MEASUREMENT OF THE RARE TRANSITION $b \to s\gamma$ USING THE COMPLETE *BABAR* DATA SET.

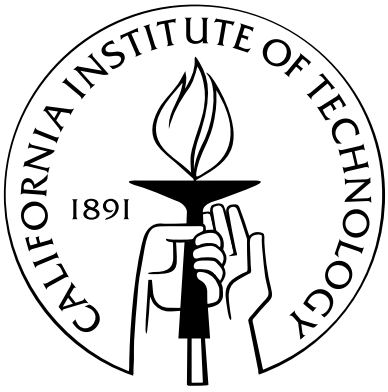
Thesis by

David Andrew Doll

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2011

(Defended May 26, 2011)

Thank you mom and dad for all of the encouragement.

Acknowledgements

I'd like to acknowledge all the folks who helped make this thesis possible: Dr. David Hitlin for his advisement; Dr. Frank Porter for answering occasionally elementary question without judgement; Piti Ongmongkolkul for the insights, the rest of the Caltech *BABAR* group, Heywood Tam, Yousi Ma, Zeeshan Ahmed, Mathew Matheny, Moira Gresham, Nicole Czakon, Riccardo Schmid, David Lopez Mateos, other class of 2005 Physics folks, Walter Herz, Daniel Helias, James Kalagher, everyone in the Graduate Studies Office for the coffee and conversations, the GSC, the Caltech Y, and especially Effrosyni Seitaridou for keeping me sane.

Abstract

We present the results of a measurement of the total rate and photon energy spectrum in $b \to s\gamma$ transitions using the entire *BABAR* data set, $429\,\mathrm{fb}^{-1}$. These results use a "sum of exclusives" approach in which we reconstruct a subset of the final states of the s-quark system and correct for the final states that are missing. We find $\mathcal{B}(\overline{B} \to X_s\gamma) = (329 \pm 19 \pm 48) \times 10^{-6}$ for $E_\gamma > 1.9\,\mathrm{GeV}$. We also measure the mean and variance of the photon spectrum and find $\langle E \rangle = 2.346 \pm 0.018^{+0.027}_{-0.022}$ and $\langle E^2 \rangle - \langle E \rangle^2 = 0.0211 \pm 0.0057^{+0.0055}_{-0.0069}$. Finally we fit two classes of models for the photon spectrum and extract their respective HQET parameters.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

BABAR, a particle physics collaboration and detector, located at the Stanford Linear Accelerator Center (SLAC), studies flavor physics through weak decays of the B meson by colliding positrons and electrons with a center-of mass (CM) energy corresponding to the $\Upsilon(4S)$ resonance. This bottomonium resonance decays almost exclusively to $B\overline{B}$ pairs, which then decay to final states consisting of lighter hadrons, leptons, or photons. Operating at this resonance in "factory mode" for roughly 9 years gave the *BABAR* collaboration a large data set of B mesons, which may then be used to constrain different aspects of the standard model (SM).

The primary purpose of *BABAR* was to study CP violation in the b sector, potentially responsible for the matter-antimatter asymmetry we see in nature. The large amount of B meson data acquired, however, allows for other precision tests of the SM as well. In this analysis, we perform a measurement of the transition rate of $b \to s\gamma$ and spectrum of the photon using the entire *BABAR* data set. The results presented within are an update of a previous *BABAR* analysis using roughly one-fifth the data [1].

The procedure for this analysis is similar to many other *BABAR* analysis. We reconstruct the B mesons in the recorded events based on the tracks and neutral particles detected. We then remove the events that are background to our $b \to s\gamma$ measurements while striving to maintain adequate efficiency for the $b \to s\gamma$ events themselves. To do this, we use many of the kinematic variables associated with the detected particles, variables associated with the reconstructed intermediate states, and overall event topology variables, and feed them into a variety of multivariate classifiers. Variables of primary importance to many *BABAR*

analysis are:

$$\Delta E \equiv E_B^* - \sqrt{s}/2, \tag{1.1}$$

$$m_{ES} \equiv \sqrt{(\sqrt{s}/2)^2 - P_B^{*2}}, \tag{1.2}$$

where E_B^* and P_B^* are the reconstructed B meson candidate's energy and 3-momentum in the CM system. These variables should peak at 0 and the B meson mass respectively for a correctly reconstructed B meson. We incorporate ΔE, as well as several other variables, in our multivariate classifiers, and extract the signal yield by fitting the m_{ES} distribution. We choose to use m_{ES}, the beam-energy-constrained mass, instead of the closely related variable $m_{inv.}$, or the invariant mass of the reconstructed B based entirely on the measured momenta and energies of the final state particles, because of the better precision of m_{ES} and this variable's minimal correlation with ΔE. We extract our signal yield in bins of the X_s-hadron mass, m_{X_s}, and use the measured values of the partial branching fractions in each of these bins to determine a total transition rate for the decay $b \to s\gamma$, and then extract several heavy quark effective theory (HQET) parameters.

Chapter 2

Theory

The SM describes the interactions of particles at the most fundamental level through parameterizations of three of the four known forces (the SM excludes gravity). Within the context of the SM, the fundamental particles may be classified as either quarks and leptons, spin-1/2 fermions, or force-carrying spin-1 bosons. There are six quarks and six leptons, arranged in three generation pairs. For the quarks these are, in order of increasing approximate mass, "up-down," "charm-strange," "top-bottom," and for the leptons these are electron, muon, and tau along with the respective flavor of neutrino. The electroweak and strong force carriers are the photon (γ), $W^{\pm}$, Z, and gluon.

The electroweak force, in particular through its $W^{\pm}$ mediators, is responsible for generation changes (or "flavor changes") within the SM, resulting in up-and-down-type transitions for the quarks. In the quark sector, the coupling between the generations may be succinctly described by the Cabbibo-Kobayashi-Maskawa (CKM) matrix [2], which parameterizes the mixing that results from the mass eigenbasis (what we measure) not being identical to the weak eigenbasis (what is produced). There are several ways one may parameterize this matrix, a more popular one is the Wolfenstein parameterization [3]:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix},$$

which uses λ^2 as an expansion parameter (λ itself is already small, $\mathcal{O}(0.2)$). The constraint

of unitarity on the CKM matrix implies both:

$$\sum_k |V_{ik}|^2 = 1, \tag{2.1}$$

and

$$\sum_k V_{ik} V_{jk}^* = 0. \tag{2.2}$$

This latter constraint allows for a graphical representation in the form of three "unitarity triangles"; the one most relevant to *BABAR* is shown in Figure 2.1.

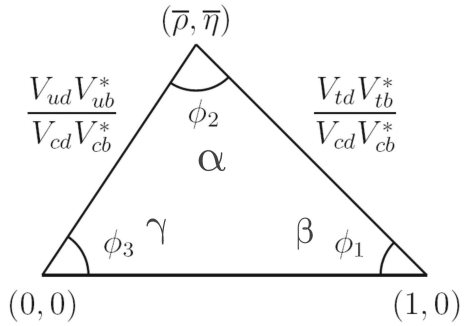


Figure 2.1: One of the three unitarity triangles with the Wolfenstein parameters shown.

2.1 $b \to s\gamma$ Transition Rate

The flavor changing neutral current (FCNC) transition $b \to s\gamma$ first occurs at the loop level within the SM (this is true in general of FCNCs in the SM), and has for a long time been recognized as an important probe for new physics (NP). Consequently, much work has gone into calculating the transition rate of $b \to s\gamma$, and the current predictions for this transition rate use the next-to-next-to leading order (NNLO) precision in the perturbative component [4]. The current world average of the measured branching fraction with a cut on $E_\gamma > 1.6\,\mathrm{GeV}$ in the $\overline{B}$-rest frame is [5]:

$$\mathcal{B}(\overline{B} \to X_s\gamma)^{exp}_{E_\gamma > 1.6\,\mathrm{GeV}} = 3.55 \pm 0.24 \pm 0.09 \times 10^{-4}, \tag{2.3}$$

where the first error is the statistical and systematic errors combined and the second is due to extrapolating down to a photon energy of 1.6 GeV based on the input shape function (the high photon energy limit being equal to $m_B/2$). The total error on this measurement is about 7.2% of the central value. However, the relation

$$\Gamma(\overline{B} \to X_s\gamma) \simeq \Gamma(b \to X_s^{parton}\gamma) \tag{2.4}$$

is valid up to non-perturbative corrections, which are on the order of 5% (and are currently the largest component of the error in the theoretical predictions at NNLO in the perturbative expansion discussed below [4]).

Following the NNLO calculation described in [4], [6], and [7], the branching ratio can be expressed as:

$$\mathcal{B}(\overline{B} \to X_s\gamma)_{E_\gamma > E_0} = \mathcal{B}(\overline{B} \to X_c e\nu) \left|\frac{V_{ts}^* V_{tb}}{V_{cb}}\right|^2 \frac{6\alpha_{em}}{\pi C}(P(E_0) + N(E_0)), \tag{2.5}$$

where $N(E_0)$ denotes the non-perturbative correction differentiating the *l.h.s.* from the *r.h.s.* in equation 2.4. The perturbative contribution to this calculation, $P(E_0)$, is taken from the ratio

$$\frac{\Gamma(b \to X_s\gamma)_{E_\gamma > E_0}}{|V_{cb}/V_{ub}|^2 \Gamma(b \to X_u e\nu)} = \left|\frac{V_{ts}^* V_{tb}}{V_{cb}}\right|^2 \frac{6\alpha_{em}}{\pi} P(E_0), \tag{2.6}$$

and it is this quantity, $P(E_0)$, for which one calculates the NNLO QCD corrections. Since equation 2.6 is normalized to the charmless semileptonic rate, the quantity C is effectively a "non-perturbative semileptonic phase-space factor," given by:

$$C = \left|\frac{V_{ub}}{V_{cb}}\right|^2 \frac{\Gamma(\overline{B} \to X_c e\nu)}{\Gamma(\overline{B} \to X_u e\nu)}. \tag{2.7}$$

With the denominator on the *l.h.s* of equation 2.6 already calculated to NNLO in [8], the procedure for extracting the value of $P(E_0)$ proceeds in three steps:

1. Write down an effective Lagrangian in terms of Wilson coefficients, $C_i(\mu_0)$, and corresponding operators, Q_i (typical examples of the operators are shown in Figure 2.2).

The C_i are effectively coupling constants at the flavor-changing vertices, Q_i. The C_i account for the full QCD theory and are evaluated at some higher-mass renormalization scale, taken to be $\mu_0 \sim M_w, m_t$, by requiring equality between the effective theory and the SM to leading order in (external momenta)/(M_w, m_t). The effective Lagrangian for the NNLO calculation is taken to be [4]:

$$\mathcal{L} = \mathcal{L}_{QCD\times QED}(u,d,s,c,b) + \frac{4G_F}{\sqrt{2}}\left[V_{ts}^* V_{tb}\sum_{i=1}^{8} C_i Q_i + V_{us}^* V_{ub}\sum_{i=1}^{2} C_i^c (Q_i^u - Q_i)\right], \tag{2.8}$$

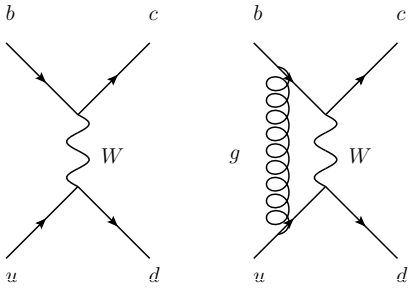
where

$$\begin{aligned}
Q_1^u &= (\overline{s}_L\gamma_\mu T^a u_L)(\overline{u}_L\gamma^\mu T^a b_L),\\
Q_2^u &= (\overline{s}_L\gamma_\mu u_L)(\overline{u}_L\gamma^\mu b_L),\\
Q_1 &= (\overline{s}_L\gamma_\mu T^a c_L)(\overline{c}_L\gamma^\mu T^a b_L),\\
Q_2 &= (\overline{s}_L\gamma_\mu c_L)(\overline{c}_L\gamma^\mu b_L),\\
Q_3 &= (\overline{s}_L\gamma_\mu b_L)\Sigma_q(\overline{q}\gamma^\mu q),\\
Q_4 &= (\overline{s}_L\gamma_\mu T^a b_L)\Sigma_q(\overline{q}\gamma^\mu T^a q),\\
Q_5 &= (\overline{s}_L\gamma_{\mu 1}\gamma_{\mu 2}\gamma_{\mu 3} b_L)\Sigma_q(\overline{q}\gamma^{\mu 1}\gamma^{\mu 2}\gamma^{\mu 3} q),\\
Q_6 &= (\overline{s}_L\gamma_{\mu 1}\gamma_{\mu 2}\gamma_{\mu 3} T^a b_L)\Sigma_q(\overline{q}\gamma^{\mu 1}\gamma^{\mu 2}\gamma^{\mu 3} T^a q),\\
Q_7 &= \frac{e}{16\pi^2} m_b(\overline{s}_L\sigma^{\mu\nu} b_R)F_{\mu\nu},\\
Q_8 &= \frac{g}{16\pi^2} m_b(\overline{s}_L\sigma^{\mu\nu} T^a b_R)G^a_{\mu\nu}.
\end{aligned}$$

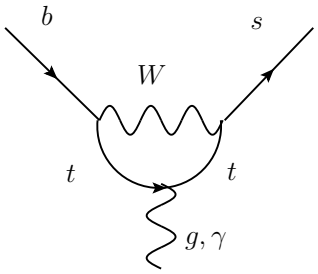
For completeness, it should be mentioned that the contribution of the last term in the bracket in equation 2.8 is minimal in the NLO calculation, and expansion to NNLO in [4] is neglected.

2. Determine how the operators mix under renormalization and evolve the corresponding $C_i(\mu)$ from μ_0 down to the appropriate energy scale for the process, taken to be $\mu_b \sim m_b$.

3. Evaluate the $b \to X_s^{parton}\gamma$ amplitudes at $\mu_b \sim m_b$.



(a) The current-current operators, $Q^c_{1,2}$ (c referring to the bottom quark transitioning to a charm quark, there are similar diagrams for the subsequent charm to strange transition)



(b) The magnetic penguin operators, $Q_{7(\gamma),8(g)}$

Figure 2.2: Typical diagrams in the full theory that lead to the different operators used in calculating the $b \to s\gamma$ branching fraction.

For comparison below to different new physics (NP) models, we summarize one last aspect of the SM calculation for this transition rate, and that is to point out that the expression for $P(E_0)$ can be written in terms of its perturbative expansion:

$$\begin{aligned} P(E_0) =& P^{(0)}(\mu_b) + \tilde{\alpha}_s(\mu_b)\left[P_1^{(1)}(\mu_b) + P_2^{(1)}(E_0,\mu_b)\right] \\ &+ \tilde{\alpha}_s^2(\mu_b)\left[P_1^{(2)}(\mu_b) + P_2^{(2)}(E_0,\mu_b) + P_3^{(2)}(E_0,\mu_b)\right] \\ &+ \mathcal{O}(\tilde{\alpha}_s^3), \end{aligned} \tag{2.9}$$

where

$$\tilde{\alpha}_s \equiv \frac{\alpha_s^{(5)}(\mu_b)}{4\pi}, \tag{2.10}$$

is the strong coupling constant evaluated at the energy scale appropriate for this process, $\mu_b \sim m_b$. The zeroth order term (as well as some of the higher order terms, see equations 2.8 and 2.11 in [4]) may be fully expressed in terms of a single Wilson coefficient, $C_7^{(0)eff}$:

$$P^{(0)}(\mu_b) = \left(C_7^{(0)eff}(\mu_b)\right)^2, \tag{2.11}$$

where the zeroth order in $\tilde{\alpha}_s$ effective Wilson coefficient, $C_7^{(0)eff}$, is defined in equations 2.7 and 2.8 in [4].

The current NNLO calculated value in the SM for $\mathcal{B}(\overline{B} \to X_s\gamma)$ is $(3.15\pm0.23)\times10^{-4}$ [4], with a minimum photon energy cutoff of $E_{min} > 1.6\,\text{GeV}$. The world-average experimental result: $\mathcal{B}(\overline{B} \to X_s\gamma) = (3.55\pm0.24\pm0.09)\times10^{-4}$ [5], consists of measurements that generally use a photon energy cutoff around $E_{min} > \{1.8 - 2.0\}$ to reduce background from other B events, and extrapolate the result to the cutoff given above.

2.1.1 NP Models and $b \to s\gamma$

There are many NP models that can either enhance or suppress the transition rate of $b \to s\gamma$; several are summarized in Table 2.1 (this summary is taken from [9], see Table I therein for further references). We will describe in more detail the effects of two of the more popular of these models: the two Higgs doublet model (2HDM) type II ("type" referring to the nature of the couplings of the doublets to the different types of quarks), and a general minimally super symmetric SM (MSSM). Typical diagrams reflecting the contributions from these models to the $b \to s\gamma$ transition rate are shown Figure 2.3.

Table 2.1: Different NP models and their effects (relative to the SM) on the transition rate of $b \to s\gamma$ ($\Uparrow$ indicating an enhancement in the rate), as well as the precision in the respective calculation of the model's effect. Please see [9] for more details and other model effects.

Model	Precision	Effect
THDM type II	NLO	$\Uparrow$ - for low m_{H^+}
general MSSM	LO	$\Updownarrow$
MFV MSSM	NLO	$\Updownarrow$
MFV SUSY GUTs	NLO	$\Downarrow$
UED 5 or UED 6	LO	$\Downarrow$

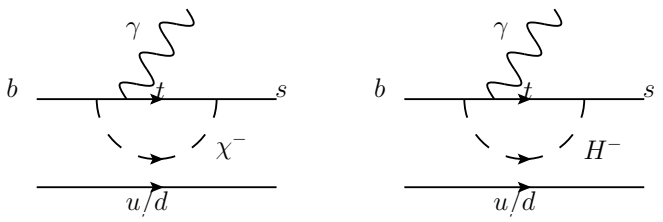


Figure 2.3: Typical NP contributions to the transition rate of $b \to s\gamma$ (chargino left, charged Higgs right).

As with most NP models' effects on $b \to s\gamma$, the type-II 2HDM would enhance the transition rate by introducing another amplitude with a charged boson, the H^-, replacing the W^- in the loop, as shown in Figure 2.3. The type-II 2HDM (type-II corresponding to one doublet, referred to as ϕ_1 in [10], coupling to the down-type quarks, and the other doublet, ϕ_2, coupling to the up-type quarks and leptons), would enhance the value of $C_7^{(0)eff}$, given above in equation 2.11. The change due to NP to the $b \to s\gamma$ transition rate at LO has the form [10]:

$$P + N = (C_{7,SM}^{(0)eff} + B\Delta C_{7,H^-}^{(0)eff})^2 + A, \tag{2.12}$$

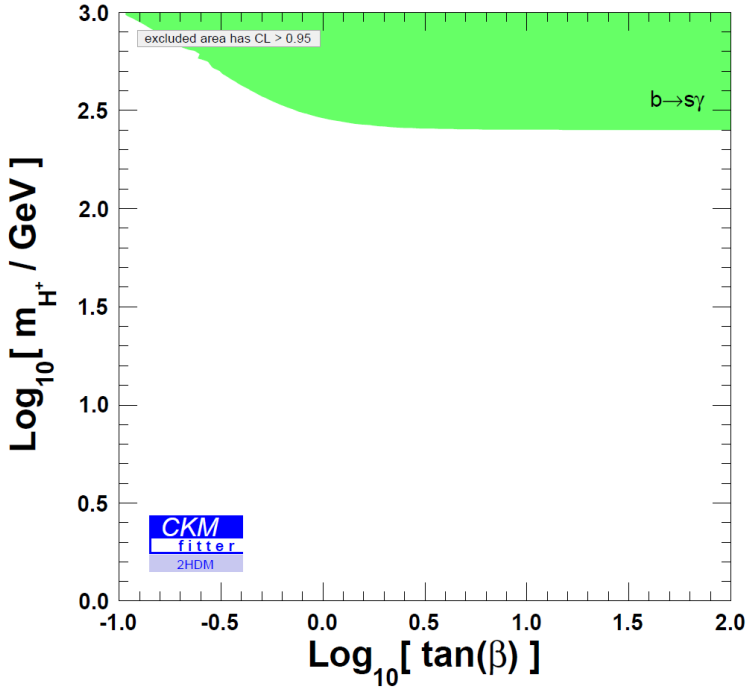
where the $C_7^{(0)eff}$ coefficients reflect either the SM Wilson coefficient, or the change introduced by the 2HDM. The functions A and B (B being positive) are independent of the relevant 2HDM parameters m_{H^-} (the mass of the new charged Higgs) and $\tan\beta$ (the ratio of the two vacuum expectation values, v_2/v_1). The quantity $\Delta C_{7,H^-}^{(0)eff}$ does depend on these parameters (it is related to different powers of the inverse of both of them, and has the same sign as its SM counterpart, effectively enhancing the transition rate). The measured rate of $b \to s\gamma$ therefore provides very strict limits on their values.

The impact of the 2HDM on the transition rate of $b \to s\gamma$ has been calculated to NLO (see, for instance, [7] or [10]), and, even with the experimental branching ratio of $\overline{B} \to X_s\gamma$ being noticeably higher than the SM prediction, this mode provides one of the strongest tests of the available parameter space for this model, as shown in Figure 2.4.

