Leptons				
Electron e	Muon μ	Tau τ	-1	EM, weak
Electron Neutrino ν_e	Electron Neutrino ν_μ	Tau Neutrino ν_τ	0	weak

Table 1.2: Properties of the Standard Model quarks and leptons.

In the Standard Model, matter is made up of quarks and leptons that can be organized into three generations (Table 1.2). Each lepton generation contains a left-handed doublet and one right-handed singlet, while each quark generation contains one more right-handed singlet than leptons. The neutrinos, because they are massless, are purely left-handed. The doublets and singlets, listed below, are unified into the

$SU(2)_L \times U(1)_Y$ symmetry group. The quarks also carry one of three color charges of the strong interaction, and are triplets under the $SU(3)_C$ group. Each elementary particle in the Standard Model has an associated anti-particle carrying conjugated internal quantum numbers, such as electromagnetic charge.

$$\begin{pmatrix} \nu_e \\ e_L^- \end{pmatrix} \quad e_R^- \qquad \begin{pmatrix} \nu_\mu \\ \mu_L^- \end{pmatrix} \quad \mu_R^- \qquad \begin{pmatrix} \nu_\tau \\ \tau_L^- \end{pmatrix} \quad \tau_R^-$$

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad u_R, d_R \qquad \begin{pmatrix} c_L \\ s_L \end{pmatrix} \quad c_R, s_R \qquad \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad t_R, b_R$$

One last particle predicted by the Standard Model but has not been observed is the Higgs boson. The Higgs boson plays a unique role in the Standard Model: they are critical in explaining the origins of the mass of other elementary particles. To date (June 2008), no experiment has directly detected the existence of the Higgs boson. Indirect constraint on the mass of the Higgs boson is based on precision measurements of the masses of $W^\pm$ (Z^0) bosons and t quark. As of now, an upper limit of 182 GeV/c^2 (at 95% C.L.) on Higgs mass is obtained [37].

1.1.2 Physics beyond the Standard Model

To date, almost all experimental tests of the three forces described by the Standard Model agree with its predictions. However, there are fundamental questions that make us believe that there exists some more complete theory. The most outstanding challenges to the Standard Model are:

1. the hierarchy problem. A hierarchy problem occurs when the fundamental parameters (couplings or masses) of some Lagrangian are vastly different (usually larger) than the parameters measured by experiment. As mentioned above, indirect constraints indicate that the Standard Model Higgs boson should have a mass of about 182 GeV/c^2 or less [37]. The quantum corrections to the Higgs mass (Figure 1-1 (a)), however, are quadratically divergent and can be many orders of magnitude larger than this.

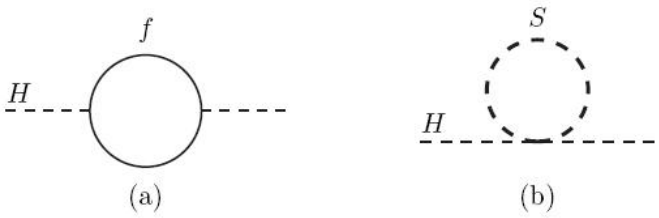


Figure 1-1: Cancelation of the Higgs boson quadratic mass between the quark(f) loop and scalar S squark Feynman diagrams in SUSY [34].

2. unification of forces. The strengths of the strong, weak, and electromagnetic gauge couplings evolve as a function of the energy scale. In the Standard Model these couplings do not unify at any energy scale.

Among new physics theories, Supersymmetric(SUSY) models are the most well-motivated and thoroughly studied. It introduces a new set of superpartner bosons for each of the Standard Model fermions, and vice versa. The Standard Model quarks and leptons are paired with squarks ($\tilde{q}$) and sleptons ($\tilde{l}$); the gluons are paired with gluinos ($\tilde{g}$).

In SUSY there is no hierarchy problem. The divergence of the Higgs boson mass is tremendously reduced by the cancelation between fermionic and bosonic superpartner interactions with Higgs (Figure 1-1) [34]. SUSY also provides a framework to unify electroweak and strong interactions. The superpartners modify the observed strengths of the interactions in such a way that their inherent strengths become equal at very short distances (grand unification) [34].

SUSY is expected to exist close to the TeV energy scale. Much work has been put into the search for new physics. As of now there is no conclusive experimental evidence for it. Study of CP violation in the B meson system provides a unique way to test the Standard Model and probe new physics. We explain the CP violation and its origin within the Standard Model in the next section.

1.2 CP Violation and the CKM Matrix

CP is the product of two symmetries: C, the charge conjugation and P, the parity. We start this section by discussing these discrete symmetries.

1.2.1 Discrete symmetries

In the Standard Model of particle physics there are three related discrete symmetries.

Parity (P)**.** Parity transformations flip the sign of all spatial coordinates. In a three dimensional space, it is also commonly described as $P(x, y, z) \rightarrow (-x, -y, -z)$.

Charge Conjugation (C)**.** Charge conjugation replaces of each particle in a process by the corresponding antiparticle. This transformations do not alter the chirality of particles.

Time reversal (T)**.** The time reversal operator leaves the spatial direction unchanged, but reverses the momentum and angular momentum. Time invariance implies that the physics does not change if the direction of time is made to run backwards instead of forward.

Among the above three discrete symmetries, charge conjugation has been known to be violated from the beginning because people found no left-handed anti-neutrinos, which are the charge conjugate partners for the left-handed neutrinos.

Parity conservation, on the contrary, was believed to be one of the fundamental conservation laws along with conservation of energy and momentum. This changed in 1956, when the study of beta decay of Cobalt-60 nuclei experiment, carried out by a group led by Chien-Shiung Wu, demonstrated conclusively that the parity symmetry is violated. They found that when a specific nucleus was placed in a magnetic field, electrons from the beta decay were preferentially emitted in the direction opposite to that of the aligned angular momentum of the nucleus. Under the Parity transformation, however, it would look like the electrons prefer to be aligned with angular momentum, since the Parity transformation flips the direction of the electron but not that of the angular momentum. As a result, parity is not conserved.

Although charge conjugation and parity are violated separately, the symmetry of a quantum-mechanical system can still be restored if the combined symmetry *CP* remains unbroken. For example, a left handed neutrino, under the *CP* transformation, will be turned into a right handed antineutrino, which exists. Therefore *CP* symmetry was proposed in 1957 by Lev Landau as the true symmetry between matter and antimatter.

Until 1964 it was thought that the combination *CP* was a valid symmetry of the universe. That year, Christenson, Cronin, Fitch, and Turlay observed the decay of the long-lived neutral K meson to 2 pions and thus proving that *CP* is violated in the kaon system [9]. In the 1990s, two B-factories, PEP-II at the Stanford Linear Accelerator Center (SLAC) and KEKB at KEK, were designed for the purpose of studying *CP* violation in the B system. Since 1999, a large number of *CP* violation processes in B meson decays have been observed [37].

Although C, P, and *CP* are proven to be broken symmetries, the product CPT is firmly believed to be conserved. It is the fundamental property of the Quantum Field Theory.

1.2.2 The CKM Matrix

In the Standard Model, *CP* violation originates from a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which is used to describe the quark mixing. In Table 1.2, we show the quarks experience both weak and the strong force. However, in the Standard Model, their weak interaction eigenstates (d',s',b') are different from their strong interaction eigenstates (d,s,b), and two are related by the CKM matrix (V_{CKM})

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

The CPT invariance guarantees that V_{CKM} is a unitary matrix, i.e. $VV^{\dagger}$=1. A 3 $\times$ 3 unitary matrix can be parameterized in terms of three angles and an imaginary

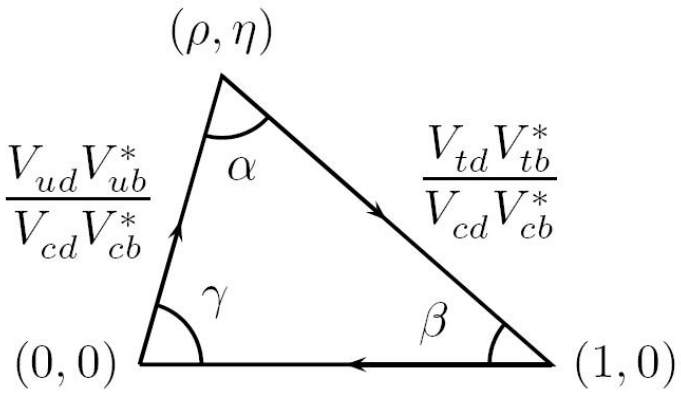


Figure 1-2: The unitarity triangle.

phase. Within the Standard Model, this imaginary phase is the only source of *CP* violating asymmetries in the quark sector. The hierarchy between elements of the CKM matrix is more apparent in the Wolfenstein parameterization [40]:

$$\begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

Here $\lambda \equiv |V_{us}| \equiv \sin(\theta_C)$, with θ_C being the Cabibbo angle, is used as an expansion parameter.

One of the six unitarity conditions of the CKM matrix can be expressed in terms of the matrix elements as:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.1}$$

which is particularly useful for describing *CP* violation in the B-meson system. After scaling, Equation 1.1 describes a triangle in the complex plane with an apex at the point (ρ, η), (Figure 1-2). The parameter λ is measured from $K^+ \to \pi^0 l^+ \nu$ decays. A is related to V_{cb} and can be determined from the study of semileptonic B decays such as $B \to D^* l\nu$. The parameters ρ and η are measured in the study of *CP* violation.

The experimental results of all the CKM matrix elements as of now are [37]:

$$\begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272^{+0.0010}_{-0.0010} & (3.96^{+0.09}_{-0.09}) \times 10^{-3} \\ 0.2271^{+0.0010}_{-0.0010} & 0.97296^{+0.00024}_{-0.00024} & (42.21^{+0.10}_{-0.80}) \times 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \times 10^{-3} & (41.61^{+0.12}_{-0.78}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix}$$

The experimental determination of all sides and angles of the Unitarity Triangle is a major ongoing activity in particle physics. To date all measurements are consistent with the hypothesis that the CKM mechanism is the only source of *CP* violation in the quark sector. In this document, we study of the *CP* violation in the $b \to s\gamma$ decays to provide another constraint to the CKM matrix.

1.3 *CP* Violation in the $b \to s\gamma$ Decays

The $b \to s\gamma$ decays are flavor changing neutral currents. In the Standard Model, they are forbidden at the tree level. The leading order diagrams proceed via a virtual quark and a W boson in the loop, with t quark loop dominating over those with lighter quarks (Figure 1-3). Both the decay rates and the asymmetry are important in testing the Standard Model.

- The $b \to s\gamma$ inclusive branching fraction is theoretically calculated [6, 35] and experimentally confirmed [27] at the NNLO. Therefore it is capable of providing important constraints on the Standard Model and many theories of new physics with very high precision. It also provides a direct estimation of the ratio of the CKM matrix $|V_{td}/V_{ts}|$ from the ratio of the branching fraction of the processes $b \to d\gamma$ and $b \to s\gamma$.

- The *CP* asymmetry between $\overline{b} \to \overline{s}\gamma$ and $b \to s\gamma$ provides an independent test of the Standard Model. It is predicted to be 0.01 in the Standard Model [30]. However, in extensions of the Standard Model with new particles and *CP*-violating coupling constants, large *CP* asymmetries are possible without conflicting the branching fraction [30, 31].

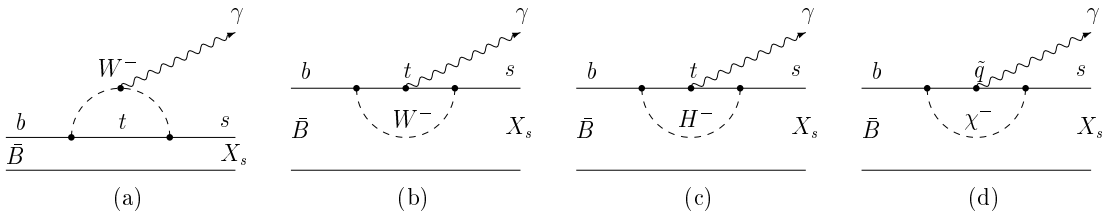


Figure 1-3: Leading order Feynman diagrams for $b \to s\gamma$ decays in the Standard Model (a,b) and for New Physics contributions from a charged Higgs (c) and a chargino predicted by supersymmetric theories (d).

In this thesis, we will focus on studying the *CP* asymmetry in the $b \to s\gamma$ decays.

1.3.1 Theoretical Predictions

The physics of heavy quark decays is often described using the Operator Product Expansion (OPE) technique [39], which aims to replace the product of operators with a linear combination of local operators. In the case of $b \to s\gamma$ transition, the effective Hamiltonian at the energy scale $\mu \sim m_b$ is given by [8]:

$$\mathcal{H}_{eff}(b \to s\gamma) = -\frac{G_F}{\sqrt{2}} V_{ts}^* V_{tb} [\sum_{i=1}^{8} C_i(\mu) \cdot O_i(\mu) + \epsilon_s \sum_{i=1}^{2} C_i(\mu) \cdot (O_i - O_i^u(\mu)], \quad (1.2)$$

where $\epsilon_s = |V_{us}^* V_{ub} / V_{ts}^* V_{tb}| \approx \lambda^2(i\eta - \rho) = \mathcal{O}(10^{-2})$. The G_F and $V_{ts}^* V_{tb}$ terms are the Fermi coupling constant and the CKM matrix elements associated with the $b \to t \to s$ transition, respectively. $O_i(\mu)$ denotes the local operators generated by strong and electroweak interactions, and $C_i(\mu)$ are the Wilson coefficient functions which now contain the information of the physics. The most relevant operators to $b \to s\gamma$ decays are:

$$O_2 = \overline{s}_L \gamma_\mu c_L \overline{c}_L \gamma^\mu b_L \quad , \quad O_7 = \frac{e m_b}{4\pi^2} \overline{s} \sigma_{\mu\nu} F^{\mu\nu} b_R$$
$$O_2^u = \overline{s}_L \gamma_\mu u_L \overline{u}_L \gamma^\mu b_L \quad , \quad O_8 = \frac{g_s m_b}{4\pi^2} \overline{s} \sigma_{\mu\nu} G^{\mu\nu} b_R.$$

CP violation arises from the interference of non-trivial weak phases, contained in

the CKM matrix elements (or Wilson coefficients in new physics), with strong phase provided by the imaginary parts of the local operators $(O_1 - O_8)$[5]. The imaginary strong phase is first seen at $\mathcal{O}(\alpha_s)$ from charm, light quarks or gluons loop diagrams [30].

Using the effective heavy quark theory, A_{CP} in $b \to s\gamma$ decays follows from the above Hamiltonian and can be expressed as [31]:

$$\begin{aligned} A_{CP}^{b\to s\gamma} &= \frac{\Gamma(\overline{B} \to X_s\gamma) - \Gamma(B \to X_{\overline{s}}\gamma)}{\Gamma(\overline{B} \to X_s\gamma) + \Gamma(B \to X_{\overline{s}}\gamma)} \\ &\simeq \frac{4\alpha_s(m_b)}{9|C_7|^2}\{0.12\mathrm{Im}[C_2C_7^*] - \mathrm{Im}[C_8C_7^*] \\ &\quad +0.01\mathrm{Im}[C_2C_8^*] - 0.99\mathrm{Im}[\epsilon_s C_2(C_7^* - 0.013C_8^*)]\} \end{aligned}$$

where the strong coupling constant $\alpha_s(m_b)$ is taken as 0.214 [30].

In the Standard Model, the Wilson coefficients take real values $C_2 \approx 1.11$, $C_7 \approx -0.31$ and $C_8 \approx -0.15$ [30]. The imaginary part of the ϵ_s is therefore the only source of *CP* violation. The latest estimate of the Standard Model *CP* violation in $b \to s\gamma$ decays is [28]:

$$A_{CP} = (0.0042 \pm 0.0008_{(m_c/m_b)} \pm 0.0003_{CKM}{}^{+0.0015}_{-0.0008}|_{scale}). \tag{1.3}$$

The small value of the *CP* asymmetry is a result of three suppression factors:

1. $\alpha(m_b)$ arising from the strong phase;

2. λ^2 representing CKM suppression, reflected in ϵ_s;

3. m_c/m_b resulting from the GIM suppression.

However, amplification of *CP* violation may arise from new physics contributions in the imaginary part of the Wilson coefficients, which are also seen at $\mathcal{O}(\alpha_s)$ from loop diagrams containing light quarks or gluons, competing with the Standard Model contribution. In SUSY, possible *CP* violations of up to 0.15 would be seen in radiative decays like $b \to s\gamma$ in the presence of chargino loops (Figure 1-3), or if non-CKM-

like intergenerational squark mixing is allowed [10, 32]. Large asymmetries are also suggested by SUSY with squark mixing or R-parity violation [31].

1.3.2 Current Experimental Results

The first measurement of the CP asymmetry in $b \to s\gamma$ decays was performed by CLEO [12]. They used two methods to reconstruct the $b \to s\gamma$ decay. The first one is the "inclusive" method, where they accept events with a high energy photon on the signal B side, and use a high momentum lepton from the other B to determine the flavor of the B meson. This technique is explained in details in Section 5.1.1. This method does not distinguish between $b \to s\gamma$ or $b \to d\gamma$. The second one is the "sum-of-exclusive" method, where they reconstruct the decays using several $b \to s\gamma$ exclusive modes. The flavor of the signal B can thus be determined from the charge of the reconstructed final state (B^-/B^+) or the charge of the kaon ($\overline{B}^0/B^0$). This method measures the $b \to s\gamma$ decays only. Therefore their final CP asymmetry result is a weighted average (Table 1.3).

Since then, the B factories have accumulated data sets in excess of 300 times that used by CLEO. In 2003 Belle published their $A_{CP}^{b\to s\gamma}$ result (Table 1.3) based on 152M $B\overline{B}$ decays [7]. In this analysis they used the sum-of-exclusive method to reconstruct the hadronic system by combining one charged or neutral kaon with one to four pions, where at most one pion can be neutral. *BABAR* also measured this variable using various techniques. In 2004, we published result base on 89M $B\overline{B}$. Using the sum-of-exclusive method, we measured $A_{CP} = 0.025 \pm 0.050(\text{stat}) \pm 0.015(\text{syst})$ in the 12 exclusive modes we reconstructed [15]. The result with the "inclusive" method was published in 2006 as $A_{CP} = -0.110 \pm 0.115(\text{stat}) \pm 0.017(\text{syst})$ [18].

These results are summarized in Table 1.3. As shown, all measurements are limited by the statistics and the best measurements so far have reached a precision of about 0.05. While many extreme new physics scenarios have already been ruled out, we have not yet reached the precision needed to test the Standard Model prediction, which is at 0.01 level. In particular, some flavor-violating effects from physics beyond the Standard Model have proven to be within the limit set by the previous measurements.

Experiment/Method	$A_{CP}(b \to s\gamma)$
CLEO/Inclusive (10M $B\overline{B}$) [12]	$(-0.079 \pm 0.108 \pm 0.022)(1.0 \pm 0.030)$
Belle/Sum-of-exclusive (152M $B\overline{B}$) [7]	$0.002 \pm 0.050 \pm 0.030$
BaBar/Inclusive (89M $B\overline{B}$) [18]	$-0.110 \pm 0.115 \pm 0.017$
BaBar/Sum-of-exclusive (89M $B\overline{B}$) [15]	$0.025 \pm 0.050 \pm 0.015$

Table 1.3: Current results on the measurement of *CP* asymmetry (A_{CP}) in $b \to s\gamma$ decays. The first error is statistical and the second error is systematic.

For example, new physics with CKM unitarity violation only modifies $A_{CP}(b \to s\gamma)$ to be ~ 0.02 [30], and with the Minimum Flavor Violation model in SUSY, the *CP* asymmetry could be less than 0.03 [28].

As pointed out by reference [22], more data is needed to improve the precision of this measurement. This work describes the measurement of *CP* asymmetry in this penguin decay as close to the level of the Standard Model as it is possible with the recent *BABAR* data sample.

Chapter 2

The *BABAR* Experiment

The *BABAR* experiment is located at Stanford Linear Accelerator Center (SLAC) in Menlo Park, California. It makes use of the asymmetric energy e^+e^- collider PEP-II at SLAC. The *BABAR* experiment started taking data in 1999 and ceased operating in April, 2008, delivering a total integrated luminosity of over 540 fb^{-1} (Figure 2-1). The data has been collected in seven run periods, interrupted by shutdowns for maintenance and upgrade.

2.1 PEP-II

The PEP-II collider consists of two storage rings. The high energy ring stores a 9 GeV electron beam while the low energy ring stores a 3.1 GeV positron beam. The PEP-II collider therefore operates at the center of mass energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance.

The e^+e^- collisions at the $\Upsilon(4S)$ resonance provide a very clean channel for B meson production, with a favorable ratio of $b\overline{b}$ production compared to lighter quark pairs ($\sigma(e^+e^- \to b\overline{b})/\sigma(e^+e^- \to q\bar{q}) \approx 0.28$, c.f. Table 3.1). The asymmetric beam energy provides a boost of 0.56 to the B meson pairs in the center of mass (CM) frame with respect to the lab frame. With this boost, the average separation along the z axis between the two B vertices about 270 μm, allowing the measurement of the B and $\bar{B}$ decay times, which is critical for studying time-dependent CP violation.

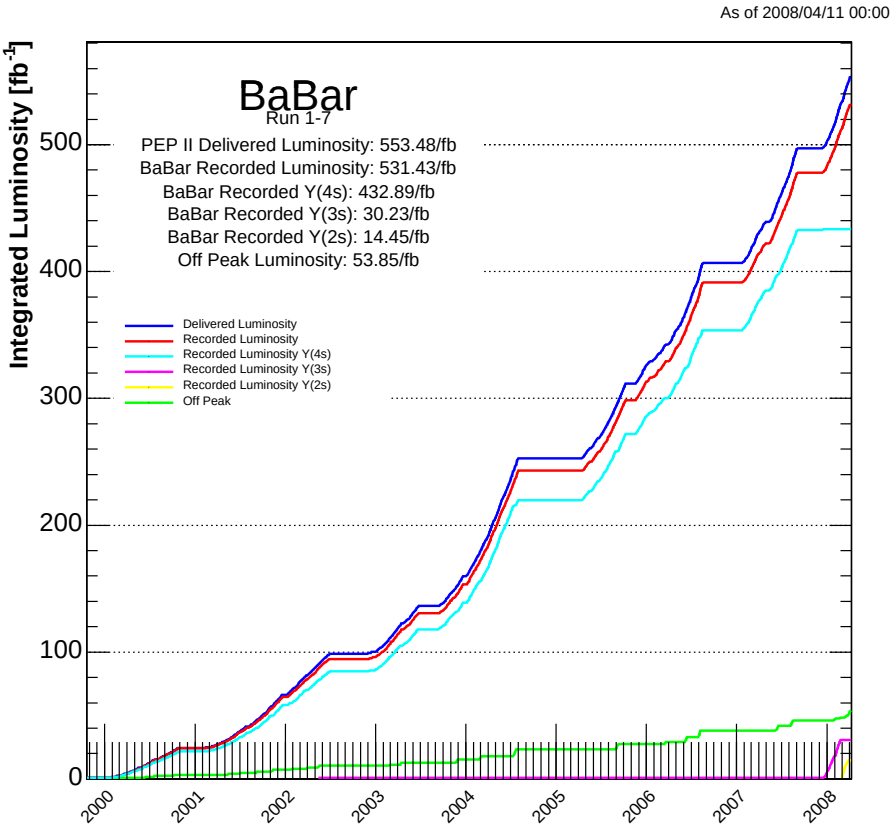


Figure 2-1: Integrated luminosity as a function of time.

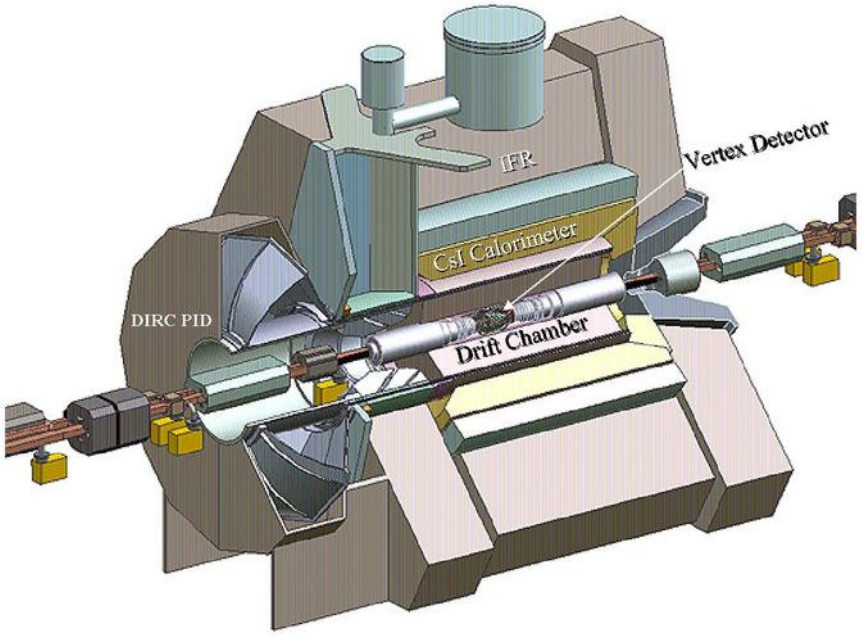


Figure 2-2: A cartoon of the *BABAR* detector.

This concept of asymmetric B-factories was first proposed by Pier Oddone in 1987.

2.2 The *BABAR* Detector

The *BABAR* detector consists of five concentric subsystems, as detailed in the following sections. It is asymmetric with respect to the collision point by design because of the B mesons are boosted along the beam line. Figure 2-2 shows a cartoon of the *BABAR* detector. The five subsystems (from innermost) are: a silicon vertex tracker (SVT), a drift chamber (DCH), the detector of internally reflected Cherenkov light (DIRC), a CsI electromagnetic calorimeter (EMC), and an instrumented flux return (IFR). Between the EMC and the IFR there is a super-conducting solenoid, which creates an uniform 1.5 Tesla magnetic field. Figure 2-3 shows a longitudinal view of the *BABAR* detector. In the following chapters the z-direction is defined to be along the beam line, while the ϕ and the θ are the polar and azimuth angle, respectively.

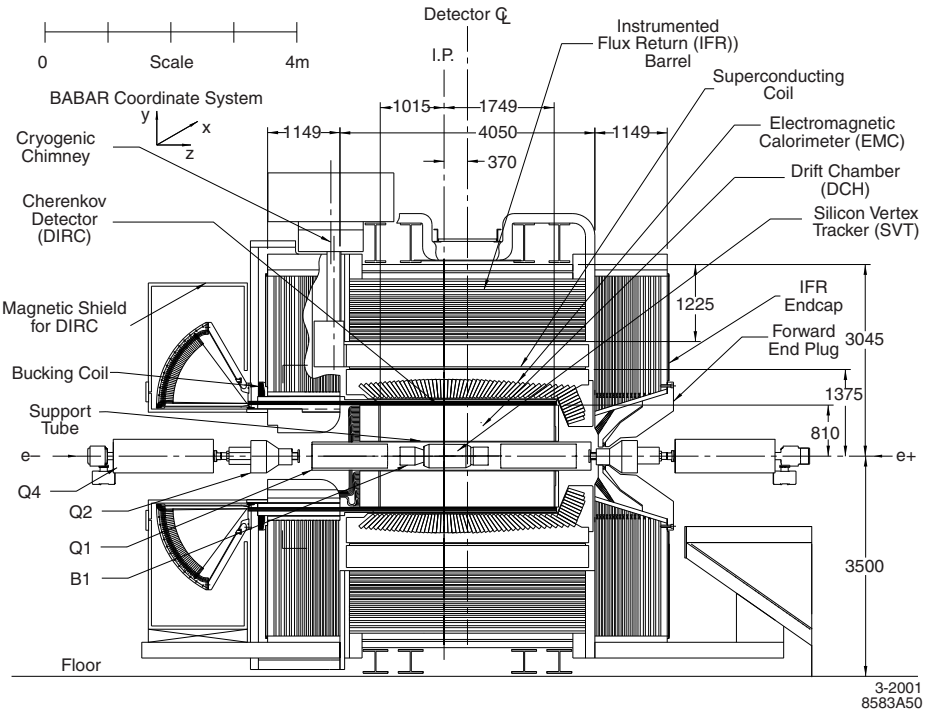


Figure 2-3: *BABAR* detector longitudinal section.

2.2.1 Silicon Vertex Tracker (SVT)

The silicon vertex tracker is the innermost part of the *BABAR* detector. It consists of five layers of double-sided silicon strip sensors which are placed in a cylindrical geometry around the beam axis at the interaction point, as shown in Figure 2-4. The silicon layers are positioned at 32, 40, 54, 91-127, and 114-144 millimeters away from the interaction point respectively.

The strips on the opposite sides of each sensor are oriented orthogonal to each other, with the ϕ measuring strips running parallel to the beam on the outer side of the sensor, and the z measuring strips laying transverse to the beam axis on the inner side. The distance between the sensor and the interaction point provides the third coordinate of the particle position in addition to two measured coordinates in the z an ϕ directions.

The SVT has full coverage in the azimuthal direction and extends from $-29.8°$ to $20.1°$ in polar angle. This covers about 90% of the solid angle in the center of

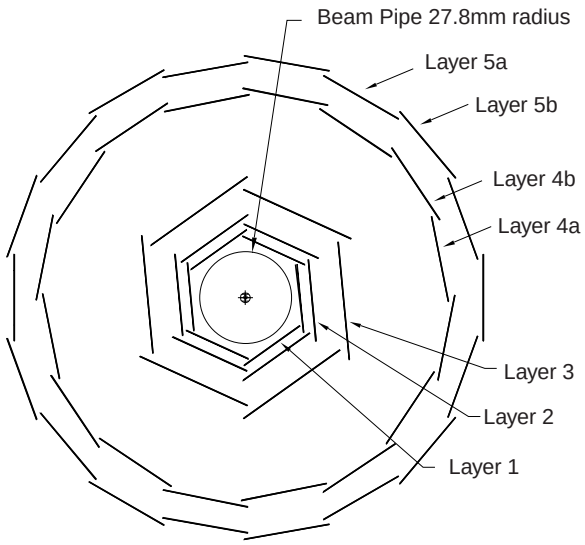


Figure 2-4: *BABAR* Schematic view of SVT: transverse section.

mass or $\Upsilon(4S)$ frame. It provides a resolution better than 40 μm in the z (Figure 2-5(a)). which is crucial in resolving the separation of the vertices for the two B-mesons, whose average separation along z axis is estimated at 270 μm as mentioned above. The resolution in the ϕ direction is generally better than 40 μm as well (Figure 2-5(b)).

2.2.2 Drift Chamber (DCH)

The drift chamber is designed to provide high precision measurement of the charged particle position and its momenta. For low momentum particles, the DCH also provides particle identification by measurement of ionization loss (dE/dx).

The DCH is a 280 cm long cylinder with an inner radius of 23.6 cm and an outer radius of 80.9 cm (Figure 2-6). It is divided to into 10 super-layers with 4 layers of hexagonal cells in each of them. Each hexagonal cell consists of a sense wire in the center of six field wires. The sense wires are kept at a voltage of 1960 Volts. The stereo angles of the super-layers alternate between axial (A) and stereo (U,V) pairs, in the order of AUVAUVAUVA, in purpose of achieving a three-dimensional track reconstruction. U and V represents super-layers titled in the opposite direction with

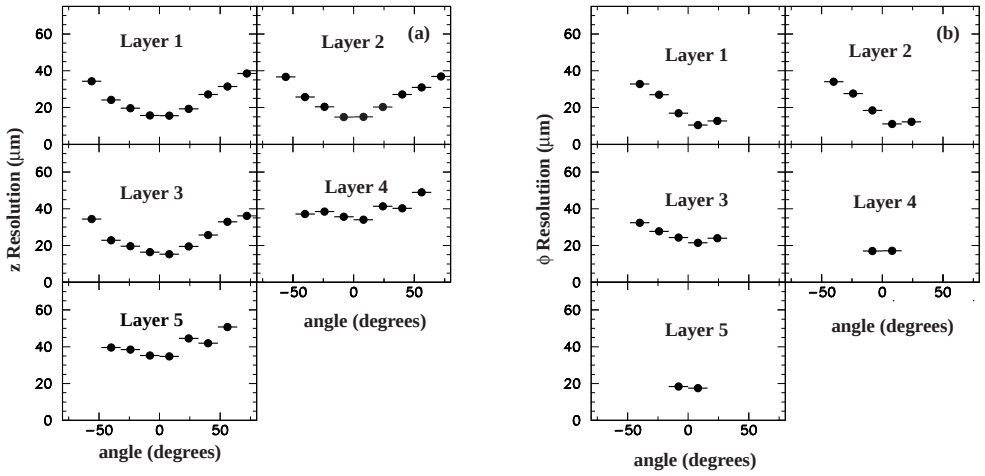


Figure 2-5: SVT hit resolution in the a) z and b) ϕ coordinate in microns, plotted as a function of track incident angle in degrees. Each plot shows a different layer of the SVT. There are fewer points in the ϕ resolution plots for the outer layers as they subtend smaller angles than the inner layers.

respect to the axial direction, which is the beam line. The volume of the DCH is filled with an 80:20 helium:isobutane gas mixture.

When a charged particle enters the DCH, it deposits energy and causes ionization in the immediate environment. The ionization electrons will drift toward the sense wires, around which they undergo avalanche multiplication. The energy deposited by the charged particle is proportional to the number of primary electrons ionized.

The DCH spatial resolution is better than 140 μm. The resolution in the transverse momentum is found to be [14].

$$\frac{\sigma_{p_t}}{p_t} = (0.13 \pm 0.01)\% \cdot p_t + (0.45 \pm 0.03)\% \tag{2.1}$$

The dE/dx of the charged particle is derived from the measurement of the total charge deposited in each drift cell. The momentum is calculated by extracting the radius of the curvature of the charged particle in the DCH caused by the Lorentz force. The drift chamber has a dE/dx resolution of 7.5% for separation of low momentum particles; this corresponds to a separation better than 2σ for charged kaons and pions

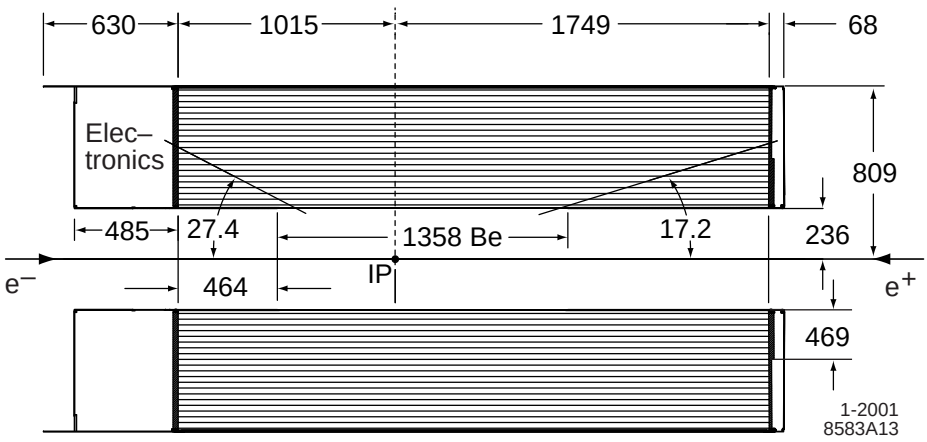


Figure 2-6: Longitudinal section of the DCH with principal dimensions; the chamber center is offset by 370 mm from the interaction point (IP).