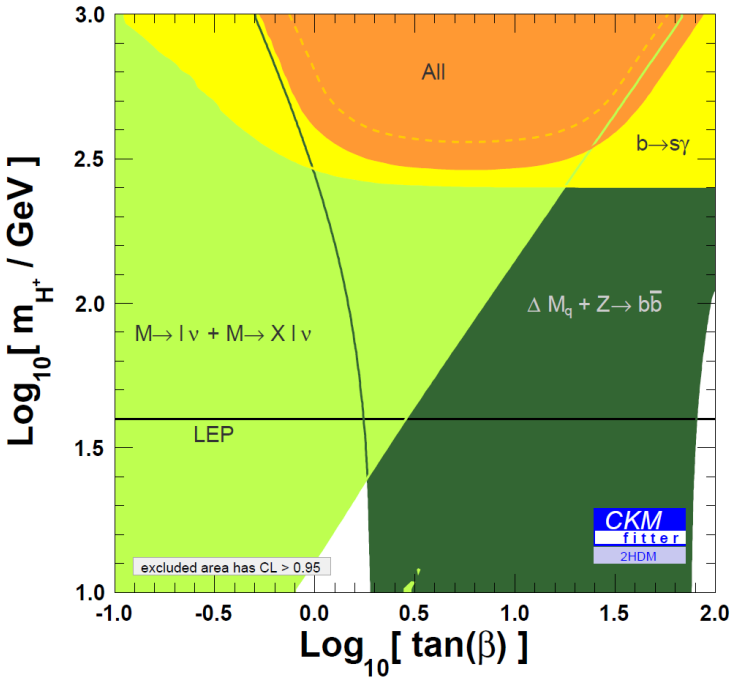
However, the 2HDM does not need to be the only source of NP. Indeed, typical SUSY models imply the presence of a 2HDM component, and an MSSM with minimal flavor violation (MFV, meaning the FCNC processes are suppressed by the same CKM elements as in the SM, thus avoiding models that predict increased rates of FCNCs) will have contributions to the $b \to s\gamma$ transitions from other sources as well.

The impact of the different sources of NP on the appropriate Wilson coefficients in an MFV MSSM can be written as (neutralino omitted since its contribution is expected to be small):

$$C_{7,8}^{eff} = C_{7,8}^{SM} + C_{7,8}^{H^+} + C_{7,8}^{\tilde{\chi}^+} + C_{7,8}^{\tilde{g}}, \tag{2.13}$$



(a) The exclusion region of m_{H^+} vs. $\tan\beta$ (white region excluded at 95% confidence) based on the current experimental value of $\mathcal{B}(\overline{B} \to X_s\gamma)$. Taken from [10].



(b) The allowed region of m_{H^+} vs. $\tan\beta$ (the region allowed by all measurements is given in orange) based on the measured $\mathcal{B}(\overline{B} \to X_s\gamma)$ as well as other measurements. Taken from [10].

Figure 2.4: The impact of the 2HDM on the transition rate of $b \to s\gamma$, and the impact of the measured $\mathcal{B}(\overline{B} \to X_s\gamma)$ on the parameters of the 2HDM.

which explicitly refers to the SM contribution, the charged Higgs contribution, the chargino contribution, and the gluino contribution. In such a model, the charged Higgs contribution still constructively interferes with the SM contribution, but there is no such *a priori* requirement on the other terms. Therefore, while in the context of a bare 2HDM, $b \to s\gamma$ places strong constraints on the available parameter space, in a typical MSSM with MFV, the potential for destructive interference between the competing diagrams limits the impact of $b \to s\gamma$ in parameter determination [11]. Nevertheless, in the context of other measurements, such as the anomalous magnetic moment of the muon, $g_\mu - 2$, which can also be strongly impacted by the existence of supersymmetry (and is currently about 2-3σ different than the SM expectation), regions of parameter space in this class of models can still be favored more than others. The combination of these two measurements tends to favor both a light chargino and light charged Higgs as described in [12].

2.2 $b \to s\gamma$ Photon Spectrum

Unlike the total transition rate, the photon energy spectrum is not expected to be heavily influenced by NP. Nevertheless, measurement of the photon spectrum, and specifically the moments of this spectrum provides constraints that improve other SM calculations.

At the quark level, the transition $b \to s\gamma$ emits a photon kinematically fixed to $E_\gamma = m_b/2$ in the b quark rest frame, where m_b is the mass of the b quark. However, two separate processes contribute to translating this delta function into a photon spectrum in $\overline{B} \to X_s\gamma$. First, what we detect is not $b \to s\gamma$, but rather $b \to X_s^{parton}\gamma$, and when one or more gluons or a $q\overline{q}$ pair is emitted, the photon is detected with $E_\gamma < m_b/2$. Second, there is a "smearing" of the photon energy due to the Fermi motion of the quark within the B meson. This smearing causes the photon to have a final energy both above and below the kinematicaly expected value above. The first process may be described by perturbative QCD, while the second must be modeled by a shape function of the Fermi motion [13]. This shape function is universal, in that the same shape function describes the non-perturbative photon spectrum in $\overline{B} \to X_s\gamma$ and the lepton energy spectrum near the endpoint in $\overline{B} \to X_u l\overline{\nu}$ decays [14], or any other decays with essentially massless partons

in the final state (as supposed to $\overline{B} \rightarrow X_c l \overline{\nu}$). Because of momentum conservation, the photon has a maximum value of $E_\gamma^{max} = M_B/2$, where M_B is the mass of the B meson, and the shape function used to model the Fermi motion of the b quark is necessary to describe the spectrum in a range $\frac{m_b+(M_B-mb)}{2} > E_\gamma > \frac{m_b-(M_B-m_b)}{2}$. As the energy of the photon gets lower, the perturbative QCD modification of the photon spectrum becomes more important. For a quark mass of m_b=4.65 GeV/c^2, this puts the lower end of the Fermi motion's impact around $E_\gamma \approx 2\,\mathrm{GeV}$ in the B meson rest frame, which is the recommended maximum experimental cutoff value given in [15] (*i.e.* experimental measurements should use this value or lower for the E_γ minimum cutoff). Others [16] recommend placing this cutoff no higher than $E_{cut} \geq 1.85\,\mathrm{GeV}$. Experimentally, however, lower photon energy cutoffs introduce larger background levels (in particular, backgrounds from other B meson decays). For this analysis, we place a minimum photon energy cut at $E_\gamma = 1.9\,\mathrm{GeV}$.

As mentioned above, relevant HQET parameters may be extracted from the spectrum of the photon energy. The parameters m_b and μ_π^2 (also called $-\lambda_1$ in the literature) may be extracted through accurate measurement of the first and second moments. The presence of a minimum photon energy cut, however, influences the measured values of these parameters.

There are several reasons for placing a cut on the photon energy as low as possible. First, there are many ways to parameterize the shape function of the Fermi motion of the b quark; two popular ones include a momentum-dependent product of a power law and an exponential, and a momentum-dependent product of a power law and a Gaussian (like the shape functions used in [15], [16], and [17]). The lower the photon energy cutoff, the less the dependence on the choice of this shape function enters into the extrapolation to $E_\gamma > 1.6\,\mathrm{GeV}$ used in theoretical calculations. By measuring the moments of the photon spectrum, theorists will be able to determine values to the parameters in whichever shape function *ansatz* they choose, and then perform this extrapolation. Second, the lower the photon energy cutoff, the more accurately the measured moments reflect the actual moments in the absence of a minimum photon energy cut, and the less biased these measured moment values become. For instance, in [16], they predict the bias on the moment values based on different minimum photon energies, and predict a resulting bias in $m_b = 0.024\,\mathrm{GeV}$ and

$\mu_\pi^2 = 0.0790\,\mathrm{GeV}^2$ with a photon energy cut at $E_\gamma = 1.9\,\mathrm{GeV}$ (assuming central values of $m_b(1\,\mathrm{GeV}) = 4.61\,\mathrm{GeV}$ and $\mu_\pi^2(1\,\mathrm{GeV}) = 0.41\,\mathrm{GeV}^2$). Finally, as one lowers the photon energy cutoff, specific X_s resonances become less important and $b \to s\gamma$ transitions may be approximated by generic X_s hadronization models. Indeed, as recommended by [15], since the widths of X_s resonances above the $K^*(892)$ exceed their spacing, we will only consider this specific resonance and model the remainder of the $b \to s\gamma$ transitions with a generic hadronization model.

Chapter 3

Analysis Procedure

There are several ways to perform a measurement of the parton transition $b \to s\gamma$. For this analysis, we use the method called "sum-of-exclusives," referring to the fact that we reconstruct the s-quark system in one of 38 final states, given in Table 3.1. These 38 modes do not represent an exhaustive list of the final states of the s quark; rather they reflect that subset of the final states for which we have reasonable detection efficiency and signal/background separation.

Because we are reconstructing the X_s-system (the final hadronic state of the s quark), we have direct access to the invariant mass of the X_s-system (to within detector resolution), which we may then translate into a measurement of the photon energy in the $b \to s\gamma$ transition in the B meson rest frame (we refer to this as the "transition photon" throughout). The photon energy is related to the invariant mass of the X_s-system through:

$$E_\gamma^B = \frac{m_B^2 - m_{X_s}^2}{2m_B}, \tag{3.1}$$

where E_γ^B is the transition photon's energy in the B rest frame, m_B is the mass of the B meson, and m_{X_s} is the invariant mass of the X_s hadronic system. Therefore, by measuring the spectrum of m_{X_s}, we will be able to measure the E_γ spectrum. Since the B mesons are not generated at rest in either the laboratory frame, or the $\Upsilon(4S)$ frame, an accurate measurement of the photon spectrum in the B rest frame is not as experimentally feasible with a more "inclusive-measurement" approach. In order to remain as spectrum-model independent as possible, we performed the majority of this analysis with simulated signal

Table 3.1: The 38 modes we reconstruct in this analysis; BiType identifies the numeric value we assign to each mode for bookkeeping; charge conjugation is implied.

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

events, or signal monte carlo (MC), that are generated with a flat photon spectrum. We thus obtain a spectrum in m_{X_s} that may be translated to a photon spectrum, and used to fit to the best parameter values with any desired *ansatz* for the photon spectrum. We perform this analysis using 18 bins of m_{X_s}, between 0.6-2.8 GeV/c^2, the first 14 being 100 MeV/c^2 wide, the last 4 being 200 MeV/c^2 wide. The photon-energy equivalent of the bin boundaries is given in Table 3.2.

The relevant quantities needed to extract the values of different HQET parameters are the moments of the photon distribution, as well as the branching fractions in bins of hadronic mass. We will also report the sum of these branching fractions to derive a total branching fraction for $B \to X_s\gamma$ with $E_\gamma > 1.9$ GeV. Finally, we fit the parameters in the specific models created by [16] and [17] based on our measured m_{X_s} spectrum to determine values for m_b and μ_π^2 used by [16] (or $\overline{\Lambda} \equiv (M_B - m_b)_{lim.\ m_b\to\infty}$ and μ_π^2 for [17]). It should be stressed, however, that the values we quote for these HQET parameters are specific to

Table 3.2: Numerical conversion from hadron system mass to photon energy.

m_{X_s} (GeV/c^2)	E_γ^B (GeV)
0.6	2.606
0.7	2.593
0.8	2.579
0.9	2.563
1.0	2.545
1.1	2.525
1.2	2.503
1.3	2.480
1.4	2.454
1.5	2.427
1.6	2.397
1.7	2.366
1.8	2.333
1.9	2.298
2.0	2.261
2.2	2.181
2.4	2.094
2.6	1.999
2.8	1.897

these models. For the models in [16], the bias in our photon spectrum's moments have been evaluated, normalizing the heavy quark parameters at $\mu = 1\,\mathrm{GeV}$. For the models from [17], these quantities have been evaluated at $\mu = 1.5\,\mathrm{GeV}$, and assume a photon energy cutoff of $1.8\,\mathrm{GeV}$. For both of these reasons, the resulting quantities are not immediately comparable. The branching fractions and moments we report, however, are independent of a particular photon-spectrum model.

Chapter 4

PEP-II and the *BABAR* Detector

In this chapter we will briefly describe the SLAC B Factory complex, consisting of the PEP-II collider, used to produce a large sample of $B\overline{B}$ pairs in e^+e^- collisions, and the *BABAR* detector, used to detect the final state particles of these collisions and reconstruct the decaying B mesons.

4.1 PEP-II

The PEP-II electron-positron collider [18] was a two-ring, high-luminosity collider located at SLAC (see Figure 4.1), operating at the $\Upsilon(4S)$ resonance. The two-ring system consists of a high-energy storage ring (HER), 2.2 km in circumference, that delivers electrons at an energy of 9.0 GeV, and a low-energy storage ring in the same tunnel, that delivers positrons at an energy of 3.1 GeV. Together, these result in a center of mass (CM) energy at Interaction Region 2 (where the *BABAR* detector is located) of $\sqrt{s}$ =10.58 GeV, the mass of the $\Upsilon(4S)$ resonance.

The 3km long SLAC linac provides electron and positron beams to fill the 1658 bunches at an injection rate of up to 60 Hz. Due to the need for 'factory' operating conditions, consistency in beam intensity is also achieved through continuous injection, or "trickle charge" mode. The peak instantaneous luminosity achieved by PEP-II during the run of the experiment was $\sim$12 nb^{-1} s^{-1}, about 4$\times$ the designed luminosity.

For the majority of the *BABAR* experiment, PEP-II ran at the $\Upsilon(4S)$ resonance, just above the $B\overline{B}$ threshold, which has a $b\bar{b}$-resonance production cross-section of $\sigma(e^+e^- \rightarrow$

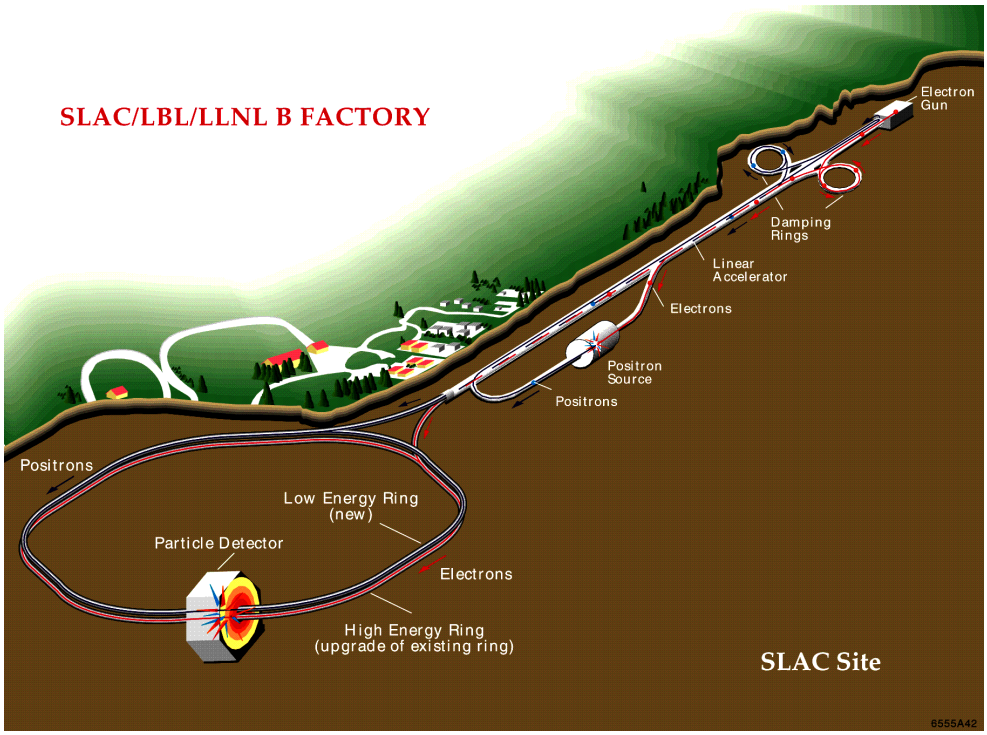


Figure 4.1: The PEP-II storage-ring facility located at SLAC.

$\Upsilon(4S)) = 1.1$ nb. At this CM energy, however, several other physics processes occur, including lighter quark pair production: $\sigma(e^+e^- \to u\overline{u}) = 1.39$ nb, $\sigma(e^+e^- \to d\overline{d}) = 0.35$ nb, $\sigma(e^+e^- \to s\overline{s}) = 0.35$, and $\sigma(e^+e^- \to c\overline{c}) = 1.30$ nb; lepton pair production: $\sigma(e^+e^- \to \mu^+\mu^-) = 1.16$ nb, and $\sigma(e^+e^- \to \tau^+\tau^-) = 0.94$ nb; and Bhabha scattering, $\sigma(e^+e^- \to e^+e^-) \approx 40$ nb, within the acceptance of the *BABAR* detector [19].

In general, these non-$b\overline{b}$ event types serve as background in most analyses involving at B mesons with the *BABAR* data set. Therefore, to study this background, a fraction of data is taken 40 MeV below the $\Upsilon(4S)$ peak, which is about 20 MeV below $B\overline{B}$ threshold. The data taken at this lower energy, "off peak" data, may then be used to characterize the backgrounds from the competing background processes with no B meson contamination.

Though it is not the focus of this analysis, a brief motivation for the asymmetric ring energies should be included. The $\Upsilon(4S)$ rest frame, with 9 GeV e^- and 3.1 GeV e^+, has a Lorentz boost of $\beta\gamma = 0.56$ with respect to the laboratory frame [20]. This boost is necessary to improve discrimination between the decay vertices of the two B meson decay

products of the $\Upsilon(4S)$ (the $\Upsilon(4S)$ decays to a $B\overline{B}$ pair $> 96\%$ of the time [21]). The average flight distance in the laboratory frame, Δz, of a B meson, having a mean lifetime of $\tau = 1.525 \times 10^{-12}$ s, is $\Delta z = c\beta\gamma\tau \approx 250\mu$m, well within the resolution of the silicon vertex tracker (SVT), as discussed below (section 4.2.1). This resolution in decay vertices allows for the study of time-dependent CP asymmetries between the two, distinguishable, B mesons, and, using decays such as $B^0 \to c\bar{c}K^{(*)0}$, yields very precise measurements of quantities such as sin2β and $|\lambda_f|$ [22] (β was introduced in Figure 2.1 and λ_f relates to the ratio of the CP-conjugate amplitudes).

The PEP-II facility also ran for extended periods of time at the lower-mass resonances $\Upsilon(3S)$ and $\Upsilon(2S)$, as shown in Figure 4.2, as well as, for a brief period, scanning energies above the $\Upsilon(4S)$. Ultimately, PEP-II delivered a luminosity of 553.48 fb^{-1}, of which 531.43 fb^{-1} were recorded by the *BABAR* detector. 432.89 fb^{-1} were recorded at the $\Upsilon(4S)$ resonance and 53.85 fb^{-1} were recorded at the off-peak energy. This analysis uses 429 fb^{-1} recorded at the $\Upsilon(4S)$ (on-peak data), corresponding to $(471.0 \pm 1.3) \times 10^6$ produced $B\overline{B}$ pairs, and 44.81 fb^{-1} of off-peak data. This luminosity corresponds to *BABAR* data collection runs 1-6, collected between 1999-2008.

4.2 The *BABAR* Detector

The *BABAR* detector [20] is a multi-system particle detector, optimized for the study of B meson decays, using the high-luminosity asymmetric energy B Factory at PEP-II. The *BABAR* detector is composed of five subdetector systems. These are, in order of increasing radial distance from the interaction point (IP): the SVT, the drift chamber (DCH), the detector of internally reflected Cherenkov light (DIRC),the electromagnetic calorimeter (EMC), and the instrumented flux return (IFR). The first four of these systems are located within a superconducting solenoid that provides a 1.5 T magnetic field. A schematic of the detector is shown in Figure 4.3. The data acquisition system (DAQ) provides event triggering, data readout, and detector control and monitoring.

The *BABAR* coordinate system is right-handed, with an origin at the nominal IP. The z-axis is along the axis of the DCH in the direction of the higher-energy, electron beam; the

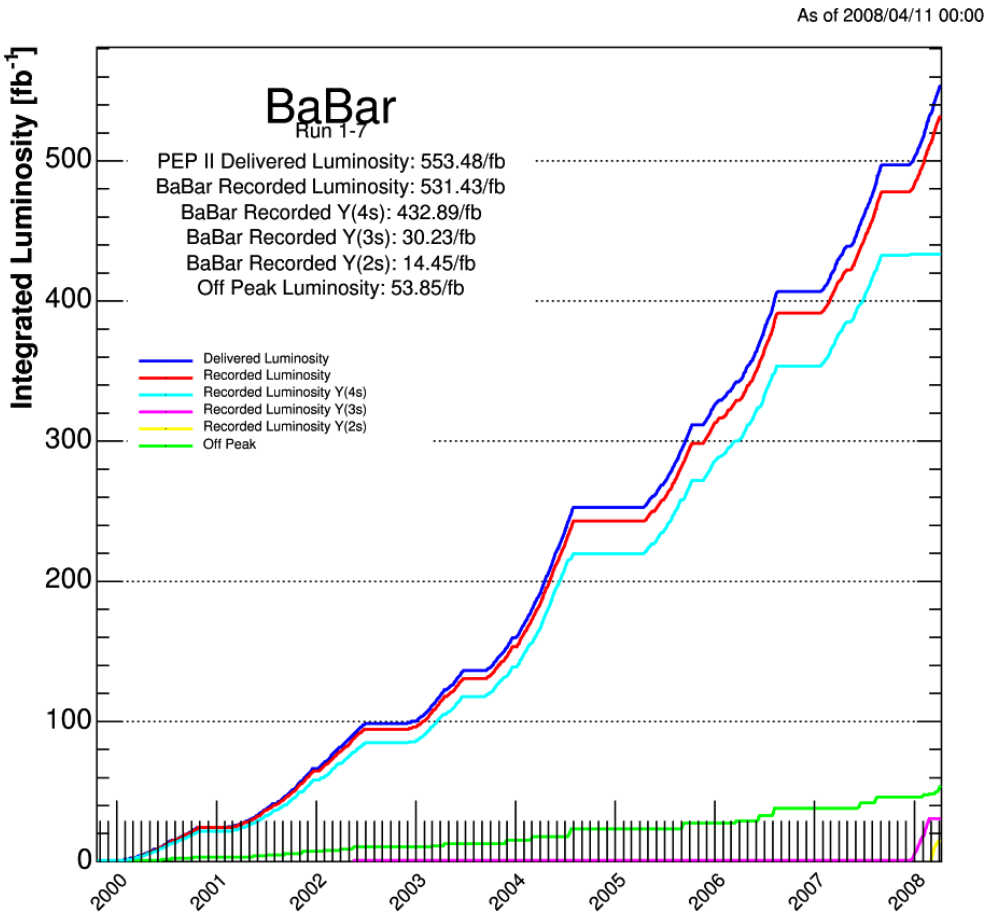


Figure 4.2: The total integrated luminosity delivered by PEP-II and recorded by *BABAR*.