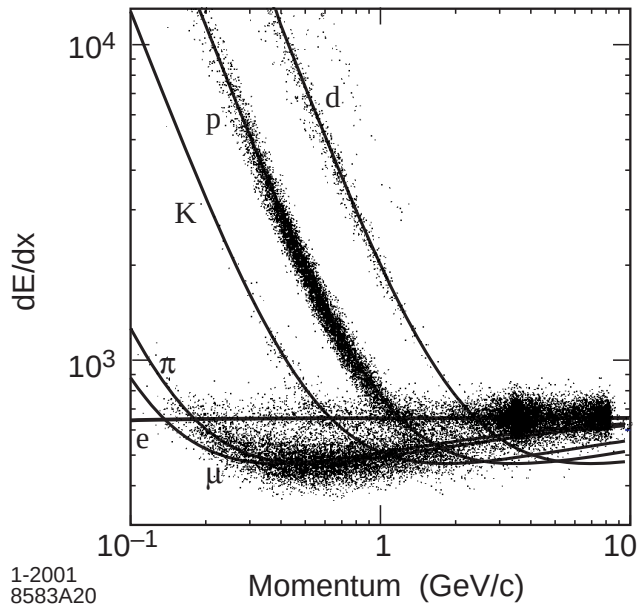


Figure 2-7: Measurement of dE/dx in the DCH as a function of track momenta. The data include large samples of beam background triggers, as evident from the high rate of protons. The curves show the Bethe-Bloch predictions derived from selected control samples of particles of different masses.

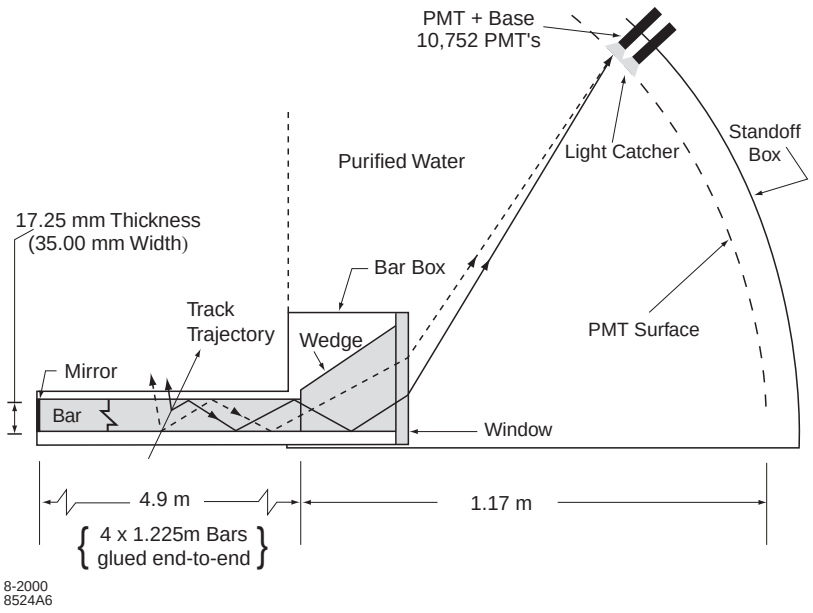


Figure 2-8: Schematics of the DIRC fused silica radiator bar and imaging region.

at momenta up to 0.7 GeV/c (Figure 2-7).

2.2.3 Detector of Internally Reflected Cherenkov Light (DIRC)

The study of CP violation requires the ability to tag the flavor of one of the B mesons through the hadronic decays. Excellent pion/kaon separation plays a crucial role in this process because when a kaon is mis-identified to be a pion (or vice versa), it is quite possible to flip the flavor of the B meson and therefore dilute the measured CP asymmetry. A good example to illustrate the above mentioned scenario would be the decay of $B^0 \to K^+\pi^-$. When the K^+ is mis-identified to be a π^+, and the π^- is recognized to be a K^-, the B meson will be mis-reconstructed to be $\overline{B}^0$.

The DIRC provides excellent particle identification through cherenkov radiation. It is a ring imaging Cherenkov detector based on total internal reflection and uses long, rectangular bars made from synthetic fused silica ("quartz") as both radiator and light guide. Figure 2-8 shows the upper half of the layout of the fused silica radiator bar and imaging region seen from the longitudinal side. In order to avoid instrumenting both ends with photomultiplier tubes (PMT), the bars are connected

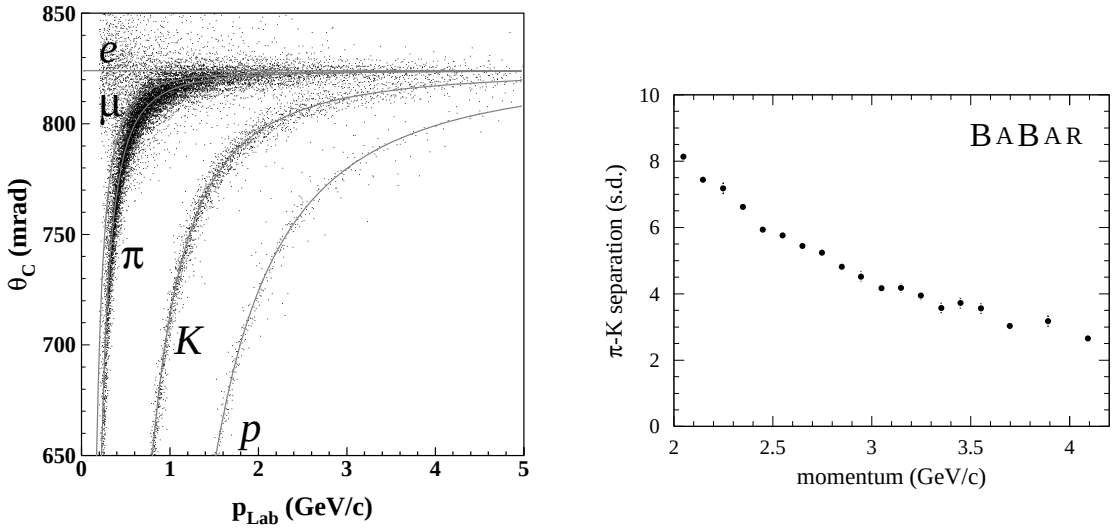


Figure 2-9: Performance plots of the DIRC: (a) Cherenkov angle θ_C versus the momentum of the charged tracks at the entrance to the DIRC. The solid lines correspond to the fitted values for the different particle species; (b) DIRC $\pi^{\pm}/K^{\pm}$ separation versus track momentum measured in $D^0 \to K^-\pi^+$ decays selected kinematically from inclusive D^* production.

to a perpendicular mirror on one end and a water-filled container on the other. The container is covered by an array of PMTs for photon detection.

The DIRC Cherenkov angle resolution is 2.5 mrad [14], as evaluated using samples of $e^+e^- \to \mu^+\mu^-$. The particle identification performance of the DIRC is evaluated using samples of well identified hadrons from $K^0_S \to \pi^+\pi^-$, $\Lambda \to \pi p$, and $D^{*+} \to \pi^+(D^0 \to K^-\pi^+)$ decays. The DIRC is particularly effective at separating charged kaons and pions with higher lab momenta, as illustrated in Figure 2-9.

2.2.4 Electromagnetic Calorimeter (EMC)

The electromagnetic calorimeter is designed to measure electromagnetic showers over an energy range from 20 MeV to 9 GeV. The calorimeter consists of 6580 Thallium doped Caesium-Iodide crystals (CsI(Tl)) and is split into a barrel and a conical endcap section. The EMC has a full coverage in azimuth and extends in polar angle from 15.8° to 141.8°, corresponding to a solid-angle coverage of 90% in the CM frame. The barrel contains 5760 crystals arranged in 48 rows in θ and 120 in ϕ, whereas the

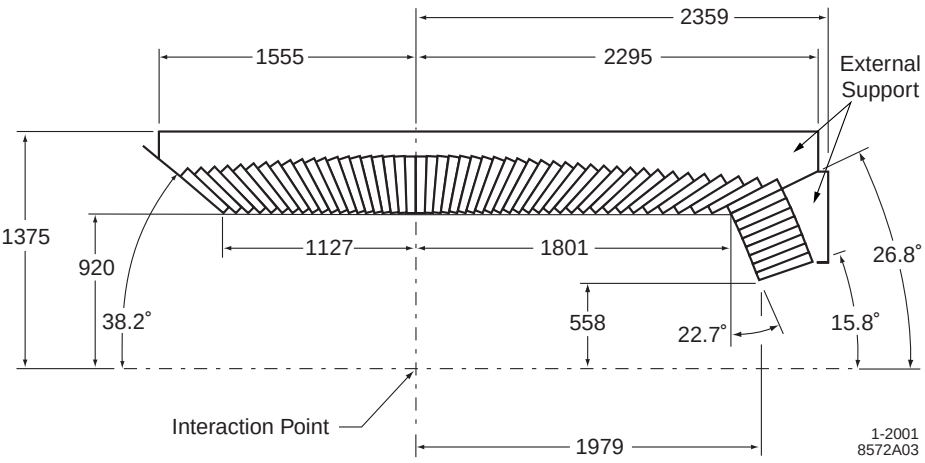


Figure 2-10: A longitudinal cross section of the EMC (only the top half is shown) indicating the arrangement of the 56 crystal rings. The detector is axially symmetric around the z-axis. All dimensions are given in mm.

endcap holds 820 crystals in 8 rings, with 80 or 120 crystals in each (Figure 2-10). Two photodiodes are mounted at the rear end of each crystal. The crystals act as a total absorption scintillating material, where the scintillating light is proportional to the deposited energy and will be converted into a measurable electric pulse by the photodiodes.

The EMC angular and energy resolutions are very important for reconstructing the correct photon energy and hadron mass in this analysis. CsI(Tl) has a high light yield and small Molière radius that allows for excellent energy and angular resolution.

Different control samples are used to derive the energy resolution in different energy regimes. At high energy (greater than 3 GeV), it is determined from Bhabha scattering, where the energy of the detected shower can be predicted from the polar angle of the $e^{\pm}$. Below 2 GeV, the mass resolution of π^0 and η mesons decaying into two photons of approximately equal energy is used to infer the EMC energy resolution. The $\chi_{c1} \to J/\psi\gamma$ decays provides a measurement at an average energy of about 500 MeV. The energy resolution of a homogeneous crystal calorimeter can be described empirically in terms of a sum of two terms added in quadrature:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus b. \tag{2.2}$$

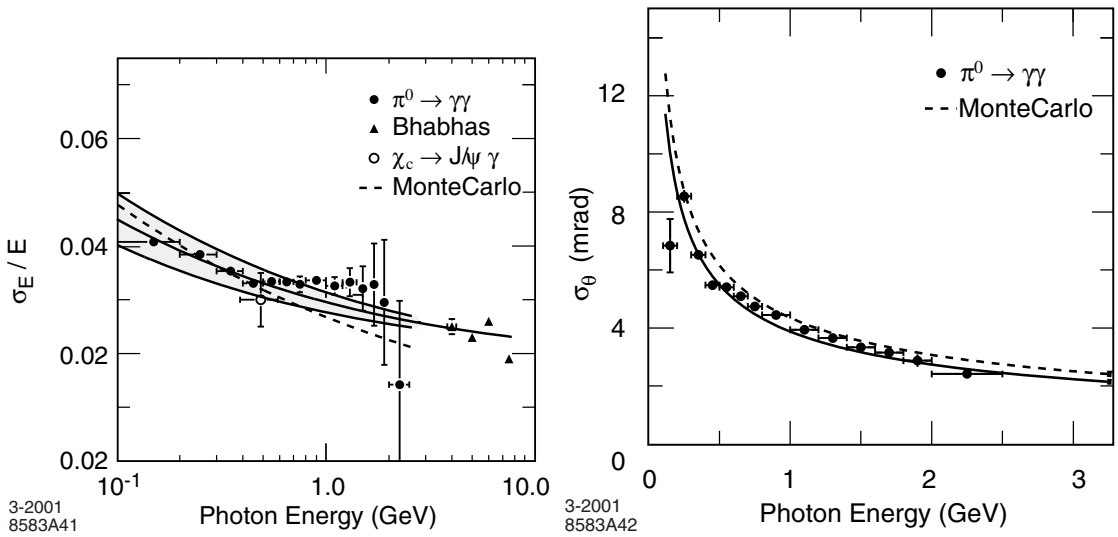


Figure 2-11: Performance plots of the EMC: (a) The energy resolution for the EMC measured for photons and electrons from various processes. The solid curve is a fit to Equation 2.2 and the shaded area denotes the rms error of the fit; (b) The angular resolution of the EMC for photons from π^0 decays. The solid curve is a fit to Equation 2.3.

A fit to the above energy dependance yields: $a = (2.32 \pm 0.30)\%$ and $b = (1.85 \pm 0.12)\%$ [14].

The angular resolution is measured using π^0 and η decays to two photons of approximately equal energy. The resolution varies between 3 mrad at high energies, to 12 mrad at low energies. A fit to the empirical parameterization of the energy dependance results in:

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} + d \tag{2.3}$$

with $c = (3.87 \pm 0.07)$ mrad and $d = (0.00 \pm 0.04)$ mrad [14].

The resolution measured in data control samples and in simulation as a function of photon energy is shown in Figure 2-11.

2.2.5 Instrumented Flux Return (IFR)

The IFR is designed to provide particle identification to muons and detect neutral hadrons (mainly K_L^0 and neutrons), as well as to provide the magnetic flux return

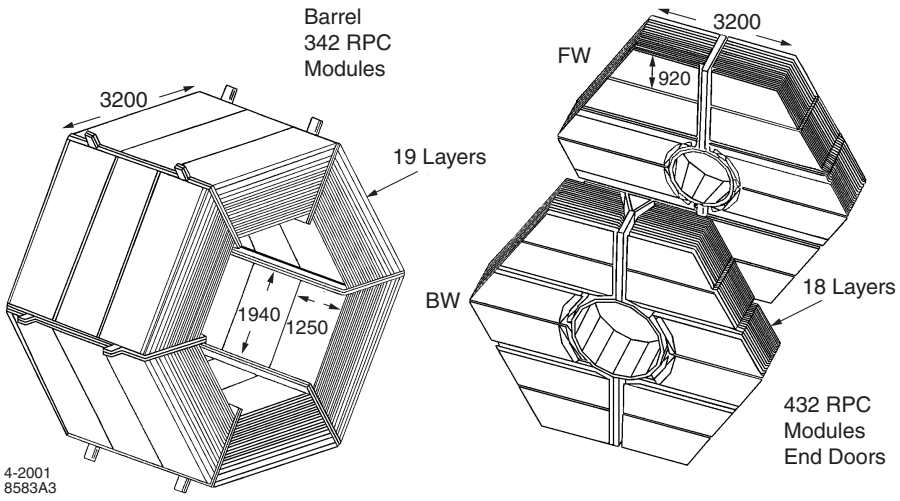


Figure 2-12: Overview of the IFR, Barrel sectors and forward (FW) and backward (BW) end doors.

in its large steel structure. The steel is segmented into 18 plates, and in between the gaps are a 19-layer resistive plate chambers (RPC) in the barrel region and two 18-layers RPC in the end caps, as shown in Figure 2-12.

A sideview of an RPC can be found in Figure 2-13. The planar RPCs consists of two bakelite sheets, 2 mm-think and separated by a 2 mm gap filled with a gas mixture of isobutane, argon, and freon. The gap is enclosed by polycarbonate spacer glued to the bakelite. RPCs detect streamers from ionizing particles. When an ionizing particle passes through the RPC, the resultant streamer depletes the local charge on the bakelite, and is detected by the readout strips.

The IFR was originally fully instrumented with (RPCs). However, the efficiencies of the RPCs were found to be quickly degrading. As a result, a decision was made to replace the barrel RPCs were replace with limited streamer tubes (LST) to improve the detection efficiencies. In the summer of 2004, the first two sextants (top and bottom) of the PRCs were replaced with LSTs. The remaining four sextants followed in the summer and fall of 2006. In this analysis, we use the *BABAR* data before summer 2006, therefore about 40% of the data were not affected by this change, and rest 60% of the data were taken with 2 sextants of LSTs.

The cross section of a LST detector is shown in Figure 2-14. Each unit consists

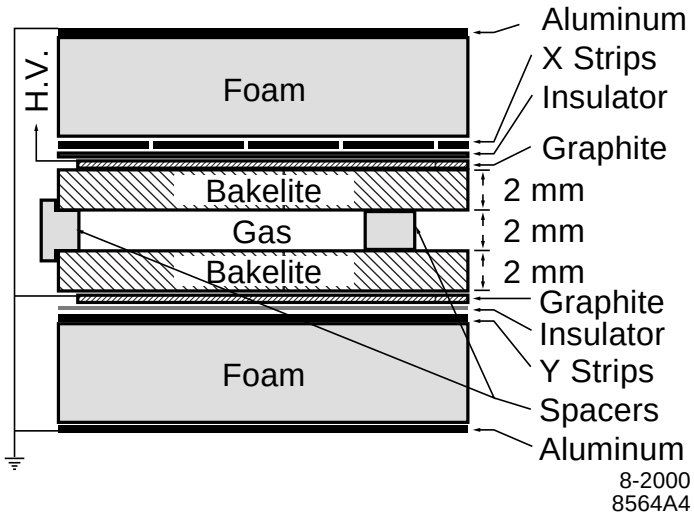


Figure 2-13: Cross section of a planar RPC with the schematics of the high voltage (HV) connection.

of 7 or 8 rectangular cells made of graphite-painted PVC (cathode). Each cell is 17 mm wide, 15 mm high and 380 mm long. In the middle of each cell is a gold-plated anode wire kept at a high voltage of 5500 V. The LSTs are filled with a (89:3:8) gas mixture of CO_2, Argon and iso-butane. This is non-flammable mixture with good quenching properties.

The physical principle of LSTs is quite simple. Each gas-filled cell has a single wire at High Voltage (HV). If a charged particle passed through the cell, the gas is ionized and a streamer builds up, which can be readout from the wire. Simultaneously a signal will be induced on a plane, which is mounted below the tube. The charge on the wire is used as the ϕ coordinate and the induced charge on the plane is detected using strips perpendicular to the wire direction, giving a z-coordinate. The r-coordinate is taken from the layer information. Together this gives a 3 dimensional information of the hit.

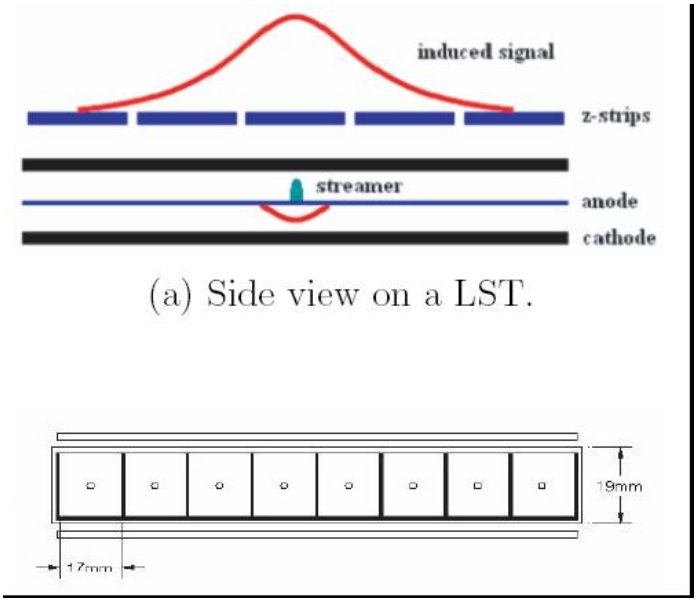


Figure 2-14: The *BABAR* Limited Streamer Tubes. They are mounted in the gaps between the iron plates of the barrel IFR. Plot (a) shows a side view of a single tube with a schematic working of the signal. Plot (b) shows a 8 cell layout, looking along the direction of the tube.

Chapter 3

Data and Monte Carlo Samples

3.1 *BABAR* Data Samples

This work uses the data collected at the *BABAR* detector between 1999 and summer 2006. This sample consists of 383×10^6 $B\overline{B}$ pairs (corresponding to $346\,\text{fb}^{-1}$) collected at the $\Upsilon(4S)$ resonance (c.f. Figure 2-1). An additional $36.3\,\text{fb}^{-1}$ off-resonance data, taken at the center-of-mass energy 40 MeV below the $\Upsilon(4S)$ resonance, is used to study the non-B decays ($e^+e^- \to f^+f^-$ where $f = u, d, c, s$) background. Table 3.1 lists the production cross-sections at $\sqrt{s} = 10.58\,\text{GeV}$ for $b\overline{b}$, other light quarks and leptons, as well as the expected number of events for each category.

Process	Cross-section (nb)	Expected number of events
$e^+e^- \to b\overline{b}$	1.05	363,300
$e^+e^- \to c\overline{c}$	1.30	449,800
$e^+e^- \to s\overline{s}$,$u\overline{u}$,$d\overline{d}$	2.09	723,140
$e^+e^- \to \tau^+\tau^-$	0.94	325,240
$e^+e^- \to \mu^+\mu^-$	1.16	401,360
$e^+e^- \to e^+e^-$	∼40	∼13,840,000

Table 3.1: Table of production cross-sections at $\sqrt{s} = 10.58\,\text{GeV}$ [20] and the expected number of events for a luminosity of $346\,\text{fb}^{-1}$.

3.2 Monte Carlo Simulations

Monte Carlo (MC) samples based on EvtGen [33] and GEANT4 [25] are used to simulate the signal and background processes as well as the detector response. We use them to determine the signal selection criteria, efficiencies and fitting techniques. We did not look at on-resonance data until all the above mentioned procedures are fixed. This is referred to as a "blind" analysis. A data control sample was constructed later to cross check the Monte Carlo (explained in Section 7.1).

Several categories of Monte Carlo are used in this analysis. Generic B Monte Carlo simulates the processes $e^+e^- \to \Upsilon(4S) \to B\overline{B}$, where the B-mesons are allowed to decay to everything. Continuum Monte Carlo simulates the charm and light quark decays $e^+e^- \to f^+f^-$ where $f = u, d, c, s$. We use the generic B Monte Carlo with signal modes removed and the continuum Monte Carlo to model the background in this analysis.

The $b \to s\gamma$ signal Monte Carlo sample is generated with a photon spectrum derived from reference [29]. This model breaks the photon spectrum (hence the hadronic system mass (M_{X_s}) in light of the underlying two-body decay [1]) into two parts: the K^* resonant region ($M_{X_s} < 1.1\,\mathrm{GeV}/c^2$), and the non-resonant region ($M_{Xs} > 1.1\,\mathrm{GeV}/c^2$). The $b \to s\gamma$ signal Monte Carlo is consequently divided into two parts. For the K^* resonant region, the decay process is simulated using the exclusive $B \to K^*\gamma$ Monte Carlo. For the latter, because the photon spectrum is not precisely known, it is generated with a flat photon energy spectrum for $E_\gamma^B = [1.6, 2.6]\,\mathrm{GeV}$. This is corresponding to a M_{X_s} between [0.65 - 3.31] MeV/c^2 assuming $M_B = 5.279\,\mathrm{GeV}/c^2$ in Equation 3.1, but we only use it for M_{X_s} greater than 1.1 MeV/c^2.

The flat spectrum is re-weighted it to produce the hadronic distribution as determined in Ref. [17]. To accomplish this, we take the bin height determined from

[1] The hadronic mass M_{X_s} and the photon energy E_γ^B in the B rest frame are associated by the kinematics:

$$E_\gamma^B = \frac{M_B^2 - M_{X_s}^2}{2M_B} \tag{3.1}$$

where M_B is the B-meson mass.

the fit to the spectrum, which corresponds to the partial branching fraction in the i^{th} energy bin ($PBF(E_i)$). The formula for reweighting an event whose true B-rest frame photon energy is E_i is then:

$$w_i = \left(\frac{PBF(E_i)}{\mathcal{B}} \right) \times f_i \tag{3.2}$$

where $\mathcal{B}$ is the total branching fraction for $B \to X_s\gamma$ and f_i is the fraction of events originally generated in bin i out of the entire flat spectrum. An event that falls into bin i in the energy spectrum is therefore counts as w_i of an event.

A final note: the re-weighted flat E_γ^B Monte Carlo is used only for the region above $M_{X_s} > 1.1\,\text{GeV}/c^2$. In this sample, the hadronic $s\overline{d}$ and $s\overline{u}$ systems are fragmented into multi-body final states using the default settings of JETSET [38]. We correct it using the *BABAR*-measured fragmentation [17] later on in Section 7.4.3.

Chapter 4

B Reconstruction and Event Pre-selection

4.1 Reconstruction Modes

We use a sum-of-exclusive modes approach in this analysis, where we reconstruct a total of 38 exclusive $b \to s\gamma$ decay modes (Table 4.1). In each event, the signal B is fully reconstructed through one of the 38 modes as sum of an energetic photon and a hadronic system (X_s) containing an s quark, while the other B is allowed to decay to everything. We allow 1 or 3 kaons and up to 4 pions in the hadronic system X_s. Many CP neutral modes, such as $B \to K^0_S\pi^0\gamma$, are also reconstructed but are not used for the CP asymmetry measurement because they do not carry flavor information. They were intended for calculating the systematic errors in the beginning. In Table 4.1, we identify each final state using an integer and will refer to final states by the decay mode indices henceforth.

The fully reconstruct B-mesons are self-tagging: the flavor of the exclusive final states is determined either by the charge of the reconstructed final state (B^-/B^+) or the charge of the kaon ($\overline{B}^0/B^0$) in the X_s system.

In the following of this chapter, we describe the particle reconstruction and identification, as well as event pre-selection which is performed at the level of particle reconstruction. Background suppression and the procedure of the event selection

Mode	Final State	Mode	Final State
1	$K^0_S\pi^+\gamma$	20	$K^0_S\pi^+\pi^-\pi^+\pi^-\gamma$
2	$K^+\pi^0\gamma$	21	$K^0_S\pi^+\pi^-\pi^0\gamma$
3	$K^+\pi^-\gamma$	22	$K^0_S\pi^+\pi^-\pi^0\pi^0\gamma$
4	$K^0_S\pi^0\gamma$	23	$K^+\eta\gamma$
5	$K^+\pi^+\pi^-\gamma$	24	$K^0_S\eta\gamma$
6	$K^0_S\pi^+\pi^0\gamma$	25	$K^0_S\eta\pi^+\gamma$
7	$K^+\pi^0\pi^0\gamma$	26	$K^+\eta\pi^0\gamma$
8	$K^0_S\pi^+\pi^-\gamma$	27	$K^+\eta\pi^-\gamma$
9	$K^+\pi^-\pi^0\gamma$	28	$K^0_S\eta\pi^0\gamma$
10	$K^0_S\pi^0\pi^0\gamma$	29	$K^+\eta\pi^+\pi^-\gamma$
11	$K^0_S\pi^+\pi^-\pi^+\gamma$	30	$K^0_S\eta\pi^+\pi^0\gamma$
12	$K^+\pi^+\pi^-\pi^0\gamma$	31	$K^0_S\eta\pi^+\pi^-\gamma$
13	$K^0_S\pi^+\pi^0\pi^0\gamma$	32	$K^+\eta\pi^-\pi^0\gamma$
14	$K^+\pi^+\pi^-\pi^-\gamma$	33	$K^+K^-K^+\gamma$
15	$K^0_S\pi^0\pi^+\pi^-\gamma$	34	$K^+K^-K^0_S\gamma$
16	$K^+\pi^-\pi^0\pi^0\gamma$	35	$K^+K^-K^0_S\pi^+\gamma$
17	$K^+\pi^+\pi^-\pi^+\pi^-\gamma$	36	$K^+K^-K^+\pi^0\gamma$
18	$K^0_S\pi^+\pi^-\pi^+\pi^0\gamma$	37	$K^+K^-K^+\pi^-\gamma$
19	$K^+\pi^+\pi^-\pi^0\pi^0\gamma$	38	$K^+K^-K^0_S\pi^0\gamma$

Table 4.1: Final states used in B reconstruction. The CP conjugate states are also reconstructed but omitted in this table.

optimization can be found in the next chapter.

4.2 Particle Selection Criteria

We require all events to contain one high energy photon and at least two well reconstructed hadron tracks in reconstructing the signal B.

The reason for the photon energy requirement is discussed in details in 4.2.2. The track quality requirements include having a polar angle within the detector acceptance, a minimum of 12 hits in the DCH and a minimum transverse energy of 100 MeV. These requirements are to make sure that the tracks leaves sufficient information in the detector to be reliably reconstructed. Further requirements include a distance of closest approach (DOCA) to the e^+e^- interaction point of less than 1.5 cm in the XY plane, a DOCA of less than 10 cm along the z axis. These requirements ensure that the tracks are not due to cosmic rays or charged particles originating from an interaction of one of the PEP-II beams with the residual gas in the beam pipe or the beam pipe itself.

4.2.1 Kaon and Pion Identification

All charged tracks are considered to be pions, unless they pass the $K^{\pm}$ identification. Hadron identification algorithms combine the information from the SVT, DCH, and DIRC into a likelihood function for each particle hypothesis:

$$\mathcal{L}_{ihad} = \mathcal{L}_{ihad}^{SVT} \times \mathcal{L}_{ihad}^{DCH} \times \mathcal{L}_{ihad}^{DIRC} \tag{4.1}$$

where *ihad* is the hadron species ($\pi^{\pm}$, $K^{\pm}$, or proton). The detector quantities considered in the likelihood are:

$\mathcal{L}_{ihad}^{SVT}$, the difference between the dE/dx measured in the SVT and the expected dE/dx under the appropriate hadron hypothesis.

$\mathcal{L}_{ihad}^{DCH}$, the difference between the dE/dx measured in the DCH and the expected

dE/dx under the appropriate hadron hypothesis.

$\mathcal{L}_{ihad}^{DIRC}$, the Cherenkov angle measured, the number of observed photons and the quality of the track prior to reaching the DIRC.

The Likelihood selectors consist of different requirement on the relevant likelihood ratios and we require that $(\mathcal{L}_K/(\mathcal{L}_\pi + \mathcal{L}_K) > 0.9)\&\&(\mathcal{L}_K/(\mathcal{L}_p + \mathcal{L}_K) > 0.20)$ for charged kaons. This criteria is chosen to obtain a good separation between $K^\pm$ and $\pi^\pm$ while maintaining good $K^\pm$ efficiencies.

The $K^\pm$ likelihood selector performance as a function of momentum is shown in Figure 4-1[1]. The kaon efficiency is evaluated from *BABAR* data using a sample of kaons from the decay $D \to K\pi$, where the D is selected from the decay of a D^* [21]. The $\pi^\pm \to K^\pm$ misidentification rate is evaluated using pions from the same source. For kaons with momenta below $\sim 1.0\,\mathrm{GeV}/c$, where $K\pi$ discrimination is provided by the dE/dx measured in the DCH, the efficiency is $80 - 95\%$ with pion mis-ID rates below 1%. Above $1\,\mathrm{GeV}/c$, $K\pi$ separation is primarily due to the Cherenkov angle measurement in the DIRC. The kaon efficiency in this range starts from 70%, rises to 90% at 2 GeV and then falls down to 70%, while the pion mis-ID increases with momentum to 8% above 4 GeV/c.

We do not use the kaons outside the DIRC acceptance because they rely only on dE/dx for particle identification, which itself might introduce K^+ and K^- cross-section related false CP violation (owing to different interaction cross-section between K^- and K^+ in the detector, as explained later). We therefore require that all charged kaons lie within the DIRC acceptance $(0.78 < \cos(p_K) < 0.90)$ and have momenta greater 0.8 GeV/c in the laboratory frame, as slow kaons may not be able to reach the DIRC.

To maintain a good efficiency, no particle identification requirements are placed on the charged pions.

[1] http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Selectors/r18b/TightLHKaonMicroSelection_details.html

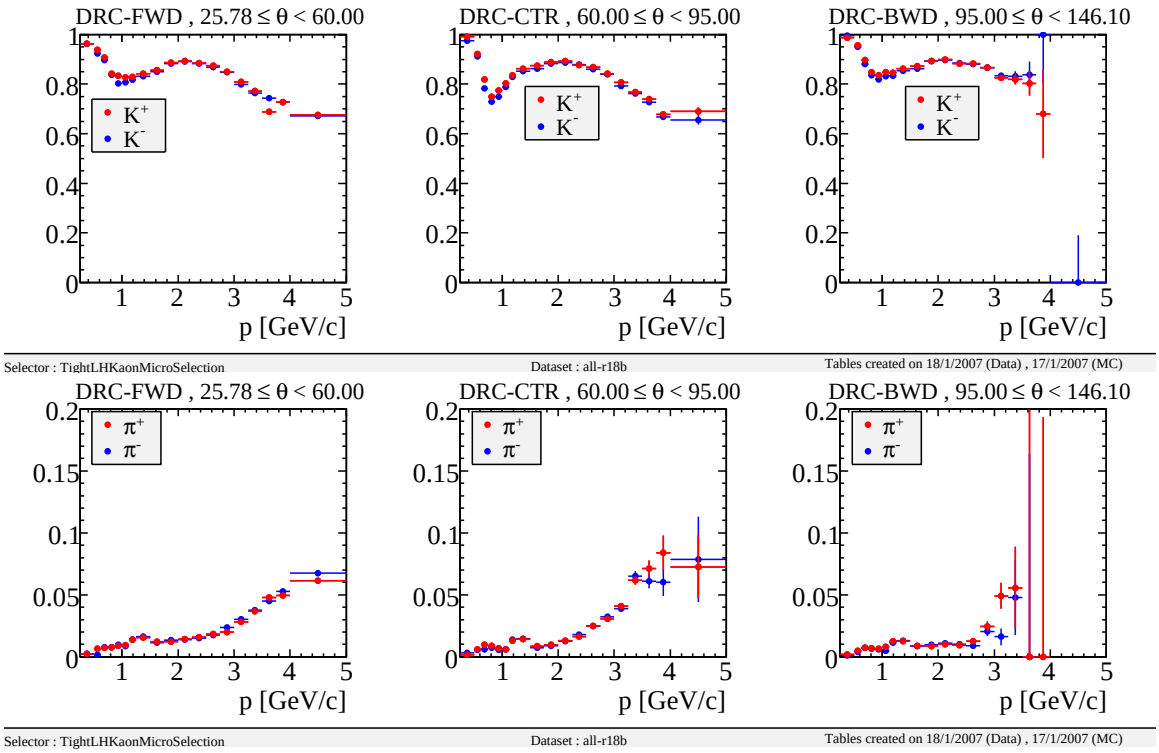


Figure 4-1: Kaon efficiencies (upper plots) and $\pi^{\pm} \rightarrow K^{\pm}$ mis-identification rates (lower plots) as a function of momentum in different detector polar angle (θ) bins.

4.2.2 Photon Selection

Photons in this analysis fall into two sub-categories: the primary photons that come directly from the signal B decay and all other photons produced in other processes, such as π^0 and η decays or initial state radiation (ISR). In this section, we first describe the selection for all the photons and, in the second part, we will discuss the additional requirements placed on the primary photons since they have rather unique signatures.

General photon selection

A photon in this analysis is defined as an isolated energy cluster which contains only a single bump with no charged-particle track associated with it. In addition, the following selection criteria apply:

- the energy deposited in the EMC be more than 30 MeV to exclude random noises from the electronics and cosmic ray background, etc;
- the lateral moment[2] of the shower be less than 0.8.

  The lateral moment is a measure of the spread of the cluster in the EMC. An electromagnetic shower typically has smaller lateral distribution than the hadronic shower. Most of the energy of an electromagnetic shower is deposited within only 2-4 crystals, while a hadronic shower spreads out over more crystals in the EMC.

[2]The lateral moment is defined as

$$L = \frac{\sum_{i-3}^{N} E_i r_i^2}{\sum_{i-2}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2} \tag{4.2}$$

where

- N is the number of crystals hit by the shower;
- E_i is the energy deposited in the $i-th$ crystal; ordered in decreasing energy starting with 1 being the highest energy;
- r_i is the lateral radius between the centroid of the shower and the $i-th$ crystal.
- $r_0 = 5\,\mathrm{cm}$, which is roughly the width of a crystal and thus the distance between two crystal centers.

Primary photon requirements

Photons coming directly from the B mesons are required to pass the above generic photon selection as well as the following additional selection:

- the photon energy in the center of mass frame (E^*_γ) is between [1.6, 3.0] GeV.

  Since the signal $B \rightarrow X_s\gamma$ events are two-body decay processes, the photon energy spectrum can be derived from kinematics:

  $$E^B_\gamma = \frac{M^2_B - M^2_{X_s}}{2M_B}, \tag{4.3}$$

  where E^B_γ is the photon energy, M_{X_s} the hadronic mass and M_B the B meson mass. M_{X_s} range is [0.6-2.8] GeV/c^2 in our analysis, thus leading to a high energy E^B_γ in [1.9-2.6] GeV in the B meson rest frame. We therefore place a cut of the CM frame photon energy at [1.6-3.0] GeV to allow for the difference in the photon energies in different reference frames.

- the polar angle is within the detector acceptance.

  The EMC has full coverage in the azimuthal direction but not in the ϕ direction. To exclude showers falling outside the EMC, we require the photon to be within $-0.74 < cos(\phi) < 0.93$ in the laboratory frame.

- the cluster status is good.

  Not all of the EMC crystals are working properly. This requirement is placed to ensure that the cluster does not contain a noisy or dead crystal.

- the cluster's Second Moment[3] be less than 0.002.

[3]The Second Moment is calculated as

$$L_2 = \sum_{crystal\ i} \frac{E_i[(\theta_i - \theta_C)^2 + (\phi_i - \phi_C)^2]}{\sum_i E_i} \tag{4.4}$$

where E_i is the energy deposited in the crystals. θ_C and ϕ_C are the polar and azimuthal angles of the centroid of the bump and θ_i and ϕ_i are the angles of the ith crystal of the bump.

The second moment a measure of how circular the shower in the calorimeter is. A cut on this rejects background from merged π^0's or η 's. If the two photons from a π^0 or η decay are merged and form a cluster this cluster is more likely to be oval shaped. Clusters from one single photon should be spherical which gives a small second moment.

4.2.3 Other Neutral Particles

π^0 selection

π^0 mesons decay to $\gamma\gamma$ with a mean lifetime $(8.4 \pm 0.6) \times 10^{-17}$, and a branching fraction 98.8% [37]. Therefore they are reconstructed by combining two photons whose invariant mass fall in a window [115-150] MeV/c^2 around the nominal π^0 mass (135.0 MeV/c^2). The photons are required to have a minimum energy of 50 MeV. At the initial reconstruction level the momentum of the π^0 is required to be at least 200 MeV/c^2. The π^0 nominal mass is consequently assigned.

η selection

We reconstruct η candidates by combining two photons, each with an energy above 50 MeV in the laboratory reference frame. The η candidates are required to have laboratory momenta greater than 200 MeV/c, and their invariant masses are required to be between 470 and 620 MeV/c^2.

K_S^0 selection

K_S^0 candidates are reconstructed by combining two oppositely charged pions with invariant mass in the mass window [0.489,0.507] GeV/c^2. $K_S^0 \to \pi^0\ \pi^0$ are not used in our analysis.

4.3 Event Pre-selection

At the B reconstruction level we perform a crude first-step event pre-selection, where we impose many loose cuts to reject some major background. This also helps us to reduce the Monte Carlo and real data to a manageable size.

4.3.1 Signal and Background Definition

The dominating background in this analysis is the continuum processes ($e^+e^- \to q\bar{q}$, with $q = u, d, s, c$) that produce a high-energy photon either by initial-state radiation or from the decay of π^0 and η mesons. In the next section we will focus on rejecting our this background.

$B\overline{B}$ background includes those B meson decays to a final state other than $X_s\gamma$. Combinations of final state particles can be mis-reconstructed to be $X_s\gamma$. This mostly happens when the reconstruction of the decays such as $B \to X_s\pi^0$ loses a photon from π^0 decay. We use the generic $B\overline{B}$ Monte Carlo with the signal modes removed to simulate this background. In addition, the $B \to X_s\gamma$ signal events can sometimes be mis-reconstructed in the wrong final state. This can occur, for instance, when we lose or add a π^0 in the reconstruction. This sample, denoted as "cross-feed", is simulated with the $b \to s\gamma$ signal Monte Carlo.

From a technical point of view, in order to separate the signal and cross-feed events in the Monte Carlo, we make use of the Monte Carlo truth-matching information[1]. To define signal, we use the following "pseudo truth-matching" criteria:

1. all the final-state particles from the B decay are matched to particles generated in the Monte Carlo event, and all those true particles come from the same B;

2. the difference between the true and reconstructed X_s mass $|\Delta M_{X_s}|$ is less than 0.050 GeV/c^2. Because the π^0 mass is around 135 MeV/c^2 [37], this requirement will effectively prevent losing or adding π^0 in the final state.

"cross-feed" is then defined as any signal Monte Carlo that fail this requirement. More discussion of the signal truth-matching can be found in Appendix A. Henceforth, the

signal Monte Carlo refers to those pseudo truth-matched signal events.

4.3.2 Pre-selection using Event Shape Variables

Our first step to reject the continuum background makes use of the event topology. Real B decay events are almost isotropic since the $\Upsilon(4S)$ resonance is only about 20 MeV/c^2 heavier than the sum of the two B meson masses and thus there is almost no excess of kinematic energy available which could go into the momenta of the B mesons. On the other hand, continuum events are much more jet-like than B-decay events since they result from hadronic fragmentation of high-momentum quarks back-to-back in the center-of-mass frame. These event shape differences provide a way to distinguish and reject the continuum background.

The second to the zeroth Fox-Wolfram moment R_2

The variable is computed in the center-of-mass frame as

$$R_2 = \frac{H_2}{H_0} \tag{4.5}$$

with H_l being the Fox-Wolfram moment [23]:

$$H_l = \sum_{i,j} \frac{|\overrightarrow{p_i}||\overrightarrow{p_j}|}{s} P_l(\cos\phi_{ij}) \tag{4.6}$$

where $|\overrightarrow{p_i}|$ and $|\overrightarrow{p_j}|$ are the momenta of two particles in the center-of-mass frame, ϕ_{ij} is the angle between these two momenta, P_l is the Legendre polynomials and s is the total center-of-mass energy squared. A lower R_2 is characteristic of the isotropic decay of signal events, as opposed to the continuum background which highly peaks at 1. R_2 is required to be less than 0.9 to reject the continuum background.

[1] An important property of Monte Carlo samples is the knowledge of the actual generated physics process in each event. Ideally, the true identity and the true four-momentum of each reconstructed particle is known. If it matches the generated one, it is said that the simulated particle is "truth-matched".

The second to the zeroth Legendre moment L_2/L_0

We also use the angular distribution of momenta in the event, encoded in the second to zeroth Legendre moment (L_2/L_0), to remove more jet-like light-quark fragmentation event. We compute L_2/L_0 by decomposing the momenta of the rest-of-event (ROE) particles along the B thrust axis. ROE is defined as all those particles not used in reconstructing the signal B candidate. The j_{th} order Legendre moment is computed as

$$L_j = \frac{\sum_{k\ in\ ROE} |p_k||\cos(\theta_k)|^j}{\sum_{k\ in\ ROE} |p_k|} \tag{4.7}$$

where θ_k is defined for the k^{th} particle in the ROE with respect to the signal B thrust axis. Continuum events tend to have most of the decay particles momenta aligned closely with the ROE thrust axis, leading to larger values of L_2/L_0. We require that L_2/L_0 less than 0.46 at the B reconstruction level.

The variable $|\cos(\theta_{\gamma,\vec{T}})|$

$\theta_{\gamma,\vec{T}}$ is the angle between the photon and the ROE thrust in the center-of-mass frame. High-energy photons in the continuum events tend to be collinear with the thrust axis formed from the ROE, leading to a large value of $|\cos(\theta_{\gamma,\vec{T}})|$. We reject such backgrounds by requiring this variable to be less than 0.85.

4.3.3 Pre-selection using Photon Background Rejection

We study the background that causes photons from decays other than B decays to enter our reconstruction and found that over 50% of such background comes from high-energy π^0 decay, where one of the photons from the decay has a significant energy and enters our B reconstruction. In addition, a significant fraction of these background events ($\sim 20\%$) comes from high-energy η decay. To reduce these backgrounds, we use two Boosted Decision Trees (BDT) to veto the photons from the π^0 (π^0 veto) and η(η veto) decays respectively.

We associate the high-energy photon candidate γ_{HE} with another photon candidate γ_2 in the event. For multiple γ_2 candidate in an event, we choose the $\gamma_{HE}\gamma_2$ pair

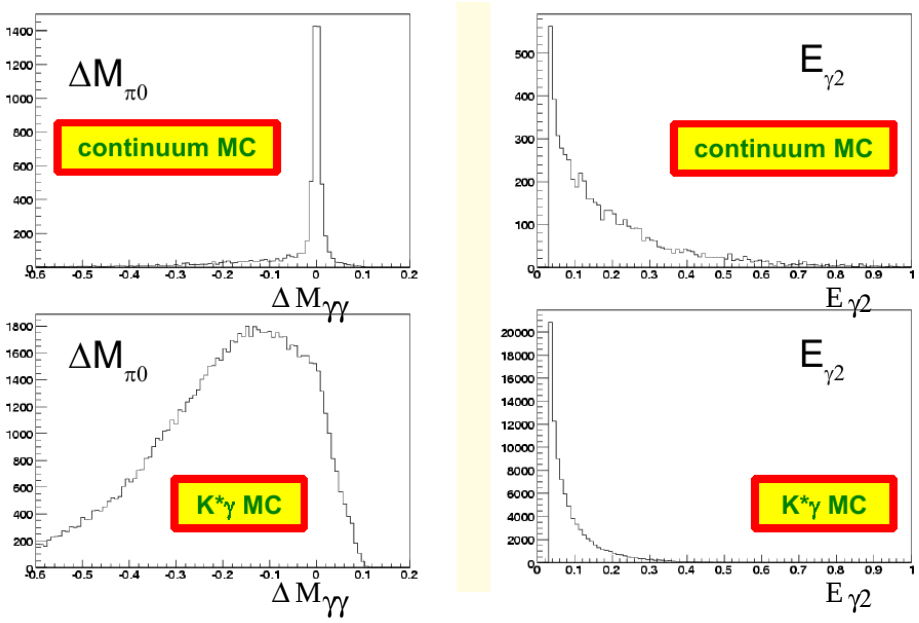


Figure 4-2: Input variables to the BDT that is used to reject photons from π^0 decay.

with invariant mass closest to the nominal π^0 mass (or η mass in case of η veto).

We use two inputs in each BDT: the difference between the two-photon invariant mass and the nominal π^0/η mass ($\Delta M_{\gamma\gamma}$) and energy of the second photon, E_{γ_2}. The distribution of these variables for photons from B decays and from π^0 decay are shown in Figure 4-2. The equivalent distributions for photons from the η decays are shown in Figure 4-3

We train the BDTs with AdaBoost algorithm (c.f. Appendix B) using a few thousand events from the off-resonance continuum Monte Carlo (background sample), and the $B \to K^*\gamma$ Monte Carlo (signal sample). The output of the BDTs, shown for on-resonance continuum Monte Carlo and $B \to K^*\gamma$, is shown in Figure 4-4 (π^0) and Figure 4-5 (η). The BDT output can be interpreted as a probability, so that events that lie below 0.5 are those with a 50% or less chance of being signal. We can see that the signal accumulates with high significance at 1.0, while the background is distributed more toward 0.0. For η's, the separation is not as strong (owing to the fact that the two-photon invariant mass distribution have a similar peaking structure in both signal and background), but a large portion of the background from this source

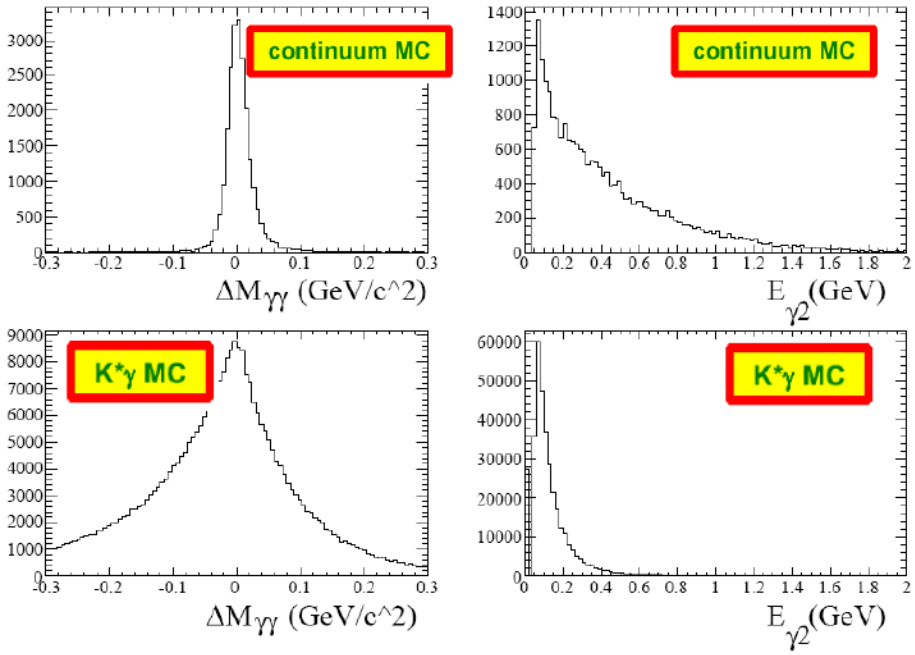


Figure 4-3: Input variables to the BDT that is used to reject photons from η decay.

can still be removed with a loose cut. We require that the output of both BDTs be greater than 0.5, removing the worse background events. We will later optimize the selection on these variables.

We find that the π^0 veto removes 35% of continuum events, while retaining 95% of signal events. We find that it also removes 50% of the $B\overline{B}$ background events. The veto removes 7% of the continuum events and 2% (8%) of B^+B^- ($B^0\overline{B}^0$) events, while retaining 99% of the signal. While the background rejection of the η veto is not quite as good as for the π^0 veto, one needs to remember that this background accounts for only about 20% of the photon background, and is expected to therefore constitute at least a 3 times smaller contribution to the overall continuum events than the π^0 background.

4.3.4 $m_{\rm ES}$ and ΔE after B Reconstruction

Fully reconstructed $b \to s\gamma$ decays are characterized by two kinematic variables:

- ΔE, defined as the difference between the reconstructed energy of the B meson

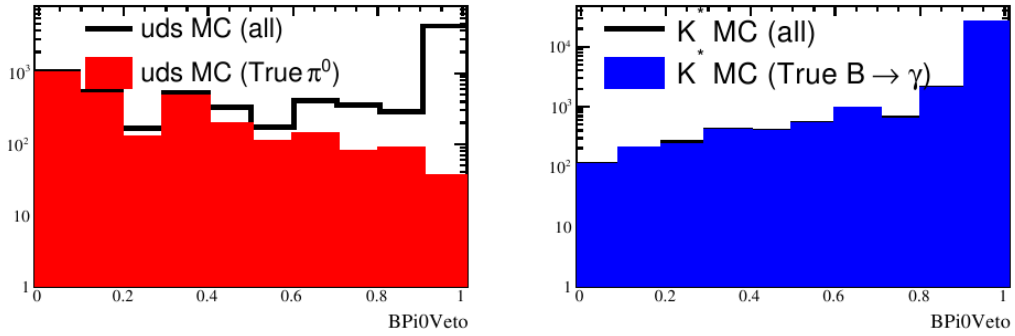


Figure 4-4: Output of the π^0-veto BDT. The left plot shows the distribution of the BDT for all continuum events (black), and just those where the γ_{HE} comes from a true π^0 decay. The right plot shows the output on signal Monte Carlo for all events (black) and just those where the photon is truth-matched to come from a B decay (blue).

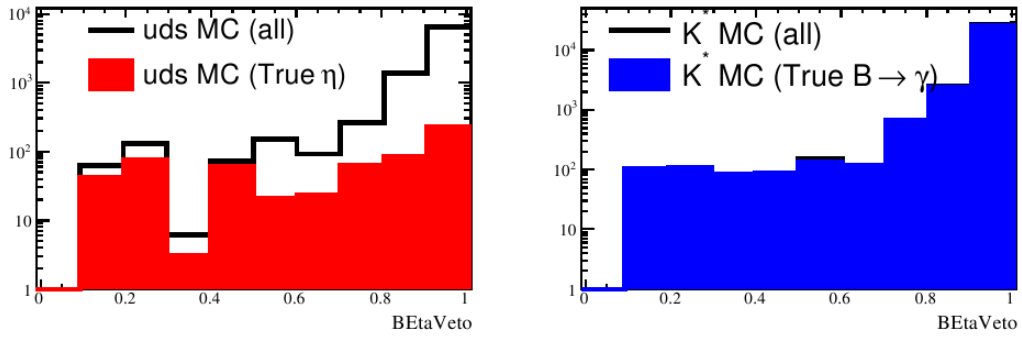


Figure 4-5: Output of the η-veto BDT. The left plot shows the distribution of the BDT for all continuum events (black), and just those where the γ_{HE} comes from a true η decay. The right plot shows the output on signal Monte Carlo for all events (black) and just those where the photon is truth-matched to come from a B decay (blue).

in the center-of-mass frame (E_B^*) and the expected energy of the B meson, which is simply $\sqrt{s}/2$, (s is the center-of-mass energy squared). Therefore we obtain $\Delta E = E_B^* - \sqrt{s}/2$.

- the beam-energy substituted mass $m_{\rm ES} = \sqrt{s/4 - {p_B^*}^2}$, where p_B^* is momentum of the B candidate in the center-of-mass frame. We choose $\sqrt{s}/2$ instead of E_B^* here because the beam energy is better known than the measured energy.

Signal events are expected to have a ΔE distribution centered near zero and a $m_{\rm ES}$ distribution centered at the mass of the B meson. At the B reconstruction level, we require that all the B candidate within $|\Delta E| < 0.4\,{\rm GeV}$ and $m_{\rm ES} > 5.21\,{\rm GeV}/c^2$. These cuts are tightened later on. Figure 4-6, shows the distributions of $m_{\rm ES}$[1] and ΔE just after the B reconstruction effort, for M_{X_s} between [0.6-2.8] GeV/c^2.

As the signal events concentrates in $m_{\rm ES} > 5.26\,{\rm GeV}/c^2$, we refer to this region as the "$m_{\rm ES}$ signal region", and $m_{\rm ES} < 5.26\,{\rm GeV}/c^2$ as the "$m_{\rm ES}$ sideband". (Sometimes, we use 5.27 GeV/c^2 as the boundary, this will be noted in the text.)

After reconstructing events using the aforementioned B reconstruction and event pre-selection criteria, we still have a significant contribution from background. In the $m_{\rm ES}$ region around the B mass, we see that a given $1\,{\rm MeV}/c^2$ bin can have $\sim 200,000$ background events, over 50% of which are continuum, compared to just 2500 signal events. (Multiple candidates from the same event are allowed to populate this plot.) In the next section, we will continue to reduce the background by combining B kinematic, event shape, and B flavor tagging variables in a multi-variable background rejection approach.

[1]Note: we use the endpoint-corrected $m_{\rm ES}$, defined as $5.29 + m_{\rm ES} - \sqrt{s}/2$, which corrects for the fact that off-resonance data is taken at 40 MeV below the $\Upsilon(4S)$ resonance.

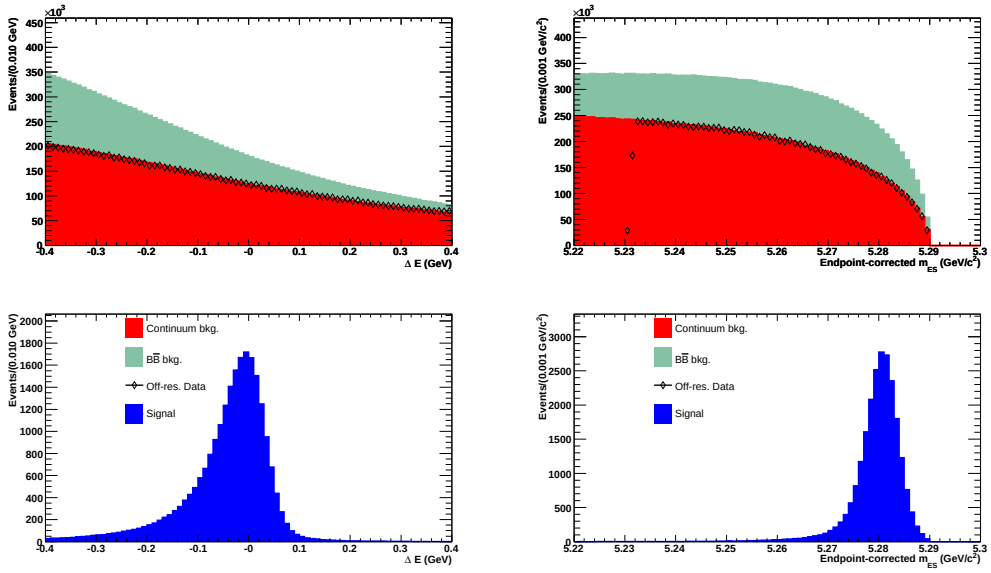


Figure 4-6: ΔE (left) and m_{ES} (right) after the B reconstruction described in this section. The off-resonance data falls off below $5.234\,\mathrm{GeV}/c^2$ in the m_{ES} plot. This is because we impose the requirement that $m_{\mathrm{ES}} > 5.21\,\mathrm{GeV}/c^2$ during the B reconstruction, but then we correct the m_{ES} for the endpoint. This leaves a gap with an irregular boundary below $5.234\,\mathrm{GeV}/c^2$. A requirement that $m_{\mathrm{ES}} > 5.234\,\mathrm{GeV}/c^2$ is made when showing ΔE.

Chapter 5

Event Selection and Optimization

After the aforementioned B reconstruction and pre-selection criteria, we still have a significant contribution from background. In this chapter, a multi-variable discriminating method to reject the remaining background is developed.

We study a suite of variables which are useful in rejecting this remaining background and combine them variables using a Boosted Decision Tree into a single-output discriminant trained to reject the continuum. Optimization of the selection criteria is consequently performed. In the next chapter, we will pursue a maximum likelihood fitting strategy to extract the signal events from the sample with all selection criteria applied.

5.1 Continuum Background Rejection with a Boosted Decision Tree

5.1.1 Discriminating Variables

These input variables to the Boosted Decision Tree are divided into three categories.

Event shape variables

The following event shape variables are incorporated in the background rejecting algorithm.

1. The second to the zeroth Legendre moments, L_2/L_0, was used in the pre-selection (Sec 4.3.2) and loosely constrained. Its shape still has good discriminating power between signal and the Monte Carlo (Figure 5-1[1]).

2. The absolute value of the cosine of the angle between the γ and the thrust of the ROE, $|cos(\theta_{\gamma,\vec{T}})|$, computed in the center-of-mass frame. This variable was also used in the pre-selection and loosely constrained (Sec 4.3.2). It still has shape discrimination between continuum and signal (Figure 5-2). The edges of the distribution have been sculpted by the requirement that was made on L_2/L_0 in the pre-selection.

3. The second to the zeroth Fox-Wolfram moment, R_2 (Sec 4.3.2). When we study this variable after the pre-selection level (Figure 5-3), we do not find a strong difference in the shape between signal and continuum. This is likely due to the fact that we have already removed the most jet-like events using pre-selection constraints on the relationship with the thrust angle and the Legendre moment.

4. The second to the zeroth Fox-Wolfram moment, R_2 ($\gamma-$recoil), computed in the frame recoiling against the high-energy photon instead of the center-of-mass frame. In events where ISR is the source of the photon, this will represent R_2 in the true e^+ e^- collision frame. In this frame, one expects the ISR events to be jet-like. There is also a weak separation between signal and continuum (Figure 5-4).

5. ROE Legendre moments (Sec 4.3.2), calculated along the photon axis. We compute the zeroth, first-to-zeroth, and second-to-zeroth order Legendre monomials of the ROE along the photon axis in the center-of-mass frame. There is better separation between signal and continuum in these variables (Figure 5-5 - Figure 5-7).

[1]For all figures in this section, a requirement is made that the $m_{\rm ES} > 5.234\,{\rm GeV}/c^2$ to remove the artificial event deficit in the off-resonance data below that point.

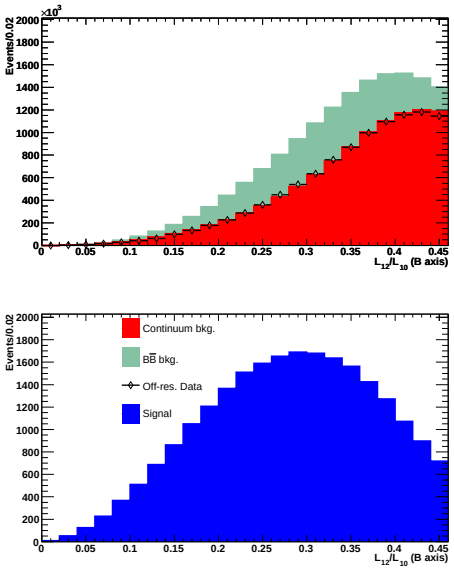


Figure 5-1: The second-to-zeroth Legendre moment, computed for the ROE around the B thrust direction in the center-of-mass frame.

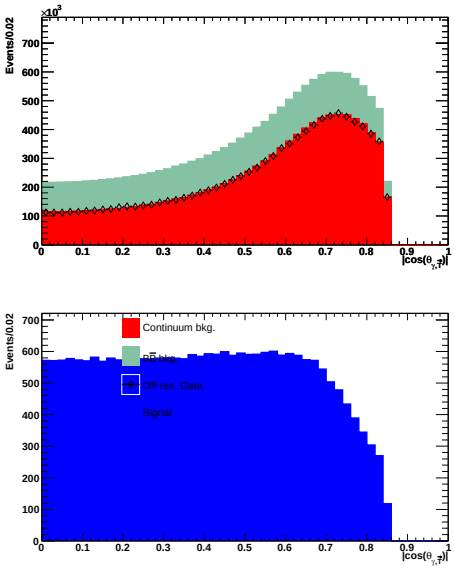


Figure 5-2: The absolute value of the cosine of the angle between the high-energy photon and the thrust of the ROE in the center-of-mass frame.

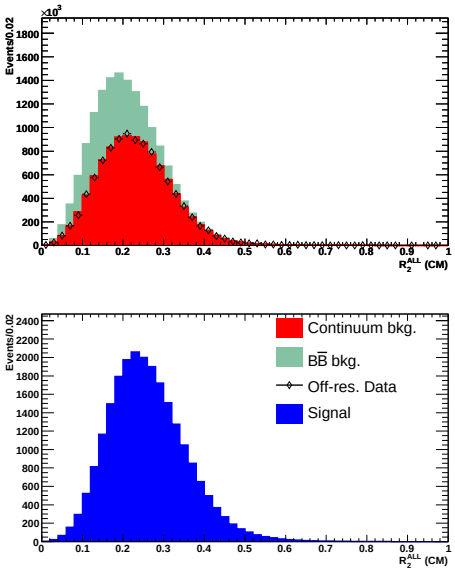


Figure 5-3: R_2, computed in the center-of-mass frame, for all particles in the event.

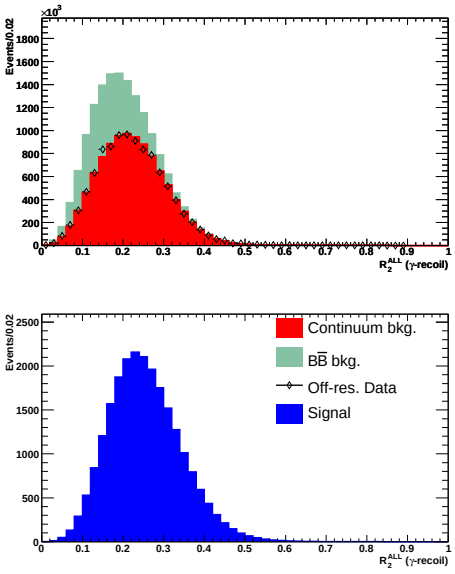


Figure 5-4: R_2, computed in the photon-recoil frame, for all particles in the event.

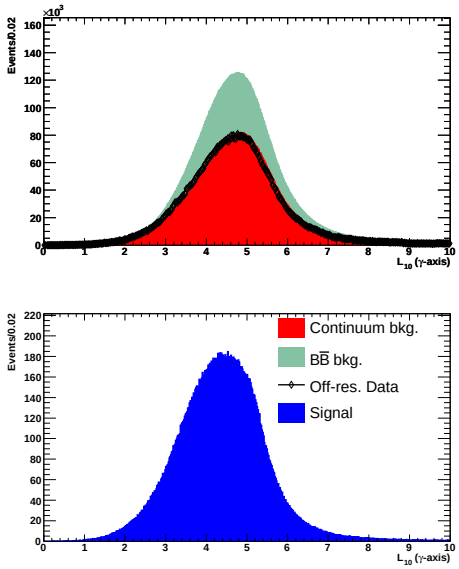


Figure 5-5: The zeroth Legendre moment, computed for the ROE along the photon axis in the center-of-mass frame.

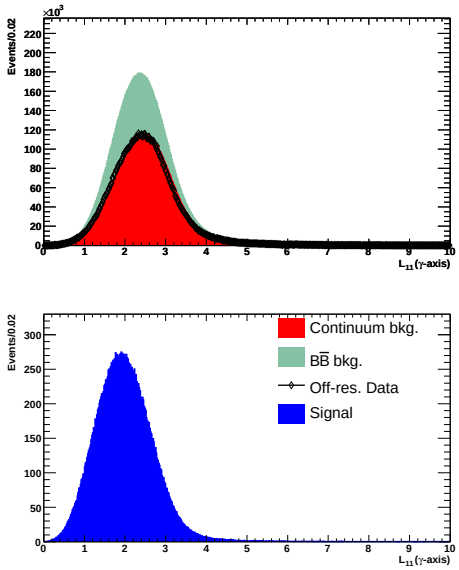


Figure 5-6: The first Legendre moment, computed for the ROE along the photon axis in the center-of-mass frame.

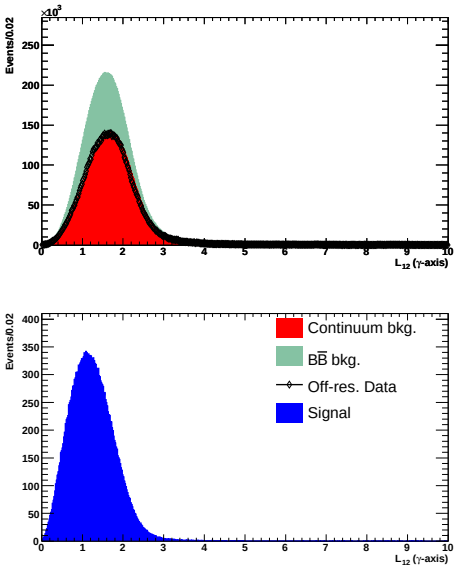


Figure 5-7: The second Legendre moment, computed for the ROE along the photon axis in the center-of-mass frame

B kinematics variables

In addition to the information about the event topology, we use the kinematics of the B itself to discriminate against the continuum background. The following variables are included in the BDT:

1. $|\cos(\theta_B)|$, the cosine of the B flight direction, computed in the center-of-mass frame with respect to the z axis. This has the expected distribution for one of a pair of pseudo-scalars (the B) having been produced from a vector meson, while it is effectively flat for background (Figure 5-8).

2. B Flight Significance (unsigned), computed as $\Delta z/\sigma(\Delta z)$ where Δz is the separation between the two B decay vertices in the z-direction and $\sigma(\Delta z)$ is the uncertainty on it. Due to the longer lifetime, B mesons tends to fly further, and with more significance, than the fake Bs that are made from continuum events (Figure 5-9).

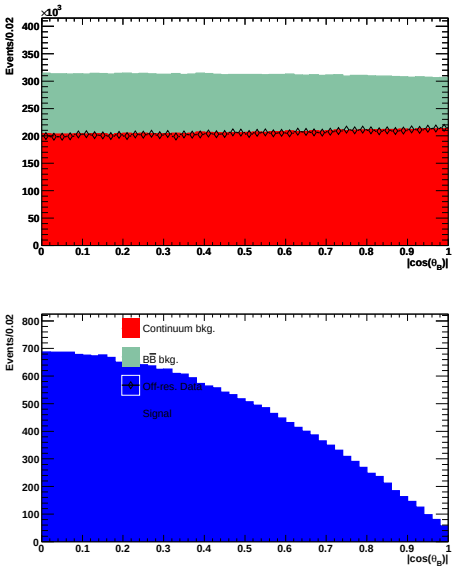


Figure 5-8: Cosine of the B flight direction, computed in the center-of-mass frame with respect to the z axis

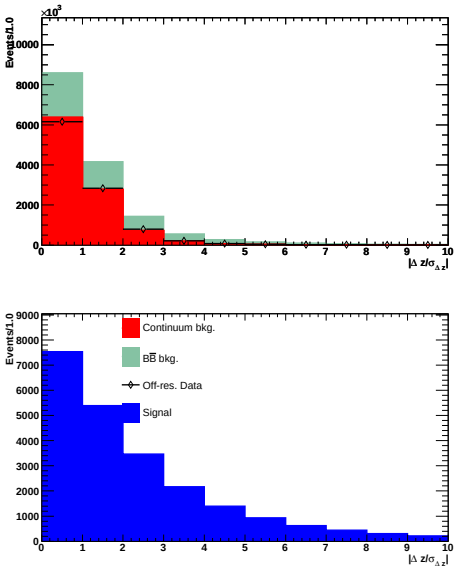


Figure 5-9: B flight significance along the z axis

B flavor-tagging variables

B flavor-tagging variables refer to the ROE variables that are used to identify the flavor of the B meson. An example is the decay $B^0 \to D^{*+} l^- \nu_l$. The charge of the lepton unambiguously identifies the decay as from a B^0 but not $\overline{B}^0$. The subsequent decays of $D^{*+} \to D^0 \pi^+$ and $D^0 \to K^- X^+$ gives a slow pion and a kaon in the final state whose charges also uniquely identify the flavor of the B. Another example is the decay $B^0 \to D^{*-}(\to \overline{D}^0 \pi^-)\pi^+$, which gives an energetic π^+ and a low-momentum π^- in the final state, and the flavor of the B can be inferred from the charge of both. Once the ROE B meson flavor is identified through the flavor tagging techniques, the signal B flavor can be uniquely determined as well.

While we do not require flavor-tagging information for the purpose of obtaining the B flavor, we can use it to discriminate between continuum events and signal: when an event contains a high-momentum lepton in the ROE, it is very likely to be from a b or c semi-leptonic decay rather than the lighter-quark continuum processes. While not every event contains such a lepton, by combining information from many sources (leptons, high-momentum tracks, kaons, low-momentum pions) we gain rejection again a wider class of backgrounds.

BABAR developed a series of neural network algorithms (sub-taggers) [16] for determining the flavor of the B mesons in the event. Useful input information includes the properties (such as charge, momentum and decay angles) of charged leptons, kaons and pions. The sub-tagger output values range between -1 and 1, which can be mapped to the probability of being a $\overline{b}$ (-1) or a b (1). Each sub-tagger is described in details below. The output of these taggers is shown in Figure 5-10 - Figure 5-14 [1].

1. Lepton sub-taggers, which include
   (a) Electron sub-tagger
   (b) Muon sub-tagger

[1] We take the absolute values of the output of the sub-taggers, because it will reduce or remove any effects that might cause the continuum BDT to cut asymmetrically on the signed output, inducing a false *CP* violation.

(c) Kinematic lepton sub-tagger

Electrons and muons produced in semileptonic B decays provide excellent tagging information. The charge of an electron (or muon) from a $b \to cl^-\bar{\nu}_l$ transition is directly associated to the flavor of the B^0 meson. Leptons occurring via the transition $b \to W^-c(\to sl^+\nu)$, carry tagging information as well: their charge is opposite to the leptons from b decay and they are characterized by a much lower momentum.

While the first two sub-taggers are require lepton presence in the event, the kinematic lepton sub-tagger is designed to recover leptons from B meson decays but did not pass either electrons or muons selection criteria. It is primarily designed to distinguish leptons from B decays from other charged particles, but it also has power to separate b from $\overline{b}$.

2. Kaon sub-tagger

   The dominant source of charged kaons is from a $b \to c \to s$ transition where the charge of the kaon tags the flavor of B. The high probability of producing a charged kaon in the decay of a B meson make the Kaon tagger a powerful source of tagging information.