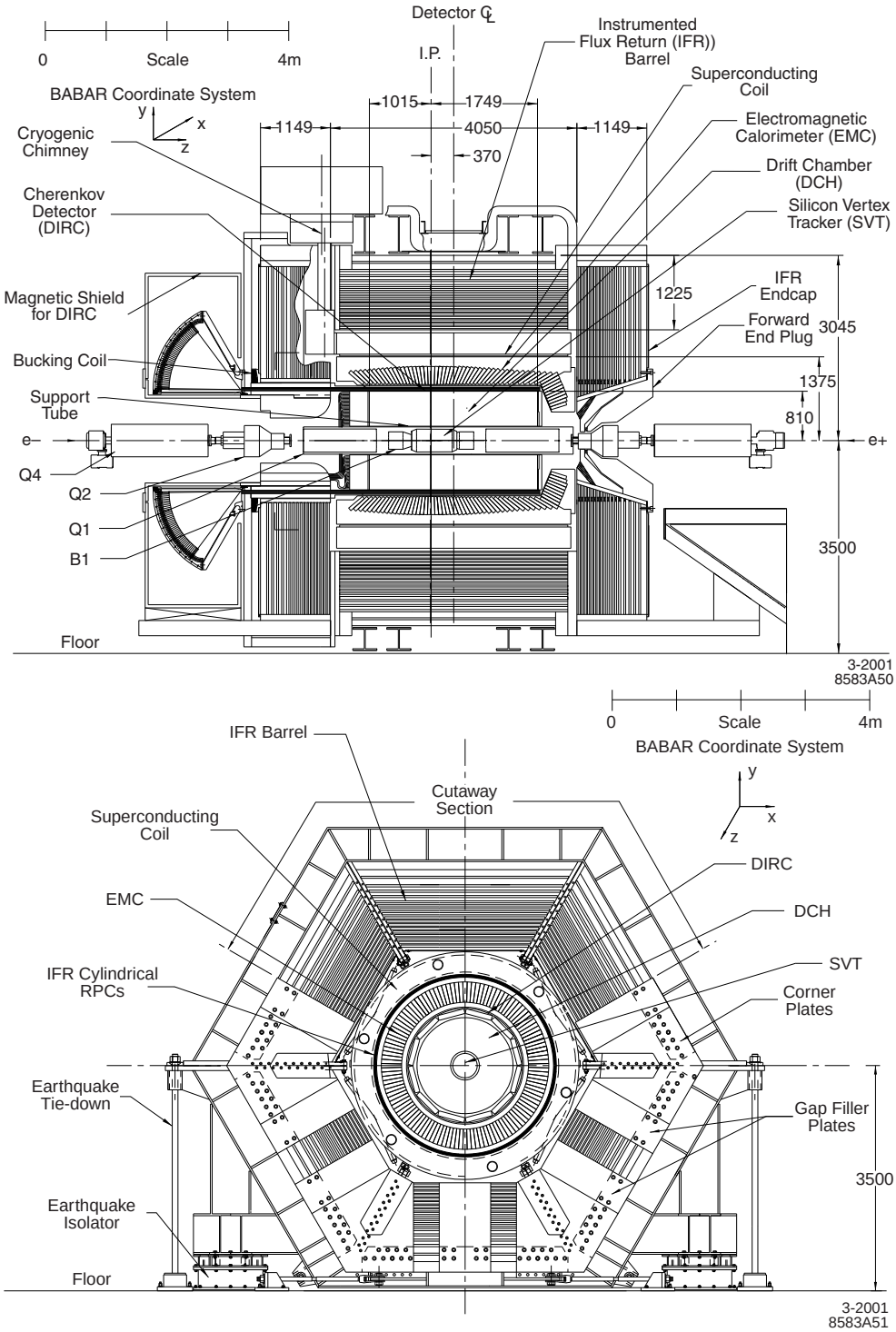


Figure 4.3: The *BABAR* detector.

y-axis points vertically upward; and the x-axis points horizontally outward from the center of the storage rings.

4.2.1 Silicon Vertex Tracker

The SVT is one of the two components that make up the charged particle tracking system, the other being the DCH (see section 4.2.2). For high-multiplicity final states of the s quark system in $b \to s\gamma$ transitions, the high efficiency of the *BABAR* particle tracking system is critical. The SVT measures the positions of charged particles and decay vertices just outside of the beam pipe using five layers of double-sided silicon strip detectors with readout at each end to minimize the amount of inactive material in the acceptance volume of the entire *BABAR* detector. The silicon strips on each side of the five layers are oriented orthogonally to each other with the ϕ measuring strips running parallel to the beam axis and the z measuring strips oriented transverse to the beam axis.

Since the SVT is the detector component closest to the IP, it must provide stand-alone tracking for low transverse momentum, p_t, particles; in the 1.5 T magnetic field of the solenoid, particles with $p_t < 120\,\mathrm{MeV}/c$ cannot be reliably measured in the DCH alone. The SVT also provides the best measurement of track angles, essential for more accurate particle identification (PID) using the Cherenkov angle in the DIRC (see section 4.2.3).

From the inner layer outward, the five layers consist of 6, 6, 6, 16, and 18 modules, respectively, as shown in Figure 4.4. In total, there are approximately 150,000 channels read out from the SVT. For this analysis, the precision of the vertex location is less crucial, but PID is very important (for example K/π identification is crucial for identifying the final state of the s quark). Since each layer of the SVT is double-sided, there is a potential for ten signals from a given charged particle, and, because the readout mechanism consists of a "time over threashold" reading that is a quasi-logarithmic function of collected charge, there is the potential for ten measurements of dE/dx per track that may be used in PID.

The resolution performance in both z and ϕ as a function of incident angle for the SVT is shown in Figure 4.5.

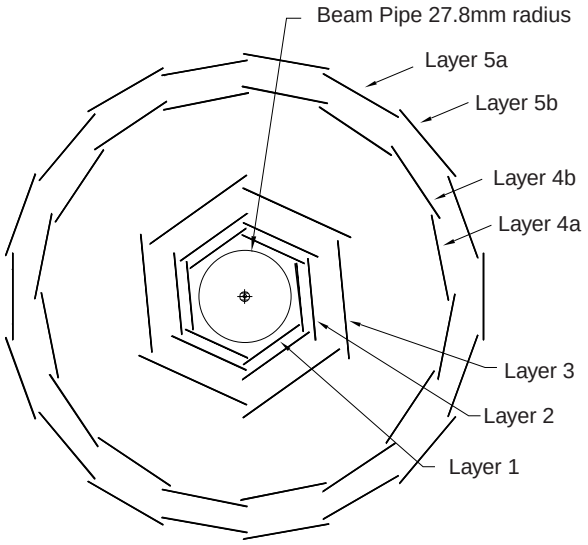
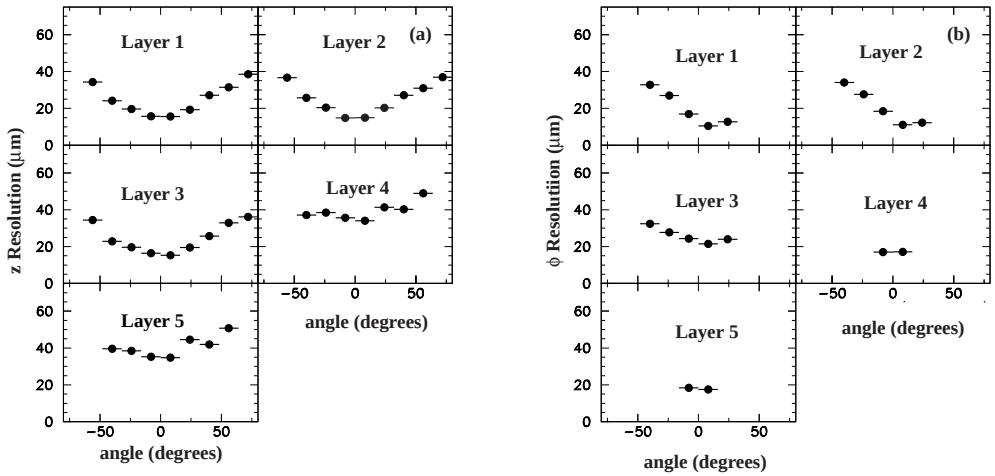


Figure 4.4: The orientation of the five SVT layers.



(a) (b)

Figure 4.5: The SVT resolution in both (a)z, and (b)ϕ as a function of incident angle.

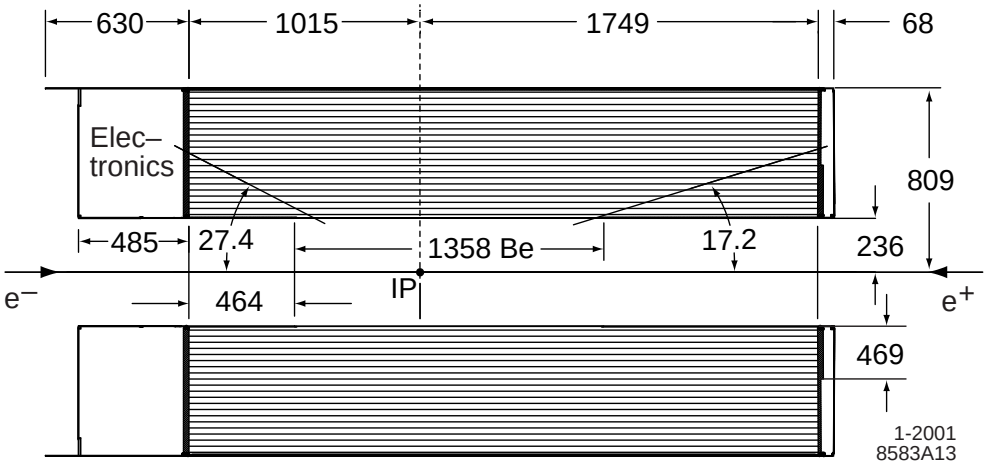
4.2.2 Drift Chamber

The *BABAR* DCH complements the SVT measurements of charged particle tracking and identification. It also is the sole source of information on particles that decay outside of the SVT, such as the K_S^0 mesons used in this analysis, and therefore measures not just transverse momenta and positions, but also longitudinal positions of the tracks.

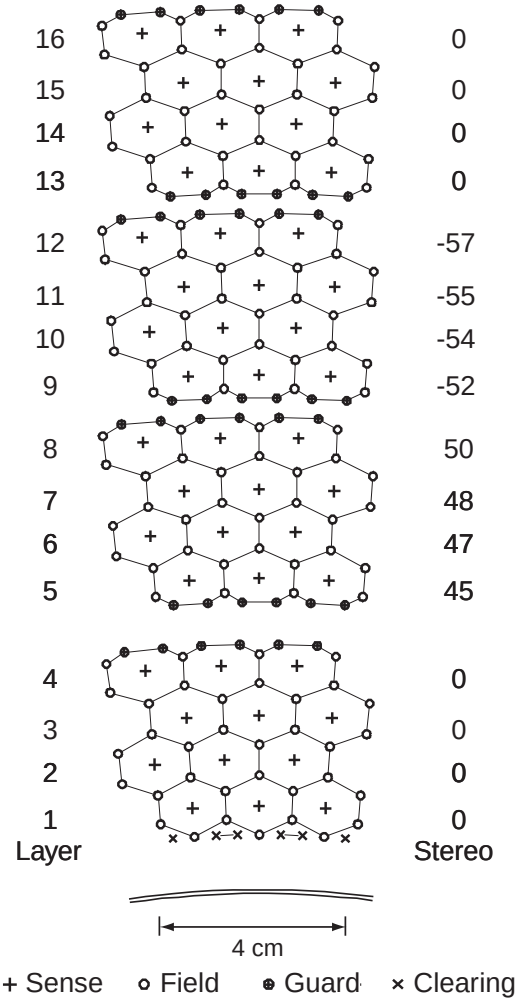
The DCH extends for almost 3m along the *BABAR* detector's z-axis (see Figure 4.6a), and is composed of 40 cylindrical layers of small hexagonal cells, extending from roughly 0.2-0.8m away from the IP, allowing for up to 40 spatial and ionization loss measurements for charged particles with transverse momentum greater than 180 MeV/c. To provide longitudinal position information, 24 of the 40 layers are oriented at small stereo angles around the z-axis. The 40 cylindrical layers are grouped by 4 to create 10 superlayers, which have either axial alignment, A, (no angle with respect to the z-axis), or a small positive or negative stereo angle with respect to the z-axis, U and V. These superlayers are arranged in the order AUVAUVAUVA, moving away from the IP (see Figure 4.6b).

In total, there are 28,768 field, sense, and guard wires, arranged to create 7,104 drift cells, enclosed in a gas mixture that is 80% helium and 20% isobutane. The chosen gas has a radiation length that is five times larger than argon-based gases, and the entire DCH apparatus accounts for less than 0.2% of a radiation length.

Together, the SVT and DCH measurements allow for the reconstruction of charged particle tracks; each track being defined by five parameters (d_0, ϕ_0, ω, z_0, tanλ), measured at the point of closest approach to the z-axis. d_0 is the distance of this point from the origin in the $x-y$ plane, z_0 is the distance from the origin along the z-axis, the angle ϕ_0 is the azimuth of the track, λ is the dip angle relative to the transverse plane, and $\omega = 1/p_t$ is its curvature. The fitting algorithm searches the hits in both the SVT and DCH layers to find tracks, then iterates over the remaining hits to determine if any of these are consistent with a charged particle track. The efficiency of the fitting algorithm for tracks in the DCH as a function of transverse momentum and polar angle are shown in Figure 4.7. The drop in efficiency for finding charged tracks at low transverse momentum is due to limited information in the DCH for low transverse momentum particles. With the DCH and SVT



(a) Longitudinal cross-section of the DCH with principal dimensions.



(b) Layout of drift cells for the four innermost superlayers (lines between filed wires have been added to aid visualization of the cell boundaries).

Figure 4.6: Cross-sections of the drift chamber.

measurements combined, the uncertainty on the transverse momentum measurement is (p_t is measured in GeV/c):

$$\sigma_{p_t}/p_t = (0.13 \pm 0.01) \times p_t\% + (0.45 \pm 0.03)\% \tag{4.1}$$

Also crucial to this analysis, both the DCH and SVT make measurements of a particle's ionization loss per unit track length, dE/dx. This variable follows the Bethe-Bloch distribution, which depends on both particle mass and momentum. The dE/dx measurement in each of these systems, shown for the DCH in Figure 4.8, is crucial for PID, in particular for low-momentum K/π identification, and can be used to compute the likelihood for different particle hypothesis.

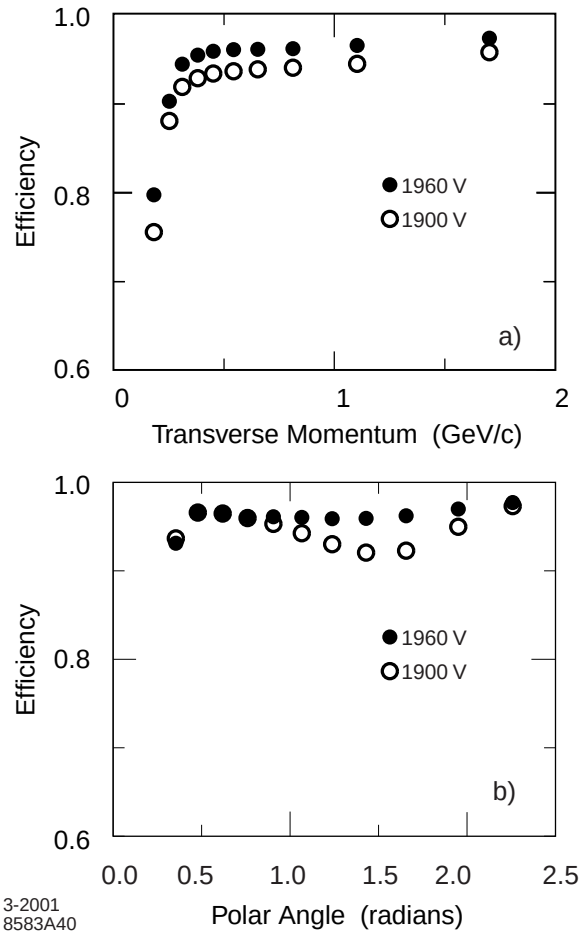


Figure 4.7: The efficiency of track reconstruction as a function of transverse momentum and polar angle at two DCH voltages.

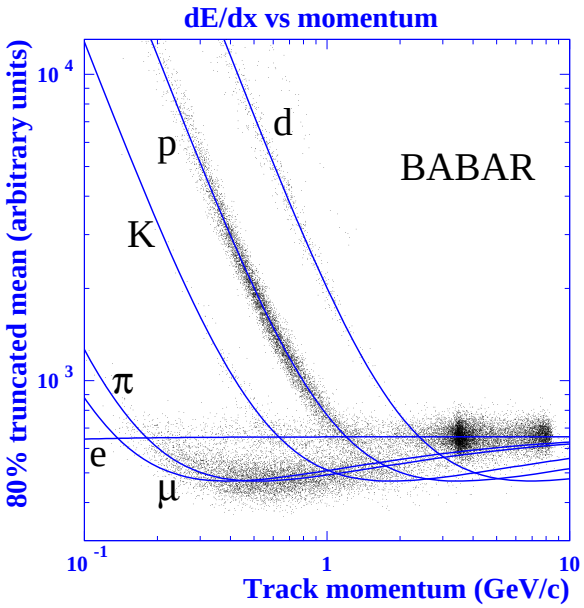


Figure 4.8: The dE/dx distribution as a function of momentum as measured by the DCH. The Bethe-Bloch curves for the different particle types are also shown.

4.2.3 Detector of Internally Reflected Cherenkov Light

The DIRC subdetector (shown schematically in Figure 4.9) is composed of a 12-sided, cylindrically symmetric, polygon barrel of 144 synthetic fused silica bars, each side of the polygon being composed of 12 bars glued together, and optically isolated from the other sides of the polygon. The DIRC provides PID for higher-momentum particles. The idea behind the DIRC relies on the Cherenkov radiation's opening angle remaining constant as the photons are internally reflected along the length of the DIRC bars until they are read out by an array of 10,752 photo-multiplier tubes (PMTs) set at the end of the detector away from the boost. As charged particles, detected by the tracking systems described above, enter the DIRC, they emit Cherenkov photons that are read by the PMT array. A time window of $\pm$300 ns around the trigger is imposed to cut down on the flat photon background detected by the PMTs. A vector, pointing from the center of the PMT to the center of the bar end, is extrapolated into the radiator bar, giving both θ_c, the Cherenkov angle of the emitted photons, and ϕ_c, the azimuthal angle of the Cherenkov photon around the track direction, up to a 16 fold ambiguity (emission on top or bottom of the bar, left or right,

forward or backward, and reflection or no-reflection in the wedge shown in Figure 4.9). By then turning to the timing information associated with the difference between the PMT detection time and expected photon arrival time, Δt, the ambiguity in photon angles is cut down from 16 to generally 3. Timing information also removes background photons that are either accelerator-induced or come from other tracks by a factor of approximately 40.

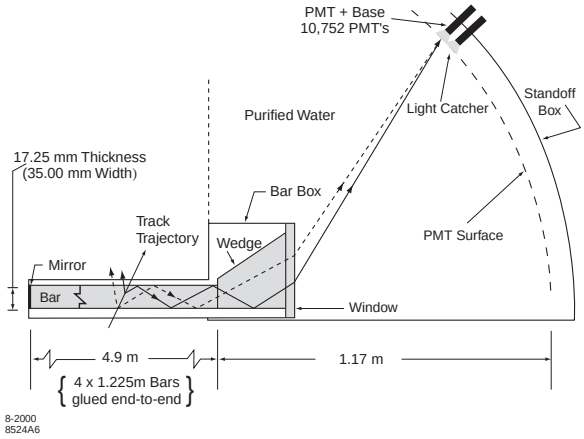


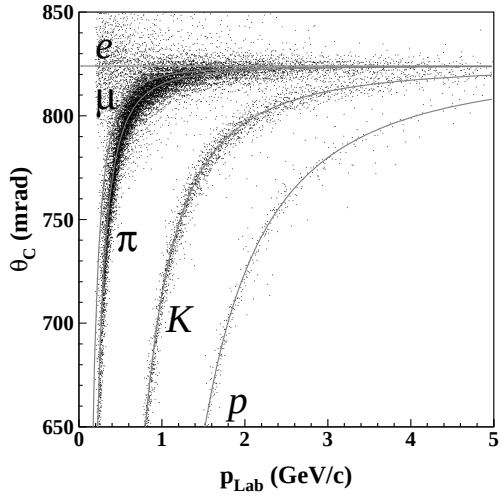
Figure 4.9: The longitudinal cross-section of the DIRC with an example particle included.