3. Slow pion sub-tagger

   Low momentum pions (slow pions, with momentum less than 190 MeV/c) from $D^{*\pm} \to D^0\pi^\pm$ decays provide another source of flavor tagging information. The charge of the D^* (therefore B meson) can be identified from the charge of the pions.

4. Kaon - slow pion correlation sub-tagger

   This sub-tagger aims to identify $D^{*\pm} \to D^0(\to K^-X^+)\pi^\pm$ decays through the angular correlation of a charged kaon and a slow pion. The $D^{*\pm}$ flavor information is provided by a pair of a kaon and a slow pion of opposite charge and emitted in approximately the same direction in the center-of-mass frame. It is trained to separate $K^\mp\ \pi^\pm$ pairs from $D^{*\pm}$ from random $K^\mp\ \pi^\pm$ combinations.

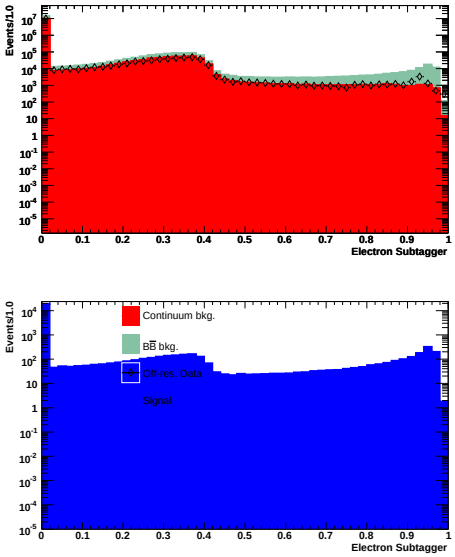


Figure 5-10: Output of the electron sub-tagger. The bin at 0 is where events with no electron in the ROE will accumulate.

5. Maximum p^* sub-tagger

   The flavor tagging information can also be obtained from the track with the highest momentum $p*$ in the $\Upsilon(4S)$ rest frame. The purpose of this sub-tagger to identify fast particles coming from the hadronization of the W boson produced in the $b \to cW^-$ decay or recover high momentum leptons missed by the leptonic tags. Direct charged hadrons or leptons are lighter than the charm decay products, so they are usually energetic in the B (hence $\Upsilon(4S)$) rest frame, and fly in the opposite direction of the charm products. We can identify these particles by associating their energy with their flight direction,

5.1.2 Boosted Decision Tree Performance

We combine the 16 variables described in the previous section using another BDT. This one is also trained using the AdaBoost method, by finding the training point which stabilizes the figure-of-merit (the Gini index, c.f. Appendix B) over an ensemble of decision trees within the BDT. We train, test, and validate the algorithm using a small fraction of the total signal and continuum Monte Carlo.

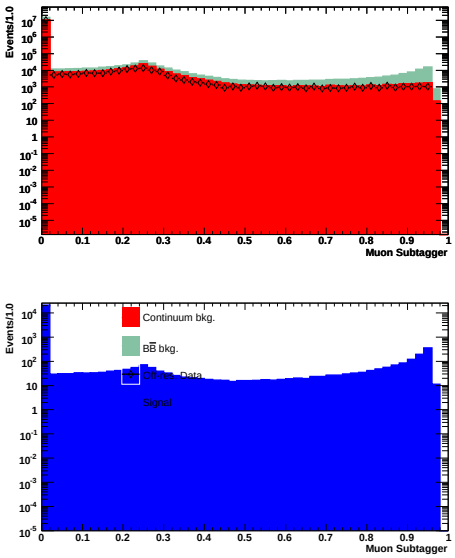


Figure 5-11: Output of the muon sub-tagger. The bin at 0 is where events with no muon in the ROE will accumulate.

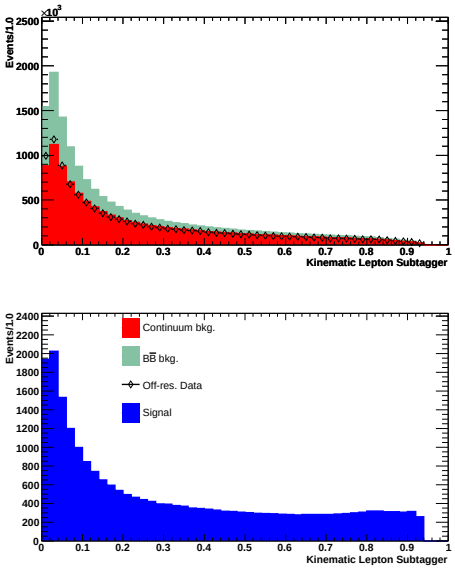


Figure 5-12: Output of the kinematic lepton sub-tagger. The bin at 0 is where events with no kinematic lepton in the ROE will accumulate.

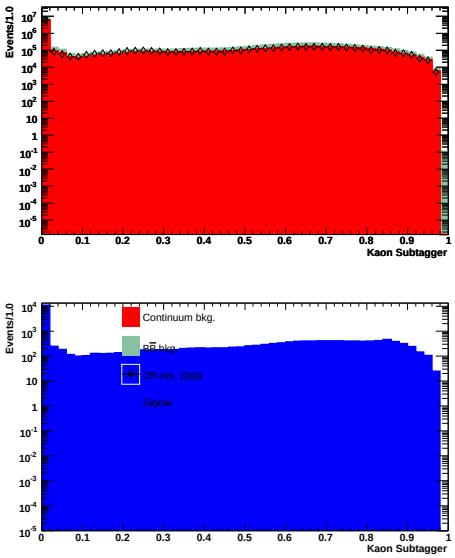


Figure 5-13: Output of the kaon sub-tagger. The bin at 0 is where events with no kaon in the ROE will accumulate.

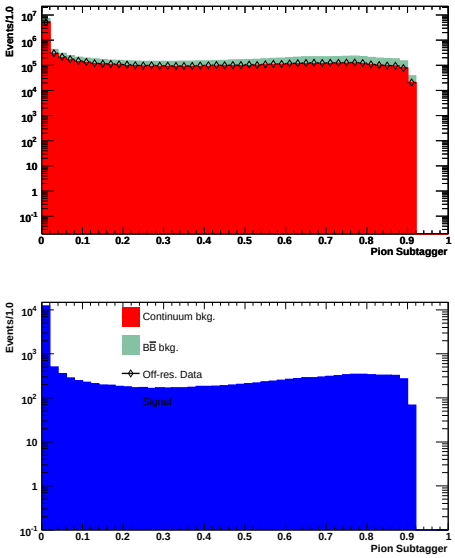


Figure 5-14: Output of the slow pion sub-tagger. The bin at 0 is where events with no slow pion in the ROE will accumulate.

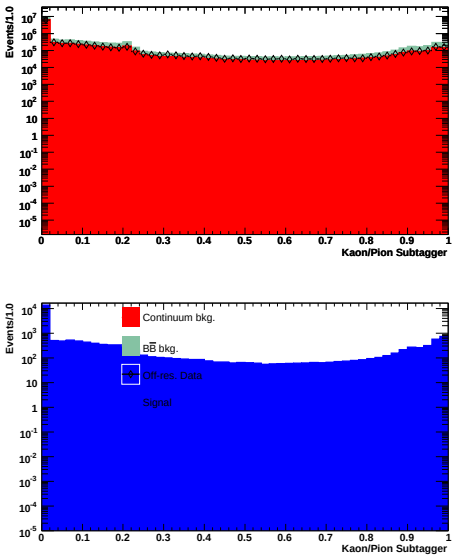


Figure 5-15: Output of the kaon-slow pion sub-tagger. The bin at 0 is where events with no kaon-slow pion pair in the ROE will accumulate.

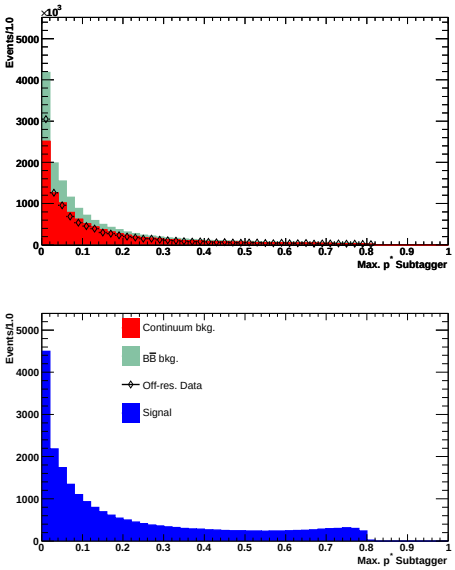


Figure 5-16: Output of the maximum p^* sub-tagger, which uses the highest-momentum track in the ROE.

We refer to this BDT as the $BDT_{continuum}$. The BDT output is shown in Figure 5-17 for signal, continuum Monte Carlo, and off-resonance data. In general, the signal does peak near 1.0, but it also has a peak near zero. Looking at this more closely, we see that events in the K^* mass region ($M_{X_s} < 1.1\,\mathrm{GeV}/c^2$) are the ones that clearly peak at 1.0 and fall off to a flat distribution toward zero. Signal events from above the K^* mass region are harder to separate from background, but also peak at 1.0 (though less strongly).

We loosely cut on the continuum BDT output at 0.2 at first, and will tighten it in the selection optimization stage.

5.2 Best Candidate Selection

After pre-selection described in the previous chapter and the loose cut of continuum BDT at 0.2, we still have many events containing more than one B candidate. In order to choose a best candidate per event, we use the ΔE variable, (c.f. Section 4.3.4). Signal events are expected to have a smaller $|\Delta E|$ than the background events. In Figure 4-6, we see the signal shows a peaking structure around zero while the background shows the falling slope shape from -0.4 to 0.4. Therefore in events with multiple B candidates, we decided to select the one with the smallest $|\Delta E|$.

5.3 Selection Optimization

With the best candidate selected in each event, we optimize the selection on the following variables: the π^0-veto BDT output, the η-veto BDT output, and the continuum BDT output. We perform the optimization in the most signal-like region of m_{ES} and ΔE ($m_{\mathrm{ES}} > 5.26\,\mathrm{GeV}/c^2$ and $|\Delta E| < 0.15\,\mathrm{GeV}$) and choose to optimize the statistical signal significance: $S/\sqrt{S+B}$, where S is the number of signal events and B is the number of background events passing the selection criteria. The optimal selection criteria over all 38 final states are as follows:

1. $BDT_{\pi^0} > 0.84$

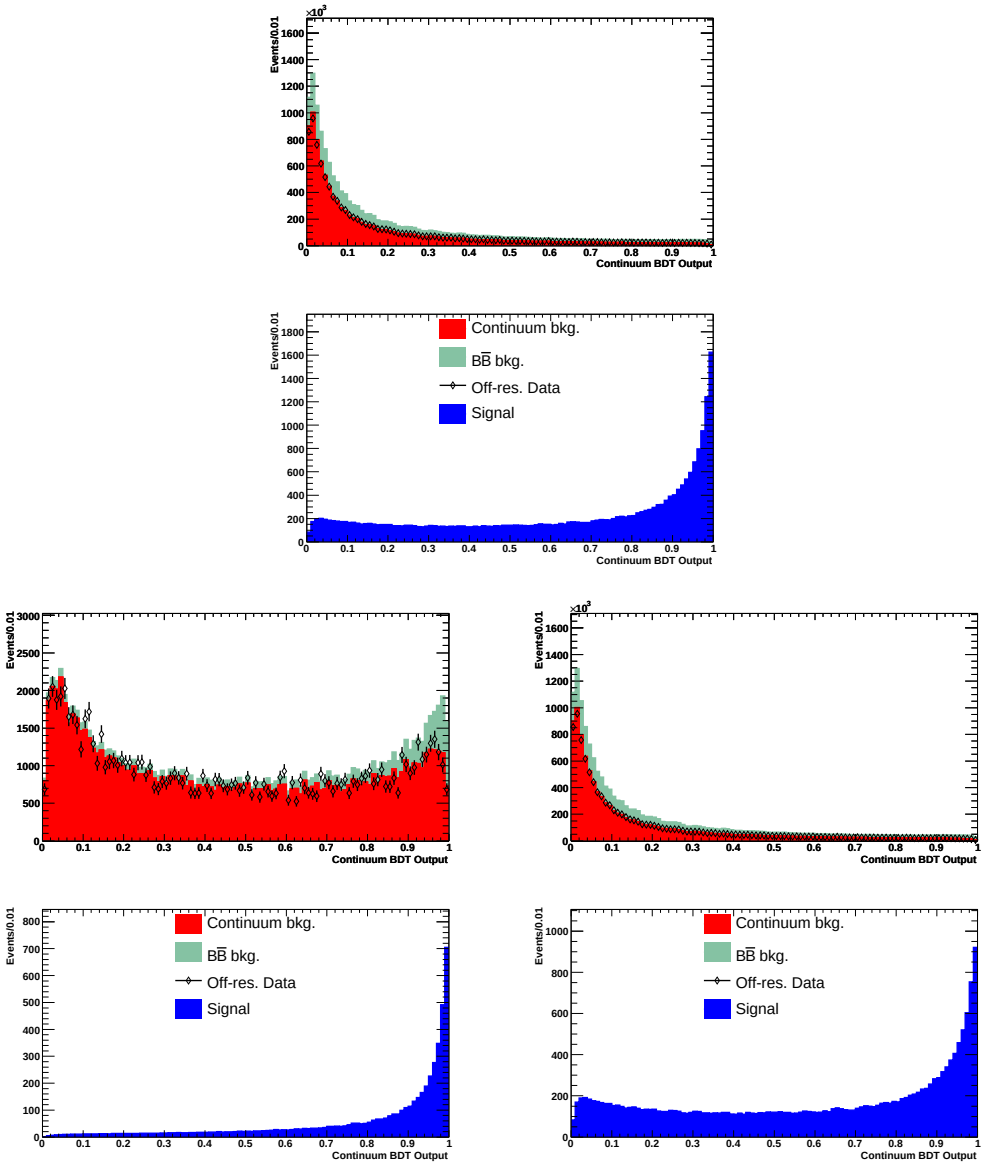


Figure 5-17: Output of the continuum BDT. Events are required to satisfy $m_{\rm ES} > 5.234\,{\rm GeV}/c^2$ and $M_{X_s} = [0.6, 2.8]\,{\rm GeV}/c^2$. The bottom two plots show the events which lie in the K^* mass region, below $1.1\,{\rm GeV}/c^2$ (left) and the region above the K^* (right).

2. $BDT_{\eta} > 0.75$

3. $BDT_{continuum} > 0.68$

In addition, we tighten the requirement on ΔE to be $|\Delta E| < 0.1\,\mathrm{GeV}$ to exclude more background.

5.4 X_s Final State Selection

Having chosen the best candidate and the optimal selection criteria, we can now study which of the 38 modes are worth retaining for the remainder of this analysis. By default, we choose the 12 modes used in the *BABAR*'s previous $b \to s\gamma$ *CP* asymmetry study using the "sum-of-exclusive" approach [15]. These are the first 12 *CP* specific modes in Table 4.1. In addition we add modes which, due to lower statistics or higher backgrounds were inaccessible in the previous analysis.

We show the signal significance for each reconstructed mode and in Table 5.1. All of the modes used in the *BABAR*'s previous analysis have a significance in the current data set of at least or about 1, except for mode 13 (0.325) $B^+ \to K^0_S\pi^+\pi^0\pi^0\gamma$. We added modes 23 ($B^+ \to K^+\eta\gamma$), 27 ($B^0 \to K^+\eta\pi^-$), 33 ($B^+ \to K^+K^-K^+\gamma$), and 37 ($B^0 \to K^+K^-K^+\pi^-\gamma$). These all have significance of about 1 or better. Modes 21, 24, and 34 also have a significance of about 1 or better, but they are flavor-blind.

Having selected a set of promising final states, we revisited the cut optimization choices. In particular, we optimize the cuts only on the above final states. Again, the optimization is performed in the most signal-like region ($m_{\mathrm{ES}} > 5.26\,\mathrm{GeV}/c^2$ and $|\Delta E| < 0.15\,\mathrm{GeV}$) and we find that the optimization prefers

1. $BDT_{\pi^0} > 0.77$

2. $BDT_{\eta} > 0.73$

3. $BDT_{continuum} > 0.66$

These are consistent with the original optimization, and suggest that once the above final states are chosen the remainder have little influence on the optimal selection

modes	S	B	$S/\sqrt{S+B}$
1	725.1	307.7	22.56
2	622.8	522.2	18.41
3	2403	693.7	43.18
4	169.7	142.5	9.603
5	729.3	1085	17.12
6	127.6	501.1	5.088
7	37.55	356	1.893
8	224.2	310.9	9.692
9	398.4	1166	10.07
10	10.72	90.26	1.067
11	38.4	313	2.048
12	122	1094	3.499
13	6.554	269.9	0.3941
14	98.09	545.4	3.867
15	40.69	308.5	2.178
16	18.86	584.8	0.7678
17	11.17	223.2	0.7295
18	9.864	292.6	0.5672
19	7.327	520.6	0.3189
10	4.521	87.6	0.471
21	20.89	465.8	0.9468
22	3.345	203.3	0.2327
23	39.92	226.2	2.447
24	11.14	51.21	1.411
25	6.545	184.5	0.4736
26	4.786	240.9	0.3053
27	20.24	409.4	0.9762
28	1.138	67.33	0.1375
29	3.635	289.3	0.2124
30	1.095	204.3	0.07643
31	1.478	88.91	0.1555
32	1.91	383.2	0.09734
33	14.41	5.565	3.224
34	2.602	0.5199	1.473
35	0.7919	4.041	0.3602
36	0.6614	3.98	0.307
37	2.006	2.096	0.9904
38	0.08711	3.195	0.04808

Table 5.1: The signal, background, and significance in each X_s final state for each mode (c.f. Table 4.1). All selection criteria are applied and the signal significance is calculated in the $m_{\rm ES}$ signal region ($m_{\rm ES} > 5.26\,{\rm GeV}/c^2$).

Table 5.2: Signal cut flow table

Selection	Expected Yield	Relative Efficiency
Pre-selection $BDT_{continuum} > 0.2$ + Best B	10192	-
modes selection	9104	0.93
$0.6 < M_{X_s} < 2.8$	8988	0.98
$BDT_{\pi^0} > 0.77$	8646	0.96
$BDT_{\eta} > 0.73$	8527	0.97
$BDT_{continuum} > 0.66$	6139	0.80
$\|\Delta E\| < 0.1$	5645	0.91
$m_{\rm ES} > 5.22$	5645	1.00
$m_{\rm ES} > 5.26$	5624	1.00

criteria. As before, we additionally require that $|\Delta E| < 0.1\,\text{GeV}$. We apply these cuts as the final set of optimized cuts.

5.5 Event Selection Efficiencies

Having determined the optimal event selection criteria, we now study the results of making these selections on the signal and background Monte Carlo. The cut flow tables below (Table 5.2 - 5.4) illustrate the expected yields and relative efficiency of each of our selection criteria. We see that most signal efficiencies are better than 95% and most continuum Monte Carlo efficiencies are only around 50%. The continuum Monte Carlo has a better efficiencies The BDT_{π} and BDT_{η} selection efficiencies for the continuum Monte Carlo are high (89% and 96% respectively). This is because we have already precut on them in the pre-selection state.

We also show the $m_{\rm ES}$ distribution after all selection criteria in Figure 5-18. The signal to background ratio is found to be better than 1:1 now near the B mass region, which is a tremendous improvement from that in Figure 4-6.

Table 5.3: Continuum Monte Carlo cut flow table

Selection	Expected Yield	Relative Efficiency
Pre-selection	1.126×10^6	-
$BDT_{continuum} > 0.2$ + Best B		
modes selection	5.776×10^5	0.51
$0.6 < M_{X_s} < 2.8$	3.076×10^5	0.53
$BDT_{\pi^0} > 0.77$	2.726×10^5	0.89
$BDT_{\eta} > 0.73$	2.608×10^5	0.96
$BDT_{continuum} > 0.66$	7.327×10^4	0.28
$\lvert\Delta E\rvert < 0.1$	3.733×10^4	0.51
$m_{\rm ES} > 5.22$	3.102×10^4	0.83
$m_{\rm ES} > 5.26$	5046	0.16

Table 5.4: Generic $B\overline{B}$ background cut flow table

Selection	Expected Yield	Relative Efficiency
Pre-selection	2.733×10^5	-
$BDT_{continuum} > 0.2$ + Best B		
modes selection	8.232×10^4	0.49
$0.6 < M_{X_s} < 2.8$	4.87×10^4	0.87
$BDT_{\pi^0} > 0.77$	4.144×10^4	0.94
$BDT_{\eta} > 0.73$	4.012×10^4	0.98
$BDT_{continuum} > 0.66$	1.65×10^4	0.50
$\lvert\Delta E\rvert < 0.1$	1.129×10^4	0.72
$m_{\rm ES} > 5.22$	1.035×10^4	0.94
$m_{\rm ES} > 5.26$	3875	0.42

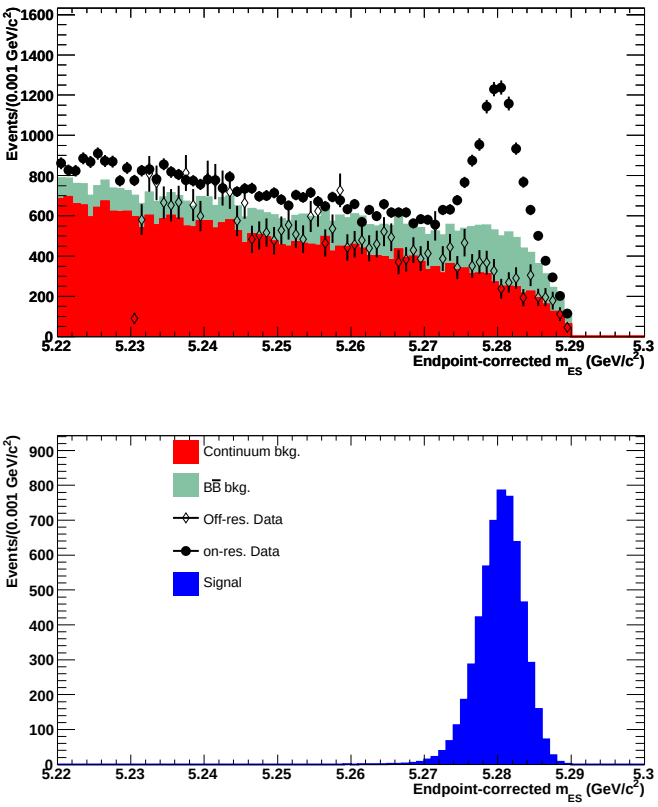


Figure 5-18: $m_{\rm ES}$ after all selection criteria are applied. All events with $0.6 < M_{X_s} < 2.8\,{\rm GeV}/c^2$ are included.

Chapter 6

Fitting of Monte Carlo and Data

We have finalized the selection criteria. To determine the amount of signal and background in the data, we will perform a maximum likelihood (ML) fit of the data. We fit, in one dimension only, the $m_{\rm ES}$ distribution between [5.22 5.29] GeV/c^2. Our strategy is to fit the data in five regions of M_{X_s}: [0.6-1.1], [1.1-1.5], [1.5-2.0], [2.0-2.8] GeV/c^2, as well as the entire region [0.6-2.8] GeV/c^2 to check whether the *CP* asymmetry has significant mass dependence.

We construct probability density functions (PDFs) for the signal, continuum, and $B\overline{B}$ background in the five regions. We use the b-quark flavor to define two flavor categories, and in each M_{X_s} region, we perform a simultaneous fit for the flavor fraction $N_{\overline{b}}/(N_b + N_{\overline{b}})$, where N_b and $N_{\overline{b}}$ are the yields for b and $\overline{b}$ respectively. We can then determine *CP* asymmetry as a function of the M_{X_s} spectrum. In the following sections, we describe how we determine the nominal PDFs used in the ML function.

6.1 The Probability Density Functions

The continuum background from $c\overline{c}$ and light quarks is combined and modeled using an ARGUS function [4], which is a threshold function capable of describing a combinatoric distribution with a kinematic endpoint. Signal events, which have a Gaussian-like peak and a low-side tail, are modeled using a Cruijff function. These

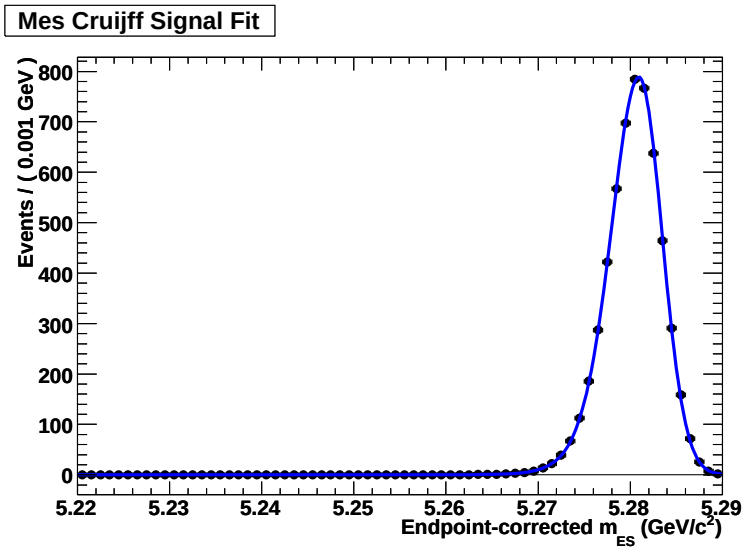


Figure 6-1: $m_{\rm ES}$ plot for the signal Monte Carlo for the entire M_{X_s} region.

PDFs are described in Appendix C.1. The main challenge in fitting the data is to describe the shapes of the generic $B\overline{B}$ background events, which tend to have lower statistics than continuum and a broad peak in $m_{\rm ES} > 5.27\,{\rm GeV}/c^2$ region. Given the difficulty of modeling this distribution with a sum of analytic functions, we choose to model the distribution using a binned PDF.

6.1.1 Signal Model

The signal Cruijff function is determined from the signal Monte Carlo. We first fitted the signal simultaneously in categories of both flavors (b and $\overline{b}$) and M_{X_s} regions. We allow the parameters of the Cruijff function to vary with flavors and M_{X_s} region (c.f. Appendix C.3). This first test shows that the PDFs in different categories are all consistent with each other, so we decide to describe the shape of all the signal events with a single Cruijff function (Figure 6-1). We find the fit quality to be unchanged after making this choice. The Cruijff function parameters are shown in Tables 6.1.

	Cruijff Function Parameters
Mean: m_0	$5.2810 \pm 1 \times 10^{-4}$
Left Width: σ_L	$3.001 \times 10^{-3} \pm 8.6 \times 10^{-5}$
Right Width: σ_R	$2.498 \times 10^{-3} \pm 6.2 \times 10^{-5}$
Left Tail: α_L	$7.863 \times 10^{-2} \pm 7.11 \times 10^{-3}$
Right Tail: α_R	0

Table 6.1: The parameters of the Cruijff function used to model the $m_{\rm ES}$ distribution for signal events. The right tail parameters (α_R) is purposely fixed at zero (c.f. Sec. C.3) while other parameter are floated in the fit. But they are not allowed to vary between the flavors.

6.1.2 $B\overline{B}$ Background Model

The $B\overline{B}$ background model is also obtained from a fit to the Monte Carlo. We did not distinguish between the generic $B\overline{B}$ Monte Carlo and the cross-feed. However, we find that at higher statistics, the $B\overline{B}$ background shape could not be very well modeled by a single or a sum of a few analytic functions. We choose to describe the shape with a binned PDF, using 1 MeV/c^2 bins. We divide the $B\overline{B}$ sample into four sub-samples according to the reconstructed M_{X_s}, and fit simultaneously for both flavors in each of the four M_{X_s} regions. The shapes are allowed to vary between b and $\overline{b}$ events in each M_{X_s} region, in order to eliminate the possibility of a false *CP* asymmetry. We also performed a simultaneous fit to the entire M_{X_s} range for the b quark flavors. Figure 6-2 shows the $m_{\rm ES}$ distribution in the four M_{X_s} regions separately while Figure 6-3 shows the $m_{\rm ES}$ distribution in the entire M_{X_s} range.

6.1.3 Continuum Background Model

For this background, we can either determine the shape from the continuum Monte Carlo or from the off-resonance data. The Monte Carlo has the benefit of providing us with enough statistics to determine its shape accurately, however, being the computer simulated events themselves, they might not provide a good description of the continuum background shape in data. Off-resonance data, on the other hand, is a true representation of the continuum background shape. The disadvantage is that the accuracy of the fit is limited by the small statistics. There are only about 3000

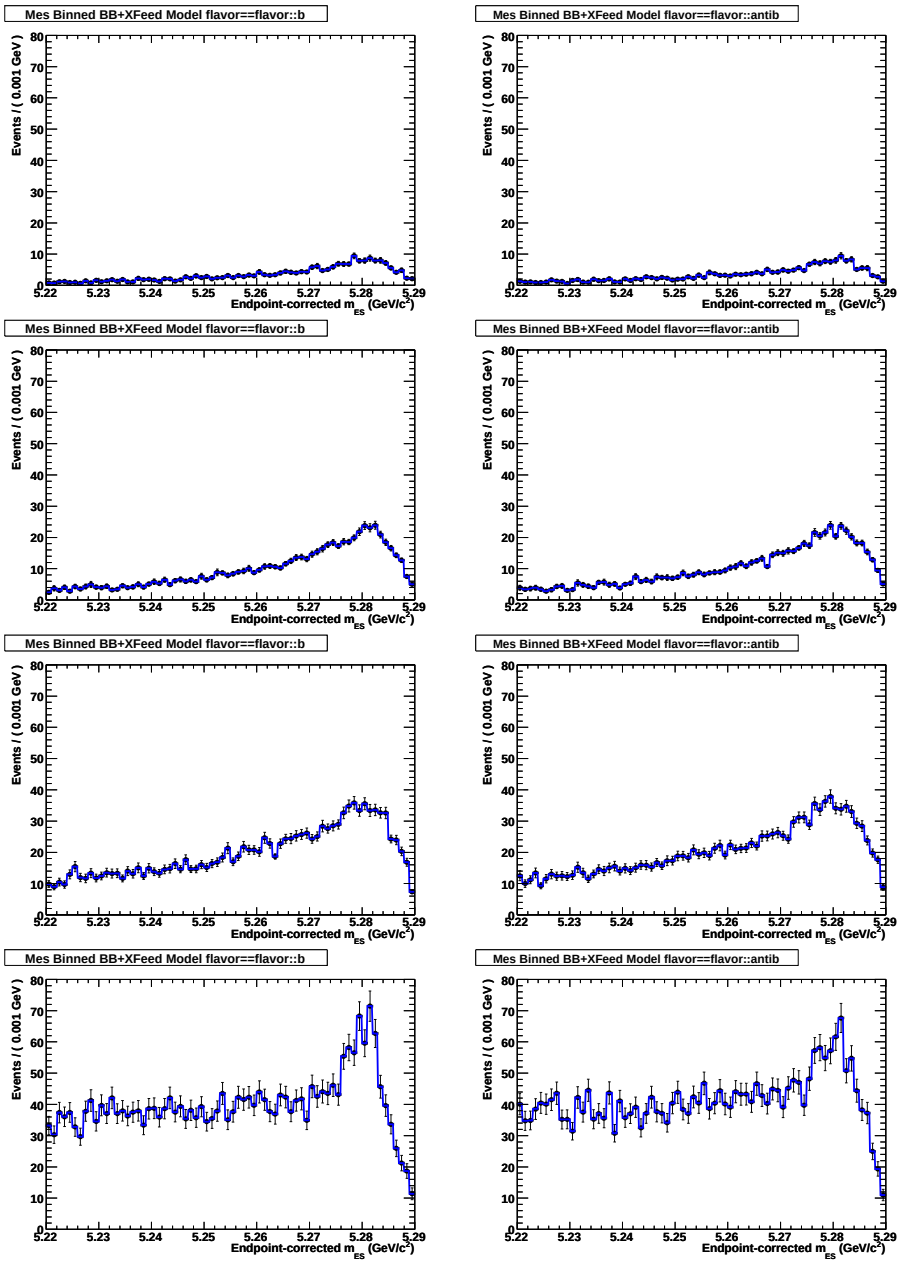


Figure 6-2: $m_{\rm ES}$ pdf shapes for the $B\overline{B}$ background events for both flavors (b on the left, $\overline{b}$ on the right) in the four M_{X_s} (GeV/c^2) regions: $0.6 < M_{X_s} < 1.1$ (first row), $1.1 < M_{X_s} < 1.5$ (second row), $1.5 < M_{X_s} < 2.0$ (third row), $2.0 < M_{X_s} < 2.8$ (fourth row).

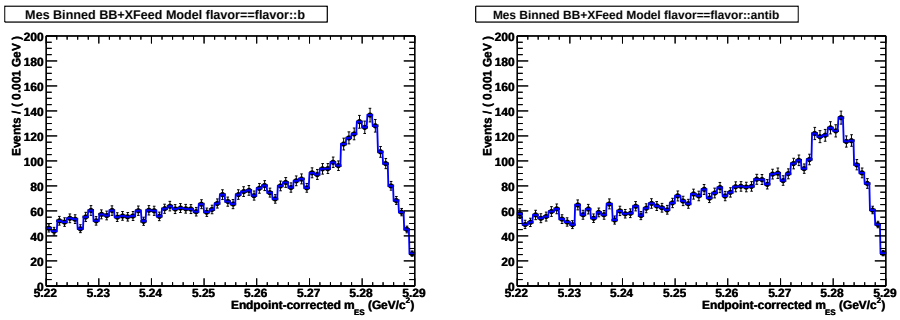


Figure 6-3: $m_{\rm ES}$ pdf shapes for the $B\overline{B}$ background for the entire M_{X_s} (GeV/c^2) region: $0.6 < M_{X_s} < 2.8$ (b on the left, $\overline{b}$ on the right).

total events in the off-resonance data sample which pass all selections. The largest concentration of these events is in the highest X_s mass region (about 50%). The other regions suffer from even lower statistics.

To determine which sample to use in describing the PDFs for the continuum background, we first compare the continuum Monte Carlo shape and the off-resonance data shape. We again divide both samples into different M_{X_s} regions and fit simultaneously between flavors (b and $\overline{b}$) in each region. We keep both the power and the slope of the ARGUS function floated in the fits, but do not allow them to vary between the flavors in each M_{X_s} region. Since we have already corrected the end point of the $m_{\rm ES}$ (c.f. Sec. 4.3.4) distribution for each events, we fixed the threshold of the ARGUS function to be 5.29 GeV/c^2. The results of the continuum Monte Carlo fits are shown in Figure 6-4. Off-resonance data is fitted in the restricted region $m_{\rm ES} = [5.24 - 5.29]\,{\rm GeV}/c^2$[1], and the results are shown in Figure 6-5. The variation of the ARGUS parameter with M_{X_s} category in Table 6.2.

In Figure 6-6 we directly compare the shapes determined from continuum Monte Carlo and off-peak data. We find that in most X_s mass regions the Monte Carlo shape is different from the data shape, with the data preferring more events in the signal region and fewer in the $m_{\rm ES}$ tail than the Monte Carlo.

Given these shape differences between the Monte Carlo and off-resonance data

[1]Looking back at Figure 4-6, we see that off-resonance data falls off below 5.234 GeV/c^2 in the $m_{\rm ES}$ plot. We only use events above 5.24 GeV/c^2 in the fit to make sure the $m_{\rm ES}$ distribution is not affected by the artificial cut-off.

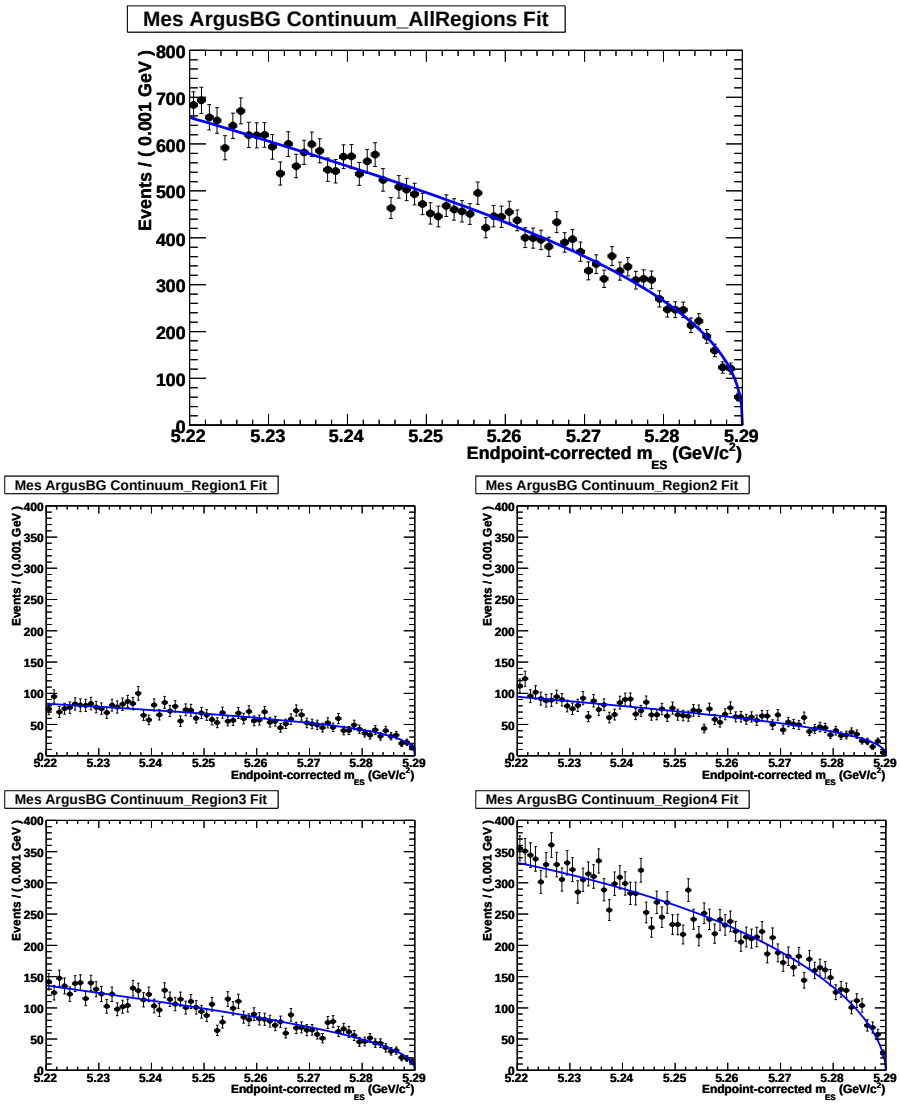


Figure 6-4: $m_{\rm ES}$ shapes for the continuum Monte Carlo for entire M_{X_s} region (top plot) and for the four M_{X_s} (GeV/c^2) sub-regions: $0.6 < M_{X_s} < 1.1$ (middle left), $1.1 < M_{X_s} < 1.5$ (middle right), $1.5 < M_{X_s} < 2.0$ (lower left), $2.0 < M_{X_s} < 2.8$ (lower right).

X_s Region	power (MC)	power (DATA)	slope (MC)	slope (DATA)
1	0.340 ± 0.049	0.500 ± 0.223	2.38 ± 5.13	6.61 ± 29.13
2	0.410 ± 0.054	0.400 ± 0.202	5.13 ± 5.30	16.71 ± 27.38
3	0.432 ± 0.047	0.500 ± 0.192	8.07 ± 4.61	6.20 ± 25.18
4	0.447 ± 0.030	0.501 ± 0.116	2.92 ± 2.89	-11.32 ± 15.72
All	0.417 ± 0.020	0.500 ± 0.084	4.77 ± 2.02	-2.70 ± 11.18

Table 6.2: Parameters of the Continuum ARGUS model, as determined from continuum Monte Carlo and off-resonance data.

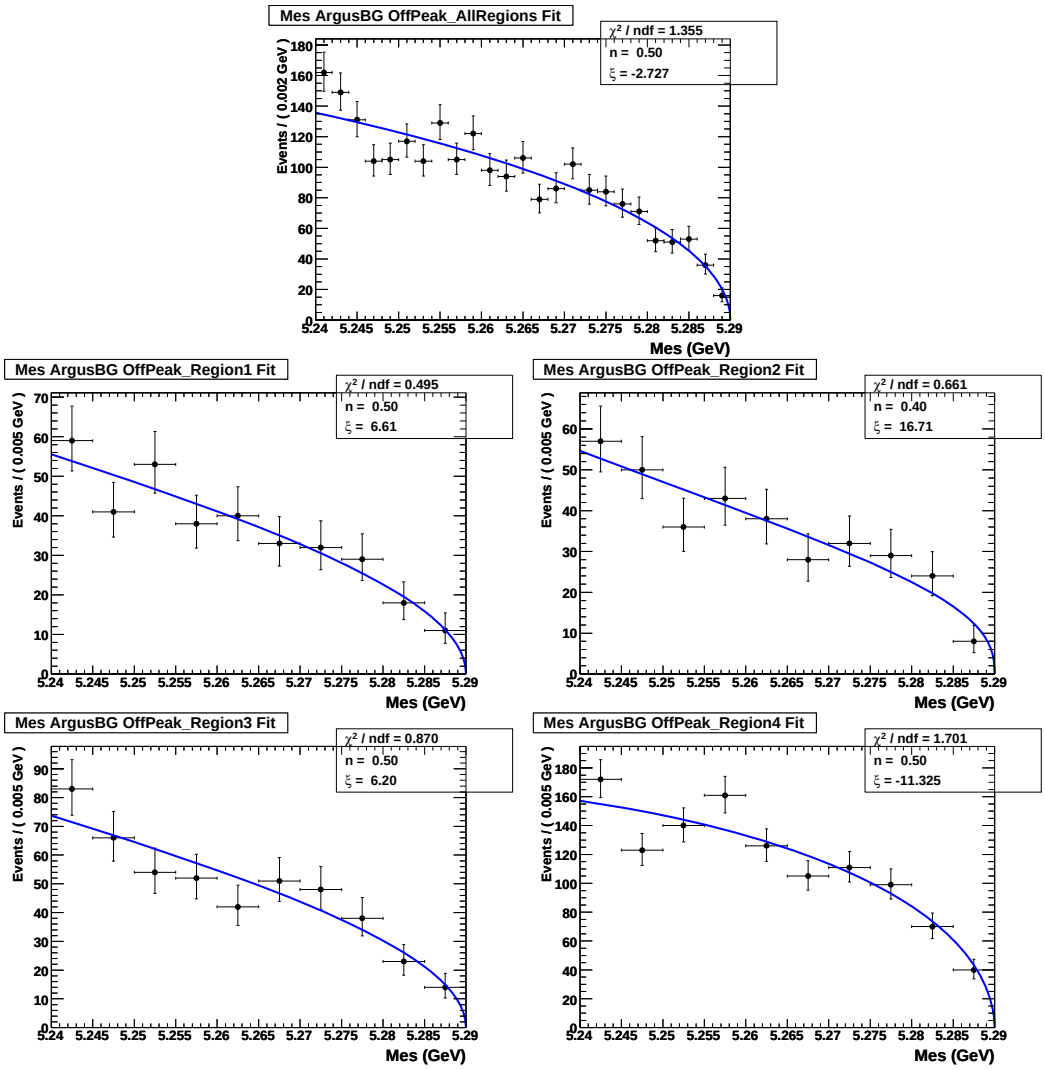


Figure 6-5: $m_{\rm ES}$ shapes for the off-resonance data for entire M_{X_s} region (top plot) and for the four M_{X_s} (GeV/c^2) sub-regions: $0.6 < M_{X_s} < 1.1$ (middle left), $1.1 < M_{X_s} < 1.5$ (middle right), $1.5 < M_{X_s} < 2.0$ (lower left), $2.0 < M_{X_s} < 2.8$ (lower right).

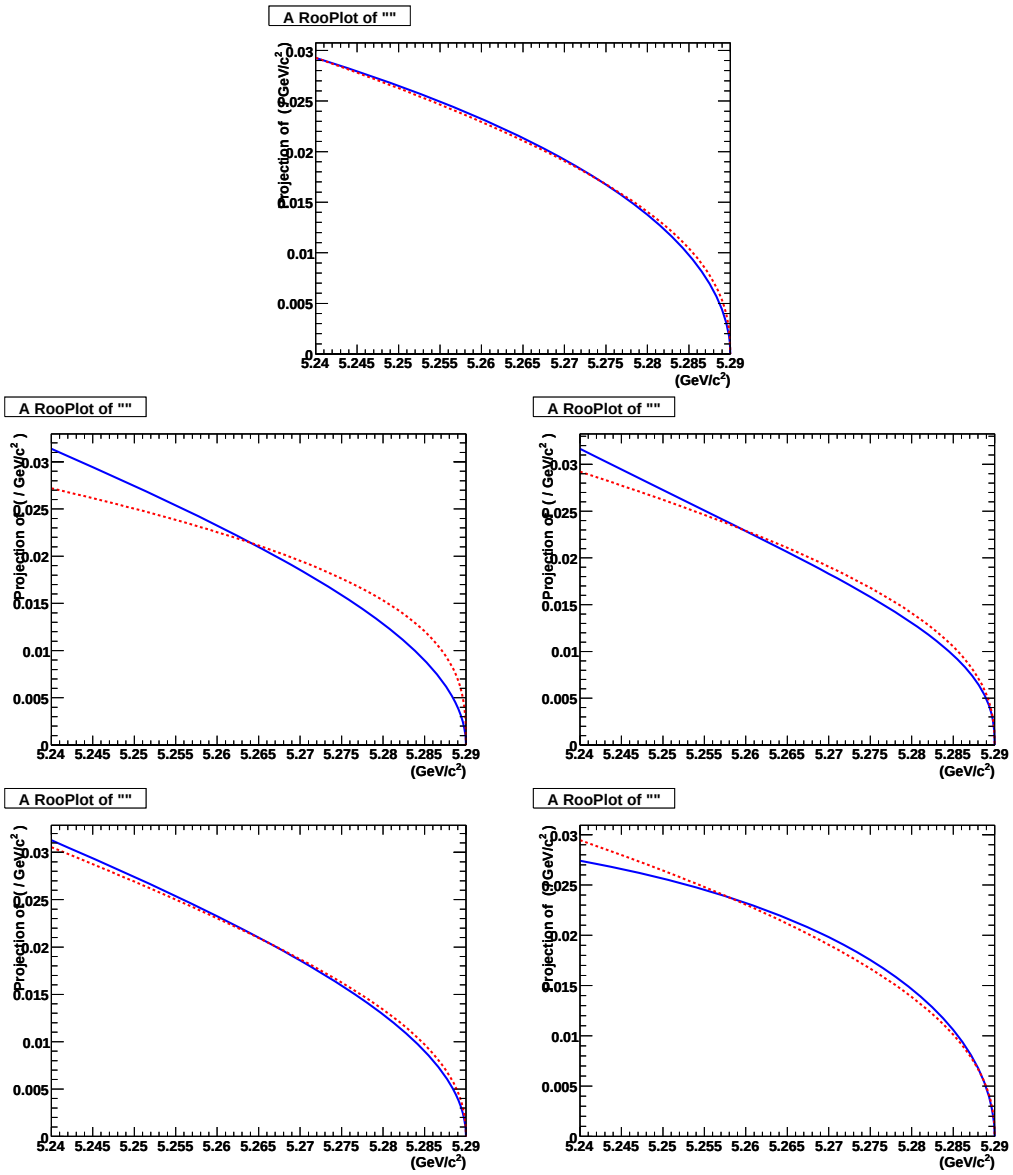


Figure 6-6: The continuum ARGUS models are compared, as determined from Monte Carlo (blue, solid) and off-peak data (red, dashed). The top plot is for all X_s mass regions, and the other four plots are for the sub-regions.

shape, we decide to take the safe approach and use the off-resonance data shape to describe the continuum background. The limitation with using this method is the low statistics of the off-resonance data, especially when we sub-divide the off-resonance data into four sub-regions. However, since the fit to the four sub-regions yield results that are well compatible with error (c.f. Table 6.2) We can use the shape that is determined from the off-resonance data in the entire M_{X_s} region to fit the four sub-regions. The larger number of events helps us to better determine the ARGUS shape and reduce the corresponding systematic error.

6.2 Tests of the Fits

To check whether our fit models have any bias, we perform a "toy" Monte Carlo study. A toy Monte Carlo study consists of a large ensemble of toy Monte Carlo experiments. In each of them, one generates the expected number of events for each category (usually allowing for Poisson fluctuations) according to the constructed PDFs and then performs a ML fit to the sample. A shift of the fit results from the input value would indicate a problem with the choice of the PDF in use.

In this analysis we generate 1000 toy Monte Carlo samples, representing 1000 experiments. In every experiment, each sub-sample (signal, $B\overline{B}$, and continuum background) is generated according to the individual PDFs. The number of events in each categories is obtained from the cut-flow tables 5.2-5.4 for $m_{\rm ES}$ greater than 5.22 GeV/c^2, allowing for Poisson fluctuations. We also allow the flavor fraction $N_{\overline{b}}/(N_b + N_{\overline{b}})$ in each sub-sample to fluctuate around a central value at 0.5. We then fit each experiment with the full extended ML function, with all PDF shape parameters fixed to their nominal values determined from the previous chapter.

A useful method to compare the generated and fitted parameter values is to calculate the pull. The pull is defined as

$$pull(x) = \frac{x_{fit} - x_{input}}{\sigma_{x_{fit}}} \tag{6.1}$$

where x_{input} is the value of the input parameter(i.e. the one used to generate the data set), x_{fit} is the value of the parameter returned from the fit and $\sigma_{x_{fit}}$ is returned error. When the PDF model correctly describes the signal shape and estimates the errors, the distribution of the pull over all 1000 toy Monte Carlo experiments should follow a Gaussian distribution with a mean of zero and a unit width.

Figure 6-7 shows the resulted distribution of the signal yields for the entire M_{X_s} regions, pull distribution of the signal yields, and the distribution for the errors of the signal yields respectively. The signal yields shape presents a nice Gaussian shape with a mean value consistent with the input number of events (Table 5.2). We also test the PDFs in the four M_{X_s} sub-region. The input number of events is obtained by sub-dividing the ones from the cut-flow tables into four M_{X_s} regions. Results are shown in Figure 6-8. We proceed to check distribution of the returned flavor fractions ($N_{\overline{b}}/(N_b+N_{\overline{b}})$) in the signal events. Figure 6-9 and Figure 6-10 show the distributions of the returned flavor fraction and their pull distribution respectively.

In the first figure, we find that except for the highest X_s mass region $2.0 < X_s < 2.8$ (lower right plot), the distribution of the flavor fraction all highly peak around 0.5. It is consistent with the input where there are similar number of b and $\overline{b}$ events. In the highest X_s mass, the distribution showed a wider width. This is due the fact that the signal events is extremely low compared to the background events in the region. We find the signal yields are negative at occasions and we have to let the fraction to go beyond the physical region (0,1) to get a sensible distribution. This effect also can be seen in the pull distribution: the lower right plot in Figure 6-10 showed a width of 0.76, while the widths of other plots in the same figure are around 1. However, this will not affect the overall performance of the fitter as we can see from the behavior of the distributions for the entire X_s mass region at the bottom of each figures mentioned above.

We conclude from the toy Monte Carlo experiments that there is a very small bias in the fitted signal yield, and within each category we find the bias to also be small.

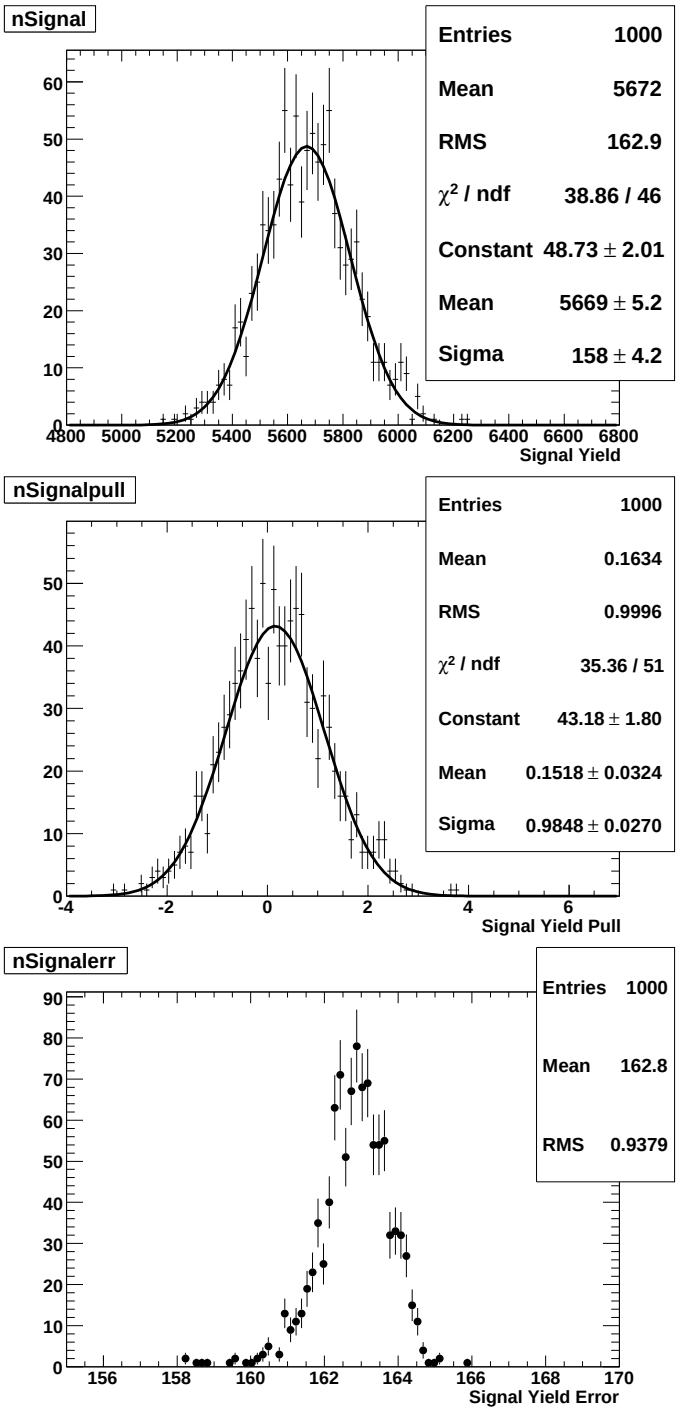


Figure 6-7: Results from the toy Monte Carlo studies. From top to bottom are the distribution of the signal yields, pull distribution of the signal yields, and the distribution for the error of the signal yields respectively. This result is for the entire M_{X_s} region.

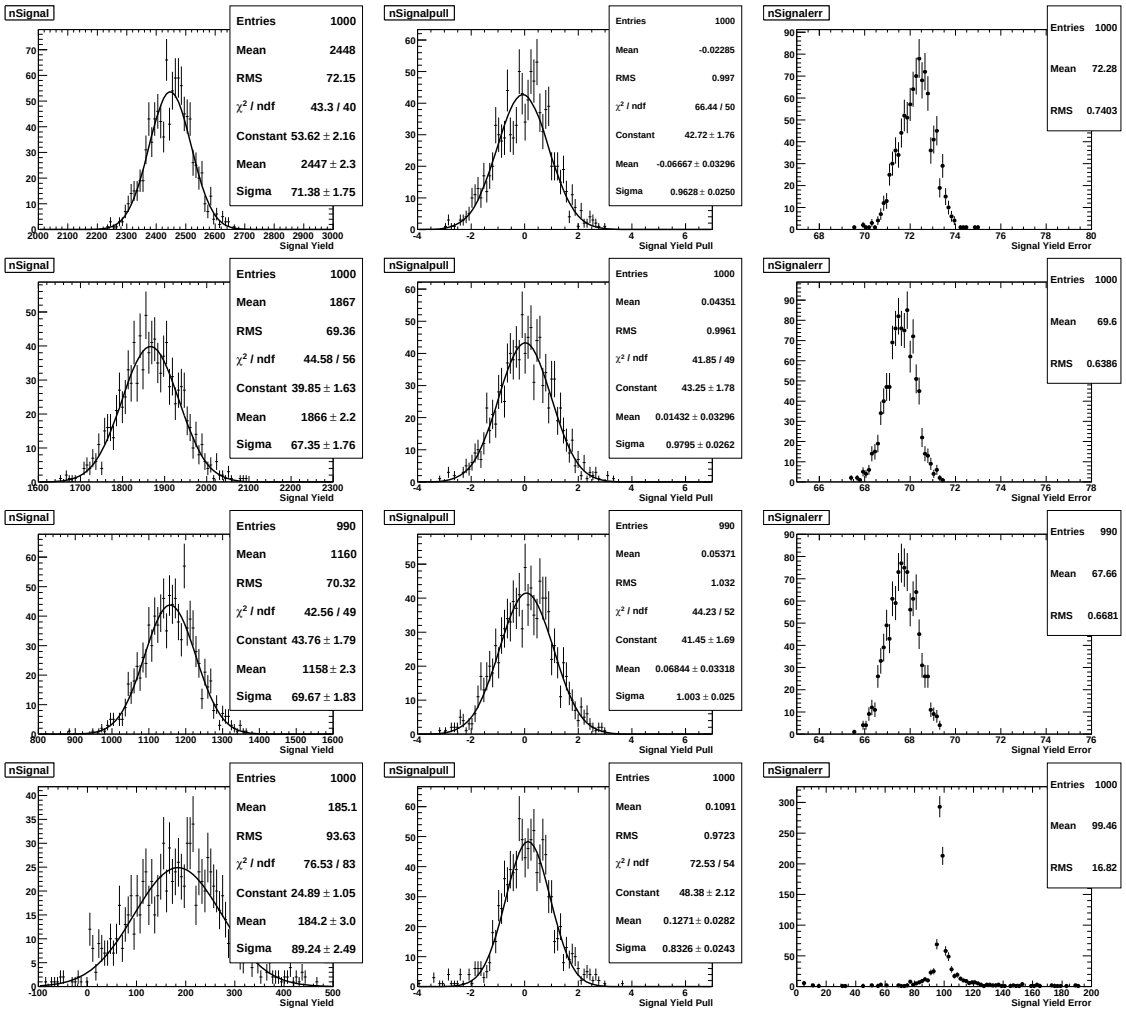


Figure 6-8: Results for the toy Monte Carlo studies for each M_{X_s} (GeV/c^2) region. From top to bottom, we show the toy Monte Carlo results for $0.6 < M_{X_s} < 1.1$, $1.1 < M_{X_s} < 1.5$, $1.5 < M_{X_s} < 2.0$, $2.0 < M_{X_s} < 2.8$ regions respectively. In each row, from left to right, we show the distribution of the signal yields, pull distribution of the signal yields and the distribution for the error of the signal yields respectively.

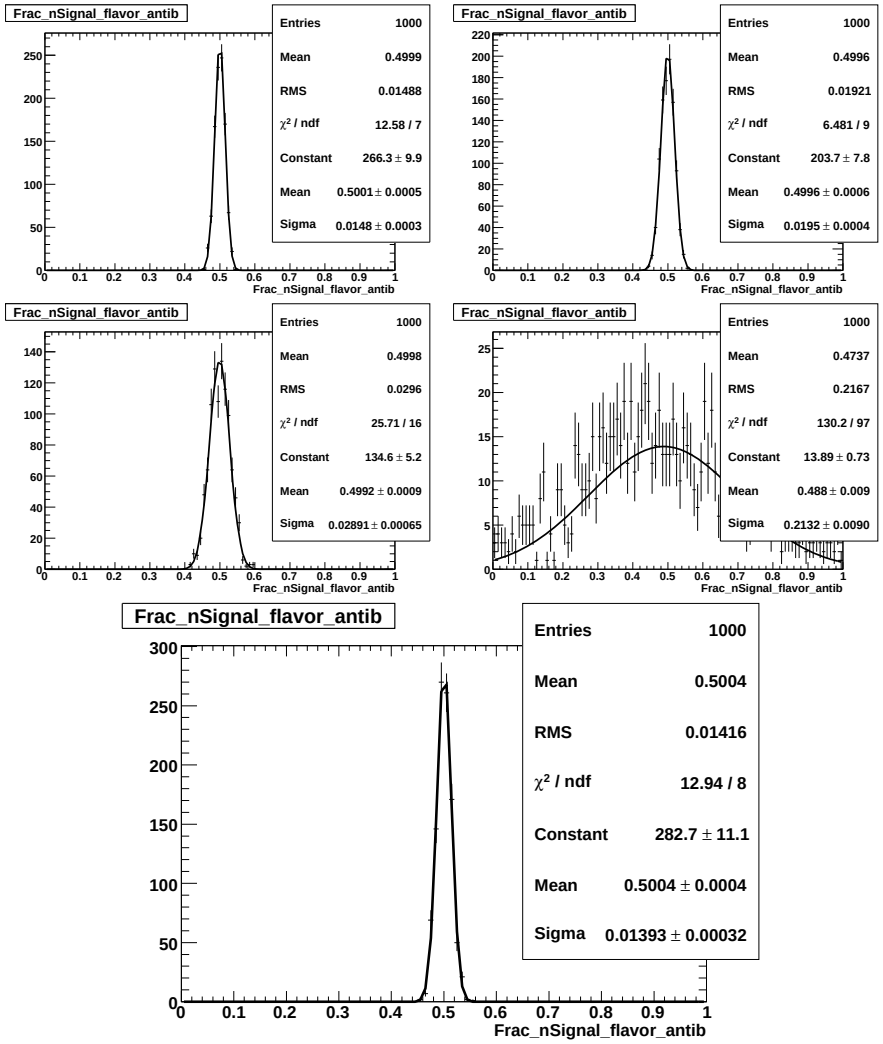


Figure 6-9: The distributions of the flavor fraction of signal events fitted in each M_{X_s} (GeV/c^2) region. The fraction of the b events are not shown as they are the normalization categories. $0.6 < M_{X_s} < 1.1$ (top left), $1.1 < M_{X_s} < 1.5$ (top right), $1.5 < M_{X_s} < 2.0$ (lower left), $2.0 < M_{X_s} < 2.8$ (lower right) and for the entire M_{X_s} region: $0.6 < M_{X_s} < 2.8$(bottom plot). The distribution of the b events are not shown as they are the normalization categories.

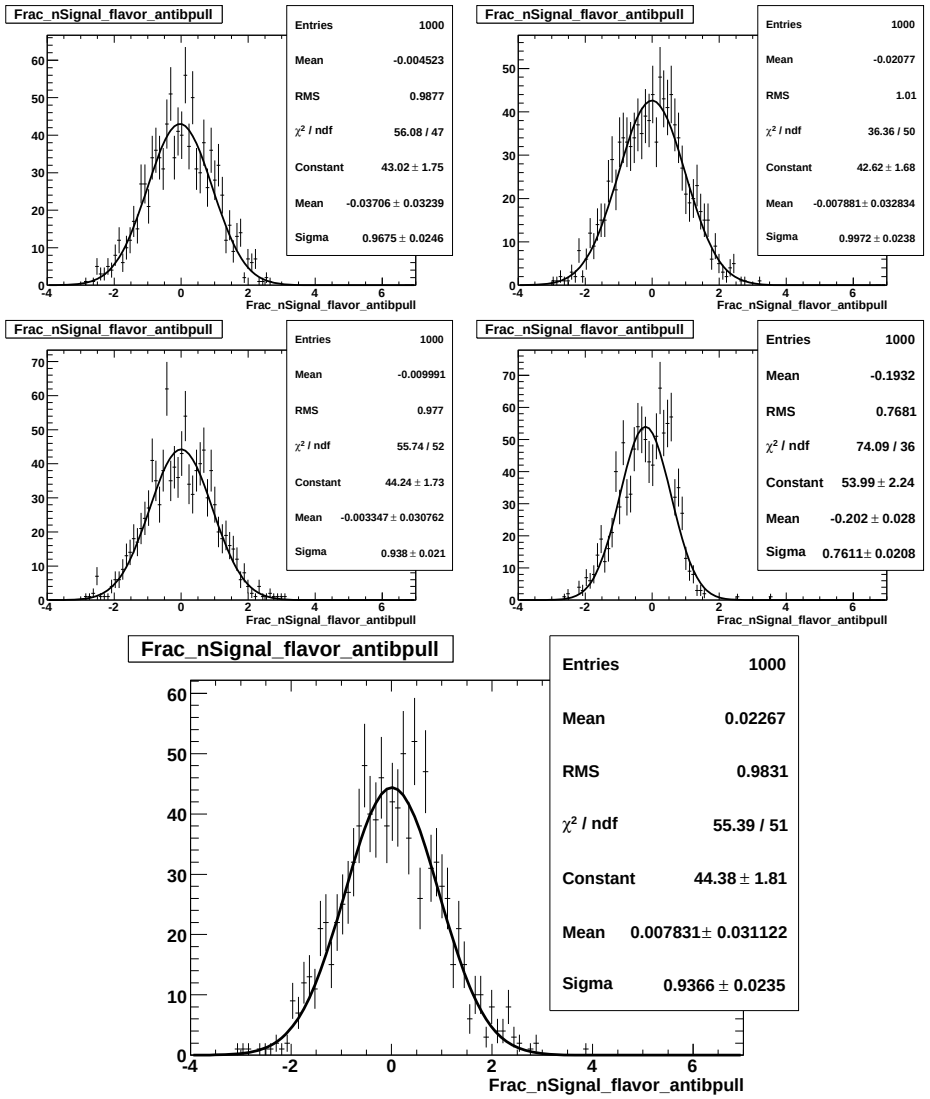


Figure 6-10: The pull distributions of the flavor fraction in signal events in each M_{X_s} (GeV/c^2) region. $0.6 < M_{X_s} < 1.1$ (top left), $1.1 < M_{X_s} < 1.5$ (top right), $1.5 < M_{X_s} < 2.0$ (lower left), $2.0 < M_{X_s} < 2.8$ (lower right) and for the entire X_s mass region: $0.6 < M_{X_s} < 2.8$(bottom plot). The pull distribution of the b events are not shown as they are the normalization categories.

6.3 Fit Results to Data

With all the PDFs established in Sec. 6.1, we perform a ML fit to data. All parameters are fixed to their PDF values except for the mean (m_0) of the signal model, which is determined from a fit to the data as described below. This is because it is observed to be shifted to the higher $m_{\rm ES}$ value in the data than in the Monte Carlo in several other related analyses.

6.3.1 Signal Model Mean m_0

We predict that the biggest bias in the fit model comes from the mean value of the signal model. We decide to fit the data in the two lower M_{X_s} bins ($0.6 < M_{X_s} < 1.1$ and $1.1 < M_{X_s} < 1.5$ GeV/c^2) for the mean value of the signal model, and then fix m_0 to be this value when we fit the other regions. The two lower M_{X_s} bins are chosen because they have the best signal to background ratios, and are the regions with largest concentration of signal events.

In the fit, the yield and the flavor fraction for each category are also floated. All other signal model parameters are fixed as well as those for the $B\overline{B}$ background and the continuum Model. We find the two regions yield compatible results (Table C.1). Compared to the mean in the Monte Carlo 5.2810 ± 0.0001, they are both shifted to the lower $m_{\rm ES}$ side, as observed in other analyses.

M_{X_s} Range (GeV/c^2)	Signal Mean m_0
0.6–1.1	$5.28060 \pm 0.8 \times 10^{-4}$
1.1–1.5	$5.28046 \pm 1.3 \times 10^{-4}$
Average [1]	$5.28056 \pm 0.7 \times 10^{-4}$

Table 6.3: Signal mean values determined from fitting the data in the two lower M_{X_s} bins.

[1] We take the weighted average of the two fitted results: the signal mean is calculated as $[(\frac{x1}{\sigma_1})^2 + (\frac{x2}{\sigma_2}^2]/[(\frac{1}{\sigma_1})^2 + (\frac{1}{\sigma_2})^2]$, with x_1 and x_2 be the central signal mean values for two regions, and σ be the corresponding error; The error on the average signal mean is $\sqrt{1/[(\frac{1}{\sigma_1})^2 + (\frac{1}{\sigma_2})^2]}$.

6.3.2 Fitted *CP* Asymmetry

With the signal mean value in hand, we proceed to find out what are the fitted *CP* asymmetries. In this step, we fixed all the PDFs parameters (the signal mean value is determined at the value calculated above, the others are determined from the nominal PDF fits), and fit for the flavor fraction (hence the fitted *CP* asymmetry A_{CP}^{Fit}) in each region. The results are shown in Table 6.4.

M_{Xs} (GeV/c^2)	Fraction of $\overline{b}$ events in signal ($f_{\overline{b}}$)	A_{CP}^{Fit} ($f_b - f_{\overline{b}}$)
0.6-1.1	0.507 ± 0.015	0.015 ± 0.029
1.1-1.5	0.502 ± 0.025	-0.003 ± 0.049
1.5-2.0	0.532 ± 0.038	-0.064 ± 0.077
2.0-2.8	0.548 ± 0.090	-0.097 ± 0.180
0.6-2.8	0.509 ± 0.015	-0.018 ± 0.030

Table 6.4: Fractions of the fitted signal $\overline{b}$ events out of the total number of fitted signal events in each M_{X_s} region. Corresponding A_{CP}^{Fit} in each M_{X_s} region is also calculated.

Chapter 7

Studies of Systematic Errors

The *CP* asymmetry (A_{CP}) is defined to be:

$$A_{CP} = \frac{N^b - N^{\overline{b}}}{N^b + N^{\overline{b}}} \tag{7.1}$$

where N^b and $N^{\overline{b}}$ are the number of b and $\overline{b}$ decays respectively. A_{CP} is related to the fitted *CP* asymmetry (A_{CP}^{Fit}) by the following formula:

$$A_{CP} = (A_{CP}^{Fit} - \Delta D)\frac{1}{\langle D \rangle} - A_{det} \tag{7.2}$$

where $\Delta D = (\bar{\omega} - \omega)$ is the difference in the wrong-flavor fraction between b and $\overline{b}$ decays, and $\langle D \rangle = 1 - (\bar{\omega} + \omega)$ is the dilution factor from the average wrong-flavor fraction. The small wrong-flavor fraction $\bar{\omega}$ (ω) is defined to be the fraction of $\overline{b}$ (b) reconstructed as the opposite flavor (detailed in Section 7.3). A_{det} accounts for the asymmetric detector responses to positively and negatively charged particles (explained in Section 7.2).

According to Equation 7.2, systematic errors arise when physics process affects the b and $\overline{b}$ final states differently, leading to a false *CP* asymmetry. Possible sources of the systematic errors that can contribute to the uncertainty on the result are: detector flavor biases, difference in the wrong-flavor fraction, and PDF modelings. In the following sections, we first take a paragraph to describe the $B \to X\pi^0$ control

sample that is constructed in the purpose of estimating the systematic errors. Then we describe the procedures that is used to quantify them.

7.1 The $B \to X\pi^0$ Control Sample

The X_s candidates in this control sample are paired with high-energy π^0 candidates to form B candidates. The reconstruction and pre-selection criteria are identical to that used in the $X_s\gamma$ reconstruction, except for the replacement of the photon by a π^0 candidate with center-of-mass energy $E^*_{\pi^0} > 1.6\,\mathrm{GeV}$. We remove the π^0 and η veto selection cuts at the reconstruction level, as they are no longer relevant to the event pre-selection. We intend to use this sample for several cross-checks of the main analysis:

1. The sample should be dominated by combinatoric candidates, since the inherent branching fractions for $B \to X\pi^0$ tend to be at the 0.1% level or smaller (some are not even measured) than $B \to X_s\gamma$. We can use this sample as a check of our understanding of the combinatoric background from both B and continuum events.

2. The sample can be used to cross-check the detector-related A_{CP} effects, since they contain the same hadronic system X_s.