Due to the asymmetric energies of the beams, the majority of Cherenkov radiation is emitted in the front end of the DIRC bars. Approximately 28 photoelectrons are expected for a $\beta = 1$ particle entering normal to the surface at the center of the bar, increasing by more than a factor of two in the forward and backward directions. The resolution on the measured Cherenkov angle for a given track, $\sigma_{C,track}$, scales as:

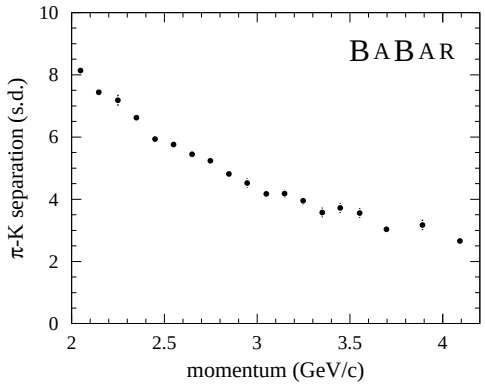
$$\sigma_{C,track} = \frac{\sigma_{C,\gamma}}{\sqrt{N_{pe}}}, \tag{4.2}$$

where $\sigma_{C,\gamma}$ is the single photon Cherenkov angle resolution, and N_{pe} is the number of detected photoelectrons. To preserve as many of the Cherenkov photons as possible, mirrors are inserted at the front end of the bars, allowing for the readout on the opposite end and greater coverage by other components (such as the EMC) in front of the boost. The measured Cherenkov angle can be combined with the measured particle momentum (from the tracking detectors discussed above) to determine a likelihood of the particle's type, as shown in Figure 4.10. This information is crucial for PID of higher-momentum particles,

and complements the tracking system's measurement of dE/dx.



(a) The expected Cherenkov angle for different particle species at different momenta.



(b) The performance of K/π separation (N_σ separation) for different momenta

Figure 4.10: The expected Cherenkov angle and K/π separation using the DIRC information.

4.2.4 Electromagnetic Calorimeter

The EMC is composed of 6,580 CsI(Tl) crystals to provide energy and angular resolution for photons and electrons, over a range of 20 MeV to 9 GeV. The properties of CsI(Tl) are given in Table 4.1. The crystals are arranged to provide full azimuthal angle coverage, and polar angle coverage from 15.8° to 141.8°, corresponding to 90% coverage in the center of mass frame. 5,760 crystals are arranged in 48 distinct rings of 120 crystals each, forming a cylinder around the interaction point, constituting the barrel region, while 820 crystals are arranged in 8 rings on a circular structure in front of the boost direction, constituting the end-cap.

The crystals range in lengths from 16-17.5 radiation lengths, depending on polar angle location, and are arranged with azimuthal symmetry. To minimize preshowering of the particles, the amount of material in front of the crystals (from other subdetector components) is minimized, and is in general between 0.3-0.6 radiation lengths. The light from the EMC crystals is detected with silicon photodiodes having a quantum efficiency of 85% for the

CsI(Tl) scintillation light. The uncertainty on the measured particle energy in the EMC is found to be:

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\%, \tag{4.3}$$

where the first term represents, in general, fluctuations in photon statistics and electronic noise, and the second term represents non-uniformity of light collection, absorption by material in front of (or between) the crystals, and calibration uncertainty.

Table 4.1: Relevant properties of CsI(Tl) crystals.

Parameter	Value
Radiation Length	1.85 cm
Moliere Radius	3.8 cm
Light Yield	50,000 $\gamma/$ MeV
Peak Emission λ_{max}	565
Signal Decay Time	680 ns (64%)
	3.34 μs (36%)

4.2.4.1 Light Yield Falloff

Over the course of the experiment, the crystals used in the EMC were subjected to intense levels of radiation, and consequently suffered radiation damage. As described in more detail in [23], the most common damage phenomenon comes from the development of absorption bands, commonly called *color centers*. These color centers reduce the light attenuation length of the crystal, and consequently the light yield of a given crystal. As shown for three different CsI(Tl) crystals in Figure 4.11, from [23][2] there is a strong correlation between the amount of absorbed radiation and the development of color centers, indicated by the decrease in transmittance at different wavelengths.

Because of the asymmetric nature of the experiment, crystals in the EMC endcap experienced higher levels of radiation than crystals in the EMC barrel, ranging from about 500 rad in the region of the barrel furthest from the boost direction, up to about 2200 rad in the endcap (this was monitored by 116 radFETs distributed throughout the EMC). The

[2]Reprinted from Publication Nucl. Instrum. Meth. **A413**, R. Zhu, Radiation damage in scintillating crystals, 297, Copyright (1998), with permission from Elsevier.

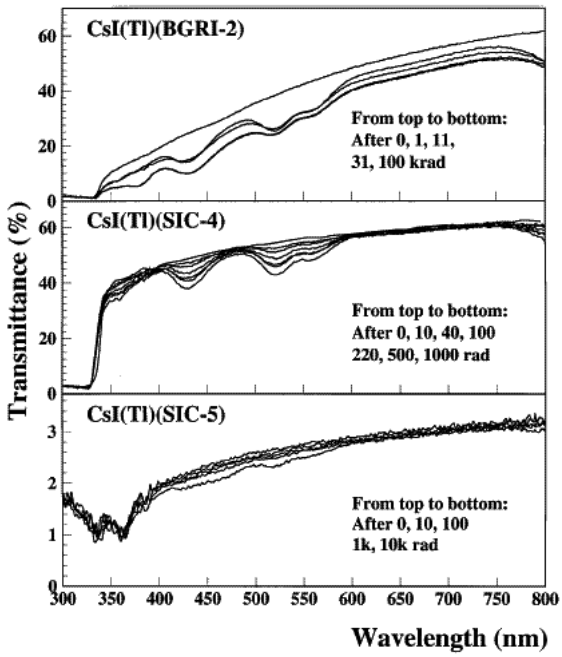


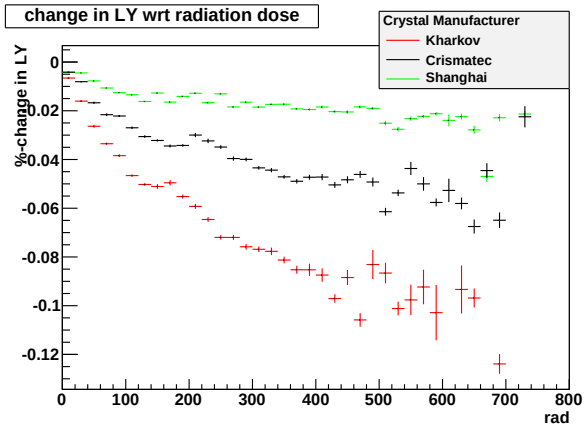
Figure 4.11: The falloff in transmittance for three crystals as a function of absorbed radiation dose.

resulting change in light yield for each crystal needed to be monitored and corrected for over the run of the experiment. Calibrations were performed at the two ends of the dynamic range of the detector: the low energy calibration coming from a 6.13 MeV radioactive photon source, and the high-energy calibration (3-9 GeV) coming from Bhabha events.

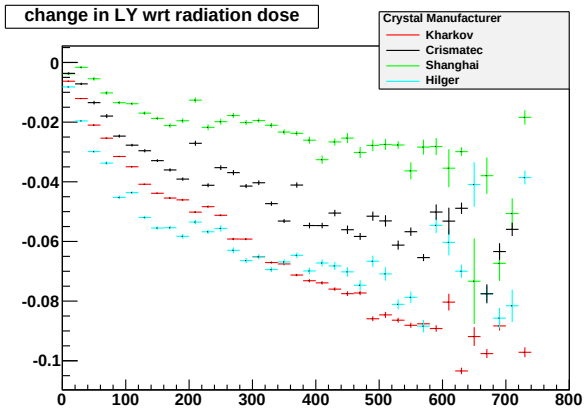
The change in light yield as a function of absorbed radiation dose, based on the low energy calibration data, is shown in Figure 4.12, categorized by crystal manufacturer. Though great care was taken to produce uniformity between the crystals, the crystals displayed various degrees of stability in their performance over the run of the experiment. Depending on the manufacturer, the total falloff in light yield ranges from a few % in the region of the barrel furthest from the boost (barrel backward), up to a few tens of % (on average) in the endcap.

4.2.5 Instrumented Flux Return

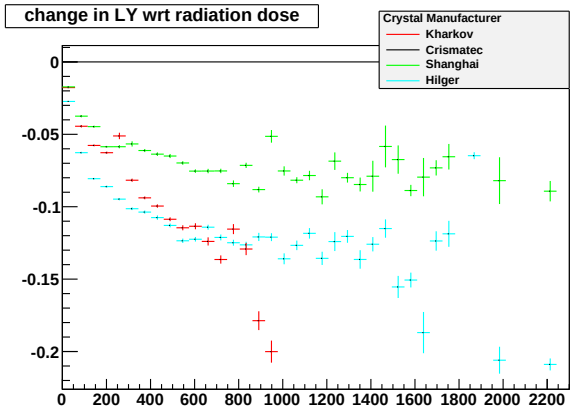
The IFR uses the *BABAR* 1.5T magnet's steel flux return as a muon filter and hadron absorber. The primary purposes of the IFR are to identify higher-momentum muons (generally



(a) Barrel Backward



(b) Barrel Forward



(c) Endcap

Figure 4.12: The percent change in light yield over the run of the experiment, plotted with respect to absorbed radiation dose (rad). The different crystal manufacturers are indicated.

down to $\sim$1 GeV/c) and to detect neutral hadrons (primarily K_L^0 and neutrons) that do not decay within the other subdetector volumes.

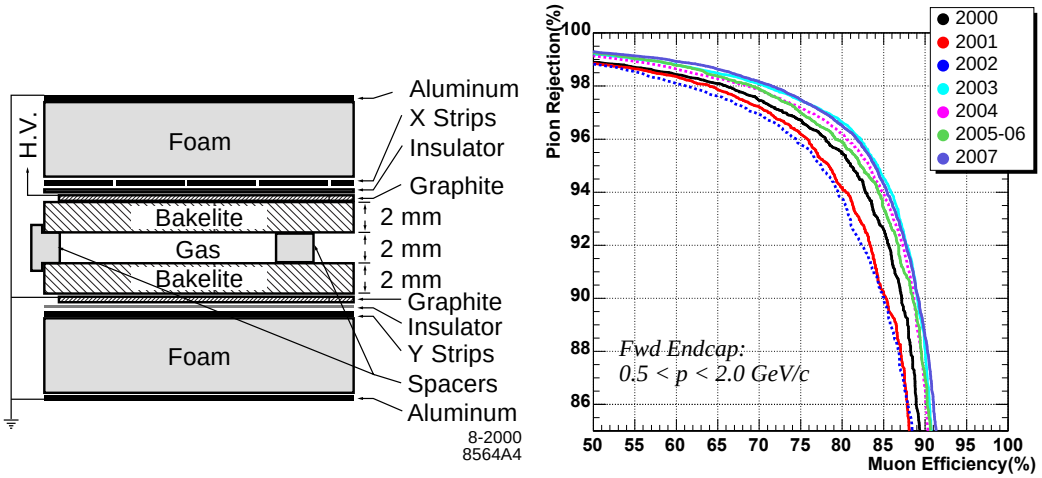
During the initial runs of the experiment, the gaps between the steel plates that make up the flux return were instrumented with resistive plate chambers (RPCs), shown schematically on the left in Figure 4.13. There are 19 instrumented layers in the barrel of the IFR, 18 in each of the two endcaps, as well as two cylindrical layers installed between the EMC and the magnet cryostat to detect particles exiting the EMC. The RPCs consisted of either two bakelite sheets with graphite coatings, read out capacitively by aluminum strips on a mylar substrate (in the barrel and endcaps), or ABS plastic and a conducting polymer (in the cylindrical layers). The RPCs were filled with a freon-argon-isobutane gas mixture, and detected streamers from ionizing particles via capacitive readout strips. As shown in the figure, each RPC within a given layer consisted of two orthogonal readout strips, allowing for readout of two coordinates in each RPC (x and y in the endcap, or ϕ and z in the barrel and cylander). The total number of channels read out from the RPC system is approximately 53,000, covering a total of $\sim$2,000 m^2.

As shown in the left plot in Figure 4.13, the performance of the IFR using RPCs showed substantial decline after the start of the experiment. This was attributed to high-temperature conditions at the beginning of the experiment, which accelerated the breakdown of the graphite and linseed oil coating. Between 2004 and 2006, most of the RPCs in the layers of the different IFR modules were replaced with limited streamer tubes (LSTs) to regain muon/pion discrimination. Details of the replacement may be found in [24]. As shown in the figure, this had a dramatic effect in improving IFR performance.

It should be noted that, since K_L^0 are considered undetected particles in this analysis, the change in performance of the IFR has limited impact.

4.2.6 Trigger

The *BABAR* triggering system selected events of interest at a total rate of $\sim$120 Hz, with the goal of being flexible for different operating conditions. The triggering system was broken into a two-level hierarchy with a Level 1 (L1) trigger from hardware followed by a Level 3



(a) The schematics of the RPCs initially used in the IFR

(b) The performance by year of the IFR at discriminating between μ/π. Both axes indicate %.

Figure 4.13: The schematics of an RPC within the IFR, and the change in performance of the IFR at distinguishing between μ/π.

(L3) trigger implemented in software. A level 2 trigger was never implemented because it was found to not be necessary (the naming convention had already been chosen, however). The L1 trigger was configured to have an output rate of typically 1 kHz during normal operation, broken down by different sources in Table 4.2 for an instantaneous luminosity of $3\times10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (compared to the record luminosity during the run of the experiment of about $12\times10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

Table 4.2: The cross-sections, production and L1 trigger rates for different physics processes at 10.58 GeV and a luminosity of $3\times10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The e^+e^- rate refers to events in which one or both leptons are found in the EMC.

Event Type	Cross-section (nb)	Production Rate (Hz)	L1 Rate (Hz)
$b\bar{b}$	1.1	3.2	3.2
other $q\bar{q}$	3.4	10.2	10.1
e^+e^-	~ 53	159	156
$\mu^+\mu^-$	1.2	3.5	3.1
$\tau^+\tau^-$	0.9	2.8	2.4

The main subdetector components of the L1 trigger were the DCH, triggering based on

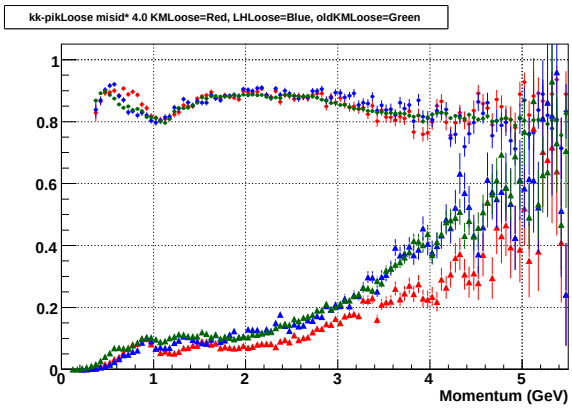
charged tracks above a threshold transverse momentum, and the EMC, triggering based on shower energy deposited. Trigger primitives from each of these subsystems were created and could combined in any number of ways to accept events of interest and prevent or prescale the recording of unnecessary background events. In general, the L1 trigger efficiency rate for $B\overline{B}$ events exceeded 99.9%.

The L3 trigger took the output from the L1 and improved both track fitting in the DCH and energy cluster sensitivity in the EMC. It filtered out events that were not of interest, either no particle with a high enough transverse momentum in the DCH or not enough energy deposited in the EMC for instance, and flaged events that passed. The combined L1 and L3 triggers had an efficiency in excess of 99.9% for $B\overline{B}$ events.

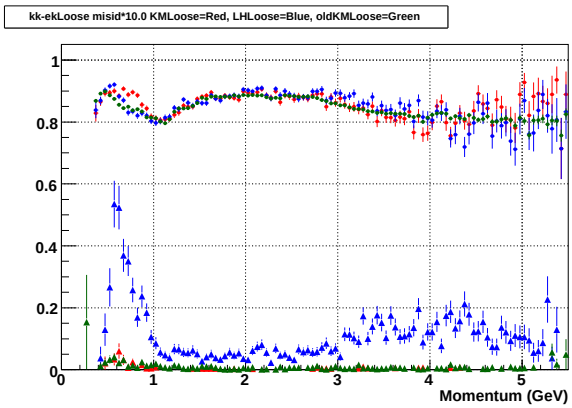
4.2.7 Particle Identification

This analysis relies heavily on the information from the particle identification; in particular on identifying kaons in the final state of the s quark. The PID algorithms at *BABAR* have gone through many changes over the run of the experiment, from simple likelihood ratios based on variables output from the subdetectors mentioned above, through more complex single neural network systems that classify particles by type using these same variables.

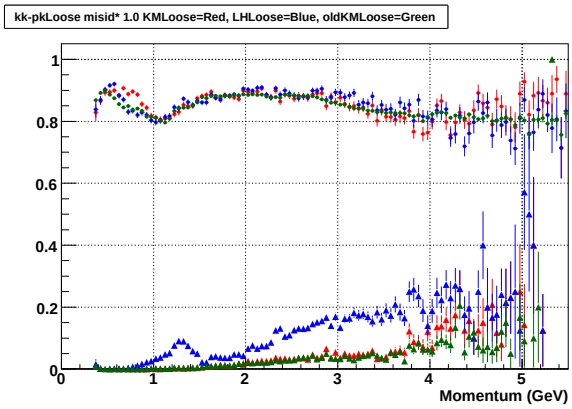
The method used for this analysis, the latest algorithm used in the experiment, is based on error-correcting output code (ECOC) [25], discussed in more detail in [26]. This algorithm is a way to use a collection of binary classifiers, such as Random Forests or Neural Networks, which by construction only separate signal from background, to create a multi-class classifier, able to separate signal A, from signal B, from signal C, *etc.*. For PID, the ECOC creates a multi-class classifier to separate kaons, pions, protons and electrons. The improvement over earlier PID techniques is demonstrated by the increased accuracy in kaon identification shown in Figure 4.14 using classifiers based on the ECOC algorithm. The most important variables for kaon-pion separation are shown, by classifier, in Figure 4.15.



(a) The efficiency of kaon identification and 4 times the pion mid-id rate.

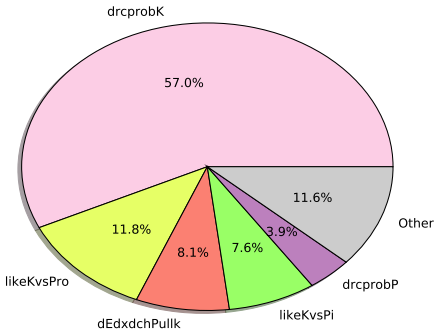


(b) The efficiency of kaon identification and 10 times the electron mid-id rate.

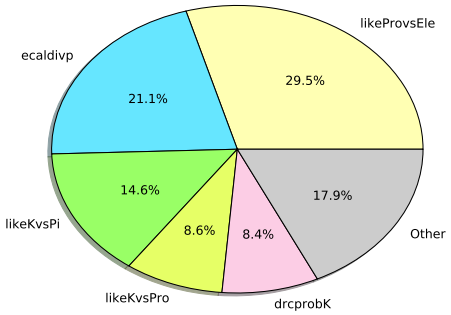


(c) The efficiency of kaon identification and 10 times the proton mid-id rate.

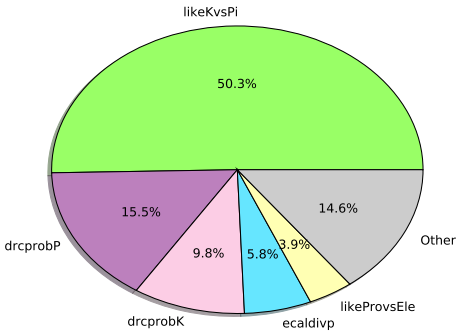
Figure 4.14: The efficiency of finding a kaon compared to mis-identifying a different type of particle as a kaon. The ECOC method is shown in red, the previous methods employing a likelihood selector and single neural network are shown in blue and green.



(a) Important variables in kaon vs. pion, electron, or proton discrimination.



(b) Important variables in kaon or electron vs. pion or proton discrimination.



(c) Important variables in kaon or proton vs. pion or electron discrimination.

Figure 4.15: The most important variables used for discriminating between the given particles. The size of the pie wedge indicates the importance of the given variables; indicated on the periphery.

Chapter 5

Preliminary Event Selection

5.1 Event Skims and B Reconstruction

Event selection and reconstruction occurs in several steps. First, the data collected by *BABAR* that passed the trigger is "skimmed" to retain events that are of interest to this analysis, while rejecting events that are not consistent with a $b \to s\gamma$ transition. The main requirements we impose at this stage are that the event must have at least three defined tracks, must have a photon with an energy $E_\gamma = [1.6, 3.0]\,\mathrm{GeV}$, in the center of mass frame, and must have a normalized second Fox-Wolfram moment [27] < 0.9. This latter requirement cuts out more jet-like events, consistent with lighter-diquark production.