We reconstruct the events as described above, and then select the best candidate without making any requirements on any of the post-reconstruction continuum veto variables listed in the cut-flow tables (Table 5.2 - 5.4). At this stage, the sample is enriched with continuum background. We show the distribution of the mass of the X_s system and m_{ES} for each event in Figure 7-1. The continuum background grows quickly at larger values of the mass.

Since we do not explicitly veto $B^0 \to \overline{D}^0\pi^0$ ($\mathcal{B} \sim 2.61 \times 10^{-4}$ [37]) decays, such events are present in the generic $B\overline{B}$ Monte Carlo, through $B^0 \to \overline{D}^0\pi^0$ and $\overline{D}^0 \to K^+\pi^-$ ($\mathcal{B} \sim 3.82\%$ [37]) and $\overline{D}^0 \to K^+\pi^-\pi^0$ ($\mathcal{B} \sim 13.5\%$ [37]) We therefore see a clear peak around around the D^0 mass. There is only a small hump shape around the K^*

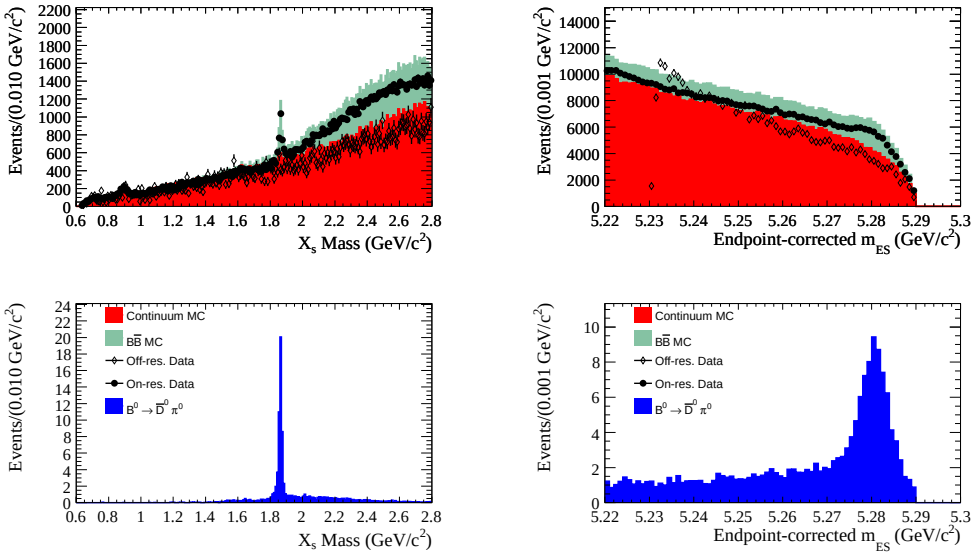


Figure 7-1: (a) X_s candidate mass, for all events after best candidate selection with $m_{\rm ES} > 5.265\,{\rm GeV}/c^2$, and (b) $m_{\rm ES}$ for all events after best candidate selection.

mass. That is because the branching fraction in this case is smaller ($\mathcal{B} \sim 3.6 \times 10^{-6}$ [37]) for $B \to K^*\pi^0$

$B^0 \to \overline{D}^0\pi^0$ behaves much like the signal in the shape the fitting variable $m_{\rm ES}$. Therefore we always show a specific $B^0 \to \overline{D}^0\pi^0$ signal Monte Carlo separate from the $B\overline{B}$ generic plots, to illustrate what it would look like once isolated. We also want to remind ourselves the existence of charm products in this control sample, therefore we choose to denote it with $B \to X\pi^0$ instead of $B \to X_s\pi^0$.

7.2 Inherent Detector Bias on A_{CP}

It is known that the detector efficiency is not symmetric for positive and negative charged tracks, therefore this asymmetry could easily produce a false CP asymmetry in the final result. We use data events in our $m_{\rm ES}$ sideband ($5.22 < m_{\rm ES} < 5.27\,{\rm GeV}/c^2$) to measure this detector asymmetry. In the $m_{\rm ES}$ sideband, most events arise from the continuum background, therefore we do not expect any physics mech-

anism to generate a flavor asymmetry. Thus A_{det} can be calculated as

$$A_{det} = (n_b - n_{\overline{b}})/(n_b + n_{\overline{b}}), \tag{7.3}$$

where n_b and $n_{\overline{b}}$ are the b and $\overline{b}$ events in the $m_{\rm ES}$ sideband respectively. Both $B \to X_s\gamma$ and the $B \to X\pi^0$ sideband data are used in this study, as they contain the same hadronic system and we expect to see similar detector responses in both samples.

Before proceeding to calculate A_{det}, we do need to verify that the the data events in the $m_{\rm ES}$ sideband have the same fragmentation as those in the signal events. Otherwise the detector responses will not be the same for the two samples, and we cannot simply subtract A_{det} obtained from the sideband data from the *CP* asymmetry calculated for signal events. Section D.4 explains the method that we use to check the fragmentation in the two samples and how we correct the fragmentation in the sideband data to have the same structure as it is in the signal events.

After this correction, we first calculate A_{det} for the $B \to X_s\gamma$ data sample. All selection criteria up to $m_{\rm ES} > 5.26\,{\rm GeV}/c^2$ in Table 5.2 - 5.4 are applied. We simply count the re-weighted number of b ($\overline{b}$) events in the $m_{\rm ES}$ sideband ($m_{\rm ES} < 5.27\,{\rm GeV}/c^2$), and the error is taken as the square root of the number of b (or $\overline{b}$) events. The charge asymmetry in the $m_{\rm ES}$ sideband is found to be -0.0059 ± 0.0062 for the entire M_{X_s} region (Table 7.1). Also listed in this table are the charge asymmetries the in $m_{\rm ES}$ sideband for $B \to X\pi^0$ data. The events are selected using the same cuts as those for $B \to X_s\gamma$ data except for the π^0 and η vetoes, and they are also re-weighted to have the same fragmentation structure as $B \to X_s\gamma$ signal events.

Given the uncertainties on both data samples, it is difficult to draw a firm conclusion on these trends. However, it's clear that for the entire M_{X_s} region, the charge asymmetries we obtained from the two data samples are consistent with each other. As a function of M_{X_s}, the pattern is difficult to discern. We decided to average the two, because they contain the same hadronic system (X_s), thus we expect the same detector effect in them. The result is the -0.0066 ± 0.0048 for the entire M_{X_s} region

M_{X_s}	$A_{det}(B \to X_s\gamma)$	$A_{det}(B \to X\pi^0)$	A_{det}(Average)
0.6–1.1	0.0014 ± 0.0173	0.0090 ± 0.0208	0.0052 ± 0.0135
1.1–1.5	0.0090 ± 0.0208	-0.0152 ± 0.0210	-0.0031 ± 0.0148
1.5–2.0	0.0014 ± 0.0134	-0.0344 ± 0.0151	-0.0161 ± 0.0101
2.0–2.8	-0.0015 ± 0.0083	-0.0015 ± 0.0082	-0.0015 ± 0.0058
0.6–2.8	-0.0059 ± 0.0062	-0.0073 ± 0.0066	-0.0066 ± 0.0048

Table 7.1: Table of A_{det} obtained from the sideband data. A_{det}(Average) is calculated as an average of results from both $b \to s\gamma$ and control sample.

as shown in Table 7.1.

In Appendix D we present a detailed study of charge asymmetry. We also verified that A_{det} does not depend on charged kaon momentum or charged track multiplicity. We split the event sample in the sideband into those reconstructed as B^0 and B^+ (charge conjugation is implied here). We find that there is no significant difference in the value of A_{det} in either of these two sub-samples. Based on this, we determine that the detector-related *CP* asymmetry is the same in Bs reconstructed either as charged or neutral. We can therefore subtract the same false A_{det} from both to determine the true A_{CP} in the $B \to X_s\gamma$ signal events.

7.3 Dilution from mis-tag rate

Signal events can be mis-reconstructed with the opposite flavor. For example, in Mode 3 $B^0 \to K^+\pi^-\gamma$, if the K^+ is mis-identified to a π^+ and the π^- is mis-reconstructed to a K^- this event will be reconstructed to be $\overline{B}^0 \to \pi^+K^-\gamma$ instead of a B^0 decay. Events reconstructed with opposite flavor are what we call "mis-tagged" events, and we define mis-tag rate $\omega(\bar{\omega})$ to be the ratio of the number of mis-reconstructed $b(\overline{b})$ events divided by the total number of reconstructed $b(\overline{b})$ events. A false *CP* asymmetry can be produced if the mis-tag rate is asymmetric. In this section, we study the false *CP* asymmetry due to the mis-tag rate.

Table 7.2 shows the signal modes where a mis-identification could lead to the opposite flavor. Note that in this table, we ignore the contribution from charge confusion, i.e. K^+ mis-identified to be K^-, or π^+ be changed to π^-, which rarely

Mode	Final State	Mis-tagged State
1	$B^+ \to K^0_S \pi^+ \gamma$	-
2	$B^+ \to K^+ \pi^0 \gamma$	-
3	$B^0 \to K^+ \pi^- \gamma$	$\overline{B}^0 \to \pi^+ K^- \gamma$
5	$B^+ \to K^+ \pi^+ \pi^- \gamma$	-
6	$B^+ \to K^0_S \pi^+ \pi^0 \gamma$	-
7	$B^+ \to K^+ \pi^0 \pi^0 \gamma$	-
9	$B^0 \to K^+ \pi^- \pi^0 \gamma$	$\overline{B}^0 \to \pi^+ K^- \pi^0 \gamma$
11	$B^+ \to K^0_S \pi^+ \pi^- \pi^+ \gamma$	-
12	$B^+ \to K^+ \pi^+ \pi^- \pi^0 \gamma$	-
13	$B^+ \to K^0_S \pi^+ \pi^0 \pi^0 \gamma$	-
14	$B^0 \to K^+ \pi^+ \pi^- \pi^- \gamma$	$\overline{B}^0 \to \pi^+ \pi^+ K^- \pi^- \gamma$
16	$B^0 \to K^+ \pi^- \pi^0 \pi^0 \gamma$	$\overline{B}^0 \to \pi^+ K^- \pi^0 \pi^0 \gamma$
23	$B^+ \to K^+ \eta \gamma$	-
27	$B^0 \to K^+ \eta \pi^- \gamma$	$\overline{B}^0 \to \pi^+ \eta K^- \gamma$
33	$B^+ \to K^+ K^- K^+ \gamma$	-
37	$B^0 \to K^+ K^- K^+ \pi^- \gamma$	$B^0 \to K^+ \pi^- K^+ K^- \gamma$

Table 7.2: Cases in which mis-tagged signal events generate the same or the opposite flavor due to pion mis-identified as a kaon. It is assumed that the signal events are not re-classified as cross-feed as a result of the mis-interpretation.

happens. Instead there is a much higher contribution from π^+ (π^-) mis-identified as K^+ (K^-)[1]. In addition, as we required the hadronic mass to be within 50 MeV/c^2 of the generated value in the "pseudo truth-matched" signal events (c.f. Section 4.3), it excludes the possibility of mis-tagging a signal event by adding or losing a pion or a kaon to the X_s final state.

We estimate the mis-tag rate using reconstructed signal Monte Carlo event. However, as the Monte Carlo does not reproduce correctly the kaon selection efficiencies, we first re-weigh each signal Monte Carlo event to have the correct kaon efficiency. At *BABAR*, the kaon identification efficiency for reconstructed tracks is measured with a control sample of kaons from D^* decays (c.f. Section 4.2.1) and presented in Figure 7-2[2]. We use the ratio of the data/MC kaon efficiency measured with this control sample to correct the signal Monte Carlo events. Then we applied the pion to kaon

[1]Note that we may easily have a kaon mis-identified as a pion ($K^+ \to \pi^+$), since there is no particle identification applied to the charged pions (i.e. we do not veto charged Kaons in the pion selection).

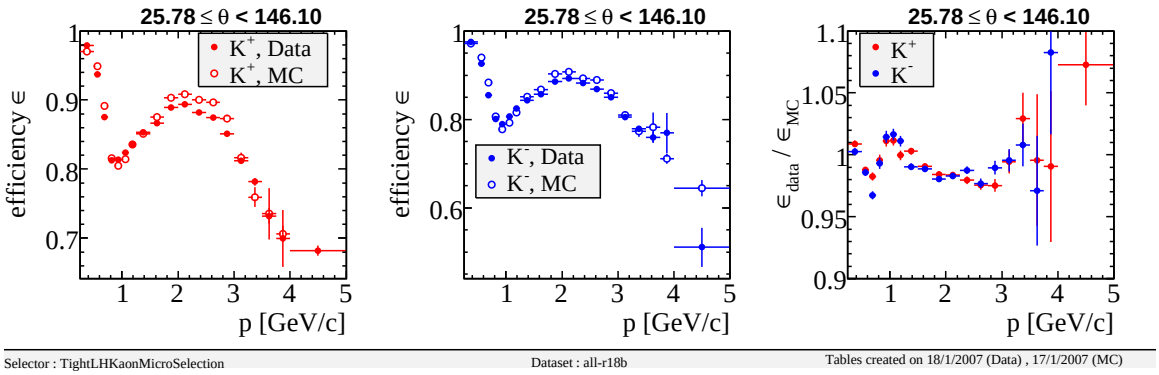


Figure 7-2: K^+ efficiencies for data and Monte Carlo (left plot), K^- efficiencies for data and MC (middle plot) and the ratio of the data/MC $K^\pm$ efficiencies (right plot).

M_{X_s}	ω	$\bar{\omega}$	$\langle D \rangle$	ΔD
0.6-1.1	0.0056±0.0001	0.0055±0.0001	0.9945±0.0001	-0.0001±0.0002
1.1-1.5	0.0050±0.0001	0.0052±0.0001	0.9949±0.0001	0.0002±0.0001
1.5-2.0	0.0056±0.0001	0.0056±0.0001	0.9944±0.0001	0.0001±0.0001
1.1-1.5	0.0053±0.0001	0.0055±0.0001	0.9946±0.0001	0.0002±0.0001
0.6-2.8	0.0053±0.0001	0.0054±0.0001	0.9946±0.0001	0.0001±0.0001

Table 7.3: Table of mis-tag rate (ω, $\bar{\omega}$) and dilution rate ($\Delta D = (\bar{\omega} - \omega)$, $\langle D \rangle = 1 - (\bar{\omega} + \omega)$) as a function of M_{X_s}.

mis-identification rate (Figure 4-1 lower plot), as a function of the pion momentum to measure the probability of the pion being mis-identified to be a kaon.

The measured mis-tag rate are reported in Table 7.3 as a function of the hadronic mass. The error on the mis-tag rate is the sum in quadrature of the statistical error. We found that both $\langle D \rangle$ and ΔD has an impact on A_{CP} only after the fourth decimal place, which is beyond the sensitivity of our analysis. Therefore we ignore them in the final calculation.

7.4 PDF Shape Systematics

The models used to constrain the signal and $B\overline{B}$ background are both taken from Monte Carlo in the nominal model. Previous experience from related analyses suggest

[2]http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Selectors/r18b/TightLHKaonMicroSelection_details.html

that we should expect deviations from the MC simulation. In the following sections, we discuss the different ways in which we will address modeling effects on A_{CP} from different PDF shapes.

7.4.1 Signal Model Systematics

The signal shape is fitted with a Cruijff function. All shape parameters are determined from the Monte Carlo except for the mean, which is determine from a fit to the data, as described in Section 6.3.1.

We explore the quantitative effect in our maximum likelihood fit by varying the Cruijff mean m_0, left width σ_L, and right width σ_R by 1σ sequentially, and calculating how much the A_{CP} changes as a consequence of these variations. The overall effect is calculated as the the sum in quadrature of the change on A_{CP} in each variation. We found that the effect on the A_{CP} is negligibly small (0.0002), which is beyond the sensitivity of this analysis.

7.4.2 Continuum Model Systematics

The continuum shape is fitted with an ARGUS function determined from a fit to the off-resonance data (Sec 6.1.3). To enlarge the statistics as to better determine the ARGUS shape parameters and reduce the corresponding systematic error, we use the shape that is determined from the off-resonance data in the entire M_{X_s} region to fit the four sub-regions. Section 6.2 shows that the parameters are very well compatible across the M_{X_s} regions.

In addition we perform a χ^2 study to check the compatibility between the ARGUS shapes. In each M_{X_s} region, we first fix the ARGUS function to be the continuum Monte Carlo shape of that region, and calculate the χ^2. Then we fix the ARGUS parameters to continuum Monte Carlo shape in merged M_{X_s} bin and recalculate the χ^2[1]. This study is done using the continuum Monte Carlo shape instead of the off-resonance shape because the latter is too low statistics to give a meaningful answer.

M_{X_s}	χ^2 with unmerged continuum MC shape	χ^2 with merged continuum MC shape
0.6-1.1	118.948	118.044
1.1-1.5	146.360	146.498
1.5-2.0	131.266	131.017
2.0-2.8	133.647	134.693

Table 7.4: Results of the χ^2 study on the on-peak data. First column shows the results where we fix the ARGUS function to be the continuum Monte Carlo shape in that M_{X_s} bin. Second column shows the results where we fix the ARGUS function to be the continuum Monte Carlo shape in the merged M_{X_s} bin.

The results are shown in Table 7.4, justifying our approach.

We proceed to study the systematic errors arising from the continuum modeling. This is done by varying the off-peak data shape within the experimental errors and check the effect of this variation on A_{CP}. The correlation between the two ARGUS parameters have been taken into account. The effect on A_{CP} from this variation is found to be 0.006 for the combined M_{X_s} region. The systematic errors arising from the continuum shape in each M_{X_s} region are listed in Table 7.7

7.4.3 $B\overline{B}$ background Modeling Systematics

The shape of the $B\overline{B}$ background, determined from generic $B\overline{B}$ Monte Carlo and cross-feed, could also be a potential source of flavor bias in the fit to the data. This background peaks broadly in the signal region, and a small shape difference as a function of flavor could create a false *CP* asymmetry in the signal events.

We studied the $B\overline{B}$ Monte Carlo shape and the cross-feed background shape separately, because the background sources are different in the two samples. $B\overline{B}$ Monte Carlo enters our reconstruction mostly when we lose a photon in the final state, while cross-feed background originates from losing or adding a π^0 to the final state. In the following of this section, we discuss the correction to the $B\overline{B}$ shape and cross-feed shape respectively.

[1]This χ^2 is calculated as a summation of χ^2 for both flavor b and $\overline{b}$, each with 70 bins.

Correction to the $B\overline{B}$ Monte Carlo shape

The high-energy $B \to X\pi^0$ control sample mentioned in Section 7.1 is used to study the uncertainty of the $B\overline{B}$ background shape. We can use the shape comparison between data and Monte Carlo in the control sample to correct our $B \to X_s\gamma$ $B\overline{B}$ Monte Carlo shape.

We produce a sample similar to our $B\overline{B}$ Monte Carlo by re-calculating $m_{\rm ES}$. To do this, we define $m_{\rm ES}{}' = \sqrt{E^{*2}_{beam} - |\vec{p}^{\,*}_B - \vec{p}^{\,*}_{\gamma_2}|^2}$, where $\vec{p}^{\,*}_{\gamma_2}$ is the momentum of the lower energy photon in the final-state π^0 in $B \to X\pi^0$. This will simulate the failure to reconstruct a photon associated with the real $B \to X\pi^0$ decay, giving us a peaking background shape in $m_{\rm ES}{}'$ to work from.

We apply all analysis cuts, except the π^0 and η vetoes. We then check whether the Monte Carlo models correctly the shape of $m_{\rm ES}{}'$ in the data. In order to do this comparison, we have to remove the contribution to the data from continuum events. Thus we subtract the $m_{\rm ES}{}'$ distribution of the continuum Monte Carlo from that of the on-resonance data. Then we are left with a pure $B\overline{B}$ shape, and we compare it with the $m_{\rm ES}{}'$ shape of the Monte Carlo $B\overline{B}$ events.

We first need to check whether there is any disagreement between different b quark flavors for the Monte Carlo and on-peak data. The shape of the Monte Carlo and the shape of the data, internally compared between b and $\overline{b}$ events, are consistent. The results are shown in Figure 7-3. Both plots are normalized to the unit area. No shape differences is observed between flavors in either Monte Carlo or data. The agreement between Monte Carlo and data is fair with the data peaking higher than the Monte Carlo in the signal region. We therefore conclude that the data/MC difference should be used as a correction to our nominal $B\overline{B}$ PDF, and either used in the default fit or as a systematic variation.

We use events from the entire M_{X_s} region and produce data/MC correction histograms as shown in Figure 7-4. Here, we use a coarse binning in the sideband region. This is because we find the plots in Figure 7-3 are too fine in binning: while the MC falls into a smooth tail into the lower $m_{\rm ES}{}'$ side, the on-resonance data presents some

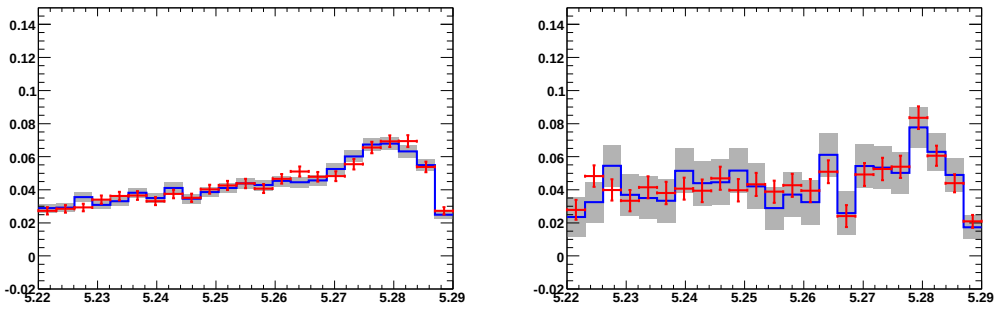


Figure 7-3: $m_{\rm ES}{}'$ distribution for b (red) and $\overline{b}$ (blue) for $B\overline{B}$ Monte Carlo (left), and on-resonance data (right) with the continuum Monte Calor subtracted. All the distribution are normalized to unit area.

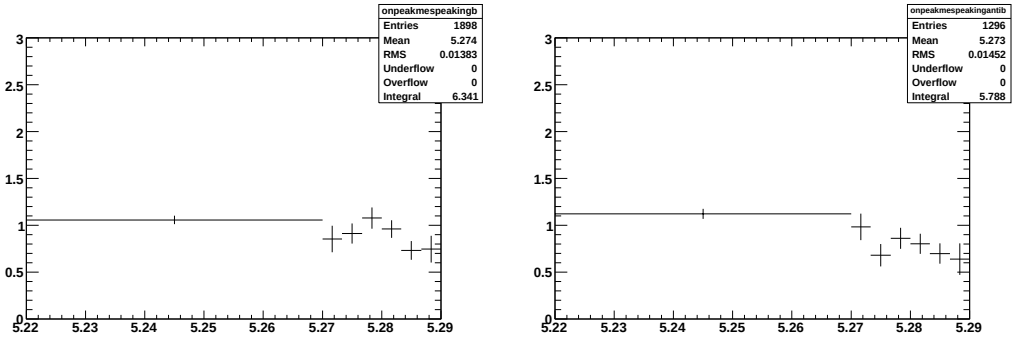


Figure 7-4: Data/MC correction histogram with coarse binning for the entire M_{Xs} region. b flavor on the left and $\overline{b}$ on the right.

statistical fluctuations in the $m_{\rm ES}{}'$ sideband. It is possible that the statistical variation will be transmitted into the new shapes of the $B\overline{B}$ background. We therefore treat the sideband $m_{\rm ES}{}' < 5.27\,{\rm GeV}/c^2$ as one bin, since the agreement is quite good overall. The peak region is binned more finely, to allow for more detailed variations.

Correction to the cross-feed shape

The biggest uncertainty in the cross-feed shape is due to the fact that JETSET does not well reproduce the observed fragmentation structure of data. We thus correct the simulation shape using the fragmentation previously determined from *BABAR* data in Ref. [17].

In reference [17], they split the data and Monte Carlo samples into several different

categories and then and find the Data/MC ratio for each of these subsets. This measures how well the JETSET fragmentation of the X_s system reproduces the data. We show the relevant ratios in Table 7.4.3 for your reference. They are applied as weights to our cross-feed events to correct for the observed fragmentation.

Final States	Data/Monte Carlo
$K^0_S\ \pi^-$ (1),$K^-\pi^+$ (3)	0.50 ± 0.07
$K^-\pi^0$ (2)	0.19 ± 0.12
$K^-\pi^+\ \pi^-$ (5)	1.02 ± 0.14
$K^0_S\ \pi^-\ \pi^0$ (6), $K^-\pi^+\ \pi^0$ (9),	1.34 ± 0.24
$K^0_S\ \pi^+\ \pi^-\ \pi^-$ (11),$K^-\pi^+\ \pi^-\ \pi^+$ (14)	2.67 ± 0.96
$K^-\pi^+\ \pi^-\ \pi^0$ (12)	1.29 ± 0.61
$K^-\pi^0\ \pi^0$ (7), $K^0_S\ \pi^-\ \pi^0\ \pi^0$ (13) $K^-\pi^-\ \pi^0\ \pi^0$ (16)	1.89 ± 1.33
$K^-\eta$(23), $K^-\eta\ \pi^+$ (27)	$0.83^{+1.00}_{-0.83}$
$K^-K^+K^-$(33),$K^-K^+K^-\pi^+$ (37)	$0.27^{+0.54}_{-0.27}$

Table 7.5: Ratios of data to Monte Carlo yields for various categories of final states obtained from Ref. [17]. They are used to adjust the fragmentation in the cross-feed. The corresponding mode indices are in parenthesis.

Systematic error from the $B\overline{B}$ background modeling

With the corrected $B\overline{B}$ Monte Carlo and cross-feed shapes in hand, we construct the new b and $\overline{b}$ binned PDFs (Figure 7-5 and Figure 7-6) with these corrected shape and fit the data a second time with them. The difference between the nominal A_{CP} and A_{CP} from this fit, is used as the systematic error from shape modeling of the $B\overline{B}$ background (Table 7.6).

7.5 Summary of Systematic Errors

To sum up the above systematic studies, we determined that

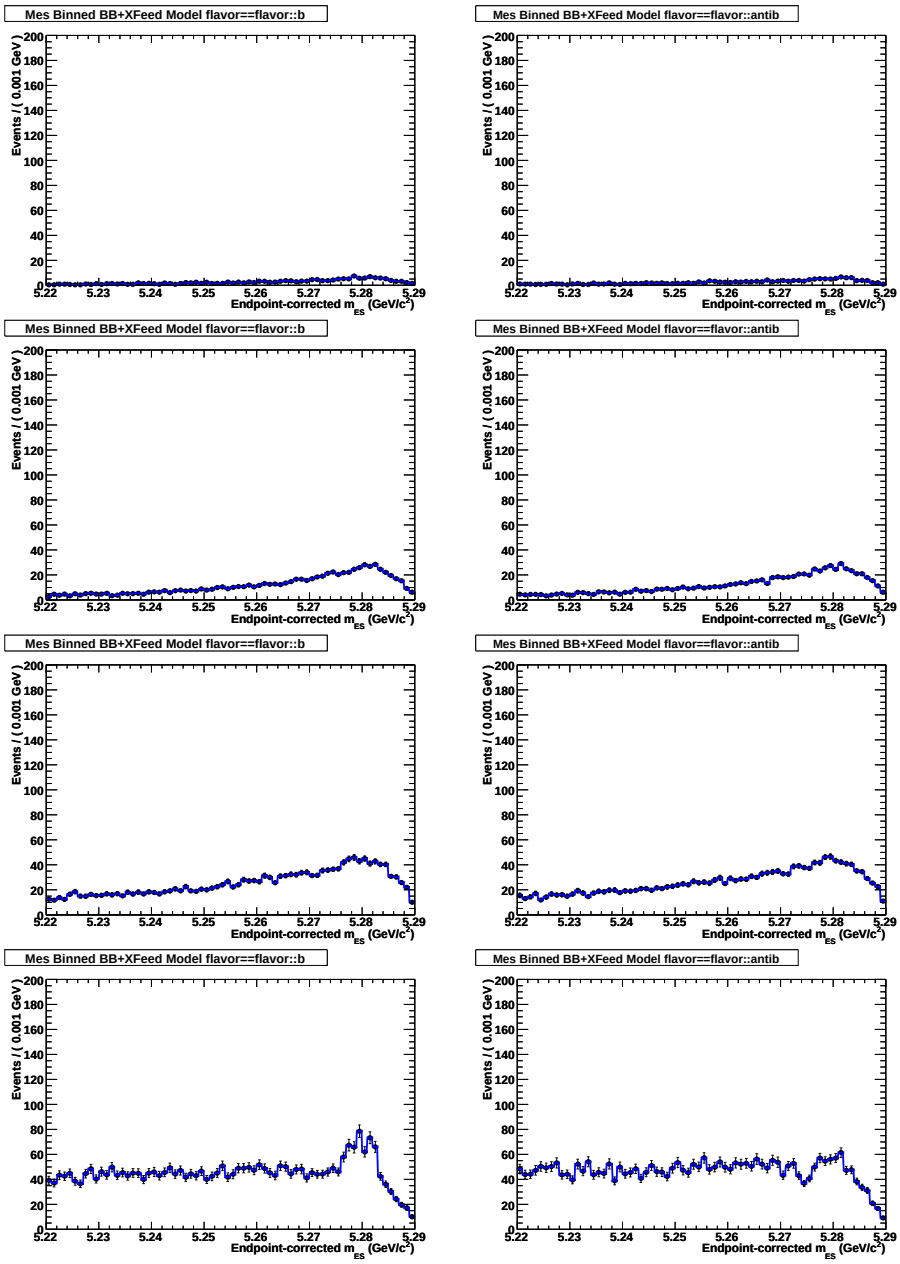


Figure 7-5: Individual m_{es} shapes for the $B\overline{B}$ background after re-weighting the $B\overline{B}$ MC and cross-feed events. Plots are for both flavors (b on the left, $\overline{b}$ on the right) and the four M_{X_s} (GeV/c^2) regions: $0.6 < M_{X_s} < 1.1$ (first row), $1.1 < M_{X_s} < 1.5$ (second row), $1.5 < M_{X_s} < 2.0$ (third row), $2.0 < M_{X_s} < 2.8$ (fourth row).

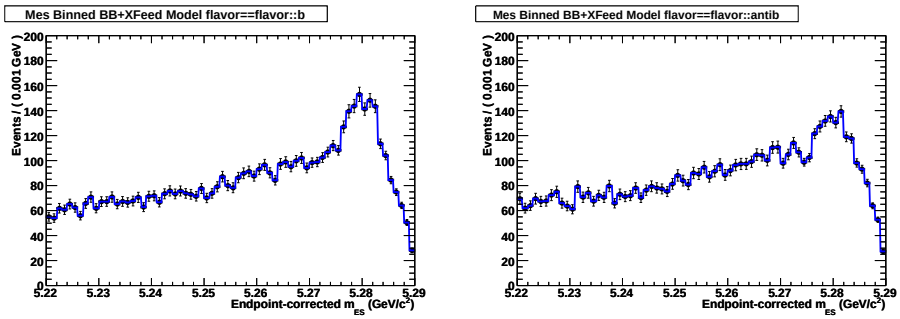


Figure 7-6: $m_{\rm ES}$ shapes for the $B\overline{B}$ background for both flavors (b on the left, $\overline{b}$ on the right) after re-weighting the $B\overline{B}$ MC and cross-feed events. Plots are for for the entire M_{X_s} (GeV/c^2) region: $0.6 < M_{X_s} < 2.8$.

M_{X_s}	Nominal A_{CP}	New A_{CP}	$\lvert\Delta A_{CP}\rvert$
0.6-1.1	0.015	0.017	0.002
1.1-1.5	-0.003	-0.006	0.003
1.5-2.0	-0.064	-0.074	0.010
1.1-1.5	-0.097	-0.167	0.070
0.6-2.8	-0.018	-0.030	0.012

Table 7.6: Table of the results from fitting the data with the nominal model and with the corrected $B\overline{B}$ background shape.

- the detector inherent bias (A_{det}) mean values will be used to correct the A_{CP} central values;
- systematic error from the mis-tag dilution and signal modeling is negligible;
- among the fit models, the systematic error on the $B\overline{B}$ background and continuum fit modeling are the major impacts.

Table 7.7 itemizes the systematic error from each source, and shows the total error in each M_{X_s} region, which is calculated as the sum in quadrature of errors on A_{det}, systematic errors arising from the continuum, $B\overline{B}$ background modeling.

M_{X_s} GeV/c^2	A_{det}	Mis-tag Dilution	Signal Model Syst.	$B\overline{B}$ bkg Model Syst.	Continuum Model Syst.	Total Syst.
0.6-1.1	0.005 ± 0.014	-	-	0.002	0.004	0.015
1.1-1.5	-0.003 ± 0.015	-	-	0.003	0.004	0.016
1.5-2.0	-0.017 ± 0.010	-	-	0.010	0.002	0.014
2.0-2.8	-0.002 ± 0.005	-	-	0.070	0.168	0.182
0.6-2.8	-0.007 ± 0.005	-	-	0.012	0.006	0.014

Table 7.7: For each M_{X_s} bin, we present the flavor-bias of the detector: A_{det}, the systematic error arising from the $B\overline{B}$ background modeling and the continuum background modeling. "-" means the effect is negligible. The last column shows the total systematic error.

Chapter 8

Results and Conclusions

In the previous chapters, we described the event reconstruction, background suppression, fitting strategies as well as the systematic error studies. In this chapter, the determination of the *CP* asymmetry as a function of the hadronic mass is performed. We first give the estimation of the expected precision on A_{CP} on Monte Carlo events and then present the results from our measurement on data. The estimation and the results are compared and discussed afterwards.

8.1 Expected Precision on A_{CP} on Monte Carlo

The so-called "toy" Monte Carlo studies described in Section 6.2 also provide a way to estimate the precision on A_{CP} fits. For each M_{X_s} region, 1000 experiments are generated according to the constructed PDFs. By allowing Poisson fluctuation to the input number of events for each sub-sample in each experiment, we can test the sensitivity of the fit parameters (signal yields in our case) with respect to the statistical fluctuations.

For the entire M_{X_s} region, we expect the mean of the yield at 5669 and the error of the signal yield at 163 (c.f. 6-7). Assuming equal number of b and $\overline{b}$, we obtain 2835 ± 115 as the yield for each flavor. With simple error propagation, this transfers into an error of 0.029 on $A_{CP} = (N_b - N_{\overline{b}})/(N_b + N_{\overline{b}})$.

Table 8.1 shows the expected signal yields and the precision on A_{CP} for each M_{X_s}

region, obtained from toy Monte Carlo studies.

M_{X_s} GeV/c^2	Expected Signal Yields from Monte Carlo	Expected Precision $A_{CP}(b \to s\gamma)$
0.6-1.1	2447 ± 72	0.029
1.1-1.5	1866 ± 70	0.037
1.5-2.0	1158 ± 68	0.059
2.0-2.8	176 ± 99	0.563
0.6-2.8	5669 ± 163	0.029

Table 8.1: The expected signal yields and the expected A_{CP} precision for each M_{X_s} bin obtained from the toy Monte Carlo studies (Figure 6-8).

8.2 Measured $A_{CP}(b \to s\gamma)$ in Data

The formula for calculating A_{CP} from experimental values was defined in Equation 7.2 as

$$A_{CP} = (A_{CP}^{Fit} - \Delta D)\frac{1}{\langle D \rangle} - A_{det}. \tag{8.1}$$

The fitted *CP* asymmetries, A_{CP}^{Fit}, along with their statistical errors are shown in Table 6.4. Table 7.7 gives the result for the detector asymmetry A_{det}, the dilution factors ΔD and $\langle D \rangle$ as well as the systematic errors arising from the PDF modelings and the total systematic errors. These values are documented in Table 8.2, along with the measured A_{CP} in data events, using Equation 7.2. The fits to the data in each of the five M_{X_s} regions are shown in Figure 8-1 and 8-2

Our most significant result is the measurement of the *CP* asymmetry for the entire M_{X_s} region ($0.6 < M_{X_s} < 2.8$ GeV/c^2). For this region, we obtain:

$$A_{CP} = -0.011 \pm 0.030_{stat} \pm 0.014_{syst}. \tag{8.2}$$

This result is consistent with zero, as predicted by the Standard Model.