The skim also reconstructs the B meson candidates in the events. The different B meson candidates are taken from combinations of different tracks and photons detected in the event. It is important to note that for any event, there may be several B meson candidates. We store all of the candidates that pass our skim requirements and select the one most consistent with a $b \to s\gamma$ transition at a later stage of the analysis.

For the charged tracks in the event, the skim procedure is very efficient and does not impose too many requirements. However, since we are trying to accurately measure the X_s spectrum, we want our resolution on the invariant mass to be as precise as possible; we therefore reconstruct K_s candidates only in charged pion final states ($K_s \to \pi^+\pi^-$). Similarly, the η mesons are reconstructed only in the $\eta \to \gamma\gamma$ final state. Other final states of the η, such as $\eta \to \pi^+\pi^-\pi^0$ are already implicitly included in our list of final states, Table 3.1, as long as the total particle multiplicity is not otherwise too high.

The reconstructed B meson candidates must also pass some event shape and kinematic requirements. The most important of these requirements in the context of this analysis is that the B meson candidate have:

$$|\Delta E| < 0.4\,\mathrm{GeV}, \tag{5.1}$$

$$m_{ES} > 5.21\,\mathrm{GeV}/c^2. \tag{5.2}$$

Indeed these two variables are often exploited by *BABAR* analysis, as they provide good discrimination between B and non-B events. The variable m_{ES} peaks at the mass of the B meson ($5.279\,\mathrm{GeV}/c^2$) for a collection of correctly reconstructed candidates and $|\Delta E|$ peaks at 0.

5.2 Data Used

For this analysis, we use the entirety of the *BABAR* data set collected at the $\Upsilon(4S)$ resonance ("on peak" data). For a control sample of our non-B background events, we also use the data collected 40 MeV below the $\Upsilon(4S)$ resonance ("off peak" data). This data was collected over several years, and is grouped in different scientific "runs." The amount of data used in this analysis, broken down by run and given in integrated luminosity, is given in Table 5.1.

Table 5.1: The run-by-run integrated luminosity of the on peak and off peak data.

	On Peak (fb^{-1})	Off Peak (fb^{-1})
Run 1	20.60	2.62
Run 2	62.07	7.03
Run 3	32.67	2.50
Run 4	100.78	10.23
Run 5	133.85	14.55
Run 6	79.03	7.89
total	429.0	44.81

We perform this analysis "blind," meaning we simulate what we expect to find when we look at the data using MC, finalize our analysis procedure, and only then do we "open

the box," or look at the data. We do this to avoid experimenter's bias, minimizing the potential to be influenced by prior experimental results or to try to construct better results than the data may actually support. In order to do this, we generate MC for the different types of physics events we expect to find in the data, generally at several times the amount we expect to find in data, in sets based on the runtime conditions of the detector. We simulate both the background events we will later try to remove using different requirements on the B candidates and the signal events. The background events are generally referred to as either "B"-background (simulated with generic B meson MC), "continuum" background (background that originates generally from $c\bar{c}$, uds, or $\tau^+\tau^-$ production), or "cross-feed" background (background that comes from an actual $b \to s\gamma$ transition, in which the final products of the s-quark hadronization are not correctly selected). As mentioned in section 2.2, we model the signal with two sets of MC: one that models exclusively the $B \to K^*\gamma$ transition, and one that we use for the X_s mass region above the K^* resonance. The background MC used in this analysis is given in Table 5.2, and the signal MC is given in Table 5.3.

Table 5.2: The run-by-run number of MC events, cross-section, equivalent luminosity, and weighting factor for each background mode.

Process	Generated Events ($\times 10^6$)	cross-section (nb)	Equivalent Luminosity (fb^{-1})	Weighting Factor (data/MC)
$e^+e^- \to u\overline{u}$, $d\overline{d}$, $s\overline{s}$	44.59	2.09	21.33	0.965
	185.90		88.949	0.698
	137.54		65.809	0.496
	409.30		195.839	0.515
	526.64		251.979	0.531
	327.03		156.475	0.505
Total $u\overline{u}$, $d\overline{d}$, $s\overline{s}$	1631.00		780.39	0.550
$e^+e^- \to c\overline{c}$	55.25	1.30	42.503	0.485
	164.72		126.709	0.490
	88.32		67.939	0.481
	267.31		205.622	0.490
	344.28		264.827	0.505
	208.66		160.511	0.492
Total $c\overline{c}$	1128.54		868.11	0.494
$e^+e^- \to \tau^+\tau^-$	19.69	0.94	20.944	0.983
	57.19		60.845	1.020
	49.00		52.130	0.627
	180.08		191.571	0.526
	237.09		252.228	0.531
	137.66		146.446	0.540
Total $\tau^+\tau^-$	680.71		724.16	0.592
$e^+e^- \to B^+B^-$	34.88	0.55	64.415	0.320
	105.56		191.929	0.323
	56.04		101.882	0.321
	166.78		303.244	0.332
	215.17		391.215	0.342
	130.34		236.975	0.333
Total B^+B^-	708.76		1289.66	0.333
$e^+e^- \to B^0\overline{B}^0$	34.94	0.55	63.527	0.324
	104.19		189.433	0.328
	57.89		105.251	0.310
	169.80		308.729	0.326
	215.95		392.642	0.341
	135.22		245.862	0.321
Total $B^0\overline{B}^0$	718.00		1305.44	0.329

Table 5.3: The run-by-run number of signal MC events generated and luminosity weighting factors.

Process	Generated Events ($\times 10^6$)	cross-section (nb)	Luminosity (fb^{-1})	Weighting Factor (data/MC)
$B^+ \to X_s\gamma$	0.632	$0.55 \times 2 \times 3.56 \times 10^{-4}$	1613.9	0.0128
	1.89	$= 0.0003916$	4816.1	0.0129
	0.99		2535.8	0.0129
	3.04		7755.4	0.0130
	4.00		10206.8	0.0131
	2.35		6006.1	0.0132
Total $B^+ \to X_s\gamma$	12.90		32934.1	0.0130
$B^0 \to X_s\gamma$	0.632	$0.55 \times 2 \times 3.56 \times 10^{-4}$	1613.9	0.0128
	1.89	$= 0.0003916$	4816.1	0.0129
	0.99		2535.8	0.0129
	3.04		7755.4	0.0130
	4.00		10206.8	0.0131
	2.35		6006.1	0.0132
Total $B^0 \to X_s\gamma$	12.90		32934.1	0.0130
$B^+ \to K^{*+}\gamma$	0.317	$0.55 \times 2 \times 4.03 \times 10^{-5}$	7150.9	0.0029
	0.94	$= 4.433 \times 10^{-5}$	21249.7	0.0029
	0.50		11188.8	0.0029
	1.52		34220.6	0.0029
	2.00		45071.1	0.0030
	1.18		26596	0.0030
Total $B^+ \to K^{*+}\gamma$	6.45		145477	0.0029
$B^0 \to K^{*0}\gamma$	0.317	$0.55 \times 2 \times 4.03 \times 10^{-5}$	7150.9	0.0029
	0.94	$= 4.433 \times 10^{-5}$	21249.7	0.0029
	0.50		11188.8	0.0029
	1.52		34220.6	0.0029
	2.00		45071.1	0.0030
	1.18		26596	0.0030
Total $B^0 \to K^{*0}\gamma$	6.45		145477	0.0029

Chapter 6

Final Event Selection

In order to further reduce our backgrounds after the skim selection, we place slightly tighter requirements on the B candidates of $|\Delta E| < 0.15$ and $5.24 < m_{ES} < 5.3$. Even after these requirements, however, many background events and false B candidates remain. This chapter will describe the procedures we use to minimize the background and select the best B candidate in an event.

6.1 π^0 Veto

After placing the preliminary requirements on our MC given above, we find a large number of background events have a high-energy photon coming from a π^0 decay, as shown in Figure 6.1. Rather than simply removing any photon that can be combined with another photon in the event to give an invariant mass consistent with the π^0, we can improve our signal efficiency if we use a multivariate classifier known as a random forest [28]. We use the random forest algorithm, SprBaggerDecisionTreeApp, developed in the StatPatternRecognition suite of tools [29].

In order to most effectively use the random forest technology to reduce our π^0 background sources, we identify the two most important variables that show good π^0 discrimination: the invariant mass of the high-energy photon combined with any other photon in the event, and the energy of the lower-energy photon. We only assign a classifier response to high-energy photons if the invariant mass of the pair falls in the window $\epsilon[0.11, 0.15]\,\mathrm{GeV}/c^2$, otherwise the classifier response defaults to a value of -1.

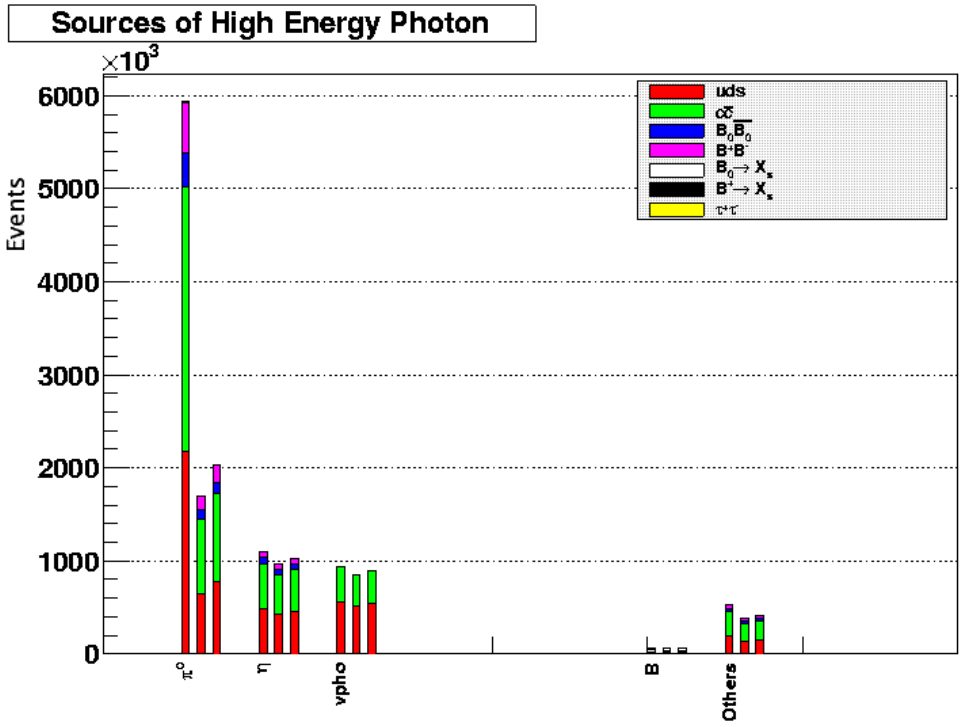


Figure 6.1: The source of the high-energy photon from our different types of background events. The y-axis indicates the number of events from these sources. The three stack plots represent: no π^0 veto, a simple mass window cut $\epsilon[0.11, 0.15]\,\mathrm{GeV}/c^2$, and a cut on the π^0 veto described in the text.

The random forest classifier is then trained on 0.25% of our MC to find a true π^0, using all of our MC types (both signal and background). Using the random forest effectively lets us recover high-energy signal photons from $b \to s\gamma$ transitions that we would otherwise remove were we to employ only a mass window requirement. The random forest response, constructed to be between 0 and 1 if the photon can be combined with another photon in the event to form an invariant mass between $\epsilon[0.11, 0.15]\,\mathrm{GeV}/c^2$, is shown in Figure 6.2. As mentioned above, the response of the classifier is -1 if the photon cannot be combined with another photon to create an invariant mass in this window.

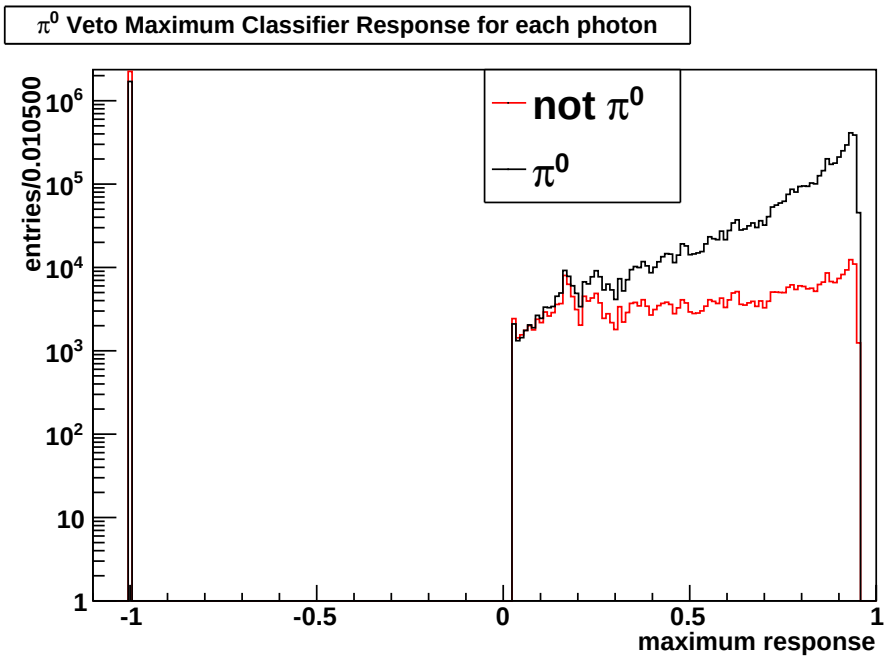


Figure 6.2: The response of the π^0 veto classifier to both true π^0s and false π^0s. The classifier is constructed to have a response between 0 and 1, 1 indicating more likely for the photon to originate from a π^0. A response of -1 indicates the high-energy photon cannot be combined with another photon to form a particle consistent with the π^0 mass hypothesis.

Rather than place a requirement on the output of the π^0 veto, we choose instead to use it as input to a more generic background-rejecting classifier (BRC) later in the analysis procedure. Also, it should be noted, that we do not have a dedicated veto for η particles that fake our high-energy photon (the second largest source of background, as shown in Figure 6.1). We find this is not necessary, as a simple mass window cut of arbitrary width decreases our figure of merit (FOM), which we take to be $S/\sqrt{S+B}$ (where S is signal and

B is background), and a classifier-based veto would likely only have marginal improvement at best (and a negative impact on the FOM at worst).

6.2 Background Rejecting Classifier, BRC

After constructing our π^0 veto described above, and having applied our preliminary requirements, we still have a lot of background events that need to be removed in order to make fitting (see below, Section 7) the signal contribution feasible. Before describing the different handles we use to reject the background events, we should first highlight what constitutes a background event at this stage of the analysis.

Every event, background or signal, in our MC sample, in general has several X_s candidates within it (potential final states containing the event's s quark). One reason for this is the multitude of final states we are reconstructing as our signal, and the fact that states of lower multiplicity are not orthogonal to states of higher multiplicity: An event that has an X_s candidate reconstructed as one kaon and four pions will have at least four potential candidates of one kaon and three pions, six potential candidates of one kaon and two pions, and four potential candidates of one kaon and one pion; all these are signal modes we reconstruct. Clearly some of these X_s candidates will not satisfy our preliminary requirements, but it is easy to see why we cannot equate the number of background candidates with the number of background events. However, in order to effectively remove background events, we need to determine the best way of removing every spurious candidate from the event.

Because one can argue that developing an algorithm for choosing the best candidate in an event (presented below in Section 6.3) will be more effective if one only focuses on more signal-like events, it is with the mind set of having multiple candidates in each event that we approach rejecting background at this point. Hence, we describe in this section the procedure we developed for rejecting background events, in particular *uds* and $c\overline{c}$ background events, before having selected the best candidate within a given event.

To build our BRC, we input several variables (see below in Section 6.2.1). We do not use variables that show a correlation with the truth matched signal MC's X_s mass, such as the normalized Fox-Wolfram moment (this correlation is at least partly due to the direct

correlation between X_s mass and final state multiplicity), even if they show good background rejection capabilities. We reserve information about the X_s mass for the signal selecting classifier detailed in Section 6.3. We deem it more important to have information about the mass of the candidates there, as the mass range will be correlated with several other variables specific to selecting the best B candidate. Also, since the majority of variables used in constructing these random forests have stronger separation power between signal and continuum background ($c\bar{c}$ and uds), we only train the random forest to find truth matched signal B candidates in our signal MC (B candidates in which the entire MC chain indicates we have reconstructed a genuine $\overline{B} \to X_s\gamma$ transition), embedded in a continuum background.

Finally, we performed a study to determine how much rejection power could be gained by training a random forest classifier in each bin of X_s mass as supposed to the whole mass range. We found the results to be comparable, and elected to only train one BRC rather than several.

6.2.1 Variables Used in BRC

Here we present the variables we included in the BRC as well as a brief description of each. The correlation between each pair of variables can be seen in Figure 6.3.

- **BmtmFlow1-17**: Momentum flow cones in the CM frame around the reconstructed B direction in 10° increments. The idea is that the distribution of energy can separate more jet-like continuum events from isotropic/spherical signal.

- **"Rest of Event" (ROE) Legendre Monomials along the Photon Axis**: The zeroth, first, and second order Legendre monomials computed in the CM frame along the primary photon's axis. The ROE corresponds to all particles not involved in signal B reconstruction.

- L_{12}/L_{10}: The ratio of the second order Legendre monomial to the zeroth when both are computed using the ROE particles (particles not involved in signal B reconstruction) along the primary photon's axis.

- $|\cos\theta^*_B|$: The absolute value of the cosine of the angle of the B flight direction with respect to the z-axis, computed in the CM frame.

- $|\cos\theta^*_T|$: The absolute value of the cosine of the angle between the thrust axis of the B candidate and the thrust axis of the ROE (particles not involved in B reconstruction), calculated in the CM frame.

- $|\cos\theta^*_{\gamma T}|$: The absolute value of the cosine of the angle between the primary photon and the ROE particles' thrust, calculated in the CM frame.

6.2.2 BRC Final Training

For the final BRC, we scan a range of settings for the SprBaggerDecisionTreeApp. First, we split our MC sample into orthogonal sets of 5% training, 5% validation, and 90% testing, in order to limit the amount of MC trained upon to a more manageable level. This separation also avoids over training the classifier: we train the classifier on the training set of events, trying many settings for the classifier, we choose the optimal setting based on the validation set's performance (the training set would, in general, have overly optimistic signal/background separation performance), and we evaluate the expectations in the data based on the testing set's performance. Prior to training, we also place the preliminary cuts:

- $m_{ES} > 5.265\,\mathrm{GeV}/c^2$

- $|\Delta E| < 0.15$

- $E_{\gamma min}$ in π^0 candidates > 0.050 GeV

- Kaons come from KaonKMLoose list

- $1.1 < m_{X_s} < 2.8$

The fourth requirement indicates our choice of PID criteria for the charged kaons, using the PID based on ECOC described in Section 4.2.7. The final requirement is used to focus the training of the background rejecting classifier on the region above the K^* resonance

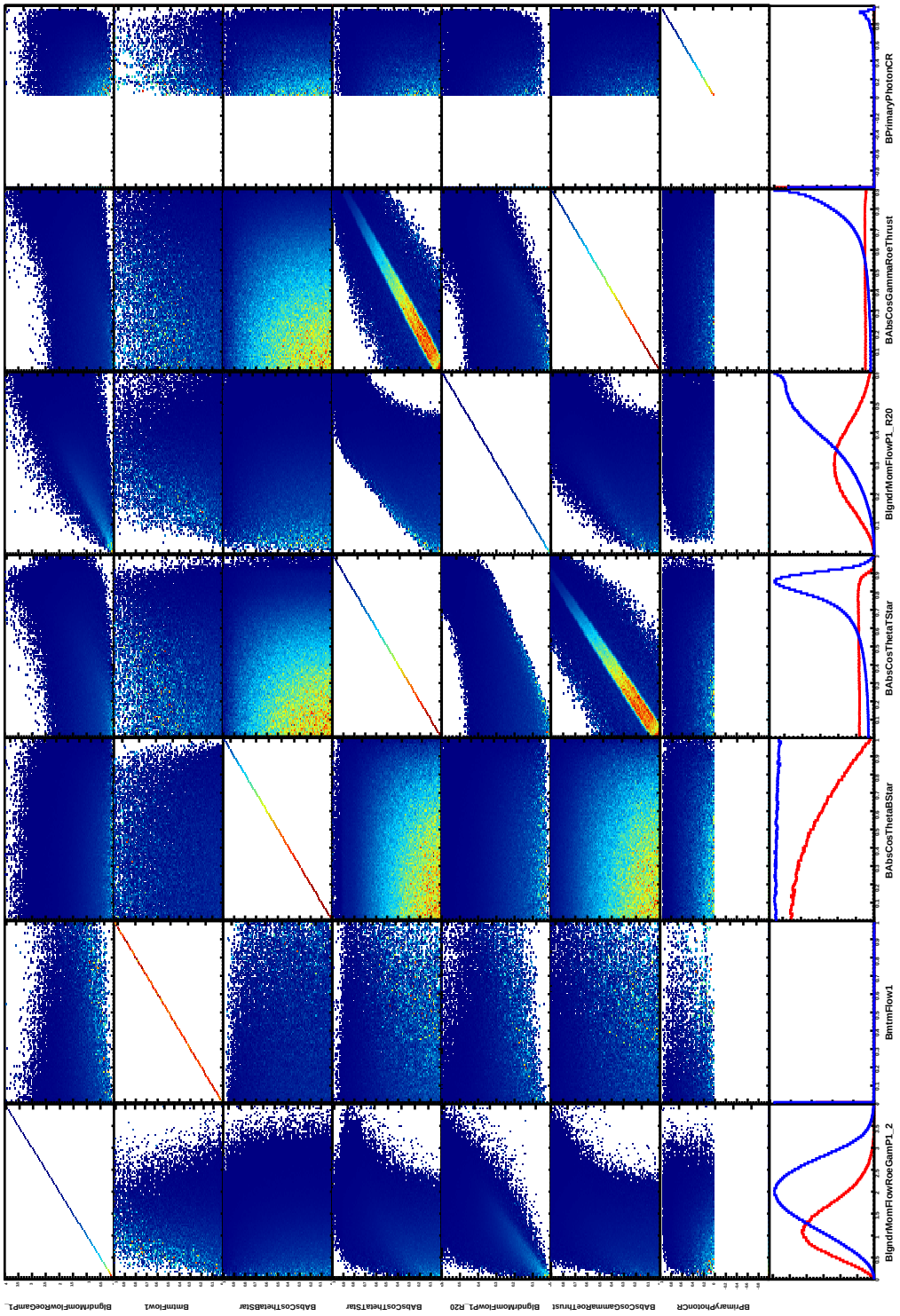


Figure 6.3: Correlation plot of variables used in BRC training. B Momentum Flow 2-17 is omitted from the plot. In the 2D histogram, red represents a higher concentration of signal events and blue represents a lower concentration of signal events. The 1D histograms show signal (red) and background (blue).

($m_{X_s} < 1.1\,\mathrm{GeV}/c^2$). Ideally, this classifier is uncorrelated with the X_s mass, but higher order correlations may be present, and we wanted to ensure that the classifier does not train on signal/background separation in this region at the expense of higher-mass regions. Finally, the signal MC used to train the classifier was weighted to model a photon spectrum consistent with a model [16] consistent with previous measurements of this decay ($m_b = 4.65, \mu_\pi^2 = 0.2, \mu_g = 0.27$); this weighting will be our default photon spectrum weighting unless otherwise specified. Based on the validation MC sample (independent of both the training MC and testing MC samples), we find the best setting for the classifier to be a random forest with a terminal leaf size of at least 80 events, sampling at most 12 variables for each decision tree. We determine this setting to be the "best" based on the validation MC sample's highest FOM (again, using the "precision" FOM$\equiv S/\sqrt{S+B}$), reweighting signal according to our usual photon spectrum model. More specifically, we trained classifiers with many settings, then applied cuts on the outputs of the resulting classifiers at many different values (from 0 to 1 with increments of 0.01); the highest FOM we found in the validation MC came for these settings when training the classifier.

We also check to ensure that the classifier is not introducing unwanted structure into the m_{ES} distribution, which we will later fit to obtain signal yield, by having different performances inside and outside the region $m_{ES} > 5.265$ over which it was trained. We demonstrate the uniform response to background for run 3 MC both inside the m_{ES} region over which the BRC is trained, and outside this region in Figure 6.4. This plot shows all candidates that pass our preliminary selections, not only the best candidate (shown in similar figures below), and each type of background source is weighted to reflect expected amount in the data and then renormalized to have unit area. The figure shows that there is negligible correlation with m_{ES} region.

6.3 Signal-Selecting Classifier, SSC

While the classifiers described to this point are effective at removing background events, we have not identified which of the multiple B candidates in an event is most consistent with the $b \to s\gamma$ hypothesis. As mentioned above, since we are reconstructing 38 final states of

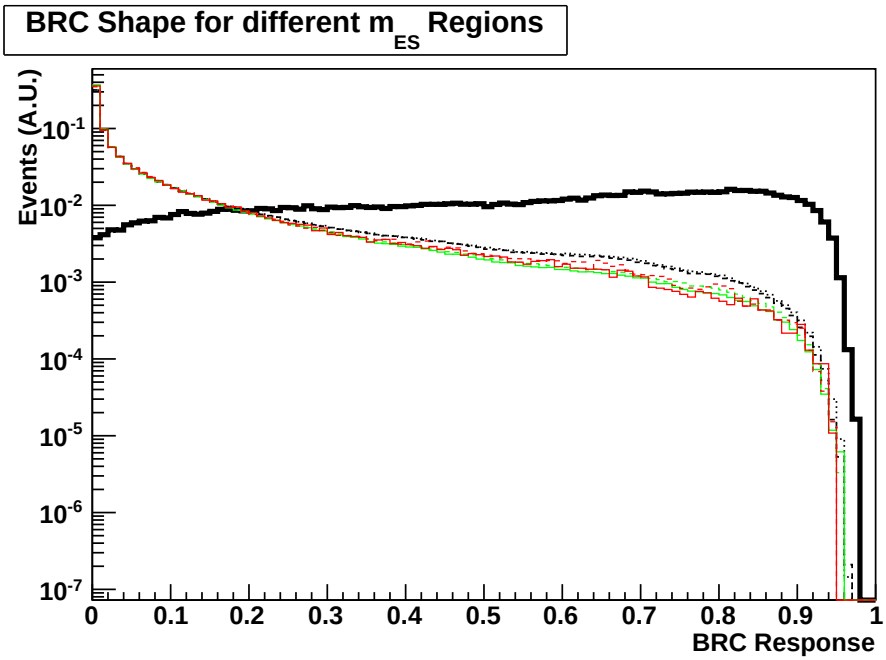


Figure 6.4: The response of the BRC (x-axis) for background candidates with $m_{ES} > 5.265$ (dashed), background candidates with $m_{ES} < 5.265$ (dotted), and signal candidates (solid line) for run 3 MC. We also compare off peak data (red) to continuum MC (green) for $m_{ES} < 5.265$ (solid) and $m_{ES} > 5.265$ (dashed). Further details about the normalization of the background are given in the text.

the B meson, there are in general many ways we can reconstruct a B candidate for a given event, and selecting the "best" one is of critical importance. To effectively do this selection, we construct a dedicated classifier, a signal selecting classifier (SSC). The variables we use in this classifier are:

- The mass of the reconstructed X_s candidate.
- Minimum π^0 momentum used in the B reconstruction (if applicable), measured in the CM frame.
- Thrust of the reconstructed B candidate.
- ΔE, defined above, normalized by the resolution: $\frac{\Delta E}{\sigma_E}$
- Zeroth Fox-Wolfram moment.
- Fifth Fox-Wolfram moment.

The correlations between these variables can be seen in Figure 6.5. We train the classifier to find truth-matched signal candidates from signal events with the cross-feed background included. Since this classifier is dedicated to finding the "best" signal candidate in a signal event, we do not include backgrounds from other sources during training.

To reduce the number of mis-reconstructed candidates in the training sample, we use only candidates which have $m_{ES} > 5.27$, a requirement that retains most of the true candidates and effectively reduces cross-feed background. After performing a training procedure similar to the BRC (separate the MC into training, validation, and test samples, try many settings of the classifier, *etc.*) we find the optimal settings for the classifier. The response of the classifier to signal and cross-feed, as well as the other sources of background events, can be seen in Figure 6.6.

The choice to use a classifier to select the best B candidate in the event is a break from how analyses have previously made this selection. Previous analyses generally take the candidate that minimizes $|\Delta E|$ to be the best B candidate, but we found very strong improvement over this method by combining this variable with others in the event into a classifier, as shown in Figure 6.7.

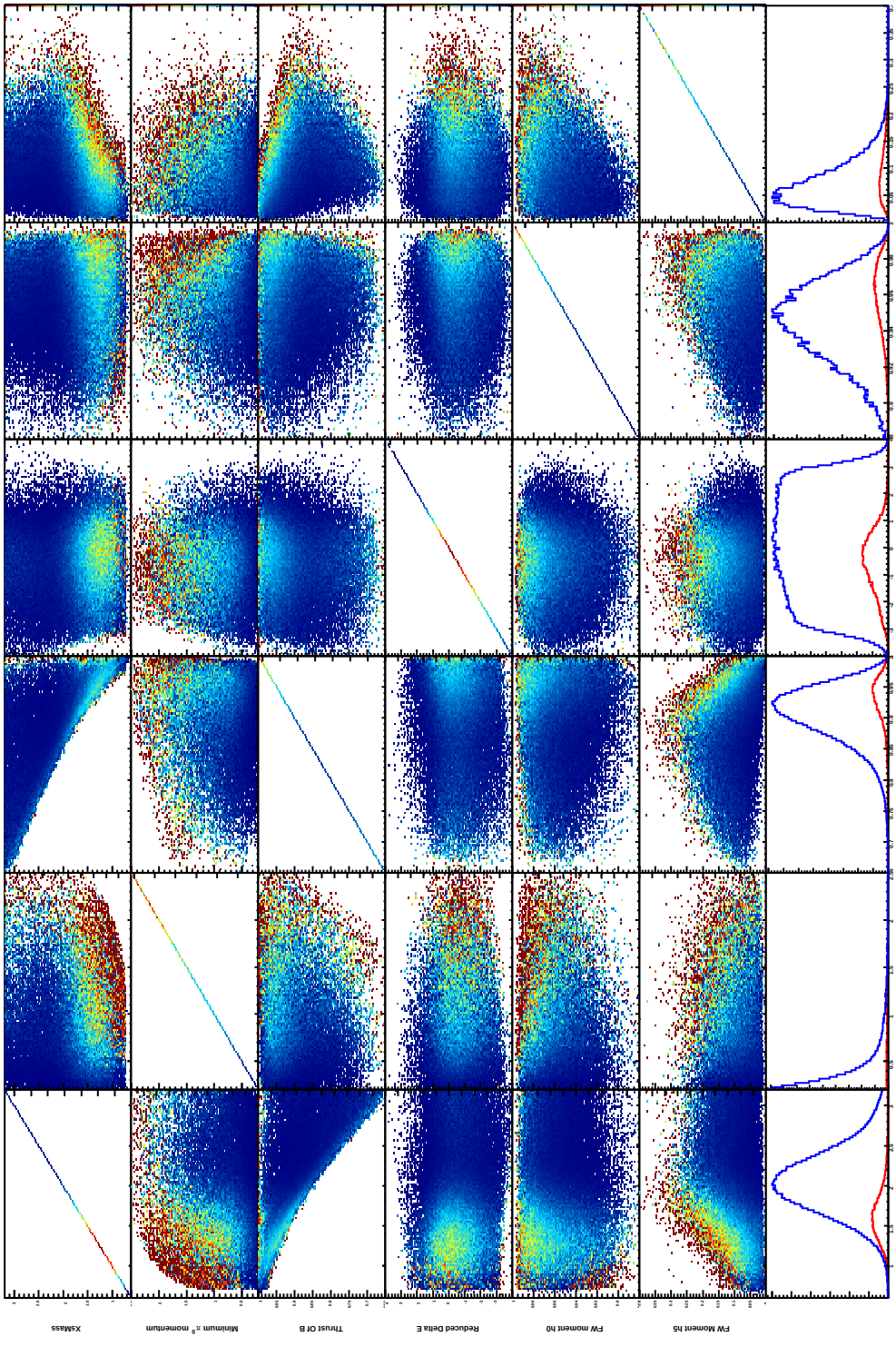


Figure 6.5: Correlation plot of variables used in SSC training. In the 2D histogram, red represents a higher concentration of true signal candidates and blue represents a lower concentration of true signal candidates.

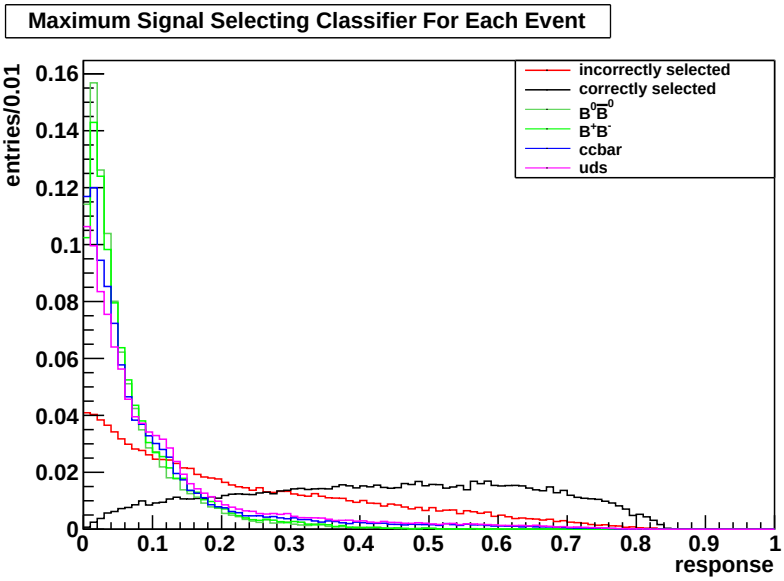


Figure 6.6: Normalized distribution of maximum SSC response for both signal and background events. We show signal MC in which the true signal candidate is selected (red) and in which the wrong candidate is selected (black).

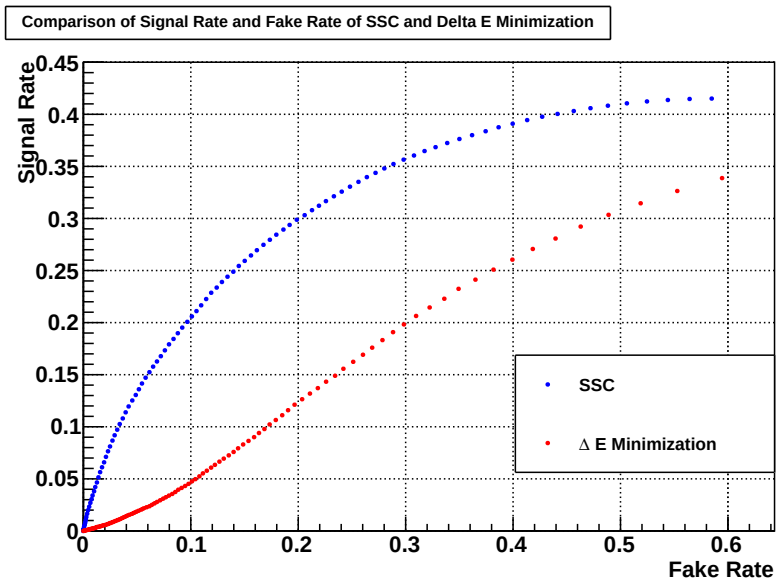


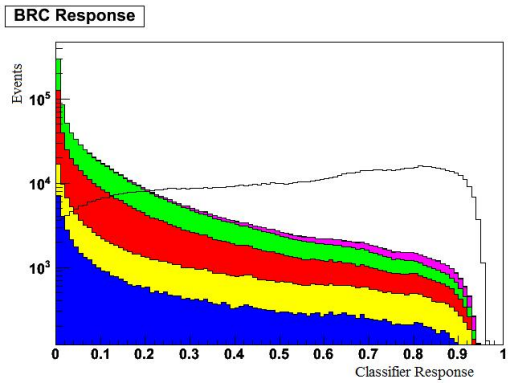
Figure 6.7: A comparison of signal efficiency and background efficiency for the two methods of choosing the best candidate in the event. We see that using the SSC method, and placing different requirements on the minimum classifier response, is more powerful (higher signal efficiency for equivalent background) than simply minimizing $|\Delta E|$ (and placing different requirements on its maximum value).

6.4 Final Optimized Cuts

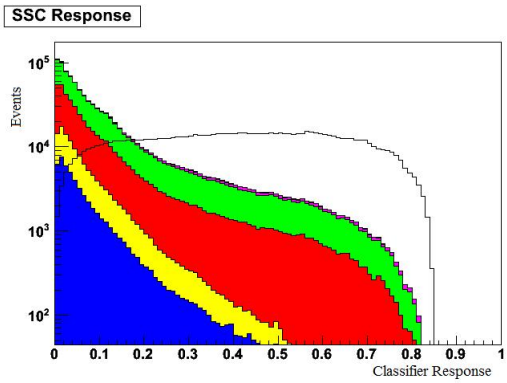
We have now constructed two classifiers: one for rejecting mostly continuum backgrounds, the BRC, and one for selecting the candidate in an event most consistent with a $b \to s\gamma$ transition, the SSC (the candidate with the highest response is taken in the event and the rest are rejected). This latter classifier also shows discrimination power with respect to $B\overline{B}$ background, so placing a tighter requirement on it, beyond simply using it to select the best candidate in an event, will improve our background rejection ability. We also identify a third signal/background discriminating variable, the distance from the detected primary photon to the next bump in the detector ("bump distance"). These three distributions are shown in Figure 6.8.

The differences in signal efficiencies between lower X_s mass bins and higher X_s mass bins means that one set of requirements for one mass region will not be optimal for a different mass region. We therefore optimize our selection requirements in four consecutive mass regions: $m_{X_s}\epsilon[0.6-1.1, 1.1-2.0, 2.0-2.4, 2.4-2.8]\,\mathrm{GeV}/c^2$ (we will use the notation "mass region" from this point forward to refer to a group of individual mass bins). To ensure that we still do not overtrain our requirements, thereby biasing our expected results, we evaluate the four sets of optimal cuts based on our validation MC set, defined above. We use a "brute force" method to optimize our requirements (try many settings for our three requirements in each region, rather than determine the best setting for one variable before optimizing the next one): We iterate over both classifiers from 0-1 in increments of 0.05 and different minimum bump distance settings, starting at 25 cm and increasing in increments of 5 cm. We then fine tune our classifier requirements by trying ± 0.05 around the value identified by our initial iteration, in increments of 0.01.

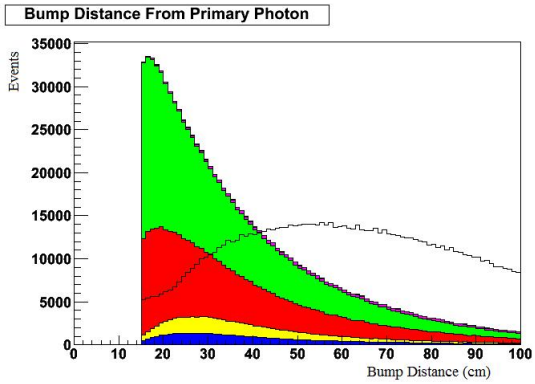
To determine the best settings, we again optimize the precision FOM ($S/\sqrt{S+B}$) based on the validation MC. We found the requirement on the photon bump distance was, in general, around 25 cm; we decided to keep it at that value in all four regions. The optimized set of requirements for our four mass regions is reported in Table 6.1. The values of our FOM are reported in Table 6.2, and the amount of background expected after applying these requirements is presented in Table 6.3.



(a) BRC distribution for the different types of events



(b) the best candidate's SSC response



(c) distance (in cm) from the primary photon to the next bump in the detector

Figure 6.8: The three variables used in optimizing our signal region requirements, having chosen the best candidate based on the SSC. In all distributions the blue is B^+B^-, the yellow is $B^0\overline{B}^0$, the red is $c\bar{c}$, the green is *uds*, the purple is cross-feed, and the unfilled black line is the signal. The background MC has been scaled to match data luminosity, the signal has been weighted to have the same area as the total background in each plot.