We split our sample into four sub-categories according to the reconstructed M_{X_s} and we show that the asymmetries in each region measured is compatible with each other (Table 8.2) and that they are all consistent with zero within $0.5\,\sigma$. This confirms

M_{X_s} GeV/c^2	A_{CP}^{Fit}	A_{det}	$A_{CP}(b \to s\gamma)$
0.6–1.1	0.015 ± 0.029	0.005 ± 0.014	$0.010 \pm 0.029 \pm 0.015$
1.1–1.5	-0.003 ± 0.049	-0.003 ± 0.015	$0.000 \pm 0.049 \pm 0.016$
1.5–2.0	-0.064 ± 0.077	-0.017 ± 0.010	$-0.047 \pm 0.077 \pm 0.014$
2.0–2.8	-0.097 ± 0.180	-0.002 ± 0.005	$-0.077 \pm 0.180 \pm 0.182$
0.6–2.8	-0.018 ± 0.030	-0.007 ± 0.005	$-0.011 \pm 0.030 \pm 0.014$

Table 8.2: For each M_{X_s} bin, we present the fitted *CP* asymmetry: A_{CP}^{Fit}, the flavor-bias of the detector: A_{det}, and the final results of the *CP* asymmetries.

the Standard Model prediction that the *CP* asymmetry should be independent of the photon energy range (hence, M_{X_s}) in use [28].

Our measurement is consistent with the previous *BABAR*[15] and Belle [7] results (Table 1.3). It is the most precise measurement of this variable so far. We reduce the statistical error by 40%, while the systematic error is comparable to the previous *BABAR* analysis [17]. However, we notice that in the previous analysis the only systematic error quoted was the one due to the detector bias A_{det}, and we do have improved the A_{det} measurement by a factor of three (Table 7.7) with the larger data sample in use. The dominating systematic error in our analysis is from the modeling of the $B\overline{B}$ background which, even with larger statistics, is still very difficult to reduce.

All systematic errors are about 0.015 except for region 4 ($2.0 < M_{X_s} < 2.8$ GeV/c^2), where the large error arises from the $B\overline{B}$ background and the continuum background shape. This region is dominated by the background events, therefore a tiny change in the background shape will affect the signal yields and can amplify the *CP* asymmetry as well.

So far, all measurements of $A_{CP}(b \to s\gamma)$ are consistent with the Standard Model prediction within the experimental resolution, but are not precise enough to exclude some new physics model. For instance, type II two-Higgs-doublet model (THDM II) [1] predicts small *CP* violating effects comparable to that of the Standard Model, although it is already constrained by the $b \to s\gamma$ Branching Fraction [35]. Therefore, with the current results, we can only elaborate and constrain possible A_{CP} phase

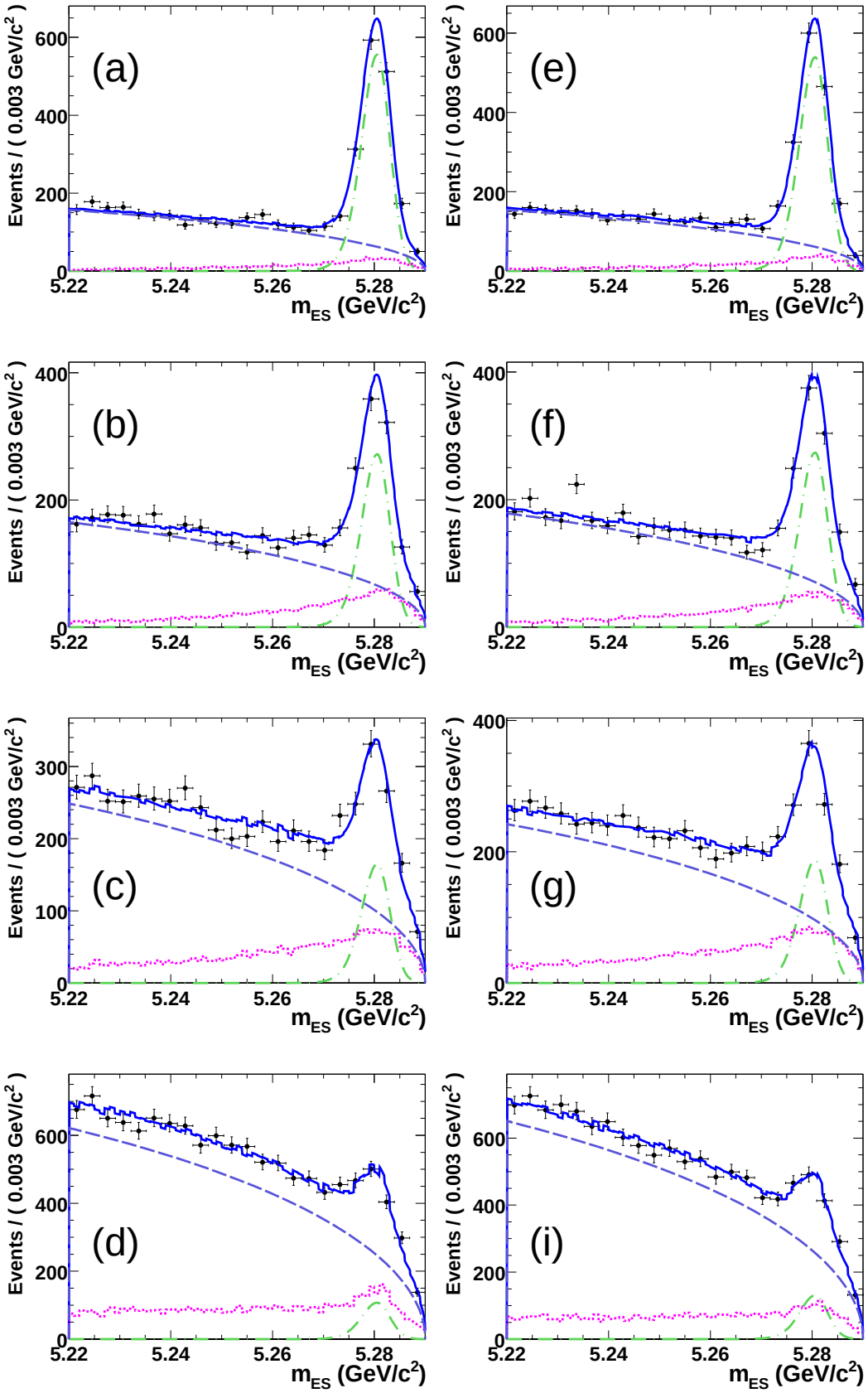


Figure 8-1: Fits to the $m_{\rm ES}$ distribution in data for $b \to s\gamma$ events in M_{X_s} (GeV/c^2) region (a) [0.6, 1.1], (b) [1.1, 1.5], (c) [1.5, 2.0], (d) [2.0, 2.8], and $\overline{b} \to \overline{s}\ \gamma$ events in M_{X_s} region (e) [0.6, 1.1], (f) [1.1, 1.5], (g) [1.5, 2.0], (h) [2.0, 2.8], The dashed line shows the shape of the continuum, dotted-dashed line shows the fitted signal shape

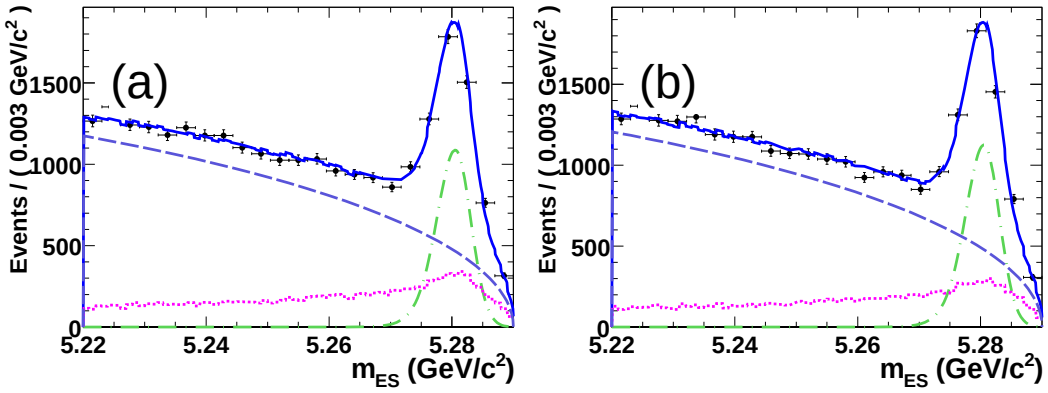


Figure 8-2: Fits to the $m_{\rm ES}$ distribution in data for (a) $b \to s\gamma$ events in and (b) $\overline{b} \to \overline{s}\ \gamma$ events in the entire M_{X_s} region. The dashed line shows the shape of the continuum, dotted-dashed line shows the fitted signal shape and the dotted line shows the $B\overline{B}$ background shape.

space in this models (Table 8.3).

Models	Predicted A_{CP} Range
SUSY with Minimum Flavor Violation [28]	~ -0.03 to $+0.03$
2 Higgs-doublet model (type II) [31]	~ 0.006
3 Higgs-doublet model [31]	~ -0.03 to $+0.03$
Supergravity [3, 26]	~ -0.10 to $+0.10$
SUSY with squark mixing [10, 32]	~ -0.15 to $+0.15$
SUSY with R-parity violation [11]	~ -0.17 to $+0.17$

Table 8.3: *CP* asymmetries in various models. Ranges are approximate using current knowledge of $b \to s\gamma$ decays.

8.3 Measured Signal Yields in Data

We find that not all the statistical errors on the fitted *CP* asymmetries match the predicted precisions on the A_{CP}(Table 8.1). We explore the reason by checking the signal yields in each M_{X_s} region.

In the first two columns (*BABAR* 2008 analysis) of Table 8.4, we show our expected number of signal events obtained from Monte Carlo and the signal yields from the fits to the data. We find that the toy Monte Carlo accurately predicted the errors on the

signal yields. It is the input number of events we obtained from the cut-flow table that does not quite match the fitted signal yields. For the entire M_{X_s} region and the K^* resonance region, the expected number of events agree with the fitted signal yields within 2 σ. For other regions, however, we see a deviation on the signal yields bigger than 3 σ. We conclude from this comparison that our PDFs are not biased, since the toy Monte Carlo studies predict the uncertainties accurately. However, we proceed to explore what causes the deviation between the expected and the fitted signal yields.

BABAR 2008 results [this work]			*BABAR* 2004 results [15, 13]		
M_{X_s} (GeV/c^2)	Expected signal yield (MC)	Fitted signal yield (Data)	M_{X_s} (GeV/c^2)	Expected signal yield (MC)	Fitted signal yield (Data)
0.6-1.1	2447 ± 72	2575 ± 76	0.6-1.1	833 ± 45	784 ± 46
1.1-1.5	1866 ± 70	1284 ± 64	1.1-1.5	308 ± 29	298 ± 32
1.5-2.0	1158 ± 68	827 ± 64	1.5-1.9	165 ± 30	263 ± 28
2.0-2.8	176 ± 99	559 ± 101	1.9-2.3	52 ± 38	168 ± 46
0.6-2.8	5645 ± 163	5208.0 ± 154.1	0.6-2.3	1380 ± 74	1556 ± 76

Table 8.4: We show the expected signal yields from Monte Carlo and the fitted signal yields from data for our analysis and *BABAR*'s previous analysis. Errors are statistical only. Input of the expected signal yields from Monte Carlo for *BABAR* 2004 analysis are from reference [13].

We find the same difference between data and Monte Carlo in the *BABAR* 2004 results [15] as well and we include their results in Table 8.4. In their analysis, they also used the sum-of-exclusive method, but as it was based on a smaller data sample (89M $B\overline{B}$ pairs), they only reconstructed 12 $b \to s\gamma$ decay modes and covers a narrower hadronic mass range (0.6-2.3 GeV/c^2) as shown in the table. They also studied the *CP* asymmetries in four M_{X_s} sub-regions, with similar M_{X_s} binning in our analysis. In their results, we see that for the two bins of $1.5 < M_{Xs} < 1.9$ and $1.9 < M_{Xs} < 2.3$ GeV/c^2, the observed signal yields is about 2-3 σ away from the Monte Calor predictions.

This signal yields deviation between data and Monte Calor could be attributed to the fact that we have limited knowledge of the $b \to s\gamma$ decays in the higher M_{X_s} regions, while the $b \to K^*\ \gamma$ decays have been very well studied. This deviation can be resolved when we

- improve the photon spectrum for the $b \to s\gamma$ decays. The photon spectrum is directly related to the branching fraction for each M_{X_s} region. However it is not precisely known. We recall that for the non K^*-resonant region ($M_{Xs} > 1.1\,\mathrm{GeV}/c^2$), the signal Monte Carlo is generated with a flat photon energy (in the B rest frame) spectrum, and we choose to re-weigh it to the M_{X_s} distribution as determined in reference [17], which was produced based on 88M $B\overline{B}$ pairs (c.f. Section 3.2). A more precise photon spectrum will help us to determine the branching fraction in each M_{X_s} better. The situation will be improved in the future: *BABAR* has an on-going analysis to measure the photon spectrum and Belle has already presented a preliminary result on photon spectrum based on $605\,\mathrm{fb}^{-1}$ data [2] at Moriond EW 2008.

- better understand the fragmentation of the hadronic system. Like the photon spectrum, the X_s fragmentation in the higher M_{X_s} regions is not precisely known. The branching fractions of many decay modes in use for this analysis are so small that they are even measured. We will have a better understanding of the fragmentation when more branching fractions of $b \to s\gamma$ decays are available with more data accumulated.

A final note: this deviation on the signal yields will only affect the precision on the measured A_{CP}, but not likely to affect the measured mean value, as it is the relative ratio of the b and $\overline{b}$ events which matters.

8.4 Summary and Outlook

This document describes the measurement of the *CP* asymmetry in the $b \to s\gamma$ decays using 383×10^6 $B\overline{B}$ pairs recorded at the e^+e^- collider at *BABAR*. Signal event yields are extracted from the exclusive reconstruction of 16 final states. The flavor of the exclusive final states is determined either by the charge of the reconstructed final state (B^-/B^+) or the charge of the kaon ($\overline{B}^0/B^0$) in the X_s system. Continuum and $B\overline{B}$ backgrounds are mainly reduced by means of three boosted decision tree variable.

The $m_{\rm ES}$ distribution is fit to extract the signal yields separately for b and $\overline{b}$. With all the events in use, we measure

$$A_{CP} = -0.011 \pm 0.030_{stat} \pm 0.014_{syst} \tag{8.3}$$

for a photon energy threshold at 1.6 GeV and the hadronic system mass between 0.6 and 2.8 GeV/c^2.

This result is consistent with Standard Model. The breakdown of *CP* asymmetries in M_{X_s} categories shows that the *CP* asymmetry in $b \to s\gamma$ decays is independent of the photon energy (hence M_{X_s}) range in use.

The measurement presented here is based on 383M $B\overline{B}$ pairs. Although we have reduced the statistical error by 40%, there is still room to improve as more data is available. Unfortunately, *BABAR* ceased operating in April 2008, but additional data will be available from the BELLE experiment as they have already accumulated in excess of 600 million $B\overline{B}$ meson pairs, and will continue running in the future. In addition, a new generation flavor physics experiment, LHCb at the Large Hadron Collider, will start operating in late 2008.

With higher statistics, dedicated detector studies, and new signal extraction techniques, future statistical error on this variable will be comparable to the systematic uncertainties. By that time, we will be able to put more stringent constraints on physics contributions beyond the Standard Model.

Appendix A

Signal Truth-Matching and Cross-feed

"Truth-matching" a reconstructed B decay means that each particle used in the B reconstruction can be matched to a corresponding Monte Carlo generated particle, with all Monte Carlo particles coming from the decay of one B meson into the same decay mode identified in the reconstruction. Truth-matching is not perfect though, because sometimes the reconstructed daughters are not close enough to the generated Monte Carlo particle. In our analysis, this "strict truth-matching" often fails because of the high multiplicity of charged tracks in the final states.

We use a "pseudo truth-matching" criteria, where we loosen the truth-matching requirement to maintain a good efficiency, and make use of the hadronic mass (M_{X_s}) to select the corresponding B candidate.

BABAR's previous analysis [22] showed that with strict truth-matching rules, ΔM_{X_s} is strictly less than 50 MeV/c^2 (Figure A-1). We thus define signal as:

1. all the final-state particles from the B decay are truth-matched to particles generated in the Monte Carlo event, and all the corresponding generated particles coming from the same B.

2. the difference between the generated and reconstructed hadronic mass, $|\Delta M_{X_s}| < 0.050\,\mathrm{GeV}/c^2$. This distribution is expected to be zero in the truth-matched

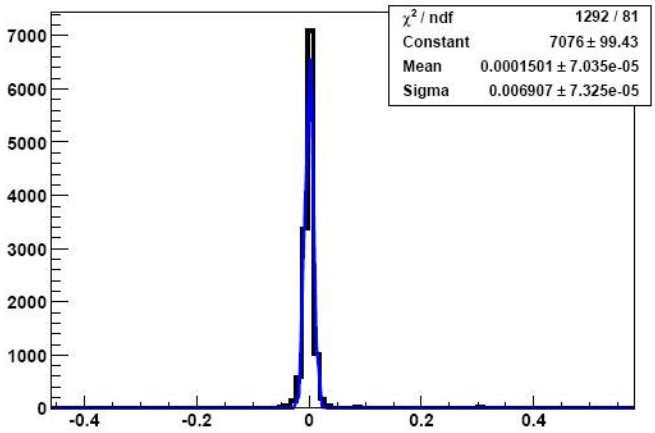


Figure A-1: ΔM_{X_s} distribution for strictly truth-matching signal Monte Carlo. This sample tells us which is our resolution on the hadronic mass [22].

event.

The first criteria does not require the reconstructed decay mode to match the generated one, and with this criteria it is likely to lose a π^0 or a charged pion in the final state. For example: a B meson decay truth-matched to Mode 3 ($B^0 \to K^+\pi^-$) might have been generated as Mode 9 ($B^0 \to K^+\pi^-\pi^0$). However, the second criteria removes this possibility.

For events in the K^* resonance region (where the true M_{X_s} lies below $1.1\,\mathrm{GeV}/c^2$, shown in Figure A-2) we find that even if events pass a truth-matching requirement there are still a few events that lie beyond $50\,\mathrm{MeV}/c^2$ from zero. A cut at 50 MeV/c^2 removes them.

For events in the higher M_{X_s} region (where the true X_s mass lies above $1.1\,\mathrm{GeV}/c^2$), we find that the peak in the truth-matched events is somewhat broader and has a more significant tail than in the K^* region. This can be attributed to the fact that the many-body final states are more likely to be missing a slow neutral particle than for the two-body states in the K^* region.

"Cross-feed" is then defined as any signal Monte Carlo events that fail this requirement (either it fails to truth-match, or it has $|\Delta M_{X_s}| \geq 0.050\,\mathrm{GeV}/c^2$).

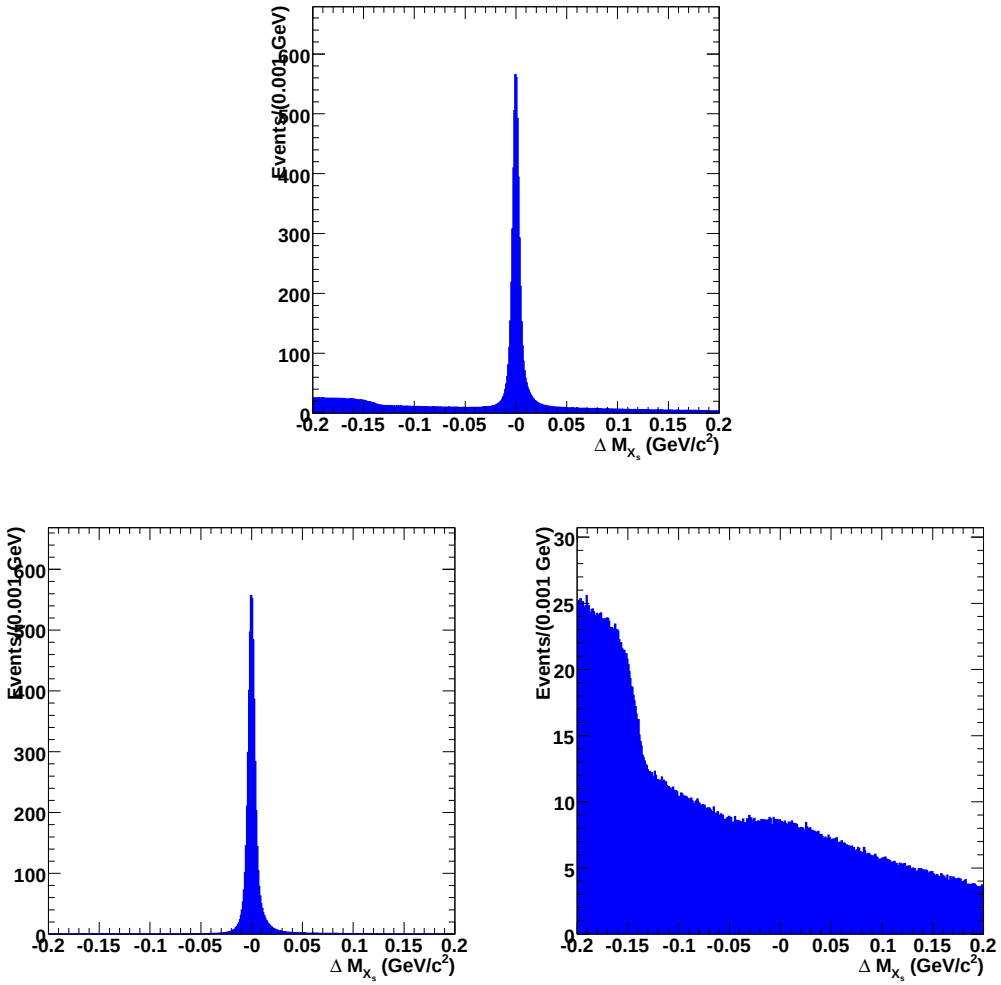


Figure A-2: ΔM_{X_s} distribution for signal Monte Carlo events where the true X_s mass is $< 1.1\,\mathrm{GeV}/c^2$. The top plot is for all events, the bottom-left plot is for events which pass our loose truth-match requirement, and the bottom-right is for events which fail the loose truth-matching requirement.

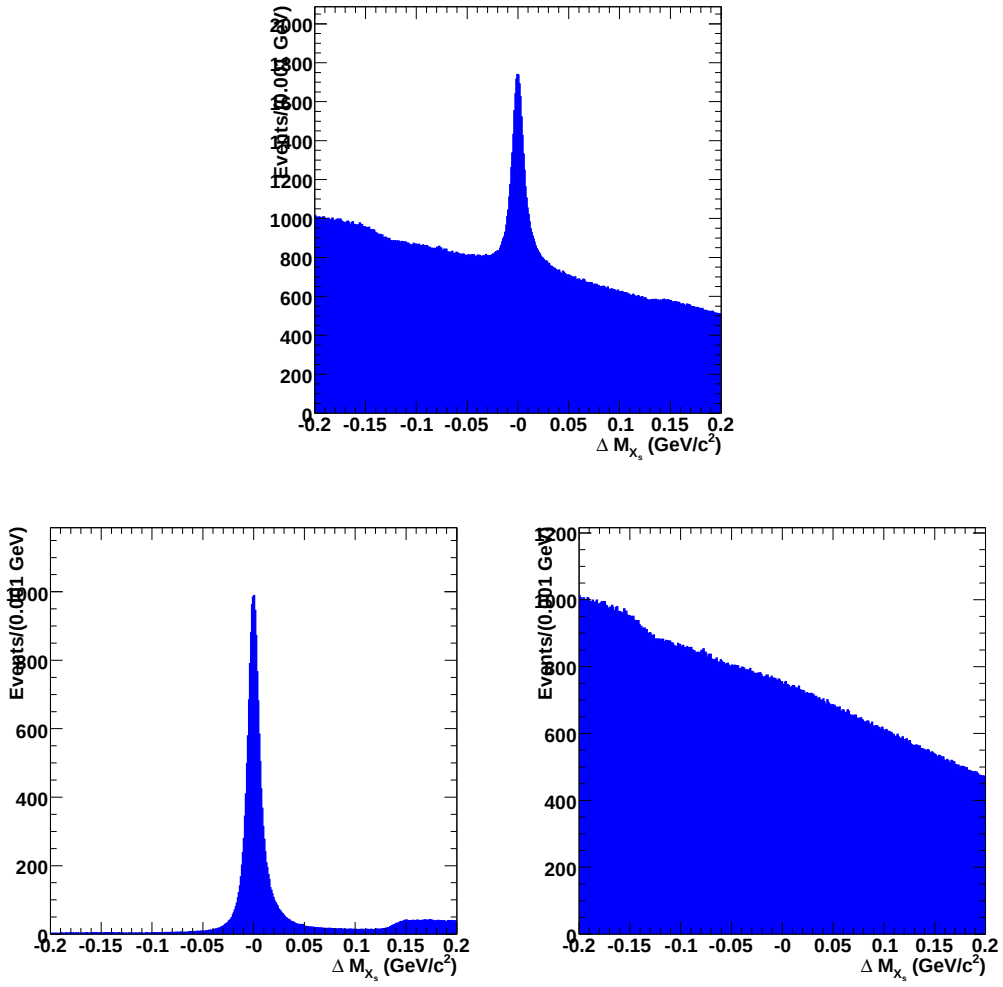


Figure A-3: ΔM_{X_s} distribution for signal Monte Carlo events where the true X_s mass is $\geq 1.1\,\mathrm{GeV}/c^2$. The top plot is for all events, the bottom-left plot is for events which pass our loose truth-match requirement, and the bottom-right is for events which fail the loose truth-matching requirement.

Appendix B

Boosted Decision Trees

A decision tree is a sequence of cuts which are applied to discriminate between classes of events. The tree has different branches, which correspond to applying cuts in different ways to select these classes. For instance, a branch that ends in a sequence that selects 90% of the signal while rejecting 95% of the background would obviously best classify events as signal.

As with a neural network, the key to using a decision tree is training it to distinguish the various event classes. We choose to use a method known as "AdaBoost" [24]. This algorithm creates a series of decision trees which, once trained, have their outputs used in a weighted "vote" on the classification of a given event. Events which are signal-like will have a weighted boosted decision tree (BDT) output ideally close to 1, and those in the background class will have a weight output ideally close to 0 (as in a neural network). Instead of using significance, the algorithm is trained to correctly classify events in each category. The figure-of-merit (FOM) is known as the "Gini Index", and describes the fraction of correctly classified events. It is related to a quantity called "exponential loss", which is to be minimized in the AdaBoost algorithm. Events which are incorrectly assigned to a node in a decision tree designed to select the target class are "boosted", or given more weight in the next training cycle. The algorithm spends more and more time on the hardest-to-classify events.

We use the `StatPatternRecognition` package developed by Ilya Narsky [36] to develop our BDTs. We train and test the three BDTs used in this analysis using a

similar procedure: we split a fraction of the available samples (signal Monte Carlo and continuum Monte Carlo) into a training and a testing sample. Training is performed by varying the two parameters of the BDT, running over the training sample and computing the FOM on the testing sample. An optimal training is one for which the FOM changes rapidly at first, and then stabilizes over time. If the FOM rises again, or if it never stabilizes (continues decreasing), the algorithm is being overtrained or is trained at a suboptimal point.

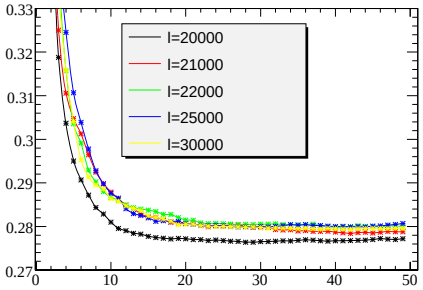
The two parameters which can be controlled are the training cycles and the minimum number of events which can be allowed in a terminal node of the tree. A terminal node is one which has no further cuts below it, and thus represents a category into which events can fall. By allowing fewer and fewer events to populate a terminal node, one can define more and more small categories into which events can fall. However, too many categories will create a situation where overtraining occurs, causing the algorithm to learn too much about the statistical fluctuations of the training sample (and thus causing the FOM on the testing sample to rise after a period of falling). If the terminal nodes have too large a minimum, then a stable separation of signal and background cannot be achieved.

We will briefly describe some of the details of the training of the π^0 and η vetos, and the continuum BDT.

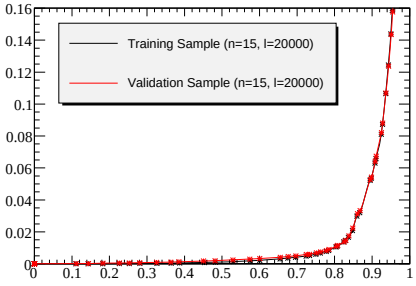
B.1 π^0 and η Veto Boosted Decision Trees

The inputs to these vetos were described in Section 4.3.3. The training sample is composed of off-peak continuum Monte Carlo events where the high-energy photon in the $b \to s\gamma$ final state is from a π^0 or η decay. The signal events are taken from signal Monte Carlo, and correspond to events where the photon truth-matches to a photon from B decay. We are interested in rejecting events where the high-energy photon arises from either π^0 or η decay.

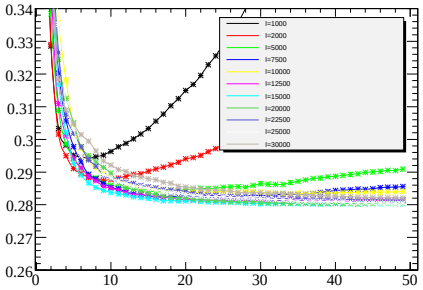
In Figure B-1, we show how the exponential loss varies as a function of both terminal node minima and training cycles. Based on the stability of the curves, we



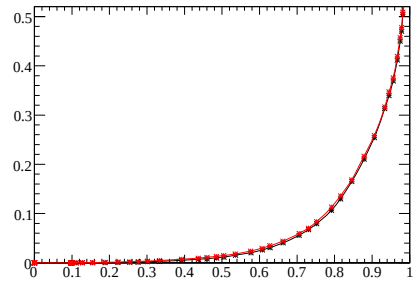
(a) Exponential Loss vs. Training Cycle for the π^0 veto



(b) Background Efficiency vs. Signal Efficiency for the the training sample (black) and the testing sample (red) in the π^0 veto



(c) Exponential Loss vs. Training Cycle for the η veto



(d) Background Efficiency vs. Signal Efficiency for the the training sample (black) and the testing sample (red) in the η veto

Figure B-1: : The exponential loss vs. training cycle for the π^0 veto (top) and η veto (bottom). On the right are the bkg. vs. signal efficiency compared between the two samples.

find that 20,000 events in a minimum terminal node are appropriate for both vetoes, while the π^0 requires 15 cycles to reach a stable point and the η veto requires 25.

B.2 Continuum Boosted Decision Tree

We described the inputs and output of the continuum BDT in Section 5.1. We apply the procedure to train the BDT as outlined in the previous section. This time, we use all continuum Monte Carlo events and truth-matched signal Monte Carlo events. The exponential loss vs. training cycle is shown in Figure B-2. We find that 50 cycles

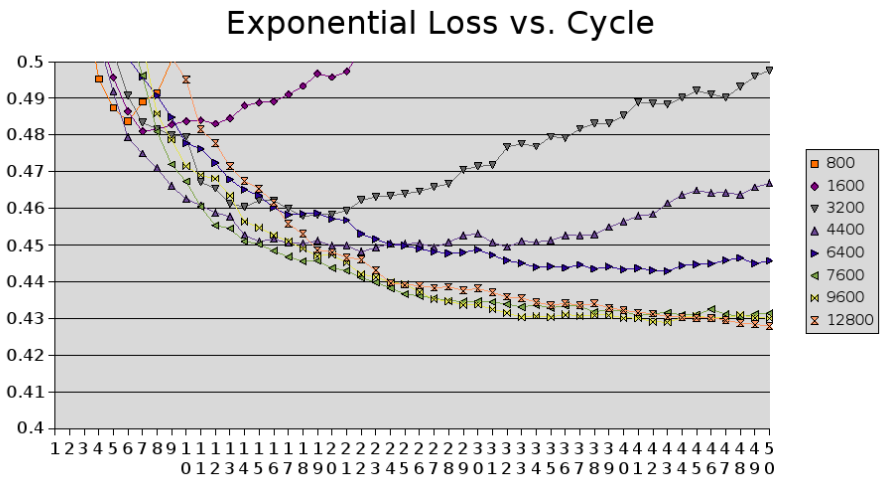


Figure B-2: : The exponential loss vs. training cycle for the continuum veto.

and a minimum number of events of 9600 in a terminal node is quite stable.

Appendix C

The PDFs

C.1 The Argus Function

The Argus [4] distribution is defined as:

$$\textbf{Argus}(m; m_0, c, p) = \frac{1}{N} \cdot m(1-(m/m_0)^2)^p \cdot \exp(c(1-(m/m_0)^2)) \cdot \theta(m < m_0) \quad \text{(C.1)}$$

where m_0 is the end-point of the distribution, p is the power parameter, and c describes the slope of the shape.

C.2 The Cruijff Function

$$\textbf{Cruijff}(m; m_0, \alpha_L, \sigma_L, \alpha_R, \sigma_R) = \frac{1}{N} \begin{cases} \exp[-(m-m_0)^2/(2\sigma_L^2 + \alpha_L(m-m_0)^2)], & m < m_0; \\ \exp[-(m-m_0)^2/(2\sigma_R^2 + \alpha_R(m-m_0)^2)], & m > m_0. \end{cases} \quad \text{(C.2)}$$

This distribution is similar to that of a Gaussian function, with m_0 being the mean of the function, but the widths on the two sides of the mean (σ_L and σ_R) are allowed to be different, and a tail (α_L or α_R) is attached to each side.

C.3 Signal PDFs

We first fitted the signal Monte Carlo events simultaneously in categories of flavors (b and $\overline{b}$) and M_{X_s} regions. All the parameters of the Cruijff function are floated. We also allow them to vary with M_{X_s} regions and flavors. The Cruijff function parameters are shown in Tables C.1-C.5. The results show that there is no difference between these PDFs in each category, so we decide to describe the signal shape using a single PDF after this study.

This first test also shows that if we keep the right tail parameter (α_R) floated, we will get negative values for them (Table C.5). Since negative tail parameters might cause the denominator in the exponentials of the Cruijff function to be zero at some value of m (Equation C.2), we decide to fix the right tail to be zero in our nominal signal PDFs.

	Cruijff Function Mean: m_0	
M_{X_s} Range	b	$\overline{b}$
0.6–1.1 GeV/c^2	$5.2809 \pm 3 \times 10^{-4}$	$5.2809 \pm 3 \times 10^{-4}$
1.1–1.5 GeV/c^2	$5.2809 \pm 3 \times 10^{-4}$	$5.2809 \pm 3 \times 10^{-4}$
1.5–2.0 GeV/c^2	$5.2808 \pm 4 \times 10^{-4}$	$5.2807 \pm 4 \times 10^{-4}$
2.0–2.8 GeV/c^2	$5.2808 \pm 9 \times 10^{-4}$	$5.2806 \pm 9 \times 10^{-4}$
0.6–2.8 GeV/c^2	$5.2809 \pm 1 \times 10^{-4}$	

Table C.1: The mean of the Cruijff function used to model the $m_{\rm ES}$ distribution in signal events as a function of flavor and M_{X_s} region

	Cruijff Function Left Width: σ_L	
M_{X_s} Range	b	$\overline{b}$
0.6–1.1 GeV/c^2	$(2.960 \pm 0.241) \times 10^{-3}$	$(2.974 \pm 0.239) \times 10^{-3}$
1.1–1.5 GeV/c^2	$(2.926 \pm 0.264) \times 10^{-3}$	$(2.963 \pm 0.263) \times 10^{-3}$
1.5–2.0 GeV/c^2	$(2.951 \pm 0.335) \times 10^{-3}$	$(2.861 \pm 0.326) \times 10^{-3}$
2.0–2.8 GeV/c^2	$(2.962 \pm 0.792) \times 10^{-3}$	$(2.797 \pm 0.793) \times 10^{-3}$
0.6–2.8 GeV/c^2	$(2.943 \pm 0.116) \times 10^{-3}$	

Table C.2: The left width of the Cruijff function used to model the $m_{\rm ES}$ distribution in signal events as a function of flavor and M_{X_s} region

	Cruijff Function Right Width: σ_R	
M_{X_s} Range	b	$\overline{b}$
0.6–1.1 GeV/c^2	$(2.639 \pm 0.275) \times 10^{-3}$	$(2.650 \pm 0.274) \times 10^{-3}$
1.1–1.5 GeV/c^2	$(2.600 \pm 0.303) \times 10^{-3}$	$(2.552 \pm 0.295) \times 10^{-3}$
1.5–2.0 GeV/c^2	$(2.636 \pm 0.371) \times 10^{-3}$	$(2.726 \pm 0.383) \times 10^{-3}$
2.0–2.8 GeV/c^2	$(2.681 \pm 0.820) \times 10^{-3}$	$(2.835 \pm 0.880) \times 10^{-3}$
0.6–2.8 GeV/c^2	$(2.634 \pm 0.138) \times 10^{-3}$	

Table C.3: The right width of the Cruijff function used to model the $m_{\rm ES}$ distribution in signal events as a function of flavor and M_{X_s} region

	Cruijff Function Left Tail Parameter: α_L	
M_{X_s}	b	$\overline{b}$
0.6–1.1 GeV/c^2	$(8.195 \pm 2.469) \times 10^{-2}$	$(7.80900 \pm 2.459) \times 10^{-2}$
1.1–1.5 GeV/c^2	$(8.641 \pm 2.705) \times 10^{-2}$	$(8.31558 \pm 2.698) \times 10^{-2}$
1.5–2.0 GeV/c^2	$(7.663 \pm 3.602) \times 10^{-2}$	$(8.64305 \pm 3.381) \times 10^{-2}$
2.0–2.8 GeV/c^2	$(7.524 \pm 8.688) \times 10^{-2}$	$(8.60299 \pm 8.354) \times 10^{-2}$
0.6–2.8 GeV/c^2	$(8.207 \pm 1.172) \times 10^{-2}$	

Table C.4: The left tail parameter of the Cruijff function used to model the $m_{\rm ES}$ distribution in signal events as a function of flavor and M_{X_s} region

	Cruijff Function Right Tail Parameter: α_R	
M_{X_s}	b	$\overline{b}$
0.6–1.1 GeV/c^2	$(-2.683 \pm 4.928) \times 10^{-2}$	$(-2.732 \pm 4.906) \times 10^{-2}$
1.1–1.5 GeV/c^2	$(-1.787 \pm 5.436) \times 10^{-2}$	$(-1.960 \pm 5.326) \times 10^{-2}$
1.5–2.0 GeV/c^2	$(-2.977 \pm 6.871) \times 10^{-2}$	$(-3.496 \pm 6.911) \times 10^{-2}$
2.0–2.8 GeV/c^2	$(-4.244 \pm 15.460) \times 10^{-2}$	$(-6.125 \pm 16.280) \times 10^{-2}$
0.6–2.8 GeV/c^2	$(-2.610 \pm 2.451) \times 10^{-2}$	

Table C.5: The right tail parameter of the Cruijff function used to model the $m_{\rm ES}$ distribution in signal events as a function of flavor and M_{X_s} region

Appendix D

Charge Asymmetry Studies

The detector efficiency is not symmetric for positive and negative charged tracks, thus could give rise to a false *CP* asymmetry. To estimate the false *CP* asymmetry we have, we study the charge asymmetry, which is defined to be $A_{cp} = (n_b - n_{\overline{b}})/(n_b + n_{\overline{b}})$ in several samples within our final selection. The samples are:

- $B \to X_s\gamma$ and $B \to X\pi^0$ Monte Carlo;
- $B \to X_s\gamma$ and $B \to X\pi^0$ on-resonance $m_{\rm ES}$ sideband data ($m_{\rm ES} < 5.27\,{\rm GeV}/c^2$);
- $B \to X_s\gamma$ off-resonance data.

We require the above $B \to X_s\gamma$ samples to pass all the cuts except for the $m_{\rm ES} > 5.26\,{\rm GeV}/c^2$ in the Table 5.2 - 5.4. For the $B \to X\pi^0$ samples, we remove the π^0 and η veto cuts.

D.1 Data/Monte Carlo Comparison For $B \to X_s\gamma$ Samples

We first compare the A_{CP} for the on-resonance data and the background Monte Carlo in the $m_{\rm ES}$ sideband region. We use a simple cut-and-count method to obtain the number of b or $\overline{b}$. Signal Monte Carlo is excluded from this study due to limited statistics. Table D.1 shows the result from this study. We find more $\overline{b}$ events than b

Samples	N_b	$N_{\overline{b}}$	A_{cp}
$B\overline{B}$ Monte Carlo	1302±31	1354±22	-1.97±1.45
cross-feed	1071±8	1105±6	-1.53±0.45
BB+cross-feed	2374±23	2459±23	-1.77±0.67
Continuum	1.071×10^4±157	1.11×10^4±112	-1.75±0.89
All Background	1.309×10^4±112	1.356×10^4±114	-1.76±0.60
Off-resonance	8993±416	8888±293	0.59±2.84
On-resonance	1.518×10^4±175	1.539×10^4±124	-0.67±0.70

Table D.1: Charge Asymmetry table for data and Monte Carlo in the $m_{\rm ES}$ sideband.

Sample	N_b	$N_{\overline{b}}$	$A_{CP}(\%)$
$B\overline{B}$ Monte Carlo	1308±31	1360±22	-1.96±1.46
cross-feed	1075±8	1108± 6	-1.52±0.45
BB+cross-feed	2383±23	2469±23	-1.77±0.67
Continuum	1.074×10^4±158	1.113×10^4±113	-1.76±0.89
All Background	1.313×10^4±113	1.36×10^4±115	-1.76±0.60

Table D.2: Charge asymmetry table for the Monte Carlo events in the $m_{\rm ES}$ sideband after correcting for the kaon efficiencies.

everywhere, which is consistent between all the samples. More importantly, we find a much smaller A_{CP} in the the on-resonance $m_{\rm ES}$ sideband data ($\sim -0.5\%$) than in all Monte Carlo samples, where A_{CP} is generally -1%–-2%. This is consistent with what has been found in the most recent *BABAR* $B \to K\pi$ analysis [19].

To confirm that the difference of A_{CP} in the data and Monte Carlo is not caused by the different kaon identification efficiencies between data and Monte Carlo, we study the kaon efficiencies corrected A_{CP}[1]. Table D.2 shows the charge asymmetry for the background Monte Carlo with kaon efficiencies corrections. Comparing it to Table D.1, the A_{CP} remains almost the same. From this study, it is clear that the kaon identification efficiency is not the major cause of the data/Monte Carlo difference. In section D.3, we study the A_{CP} in the separate sideband control sample and find the same difference between data and Monte Carlo.

Table D.3 shows the A_{CP} for all the Monte Carlo in the entire $m_{\rm ES}$ fit range

[1]At *BABAR*, the kaon identification efficiency for reconstructed tracks is measured with a control sample of kaons from D^* decays. We use the ratio of the data/Monte Carlo kaon efficiency measured with this control sample (SP-r18b/asData/run[1,2,3,4,5]-r18b/kaons[+,-]/k.Dstar.LH.Tight.ratio) to correct the Monte Carlo events.

Table D.3: Charge Asymmetry table for the entire $m_{\rm ES}$ range ($m_{\rm ES}$ = $[5.22, 5.29]\,{\rm GeV}/c^2$)

Sample	N_b	$N_{\overline{b}}$	$A_{CP}(\%)$
$B\overline{B}$ Monte Carlo	2515±43	2607±31	-1.78±1.05
cross-feed	2556±12	2619±9	-1.21±0.29
BB+cross-feed	5072±31	5225±32	-1.49±0.44
Continuum	1.518×10^4±187	1.568×10^4±133	-1.62±0.75
Background	2.025×10^4±135	2.091×10^4±137	-1.59±0.47
Signal	2776±11	2861±8	-1.50±0.24

($5.22 < m_{\rm ES} < 5.29\,{\rm GeV}/c^2$). This study is to check whether the A_{CP} is different between the entire $m_{\rm ES}$ region and the $m_{\rm ES}$ sideband region. Signal Monte Carlo is also included in this table for checking the consistency of A_{CP} between signal and background Monte Carlo. As shown in the table, the signal A_{CP} is $(-1.50 \pm 0.24)\%$, which is statistically consistent with that for the background: $(-1.60\pm0.47)\%$. All the A_{CP} calculated in this table are consistent with those for the $m_{\rm ES}$ sideband sample. We conclude from this study that there is no new significant source of A_{CP} in the signal region.

D.2 Detailed Study of A_{CP} For $B \to X_s\gamma$ Samples

We conclude from the previous studies that the charge asymmetry is different between the Monte Carlo and the data. Therefore we decide to use the $m_{\rm ES}$ sideband data to estimate the charge asymmetries. In this section, we perform detailed study to check whether the charge asymmetry in the sideband data has any dependance on parameters such as $K^\pm$ momentum, and decay modes etc. We calculate the A_{CP} for each M_{X_s} bin (Table D.4) and each mode (Table D.5) for the $m_{\rm ES}$ sideband data. The asymmetries in each M_{X_s} bin and in each mode are all statistically consistent with the average asymmetry for the $m_{\rm ES}$ sideband data.

We also calculate the A_{CP} dependance on charged kaon momentum for the $m_{\rm ES}$ sideband data and compare it with the background Monte Calor (Table D.7). we find all values agree within 1-2 σ. However, the results fluctuate around 0. To confirm

M_{X_s}	N_b	$N_{\overline{b}}$	$A_{CP}(\%)$
0.6-1.1	1889±61	1835±43	1.45±1.99
1.1-1.5	2026±65	2206±47	-4.25±1.92
1.5-2.0	3189±80	3178±56	0.17±1.53
2.0-2.8	2425±70	2509±50	-1.70±1.76
0.6-2.8	1.518×10^4±175	1.539×10^4±124	-0.67±0.70
$B^{\pm}$	8757±133	8887±94	-0.74±0.93
$B^0\overline{B}^0$	6428±114	6504±81	-0.59±1.08

Table D.4: Charge asymmetry in each M_{X_s} bin for $m_{\rm ES}$ sideband data.

Table D.5: Charge asymmetry for each mode for $m_{\rm ES}$ sideband data.

Mode	N_b	$N_{\overline{b}}$	$A_{CP}(\%)$
1	660±36	630±25	2.33±3.37
2	964±44	953±31	0.57±2.79
3	1415±54	1457±38	-1.46±2.30
4	0±0	0±0	0±0
5	2021±64	2050±45	-0.71±1.93
6	885±43	945±31	-3.28±2.91
7	738±38	733±27	0.34±3.19
8	0±0	0±0	0±0
9	2069±64	2036±45	0.80±1.90
10	0±0	0±0	0±0
11	570±33	548±23	1.97±3.63
12	2017±64	2112±46	-2.30±1.93
13	589±34	599±24	-0.84±3.57
14	1006±45	1048±32	-2.05±2.73
15	0±0	0±0	0±0
16	1158±48	1187±34	-1.24±2.55
23	304±24	307±18	-0.49±4.97
27	776±39	770±28	0.39±3.11
33	9±4	10±3	-5.26±28.84
37	4±3	6±2	-20.2±42.7

$P_{K^\pm}$(GeV/c)	Data (%)	Bkg MC A_{CP}(%)
0.8-1.0	-1.53±2.18	-0.88±1.81
1.0-1.2	-1.61±2.24	-2.65±1.87
1.2-1.4	1.53±2.24	-3.71±1.94
1.4-1.6	-0.79±2.38	-1.45±2.04
1.6-1.8	0±2.48	-2.49±2.12
1.8-2.0	-3.20±2.68	0.29±2.24
2.0-2.2	-1.99±2.87	0.36±2.46
2.2-2.4	4.52±3.02	-1.34±2.67
2.4-2.6	-5.59±3.49	-3.13±2.96
2.6-2.8	0.10±3.79	-2.52±3.34
2.8-3.0	-1.44±4.27	-4.16±3.74
3.0-4.0	-0.16±2.82	-2.43±2.64
4.0-5.0	4.55±9.02	-6.34±10.1
0.8-2.0	-0.87±0.96	-1.88±0.81
2.0-5.0	-0.52±1.32	-1.93±1.17
0.8-5.0	-0.75±0.78	-1.90±0.67

Table D.6: Charge asymmetry dependance on $K^\pm$ momentum for $m_{\rm ES}$ sideband data and background Monte Carlo. Signal Monte Carlo is excluded due to limited statistics in the region.

this is only statistical effect, we calculate the A_{CP} in each $K^\pm$ momentum bin for the signal Monte Carlo for $m_{\rm ES}$ between 5.22-5.29 GeV/c^2. We expect the statistical effect to disappear in the big signal Monte Carlo[1] sample. The results are listed in Table D.7. We also included in this table the A_{CP} calculated for background Monte Carlo in the full $m_{\rm ES}$ range to be compared with the signal Monte Carlo. Most A_{CP} in the background and signal sample lie below 0, which confirms our prediction that the fluctuation is caused by the statistical effect.

Finally, the Table D.8 lists the A_{CP} for Charge asymmetry dependance on charged tracks multiplicity. Again all the results are 0 within statistics errors, and we do not see any obvious dependance of A_{CP} on these parameters.

[1]The signal Monte Carlo is generated with ratio of $\sim 100:1$ to the on-resonance data luminosity while $B\overline{B}$ MC is available in amount of about a few times the equivalent integrated data luminosity.

$P_{K^\pm}$(GeV/c)	Signal MC A_{CP}(%)	Bkg MC A_{CP}(%)
0.8-1.0	-2.67±0.92	-1.37±1.43
1.0-1.2	-1.85±0.82	-2.90±1.46
1.2-1.4	-2.02±0.75	-2.64±1.51
1.4-1.6	-1.59±0.73	-0.17±1.57
1.6-1.8	-1.12±0.72	-0.80±1.63
1.8-2.0	-3.04±0.77	1.47±1.73
2.0-2.2	-1.64±0.82	-1.80±1.88
2.2-2.4	-2.08±0.93	-3.30±2.06
2.4-2.6	-2.88±1.06	-3.69±2.27
2.6-2.8	0.37±1.22	-3.00±2.55
2.8-3.0	-2.46±1.53	-2.83±2.81
3.0-4.0	-1.02±1.43	-1.03±2.02
4.0-5.0	-1.93±20.2	-9.88±9.12
0.8-2.0	-2.02±0.32	-1.22±0.63
2.0-5.0	-1.70±0.44	-2.55±0.89
0.8-5.0	-1.90±0.26	-1.66±0.52

Table D.7: Charge asymmetry dependance on $K^\pm$ momentum for signal and background Monte Carlo for the entire $m_{\rm ES}$ range.

Multiplicity	N_b	$N_{\overline{b}}$	A_{CP}(%)
1	2006±63	1993±45	0.33±1.93
2	5418±104	5450±74	-0.29±1.18
3	6181±112	6346±80	-1.32±1.10
> 3	1576±56	1596±40	-0.63±2.18

Table D.8: Charge asymmetry dependance on charged tracks multiplicity for $m_{\rm ES}$ sideband data.

D.3 Detailed Study of A_{CP} For $B \to X\pi^0$ Samples

We proceed to study the detector-related sources of false *CP* asymmetry in the $B \to X\pi^0$ control sample. We use these events as a cross-check of the A_{CP} results from the $B \to X_s\gamma$ samples.

The results of A_{CP} in the sideband are shown in Table D.9. It's clear that for the entire X_s mass region, the A_{CP} is consistent with the results we obtained in the $b \to s\gamma$ analysis. We also find the Monte Carlo is over predicting the charge asymmetry.

We split the sample in the sideband into those reconstructed as B^0 and B^+ (charge

M_{X_s} Selection	Monte Carlo A_{CP} (%)	Data A_{CP} (%)
0.6-1.1	-2.06 ± 1.74	1.69 ± 1.69
1.1-1.5	-1.83 ± 1.77	-1.89 ± 1.70
1.5-2.0	-0.33 ± 1.27	-3.29 ± 1.23
2.0-2.8	-2.27 ± 0.65	-0.68 ± 0.66
0.6-2.8	-1.88 ± 0.53	-1.04 ± 0.52
B Sample	Monte Carlo A_{CP} (%)	Data A_{CP} (%)
B^0	-1.95 ± 0.70	-1.08 ± 0.69
B^+	-1.79 ± 0.80	-0.98 ± 0.79

Table D.9: Detector-related fake A_{CP} in the $B \to X\pi^0$ $m_{\rm ES}$ sideband.

conjugation is implied here). We find that there is no significant difference in the value of A_{CP} in either of these two sub-samples. Based on this, we determine that the detector-related *CP* asymmetry is the same in Bs reconstructed either as charged or neutral. We can therefore subtract the same fake A_{CP} from both to determine the true A_{CP} in the $B \to X_s\gamma$ signal events.

We also split the events by the momentum of the reconstructed final-state charged kaon (for those modes with such a kaon). We find no evidence for a strong dependence of the detector A_{CP} on the kaon momentum, given the statistics of this sample (Table D.10).

Finally, we divide the events into categories based on their final-state multiplicities (Table D.11). Subdividing further by final state shows a weak pattern of improvement (return toward zero asymmetry) as more tracks appear in the final state, but the statistical uncertainty in both Monte Carlo and data obscures the true pattern. The Monte Carlo again appears to over-predict the level of detector-induced A_{CP} as a function of multiplicity, as it has in other categories.

D.4 $m_{\rm ES}$ Sideband Data re-weighting

We conclude from the above study that the charge asymmetry is different between data and Monte Carlo, with the Monte Carlo generally presents a bigger asymmetry than the $m_{\rm ES}$ sideband data. We break the sideband data into categories ($K^\pm$

K^+ Momentum Range (GeV/c)	Monte Carlo A_{CP} (%)	Data A_{CP} (%)
0.8-1.0	-4.78 ± 1.56	-1.42 ± 1.54
1.0-1.2	-1.25 ± 1.65	-2.63 ± 1.62
1.2-1.4	-2.12 ± 1.69	2.93 ± 1.69
1.4-1.6	-1.94 ± 1.77	-1.96 ± 1.75
1.6-1.8	-1.31 ± 1.84	-1.05 ± 1.81
1.8-2.0	-2.84 ± 1.92	-2.14 ± 1.94
2.0-2.2	-0.59 ± 2.16	1.33 ± 2.14
2.2-2.4	-1.24 ± 2.39	-4.03 ± 2.35
2.4-2.6	-2.26 ± 2.69	-1.55 ± 2.65
2.6-2.8	-3.52 ± 3.05	-3.43 ± 3.04
2.8-3.0	0.44 ± 3.31	-3.51 ± 3.31
3.0-4.0	-1.96 ± 2.17	-0.39 ± 2.08
4.0-5.0	-1.83 ± 7.18	-2.56 ± 6.54
0.8-2.0	-2.45 ± 0.70	-1.02 ± 0.70
2.0-5.0	-1.51 ± 1.02	-1.51 ± 1.00

Table D.10: Detector-related fake A_{CP} as a function of the momentum range of a charged kaon in the final state of the B.

Number of Tracks	Monte Carlo A_{CP} (%)	Data A_{CP} (%)
1	-3.30 ± 1.58	-2.29 ± 1.54
2	-2.02 ± 0.89	-0.86 ± 0.88
3	-1.55 ± 0.81	-1.04 ± 0.82
> 3	-1.33 ± 1.49	-0.37 ± 1.47

Table D.11: Detector-related fake A_{CP} as a function of the charged-track multiplicity of the final state. A final state with a K^0_s candidate has at least two prongs, by this definition.

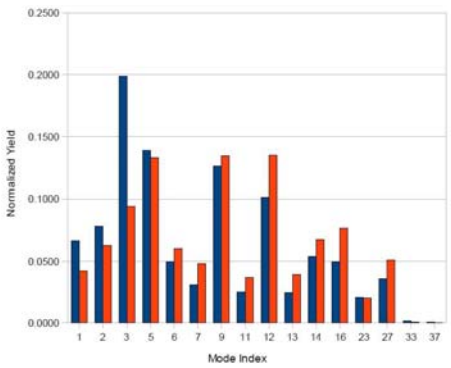


Figure D-1: The normalized yields (sum of b and $\overline{b}$ events) in each of the 16 modes for the $m_{\rm ES}$ signal region (blue) and for the $m_{\rm ES}$ sideband (red). The total yield of the 16 modes for each region is 1.

momentum, decay modes, and track multiplicity, etc), and on obvious charge asymmetry dependence is found. We decide to use the $m_{\rm ES}$ sideband data to measure the detector induced CP asymmetry in each M_{X_s} region.

Before proceeding to calculate A_{det}, we need to verify that the the X_s fragmentation in the $m_{\rm ES}$ sideband data is the same as it in the $m_{\rm ES}$ signal region ($m_{\rm ES} > 5.27\,{\rm GeV}/c^2$). We check the yield of each decay mode for both regions (Figure D-1). They are normalized in the way that the total yields in both regions are the same (1, in our case). We find that there is a small difference between the two regions: the $m_{\rm ES}$ signal region data favors the first four decay modes, especially Mode 3, $B^0 \to K^+\pi^-\gamma$.

We correct for that difference by re-weighting the events in the $m_{\rm ES}$ sideband region by $h_i^{signal}/h_i^{sideband}$, where i indicates the event decay mode, and $h_i^{signal}(h_i^{sideband})$ is normalized yield of mode i for $m_{\rm ES}$ signal (sideband) region. The charge asymmetry in the $m_{\rm ES}$ sideband data after the above mentioned reweighting is: $(-0.59 \pm 0.62)\%$ for the entire M_{X_s} region, it is consistent with the result before the re-weighting: (-0.67±0.70)% The $B \to X\pi^0$ sideband data is also re-weighted to match the $B \to X_s\gamma$ signal region fragmentation.

Bibliography

[1] L. F. Abbott, P. Sikivie, and Mark B. Wise. Constraints on charged-higgs-boson couplings. *Phys. Rev. D*, 21(5):1393–1403, Mar 1980.

[2] K. Abe et al. Improved Measurement of Inclusive Radiative B-meson decays. 2008.

[3] Mayumi Aoki, Gi-Chol Cho, and Noriyuki Oshimo. Decay rate asymmetry in $b \to s\gamma$ as a signature of supersymmetry. *Phys. Rev. D*, 60(3):035004, Jun 1999.

[4] ARGUS Collaboration, Albrecht, H. *el al.* Reconstruction of B Mesons. *Phys. Lett.*, B185:218, 1987.

[5] Myron Bander, D. Silverman, and A. Soni. CP Noninvariance in the Decays of Heavy Charged Quark Systems. *Phys. Rev. Lett.*, 43(4):242–245, Jul 1979.

[6] Thomas Becher and Matthias Neubert. Analysis of $\mathcal{B}(B \to X_s\gamma)$ at NNLO with a Cut on Photon Energy. *Physical Review Letters*, 98:022003, 2007.

[7] BELLE Collaboration, Nishida, S. and others. Measurement of the CP asymmetry in $B \to X_s\gamma$. *Phys. Rev. Lett.*, 93:031803, 2004.

[8] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. "weak decays beyond leading logarithms". *Rev. Mod. Phys.*, 68(4):1125–1244, Oct 1996.

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the $k2$ meson. *Phys. Rev. Lett.*, 13(4):138–140, Jul 1964.

[10] Chun-Khiang Chua, Xiao-Gang He, and Wei-Shu Hou. CP violating $b \to s\gamma$ decay in supersymmetric models. *Phys. Rev. D*, 60(1):014003, May 1999.

[11] Eung Jin Chun, Kyuwan Hwang, and Jae Sik Lee. CP asymmetries in radiative B decays with R-parity violation. *Phys. Rev. D*, 62(7):076006, Sep 2000.

[12] CLEO Collaboration, Coan, T. E. and others. Bounds on the CP asymmetry in $b \to s\gamma$ decays. *Phys. Rev. Lett.*, 86:5661–5665, 2001.

[13] *BABAR* Collaboration, Bard, D. J. ; Beck, T. W. ; Behera, P. K. ; Di Lodovico, F. ; Eichenbaum, A. M. ; Playfer, S. ; Pulliam, T. M. Direct CP asymmtry in $b \to s\gamma$ decays. *BABAR Analysis Document 611*, 2004.

[14] *BABAR* Collaboration, B. Aubert *et al.* *Nucl. Instrum. Methods*, A479:1–116, 2002.

[15] *BABAR* Collaboration, B. Aubert *et al.* Measurement of the direct CP asymmetry in $b \to s\gamma$ decays. *Phys. Rev. Lett.*, 93:021804, 2004.

[16] *BABAR* Collaboration, B. Aubert *et al.* A b flavor tagging algorithm for *CP* violation measurements at the babar experiment. *BABAR Analysis Document 1025*, D72, 2005.

[17] *BABAR* Collaboration, B. Aubert *et al.* Measurements of the $B \to X_s\gamma$ branching fraction and photon spectrum from a sum of exclusive final states. *Phys. Rev.*, D72:052004, 2005.

[18] *BABAR* Collaboration, B. Aubert *et al.* Measurement of the branching fraction and photon energy moments of $B \to X_s\gamma$ and $A_{CP}(B \to X_{s+d}\gamma)$. *Phys. Rev. Lett.*, 97:171803, 2006.

[19] *BABAR* Collaboration, Biesiada, J. ; Dallapiccola, C. ; Li, X. ; Olsen, J. ; Telnov, A.V. Measurement of CP Violating Asymmetries in B0 → pi+pi- and K+pi-: Run 5 update. (Observation of CP Violation in B0 → K+pi- and pi+pi-.). *BABAR Analysis Document 1551*, 2006-2007.

[20] *BABAR* Collaboration, Harrison, P. F. and Quinn, Helen R. The BaBar physics book: Physics at an asymmetric B factory. Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[21] *BABAR* Collaboration, Mclachlin, Sheila. Pid control samples at *BABAR*. *BABAR Analysis Document 1056*, 2004.

[22] Andrew Eichenbaum. Direct CP Violation in $b \to s\gamma$ Decays. *Ph.D thesis University of Wisconsin - Madison*, 2004.

[23] Geoffrey C. Fox and Stephen Wolfram. Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes. *Phys. Rev. Lett.*, 41:1581, 1978.

[24] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139, 1997.

[25] GEANT4 Collaboration, S. Agostinelli *et al.* GEANT4: A simulation toolkit. *Nucl. Instrum. Methods*, A506:250–303, 2003.

[26] Toru Goto et al. Effect of supersymmetric CP phases on the $B \to X_s\gamma$ and $B \to X_s l^+ l^-$ decays in the minimal supergravity model. *Phys. Lett. B*, 460:333, 1999.

[27] Heavy Flavor Averaging Group (HFAG) Barberio, E. and others. Averages of B-hadron properties at the end of 2005. 2006.

[28] Tobias Hurth, Enrico Lunghi, and Werner Porod. Untagged $B \to X_{s+d}\gamma$ *CP* asymmetry as a probe for new physics. *Nucl. Phys.*, B704:56–74, 2005.

[29] A. L. Kagan and M. Neubert. QCD Anatomy of $B \to X_s\gamma$ Decays. 1998.

[30] Alexander L. Kagan and Matthias Neubert. Direct CP violation in $B \to X_s\gamma$ decays as a signature of new physics. *Phys. Rev.*, D58:094012, 1998.

[31] Ken Kiers, Amarjit Soni, and Guo-Hong Wu. Direct cp violation in radiative b decays in and beyond the standard model. *Phys. Rev. D*, 62(11):116004, Nov 2000.

[32] Yeong Gyun Kim, P. Ko, and Jae Sik Lee. Possible new physics signals in $b \to s\gamma$ and $b \to sl^+l^-$. *Nucl. Phys.*, B544:64–88, 1999.

[33] D.J̃. Lange. *Nucl. Instrum. Methods*, A462:152–155, 2001.

[34] Stephen P. Martin. A Supersymmetry Primer. 1997.

[35] M. Misiak et al. The first estimate of B(anti-B $\to X_s\gamma$) at $\mathcal{O}(\alpha_s^2)$. *Phys. Rev. Lett.*, 98:022002, 2007.

[36] I. Narsky and J. Bunn. Statpatternrecognition: A c++ package for multivariate classification of high energy physics data. *Proceedings of Computing for High Energy Physics, Mumbai, India, February 13-17*, 2006.

[37] Particle Data Group, Yao, W. M. and others. Review of particle physics. *J. Phys.*, G33:1–1232, 2006.

[38] Torbjorn Sjostrand. High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4:Physics and manual. *Computer Phys. Commun.*, 82:74, 1994.

[39] Kenneth G. Wilson. Non-lagrangian models of current algebra. *Phys. Rev.*, 179(5):1499–1512, Mar 1969.

[40] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51(21):1945–1947, Nov 1983.

Bibliography

[1] L. F. Abbott, P. Sikivie, and Mark B. Wise. Constraints on charged-higgs-boson couplings. *Phys. Rev. D*, 21(5):1393–1403, Mar 1980.

[2] K. Abe et al. Improved Measurement of Inclusive Radiative B-meson decays. 2008.

[3] Mayumi Aoki, Gi-Chol Cho, and Noriyuki Oshimo. Decay rate asymmetry in $b \to s\gamma$ as a signature of supersymmetry. *Phys. Rev. D*, 60(3):035004, Jun 1999.

[4] ARGUS Collaboration, Albrecht, H. *el al.* Reconstruction of B Mesons. *Phys. Lett.*, B185:218, 1987.

[5] Myron Bander, D. Silverman, and A. Soni. CP Noninvariance in the Decays of Heavy Charged Quark Systems. *Phys. Rev. Lett.*, 43(4):242–245, Jul 1979.

[6] Thomas Becher and Matthias Neubert. Analysis of $\mathcal{B}(B \to X_s\gamma)$ at NNLO with a Cut on Photon Energy. *Physical Review Letters*, 98:022003, 2007.

[7] BELLE Collaboration, Nishida, S. and others. Measurement of the CP asymmetry in $B \to X_s\gamma$. *Phys. Rev. Lett.*, 93:031803, 2004.

[8] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. "weak decays beyond leading logarithms". *Rev. Mod. Phys.*, 68(4):1125–1244, Oct 1996.

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the $k2$ meson. *Phys. Rev. Lett.*, 13(4):138–140, Jul 1964.

[10] Chun-Khiang Chua, Xiao-Gang He, and Wei-Shu Hou. CP violating $b \to s\gamma$ decay in supersymmetric models. *Phys. Rev. D*, 60(1):014003, May 1999.

[11] Eung Jin Chun, Kyuwan Hwang, and Jae Sik Lee. CP asymmetries in radiative B decays with R-parity violation. *Phys. Rev. D*, 62(7):076006, Sep 2000.

[12] CLEO Collaboration, Coan, T. E. and others. Bounds on the CP asymmetry in $b \to s\gamma$ decays. *Phys. Rev. Lett.*, 86:5661–5665, 2001.

[13] *BABAR* Collaboration, Bard, D. J. ; Beck, T. W. ; Behera, P. K. ; Di Lodovico, F. ; Eichenbaum, A. M. ; Playfer, S. ; Pulliam, T. M. Direct CP asymmtry in $b \to s\gamma$ decays. *BABAR Analysis Document 611*, 2004.

[14] *BABAR* Collaboration, B. Aubert *et al.* *Nucl. Instrum. Methods*, A479:1–116, 2002.

[15] *BABAR* Collaboration, B. Aubert *et al.* Measurement of the direct CP asymmetry in $b \to s\gamma$ decays. *Phys. Rev. Lett.*, 93:021804, 2004.

[16] *BABAR* Collaboration, B. Aubert *et al.* A b flavor tagging algorithm for *CP* violation measurements at the babar experiment. *BABAR Analysis Document 1025*, D72, 2005.

[17] *BABAR* Collaboration, B. Aubert *et al.* Measurements of the $B \to X_s\gamma$ branching fraction and photon spectrum from a sum of exclusive final states. *Phys. Rev.*, D72:052004, 2005.

[18] *BABAR* Collaboration, B. Aubert *et al.* Measurement of the branching fraction and photon energy moments of $B \to X_s\gamma$ and $A_{CP}(B \to X_{s+d}\gamma)$. *Phys. Rev. Lett.*, 97:171803, 2006.

[19] *BABAR* Collaboration, Biesiada, J. ; Dallapiccola, C. ; Li, X. ; Olsen, J. ; Telnov, A.V. Measurement of CP Violating Asymmetries in B0 → pi+pi- and K+pi-: Run 5 update. (Observation of CP Violation in B0 → K+pi- and pi+pi-.). *BABAR Analysis Document 1551*, 2006-2007.

[20] *BABAR* Collaboration, Harrison, P. F. and Quinn, Helen R. The BaBar physics book: Physics at an asymmetric B factory. Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[21] *BABAR* Collaboration, Mclachlin, Sheila. Pid control samples at *BABAR*. *BABAR Analysis Document 1056*, 2004.

[22] Andrew Eichenbaum. Direct CP Violation in $b \to s\gamma$ Decays. *Ph.D thesis University of Wisconsin - Madison*, 2004.

[23] Geoffrey C. Fox and Stephen Wolfram. Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes. *Phys. Rev. Lett.*, 41:1581, 1978.

[24] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139, 1997.

[25] GEANT4 Collaboration, S. Agostinelli *et al.* GEANT4: A simulation toolkit. *Nucl. Instrum. Methods*, A506:250–303, 2003.

[26] Toru Goto et al. Effect of supersymmetric CP phases on the $B \to X_s\gamma$ and $B \to X_s l^+ l^-$ decays in the minimal supergravity model. *Phys. Lett. B*, 460:333, 1999.

[27] Heavy Flavor Averaging Group (HFAG) Barberio, E. and others. Averages of B-hadron properties at the end of 2005. 2006.

[28] Tobias Hurth, Enrico Lunghi, and Werner Porod. Untagged $B \to X_{s+d}\gamma$ *CP* asymmetry as a probe for new physics. *Nucl. Phys.*, B704:56–74, 2005.

[29] A. L. Kagan and M. Neubert. QCD Anatomy of $B \to X_s\gamma$ Decays. 1998.

[30] Alexander L. Kagan and Matthias Neubert. Direct CP violation in $B \to X_s\gamma$ decays as a signature of new physics. *Phys. Rev.*, D58:094012, 1998.

[31] Ken Kiers, Amarjit Soni, and Guo-Hong Wu. Direct cp violation in radiative b decays in and beyond the standard model. *Phys. Rev. D*, 62(11):116004, Nov 2000.

[32] Yeong Gyun Kim, P. Ko, and Jae Sik Lee. Possible new physics signals in $b \to s\gamma$ and $b \to sl^+l^-$. *Nucl. Phys.*, B544:64–88, 1999.

[33] D.J̃. Lange. *Nucl. Instrum. Methods*, A462:152–155, 2001.

[34] Stephen P. Martin. A Supersymmetry Primer. 1997.

[35] M. Misiak et al. The first estimate of B(anti-B $\to X_s\gamma$) at $\mathcal{O}(\alpha_s^2)$. *Phys. Rev. Lett.*, 98:022002, 2007.

[36] I. Narsky and J. Bunn. Statpatternrecognition: A c++ package for multivariate classification of high energy physics data. *Proceedings of Computing for High Energy Physics, Mumbai, India, February 13-17*, 2006.

[37] Particle Data Group, Yao, W. M. and others. Review of particle physics. *J. Phys.*, G33:1–1232, 2006.

[38] Torbjorn Sjostrand. High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4:Physics and manual. *Computer Phys. Commun.*, 82:74, 1994.

[39] Kenneth G. Wilson. Non-lagrangian models of current algebra. *Phys. Rev.*, 179(5):1499–1512, Mar 1969.

[40] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51(21):1945–1947, Nov 1983.