Table 6.1: The final cut requirements in each mass bin for selecting best candidate based on the SSC.

m_{X_s} (GeV/c^2)	Signal Classifier	min. bump distance (cm)	BRC
0.6-1.1	0.17	25	0.29
1.1-2.0	0.21	25	0.38
2.0-2.4	0.08	25	0.42
2.4-2.8	0.15	25	0.31

Table 6.2: The signal and background events in each mass bin, as well as precision, for the final cuts given in Table 6.1 using the best B candidate taken from the SSC and requiring $m_{ES} > 5.27$.

m_{X_s} bin	signal	background	precision
0.6-2.8	**17110**	**40586**	**71.2**
0.6-0.7	3.1 ± 0.1	138.3 ± 8.6	0.3 ± 0.0
0.7-0.8	86.1 ± 0.5	376.9 ± 14.1	4.0 ± 0.1
0.8-0.9	2297.1 ± 2.7	800.1 ± 18.8	41.3 ± 0.1
0.9-1.0	1918.1 ± 2.5	891.1 ± 20.2	36.2 ± 0.1
1.0-1.1	285.8 ± 1.0	1077.5 ± 22.8	7.7 ± 0.1
1.1-1.2	525.2 ± 2.7	1109.8 ± 21.5	13.0 ± 0.1
1.2-1.3	1023.2 ± 4.9	1482.1 ± 23.9	18.9 ± 0.1
1.3-1.4	1585.1 ± 7.6	2033.9 ± 27.7	26.3 ± 0.1
1.4-1.5	1966.0 ± 9.7	2520.2 ± 30.6	29.4 ± 0.2
1.5-1.6	2060.8 ± 10.6	2788.6 ± 32.3	29.6 ± 0.2
1.6-1.7	1838.9 ±10.1	2952.0 ± 33.7	26.6 ± 0.2
1.7-1.8	1283.7 ± 8.0	2301.2 ± 30.5	21.4 ± 0.1
1.8-1.9	853.5 ± 5.9	2343.5 ± 31.3	15.1 ± 0.1
1.9-2.0	441.9 ± 3.7	1318.9 ± 24.3	10.5 ± 0.1
2.0-2.2	604.3 ± 3.4	5289.8 ± 49.0	7.9 ± 0.1
2.2-2.4	207.6 ± 1.4	4645.0 ± 47.2	3.0 ± 0.0
2.4-2.6	80.5 ± 0.7	3650.6 ± 43.0	1.3 ± 0.0
2.6-2.8	48.7 ± 0.4	4866.4 ± 49.6	0.7 ± 0.0

Table 6.3: The different sources of background events, based on the test MC, scaled to reflect total luminosity, after cuts given in Table 6.1 and selecting the best B based on the SSC, $m_{ES} > 5.27$, and KaonKMLoose requirements.

mass bin	B^+B^-	$B^0\overline{B}^0$	$c\bar{c}$	uds	cross-feed	total
0.6-0.7	0.7 ± 0.5	0.0 ± 0.0	58.2 ± 5.7	69.1 ± 6.5	10.3 ± 0.3	138.3 ± 8.6
0.7-0.8	1.1 ± 0.6	4.4 ± 1.3	140.0 ± 8.8	196.2 ± 10.9	35.2 ± 0.7	376.9 ± 14.1
0.8-0.9	18.1 ± 2.6	26.3 ± 3.1	208.0 ± 10.7	361.2 ± 14.9	186.4 ± 1.3	800.1 ± 18.8
0.9-1.0	20.0 ± 2.7	26.7 ± 3.1	251.4 ± 11.7	408.2 ± 15.8	184.8 ± 2.0	891.1 ± 20.2
1.0-1.1	18.5 ± 2.6	14.3 ± 2.3	352.9 ± 13.9	508.4 ± 17.6	183.4 ± 3.3	1077.5 ± 22.8
1.1-1.2	11.8 ± 2.1	16.8 ± 2.5	326.0 ± 13.4	433.9 ± 16.3	321.2 ± 5.2	1109.7 ± 21.5
1.2-1.3	25.2 ± 3.1	20.1 ± 2.7	372.1 ± 14.3	547.6 ± 18.3	517.2 ± 8.4	1482.1 ± 23.9
1.3-1.4	31.8 ± 3.4	35.5 ± 3.6	523.6 ± 17.0	700.3 ± 20.7	742.7 ± 12.5	2033.9 ± 27.7
1.4-1.5	48.8 ± 4.3	49.0 ± 4.2	659.8 ± 19.0	818.3 ± 22.4	944.3 ± 16.2	2520.2 ± 30.6
1.5-1.6	60.3 ± 4.7	75.7 ± 5.3	743.7 ± 20.2	887.9 ± 23.3	1020.9 ± 17.8	2788.6 ± 32.3
1.6-1.7	73.3 ± 5.2	85.5 ± 5.6	808.5 ± 21.1	957.0 ± 24.2	1027.7 ± 18.0	2952.0 ± 33.7
1.7-1.8	73.6 ± 5.2	102.4 ± 6.1	640.0 ± 18.7	796.3 ± 22.1	688.9 ± 13.0	2301.2 ± 30.5
1.8-1.9	133.9 ± 7.0	360.8 ± 11.5	621.3 ± 18.5	710.1 ± 20.8	517.3 ± 10.0	2343.5 ± 31.3
1.9-2.0	95.8 ± 6.0	96.5 ± 5.9	371.0 ± 14.3	504.8 ± 17.6	250.8 ± 5.4	1318.9 ± 24.3
2.0-2.2	830.3 ± 17.5	652.9 ± 15.4	1481.5 ± 28.5	1673.8 ± 32.0	651.4 ± 10.1	5289.8 ± 49.0
2.2-2.4	923.5 ± 18.5	607.6 ± 14.9	1286.6 ± 26.6	1563.8 ± 30.9	263.5 ± 4.2	4645.0 ± 47.2
2.4-2.6	704.5 ± 16.1	383.1 ± 11.8	1120.3 ± 24.8	1354.8 ± 28.8	87.9 ± 1.3	3650.6 ± 43.0
2.6-2.8	998.6 ± 19.2	579.8 ± 14.6	1486.9 ± 28.6	1740.4 ± 32.6	60.6 ± 0.8	4866.4 ± 49.6

6.5 Review of Selection Procedure

Since the selection procedure for this analysis is quite complex, we present a summary of the steps performed to this point.

Before placing any tight requirements on events, we first apply the precuts and reconstruct the B candidates in each event as detailed in Section 5.1. After this reconstruction, we have many B candidates to choose from in a given event, as shown in Table 6.4.

Table 6.4: The average B candidate multiplicity by MC type after B candidate reconstruction. These numbers come from run 1 MC.

MC type	average B candidates
signal MC	29.1
$c\bar{c}$	16.3
uds	13.3
$\tau^+\tau^-$	1.7
B^+B^-	32.7
$B^0\overline{B}^0$	30.6

Since there is such a high multiplicity of potential B candidates in an event, we have to have an effective way of choosing one. Before doing so, we first require that the candidates have $m_{ES} > 5.24$, $|\Delta E| < 0.15$, and any photon used in the candidate have $E_\gamma > 0.06$. After applying these requirements, we choose the best B candidate in an event by choosing the candidate that has the highest response to the SSC detailed in Section 6.3.

With the best B candidate chosen, we now apply some final requirements to limit the background in each bin. These requirements are detailed in Table 6.1, and they are based on the SSC (beyond just choosing the candidate that has the highest response), the distance to the next bump in the detector around the high-energy photon in the $b \to s\gamma$ transition, and the BRC used to reject continuum (uds and $c\bar{c}$) background (detailed in Section 6.2). The expected amount of signal and background in the signal region, $m_{ES} > 5.27\,\mathrm{GeV}/c^2$, of each X_s mass bin is given in Table 6.2. The procedure for assessing the signal efficiency is:

$$N_{b\to s\gamma} = \frac{N_{yield}}{\epsilon_{38}\epsilon_{incl.}}, \tag{6.1}$$

where N_{yield} is number of signal found in a given mass bin (we fit this value based on the m_{ES} distribution, as described in Section 7), ϵ_{38} is the efficiency of finding our 38 modes within that bin, $\epsilon_{incl.}$ is the efficiency of our 38 modes to the inclusive $b \to s\gamma$ transition rate (assessing the impact of the missing modes), and $N_{b\to s\gamma}$ is the number of $b \to s\gamma$ transitions that took place in that mass bin.

The efficiency for reconstructing our 38 modes in the different mass bins, ϵ_{38}, is given for each mass bin in Table 6.5, in which we extract the amount of our 38 modes generated in the bin from a separately generated set of generator level MC (no machine response) and reweight based on the default photon spectrum.

Table 6.5: The signal efficiencies for the final cuts in each mass bin when comparing expected signal yield compared to the amount of our 38 exclusive modes generated within the bin (ϵ_{38}).

mass bin	# signal based on test MC	generated in bin	efficiency (%)
0.6-0.7	1098	7433.6±155	15.04±0.55
0.7-0.8	29993	186314.7±775	16.46±0.12
0.8-0.9	798668	4671177±3881	17.33±0.02
0.9-1.0	667859	3685046±3447	18.30±0.03
1.0-1.1	99673	513691±1287	19.55±0.08
1.1-1.2	41140	357505±1518	11.51±0.07
1.2-1.3	80390	693859±2115	11.59±0.05
1.3-1.4	123700	1162020±2737	10.65±0.04
1.4-1.5	153600	1621153±3233	9.47±0.03
1.5-1.6	160900	1917784±3517	8.39±0.03
1.6-1.7	14300	1991297±3583	7.20±0.02
1.7-1.8	100700	1822991±3429	5.52±0.02
1.8-1.9	66530	1465099±3074	4.54±0.02
1.9-2.0	34600	1061423±2616	3.26±0.02
2.0-2.2	48090	1190393±2771	4.04±0.02
2.2-2.4	16650	535161±1858	3.11±0.03
2.4-2.6	6450	281993±1348	2.29±0.03
2.6-2.8	3887	169080±1044	2.30±0.04

Chapter 7

Fitting Procedure

With our final selection requirements, given in Table 6.1, applied, we have the final amount of events we expect in the data. We choose to fit the distribution of beam energy-substituted mass, m_{ES}, to extract the signal yield in each mass bin to determine the E_γ spectrum, and to use these results to evaluate a BF for $b \to s\gamma$ over the range $0.6 < m_{X_s} < 2.8\,\mathrm{GeV}/c^2$. To do this, and to have stable fits, we have to fix many of the parameters; we will evaluate how much systematic uncertainty this introduces in Section 10.

7.1 Fitting Overview

We have three types of background in this analysis: $q\overline{q}$ ($q = c, u, d, s$) - referred to as continuum, $B\overline{B}$, and cross-feed. All three types of background have a combinatoric component that we will model with Argus functions as described below, the latter two also have a peaking component to them (referring to a peak in the m_{ES} distribution that cannot be well-modeled by an Argus function, at the same value where we expect the $b \to s\gamma$ signal to peak) that we take into account when extracting the signal yield from the data.

We fit the m_{ES} distribution using a Crystal Ball (CB) function [30] to describe the signal events, Argus functions [31] to describe the combinatorial background shapes, and Novosibirsk (Nvs) [32] functions to describe the different types of peaking backgrounds. We elect to use a Nvs for the peaking backgrounds as supposed to separate CB functions because they provide a slightly different shape, and therefore introduce a bit of orthogonality to our signal CB function. We perform fits to obtain the signal yield in each mass bin.

7.1.1 Signal Distribution

The signal MC events are fit with a CB, which accounts for the small radiative tail from the photon energy measurement; the un-normalized PDF is given by:

$$\left(\left|\frac{m_{ES}-m_0}{\sigma}\right|<\alpha\right), f(m_{ES}) = e^{\left(-\frac{(m_{ES}-m_0)^2}{2\sigma^2}\right)}, \tag{7.1}$$

$$\left(\left|\frac{m_{ES}-m_0}{\sigma}\right|>\alpha\right), f(m_{ES}) = \frac{e^{\left(-\frac{\alpha^2}{2}\right)}}{\left(\frac{n_{CB}}{\alpha}-\alpha-\frac{m_{ES}-m_0}{\sigma}\right)^{n_{CB}}}, \tag{7.2}$$

where m_0 and σ are the peak position and width, respectively. α and n_{CB} are two other parameters of the function to account for the non-Gaussian tail. The shape of the Crystal Ball comes from a χ^2 fit to the truth-matched signal MC candidates, after all of our selection requirements are applied, with a mass requirement of 0.6< m_{X_s} <2.8 GeV/c^2. The signal MC weighted according to the BBU model with m_b=4.65, consistent with the results of the previous *BABAR* analysis (our default photon-spectrum model). We then use the same signal CB shape for each mass bin. The choice for using the same signal shape is supported by observing that the central value on the fit parameters doesn't change much with each mass region in Table 7.1. The error we quote is based on the sum of weights squared for each signal event in a χ^2 fit.

Table 7.1: The different shape parameters for the signal CB when fit in each mass region vs. the entire range. We only use the final row parameter values for both mass-bin specific fits and full range fits, since the values are all consistent.

m_{X_s}	α	m_0	width ($\times 10^{-3}$)	n_{CB}
0.6 to 1.1	1.133±0.0001	5.2804±0.00001	2.819±0.001	144.9
1.1 to 2.0	1.123±0.006	5.28034±0.00001	2.844±0.006	144.8
2.0 to 2.4	1.132±0.014	5.28035±0.00002	2.813±0.014	144.9
2.4 to 2.8	1.20±0.02	5.2804±0.00002	2.72±0.02	146.24
0.6 to 2.8	**1.124±0.004**	**5.28037±0.00001**	**2.835±0.004**	**144.8±14.5**

7.1.2 Cross-feed Background Distribution

To model the cross-feed (signal events in which we choose the wrong candidate, either because of a less-than-perfect candidate selection algorithm, or because the true signal candidate had a final state that was not one of the 38 we reconstruct), we use a Nvs function for the peaking component, as well as an Argus function for the combinatoric component. The Nvs PDF is given in un-normalized form as:

$$f(m_{ES}) = e^{\left(-\frac{1}{2}\left(\ln^2[1+\Lambda\tau(m_{ES}-m)]/\tau^2+\tau^2\right)\right)}, \tag{7.3}$$

where

$$\Lambda = \sinh(\tau\sqrt{\ln 4})/(\sigma\tau\sqrt{\ln 4}), \tag{7.4}$$

where the peak is m, the width is σ, and τ is the tail parameter. The Argus function is defined as:

$$f(m_{ES}) = m_{ES}\left(1-\left(\frac{m_{ES}}{m}\right)^2\right)^{\frac{1}{2}} \times e^{\left(c\frac{m_{ES}}{m}\right)}, \tag{7.5}$$

where m is the cutoff, and c is the slope.

We fix the cross-feed Nvs parameters to different values, depending on the m_{X_s} region given in Table 7.2 for the individual mass bins. The uncertainties quoted in the table are based on a χ^2 fit in which the statistical error within each region comes from the number of unweighted MC events (sum of weights squared for error on the event), not the number of events expected in the data. As described in Section 9.2, the parametrization of the cross-feed will change when we redistribute the signal MC after our fragmentation study; we therefore present the cross-feed PDF values in this section as "starting values" for that study. We present the final values for the cross-feed PDF parameters in Section 9.2.

For the Argus component of the cross-feed, we fit the slope parameter as well as the fractional peaking contribution for the cross-feed in each mass bin, and then fix this value in the fits to data. The Argus slope and peaking fractions are given in the top (Nvs) part of Table 7.3. We also fix the cross-feed yield (both the peaking plus the combinatoric components) to the signal yield in the final fit. To do this we fix the ratio of signal to

Table 7.2: The different shape parameters for the cross-feed Nvs function used to model the peaking component in the different m_{X_s} regions.

m_{X_s} range	peak	width	tail
0.6 to 1.1	5.28043±0.00006	0.00350±0.00006	-0.160±0.017
1.1 to 2.0	5.2809±0.0001	0.0047 ±0.0001	-0.247±0.017
2.0 to 2.4	5.2815±0.0002	0.0062±0.0005	-0.36±0.08
2.4 to 2.8	5.2811±0.0005	0.0064±0.0011	-0.39±0.14

cross-feed based on the truth information in our MC sample in each bin, and then allow the yield of the total signal plus cross-feed PDF (consisting of the signal CB plus cross-feed Nvs plus cross-feed Argus) to float.

As a cross-check on our choice of a Nvs function for the peaking cross-feed, we also use a CB function; the results are presented in Table 7.3. We address the errors associated with our choice for the cross-feed in Section 10.

7.1.3 Peaking $B\overline{B}$ Background Distribution

We also model the peaking component of the generic, non-$b \to s\gamma$, $B\overline{B}$ MC with a Nvs function. We fix the parameters of the peaking $B\overline{B}$ Nvs to different values, depending on the m_{X_s} region, given in Table 7.4. The uncertainties quoted in the table are based on a χ^2 fit in which the statistical error within each bin comes from the number of unweighted MC events, not the number of events expected in the data. We also fix the yield of peaking $B\overline{B}$ events described by this Nvs within each mass bin. We get this yield by simultaneously fitting the $B\overline{B}$ MC within a given mass bin with the Nvs given in Table 7.4 and an Argus function in which the slope is allowed to float and determine what fraction of the $B\overline{B}$ background should be considered peaking. We evaluate a systematic uncertainty specific to fixing the peaking $B\overline{B}$ yield and shape through sideband studies. In these studies, discussed more in Chapter 8, we will fit the distributions of data in $2.9 < m_{X_s} < 3.0\,\mathrm{GeV}/c^2$, and in classifier sidebands, allowing the peaking $B\overline{B}$ contribution to float, and compares this to the value we would have assigned it based on the MC study.

We fix the number of peaking $B\overline{B}$ MC in the final fit to the value given in Table 7.5,

Table 7.3: The change in peaking cross-feed when using the Nvs function vs. a CB function.

Fit with Nvs				
mass bin	Argus Slope	Peaking frac.	Peaking XF	χ^2/DoF
0.6-0.7	-64.6±6.1	0.020±0.018	0.5±0.5	0.87
0.7-0.8	-94.6±4.7	0.124±0.014	8.8±1.0	1.26
0.8-0.9	-94.5±4.9	0.631±0.007	154.7±1.7	2.47
0.9-1.0	-104.8±3.5	0.411±0.007	115.1±2.1	1.38
1.0-1.1	-103.3±2.6	0.104±0.008	37.7±2.9	1.18
1.1-1.2	-116.3±2.6	0.168±0.010	97.4±5.7	1.29
1.2-1.3	-112.7±2.3	0.208±0.009	187.7±7.8	1.39
1.3-1.4	-107.2±2.0	0.205±0.008	268.9±10.0	1.27
1.4-1.5	-100.9±1.9	0.196±0.007	331.7±11.8	1.29
1.5-1.6	-99.4±1.8	0.184±0.007	341.9±12.5	1.95
1.6-1.7	-96.9±1.7	0.163±0.007	304.3±12.6	1.05
1.7-1.8	-84.4±2.2	0.161±0.008	204.2±10.2	1.33
1.8-1.9	-77.4±2.4	0.137±0.009	133.5±8.4	1.10
1.9-2.0	-61.8±3.4	0.139±0.011	68.8±5.4	0.98
2.0-2.2	-30.3±2.3	0.154±0.008	235.5±11.7	1.22
2.2-2.4	-12.0±2.8	0.097±0.009	70.3±6.1	0.37
2.4-2.6	7.1±4.4	0.122±0.013	29.4±3.1	0.90
2.6-2.8	10.3±4.4	0.119±0.013	19.6±2.1	1.58
Fit with CB				
mass bin	Argus Slope	Peaking frac.	Peaking XF	χ^2/DoF
0.6-0.7	-64.0±6.4	0.023±0.020	0.6±0.5	0.86
0.7-0.8	-91.6±5.0	0.137±0.016	9.8±1.1	1.27
0.8-0.9	-55.7±5.8	0.695±0.007	170.6±1.7	2.79
0.9-1.0	-89.9±3.9	0.457±0.008	127.9±2.3	1.49
1.0-1.1	-100.7±2.8	0.116±0.009	42.2±3.2	1.14
1.1-1.2	-119.6±2.3	0.142±0.008	82.3±4.9	1.51
1.2-1.3	-116.9±2.0	0.177±0.007	160.3±6.7	1.59
1.3-1.4	-111.6±1.8	0.175±0.007	229.2±8.6	1.64
1.4-1.5	-105.3±1.7	0.167±0.006	282.9±10.2	1.67
1.5-1.6	-104.1±1.6	0.154±0.006	286.1±10.9	3.11
1.6-1.7	-100.9±1.6	0.137±0.006	255.5±10.9	1.77
1.7-1.8	-88.4±2.0	0.137±0.007	174.4±8.7	1.28
1.8-1.9	-81.1±2.2	0.116±0.007	113.0±7.2	1.18
1.9-2.0	-66.2±3.0	0.117±0.009	57.7±4.7	1.13
2.0-2.2	-38.6±1.9	0.111±0.006	170.0±9.0	1.46
2.2-2.4	-17.6±2.4	0.071±0.007	51.0±4.8	0.56
2.4-2.6	-4.3±3.4	0.070±0.008	16.8±2.0	1.08
2.6-2.8	0.1±3.3	0.072±0.008	11.9±1.4	1.52

where the uncertainty reflects the statistical uncertainty of the fit to the fractional contribution of the Nvs in the $B\overline{B}$ MC. In fitting to the number of peaking $B\overline{B}$, we require this number to be positive, since we are trying to subtract an amount of background. We recognize that the uncertainty on our amount of peaking $B\overline{B}$ is overly optimistic, since it is based on the MC, which has approximately 3$\times$ the on peak data statistics. We address this issue in Chapter 10, and state here that the uncertainty on peaking $B\overline{B}$ in Table 7.5 reflects our precision based on simulation.

Table 7.4: The different shape parameters for the $B\overline{B}$ Nvs function used to model the peaking component in the different m_{X_s} bins.

m_{X_s} range	peak	width ($\times 10^{-3}$)	tail
0.6 to 2.0	5.28116±0.00027	4.50±0.30	-0.2607±0.064
2.0 to 2.4	5.28097±0.00028	3.96±0.37	-0.2275±0.069
2.4 to 2.8	5.28000±0.00027	3.74±0.24	-0.031±0.063

Table 7.5: The number of peaking $B\overline{B}$ events in each mass bin. We fix this in the final fit. The uncertainty reflects the statistical uncertainty of the fit.

m_{X_s} range	$N_{B\overline{B}}$-peaking from fit to MC
0.6 to 0.7	0.00 ± 4.7
0.7 to 0.8	0.00 ± 12.6
0.8 to 0.9	41.2 ± 3.7
0.9 to 1.0	38.8 ± 5.1
1.0 to 1.1	18.7 ± 5.3
1.1 to 1.2	12.0 ± 5.9
1.2 to 1.3	26.6 ± 5.9
1.3 to 1.4	17.1 ± 7.6
1.4 to 1.5	41.0 ± 8.8
1.5 to 1.6	41.1 ± 26.3
1.6 to 1.7	52.6 ± 11.0
1.7 to 1.8	43.7 ± 12.2
1.8 to 1.9	355.2 ± 17.8
1.9 to 2.0	38.8 ± 14.1
2.0 to 2.2	376.7 ± 33.2
2.2 to 2.4	501.4 ± 32.0
2.4 to 2.6	290.1 ± 24.3
2.6 to 2.8	421.3 ± 28.6

7.1.4 Combinatoric Background Distribution

Finally, to model the combinatoric background from the $q\overline{q}$ events as well as the combinatoric component of the $B\overline{B}$ background events, we use an Argus function with a slope that is allowed to float in the final fit. The two parameters that are allowed to float are therefore the slope of this final Argus function as well as the ratio of signal+cross-feed PDFs to the other PDFs (Argus+peaking $B\overline{B}$ Nvs).

7.1.5 Complete Bin Fits

In each mass bin, we combine the PDFs described above as well as an Argus PDF to describe the combinatoric background (both $c\bar{c}$,*uds*, and non-peaking $B\overline{B}$), a summary of this procedure is presented in Table 7.1.5. We allow the fractional contribution of the signal+cross-feed PDF (an Argus+Nvs for the cross-feed plus CB for signal) and the slope of the combinatoric Argus to float. The results of our fitting procedure when fit to the MC samples are shown in Table 7.7. The full fit with all of the components is shown in Figure 7.1.

To obtain partial branching fractions, PBFs, in each mass bin, we fit the signal yield within the bin, correct for signal efficiency (both ϵ_{38} and $\epsilon_{incl.}$ from equation 6.1), and divide by total number of B events. This will give us a spectrum of the X_s hadron system that can be directly converted to a photon spectrum through equation 3.1. The numerical conversion between our mass bins and their respective photon energies is given in Table 3.2.

We can then obtain a total branching fraction by summing the PBFs we have in the 18 mass bins, giving us a transition rate for $b \to s\gamma$ for the region $E_\gamma > 1.9\,\mathrm{GeV}$. We then fit different photon spectrum models to the distribution of photon energy we find, and use this to correct for the missing range $1.9 > E_\gamma > 1.6\,\mathrm{GeV}$, the lower photon energy limit often quoted in the literature.

7.2 Fit Validation - Toy Studies

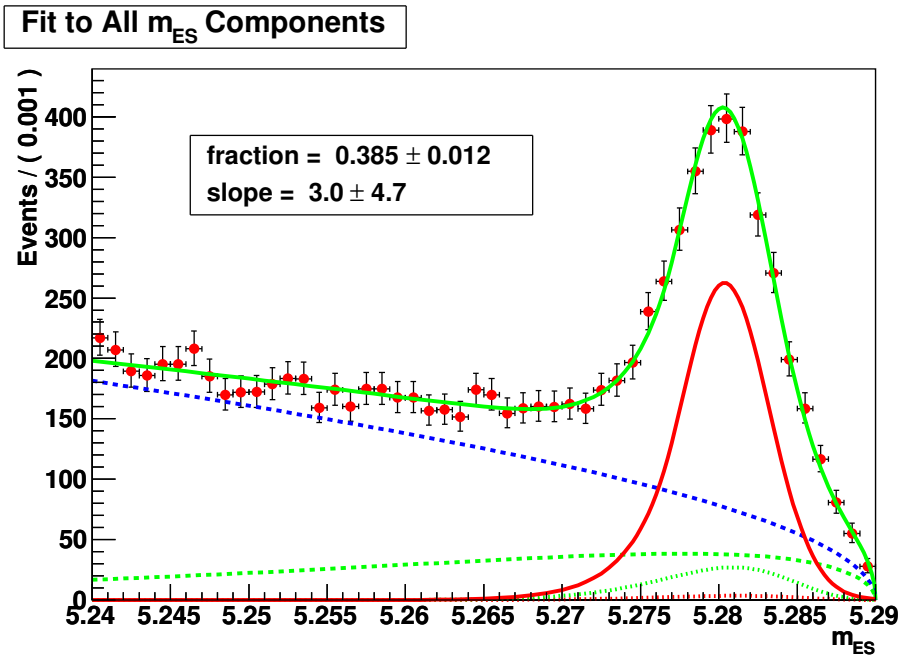
We run toy MC studies on each mass bin, testing our fitting procedure to ensure that we are not introducing a bias into either the fractional contribution of signal+cross-feed or the

Table 7.6: Summary of fitting functions and parameters; here XF stands for cross-feed. The order of the table is meant to reflect how we build up to the last fits to On Peak data. "3 or 4 mass regions" refers to groups of mass bins, as shown in other tables; "mass bins" refers to fits done on individual mass bins.

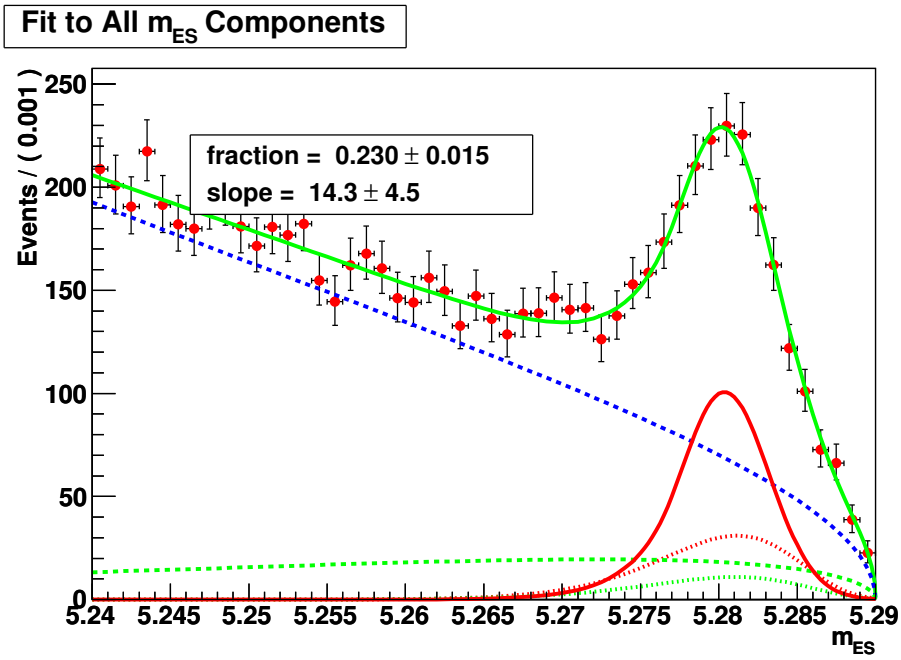
Data Set Type	Fit Function	Free Parameters
$B\overline{B}$ MC - 3 mass regions	$B\overline{B}$ Peaking Nvs + Argus	Nvs Parameters (peak, width, tail), Argus Slope
$B\overline{B}$ MC - mass bins	$B\overline{B}$ Peaking Nvs + Argus	Nvs yield, Argus Slope *we fix the Nvs to parameters found above and get peaking contribution here, which we then fix in the final fit for the bin*
XF MC - 4 mass regions	XF Peaking Nvs + Argus	Nvs Parameters, Argus Slope
XF MC - mass bins	XF Peaking Nvs + Argus	ratio Nvs:Total XF, Argus Slope *we fix the Nvs to parameters found above and get the ratio of peaking XF here, which we then fix along with the XF Argus slope found here in the final fit for the bin*
Signal MC - all bins	Signal CB	CB Parameters (α, m_0, width, n_{CB}) *we use the full mass range signal CB settings for individual mass bin fits to signal yield in the full fit for the bin*
On Peak - mass bins	Signal CB + XF Argus + XF Nvs + $B\overline{B}$ Peaking Nvs + Combinatoric Argus	Signal+XF yield, Combinatoric Argus Slope

Table 7.7: The results in each bin of our fitting procedure, the difference from the actual value, the uncertainty on the fit, and the associated χ^2.

m_{X_s} range	Actual Signal	Fit Signal	Difference	Fit Uncertainty	N_σ difference	χ^2
0.6-0.7	3.2	9.3	-6.0	11.1	-0.54	0.29
0.7-0.8	89.1	92.7	-3.5	23.1	-0.15	0.02
0.8-0.9	2348.1	2403.8	-55.7	49.1	-1.14	1.29
0.9-1.0	1956.5	1959.7	-3.1	51.0	-0.06	0.00
1.0-1.1	291.3	344.9	-53.6	36.5	-1.47	2.16
1.1-1.2	535.3	527.5	7.8	38.8	0.20	0.04
1.2-1.3	1044.9	1058.9	-14.0	46.0	-0.30	0.09
1.3-1.4	1608.9	1581.7	27.2	55.9	0.49	0.24
1.4-1.5	1995.4	2013.0	-17.5	61.6	-0.28	0.08
1.5-1.6	2094.5	2078.0	16.5	64.7	0.26	0.07
1.6-1.7	1864.5	1891.7	-27.2	63.2	-0.43	0.18
1.7-1.8	1309.2	1345.7	-36.5	55.7	-0.66	0.43
1.8-1.9	864.3	901.4	-37.1	51.2	-0.72	0.52
1.9-2.0	449.3	491.6	-42.3	39.0	-1.08	1.18
2.0-2.2	624.1	669.7	-45.6	69.8	-0.65	0.43
2.2-2.4	216.3	284.7	-68.4	63.6	-1.08	1.16
2.4-2.6	83.9	153.1	-69.2	53.9	-1.28	1.65
2.6-2.8	50.5	24.1	26.4	61.3	0.43	0.19
					total χ^2	10.02



(a) bin 1.4< m_{X_s} <1.5 GeV/c^2



(b) bin 1.8< m_{X_s} <1.9 GeV/c^2

Figure 7.1: The fit to two different mass bins in MC showing all component PDFs including: signal CB (red solid line), cross-feed Nvs (green dotted), cross-feed Argus (green dashed), peaking $B\overline{B}$ Nvs (red dotted), and combinatoric Argus (blue dashed).

Argus slope parameter. We run 1000 "pure" toy studies in each mass bin, generating each toy study from the central value of the original fits, with the number of events in each toy study consistent with the expected number of events in the given mass bin. The resulting pulls and errors on the pulls of the toy study distributions are given in Table 7.8.

Table 7.8: The results of the toy studies in each mass bin; we report the mean and sigma of a Gaussian fit to the pull distributions of the toy-fits to the fractional contribution of the signal+cross-feed and the slope parameter of the combinatoric Argus function.

m_{X_s} range	pull of fraction mean	σ of Gaussian	pull of slope mean	σ of Gaussian
0.6 to 0.7	0.131 ± 0.034	1.070 ± 0.024	-0.153 ± 0.033	1.037 ± 0.023
0.7 to 0.8	0.005 ± 0.032	1.011 ± 0.023	-0.012 ± 0.031	0.985 ± 0.022
0.8 to 0.9	0.010 ± 0.031	0.974 ± 0.022	-0.007 ± 0.031	0.985 ± 0.022
0.9 to 1.0	0.008 ± 0.031	0.977 ± 0.022	0.013 ± 0.030	0.958 ± 0.022
1.0 to 1.1	0.004 ± 0.031	0.983 ± 0.022	-0.024 ± 0.031	0.990 ± 0.022
1.1 to 1.2	0.002 ± 0.032	1.012 ± 0.023	-0.031 ± 0.031	0.977 ± 0.022
1.2 to 1.3	0.007 ± 0.032	1.025 ± 0.023	-0.024 ± 0.033	1.027 ± 0.023
1.3 to 1.4	0.004 ± 0.032	1.006 ± 0.026	-0.008 ± 0.032	1.006 ± 0.023
1.4 to 1.5	0.007 ± 0.031	0.974 ± 0.022	-0.020 ± 0.032	1.009 ± 0.023
1.5 to 1.6	0.003 ± 0.032	1.027 ± 0.023	-0.020 ± 0.033	1.056 ± 0.024
1.6 to 1.7	0.010 ± 0.032	1.015 ± 0.023	-0.018 ± 0.032	1.003 ± 0.023
1.7 to 1.8	-0.001 ± 0.031	0.980 ± 0.022	-0.024 ± 0.033	1.034 ± 0.023
1.8 to 1.9	0.003 ± 0.031	0.992 ± 0.022	-0.018 ± 0.032	1.002 ± 0.023
1.9 to 2.0	0.011 ± 0.032	1.026 ± 0.023	-0.026 ± 0.033	1.051 ± 0.024
2.0 to 2.2	0.001 ± 0.033	1.032 ± 0.023	-0.017 ± 0.032	1.012 ± 0.023
2.2 to 2.4	-0.003 ± 0.032	1.026 ± 0.023	-0.027 ± 0.033	1.030 ± 0.023
2.4 to 2.6	0.006 ± 0.032	0.999 ± 0.022	-0.029 ± 0.032	1.019 ± 0.023
2.6 to 2.8	0.000 ± 0.033	1.038 ± 0.023	-0.026 ± 0.033	1.056 ± 0.024

We see that all but the lowest mass bin (the bin with very low statistics) have means consistent with 0 and sigmas consistent with 1. An example of this type of toy study for one of our mass bins ($1.4 < m_{X_s} < 1.5\,\mathrm{GeV}/c^2$) can be seen in Figure 7.2.

Finally, we also perform embedded toy studies (studies in which we "embed" the signal and cross-feed MC, at 1× the amount expected in data, in PDF-generated background events). The PDFs used are therefore the Nvs for the peaking $B\overline{B}$, and an Argus for the combinatoric background. The results from this study, given in Table 7.9, show that in general we successfully fit the correct signal+cross-feed fraction. In this toy study, in which we analyzed 400 toy data sets for each mass bin, the amount of generated signal+cross-feed

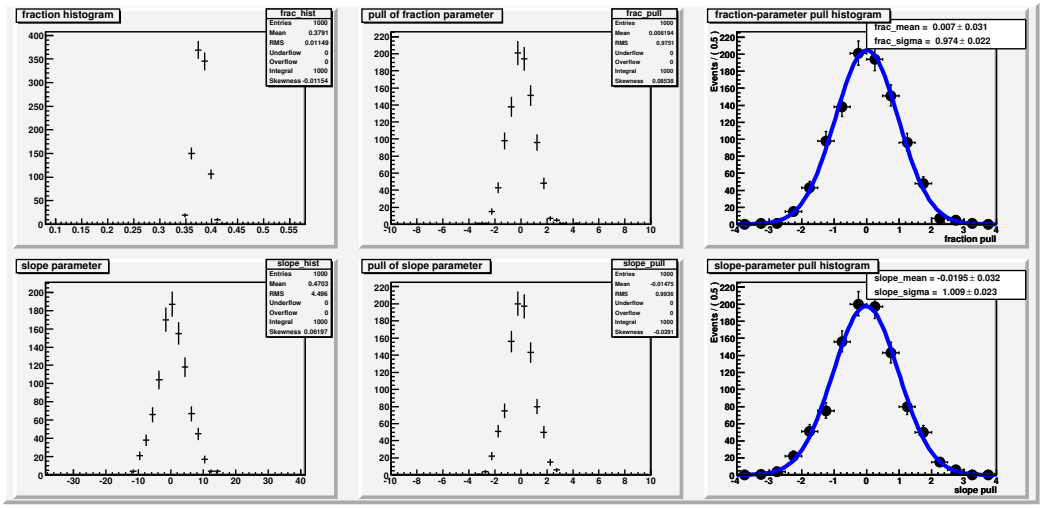


Figure 7.2: The pure toy study for bin 1.4< m_{X_s} <1.5 GeV/c^2. The top row shows the fraction of signal+cross-feed floating parameter and the bottom shows the Argus slope floating parameter. In each row, the left plot is the distribution of the values of the fit to the different parameters in each toy study, the middle shows the pull on these fits, and the right shows a zoomed-in version of the pull with a Gaussian fit.

MC is much larger than what is expected in the data. Therefore the different toy studies share a minimal amount of events and are essentially independent of one another.

Table 7.9: The results of the embedded toy studies in each mass bin in which we only embed signal and cross-feed MC in a PDF-generated background; we report the expected fraction from truth matching as well as the fitted fraction with rms error after 400 studies, and mean of a Gaussian fit to the pull distribution of the fitted fraction.

m_{X_s} range	expected fraction	fitted fraction	pull distribution mean
0.6 to 0.7	0.052	0.018 ± 0.141	0.085 ±0.054
0.7 to 0.8	0.107	0.105 ± 0.025	-0.005 ±0.049
0.8 to 0.9	0.549	0.552 ± 0.011	0.026 ±0.050
0.9 to 1.0	0.467	0.467 ± 0.012	0.009 ±0.050
1.0 to 1.1	0.166	0.164 ± 0.018	0.005 ±0.045
1.1 to 1.2	0.280	0.292 ± 0.033	0.091 ±0.056
1.2 to 1.3	0.345	0.348 ± 0.016	0.001 ±0.053
1.3 to 1.4	0.368	0.371 ± 0.014	0.004 ±0.053
1.4 to 1.5	0.381	0.371 ± 0.013	-0.0025 ±0.054
1.5 to 1.6	0.366	0.351 ± 0.014	0.005 ±0.062
1.6 to 1.7	0.336	0.322 ± 0.013	-0.005 ±0.057
1.7 to 1.8	0.297	0.287 ± 0.012	-0.0125 ±0.050
1.8 to 1.9	0.233	0.226 ± 0.013	0.009 ±0.046
1.9 to 2.0	0.186	0.184 ± 0.014	0.001 ±0.046
2.0 to 2.2	0.114	0.111 ± 0.013	0.006 ±0.051
2.2 to 2.4	0.057	0.059 ± 0.017	0.001 ±0.048
2.4 to 2.6	0.026	0.024 ± 0.018	-0.0063 ±0.052
2.6 to 2.8	0.014	0.011 ± 0.018	-0.015 ±0.050

Chapter 8

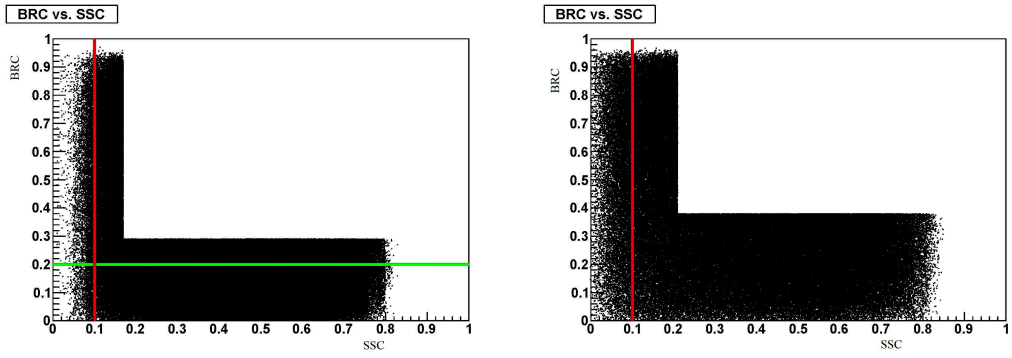
Sideband Studies

Before unblinding the signal region in data defined in Table 6.1, we identify several sidebands (regions that are very "signal like" but do not themselves contain much signal) we may use to evaluate uncertainties associated with different aspects of this analysis. In this chapter, we present these sidebands and the MC/data comparisons; we discuss how we implement these results to evaluate uncertainties on our final results in Chapter 10.

8.1 Classifier Sidebands

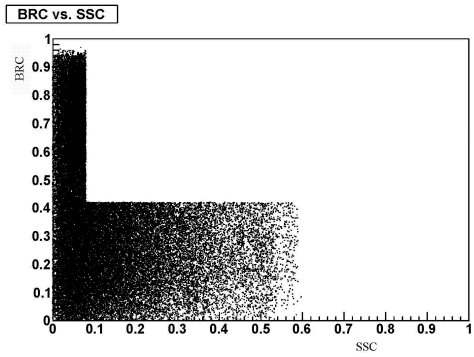
Our signal box is defined by mass-region-dependent requirements on the output of two separate classifiers. We therefore identify the inverse of each of these requirements as a sideband in which we may validate the efficiency of the requirement on the other classifier. The two-dimensional distribution of the output of each classifier for signal MC events, with the signal box removed for each mass region is shown in Figure 8.1. The motivation for shifting some of the locations further from the signal box in some of the sidebands in the lower mass regions comes from wanting to ensure that our signal constitutes a small ($< 2\%$) fraction of the data in this region.

With four mass regions and two sidebands in each, we can compare the efficiency of our cuts using eight classifier sidebands. We parse the classifier responses into bins of 0.01. To get a measurement of the efficiency of a cut in a given sideband, we compare the falloff in efficiency of the signal region's classifier cut value as compared to the number of events that are allowed if we relax the cut by five bins (*i.e.*, we compare the change in efficiency

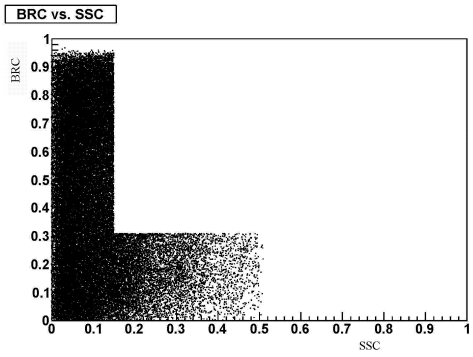


(a) mass region 0.6< m_{X_s} <1.1 GeV/c^2

(b) mass region 1.1< m_{X_s} <2.0 GeV/c^2



(c) mass region 2.0< m_{X_s} <2.4 GeV/c^2



(d) mass region 2.4< m_{X_s} <2.8 GeV/c^2

Figure 8.1: The signal MC distribution for each mass region showing the sidebands in each of the classifiers. Where indicated by a red (SSC) or green (BRC) line, we used this value of the classifier to delineate the sideband to limit the signal contribution; if there is no line drawn, the default cut of the classifier is used to define the sideband.

between the nominal cut and the nominal cut minus 0.05 for each classifier). We then compare the ratios of such efficiencies $\epsilon_{MC}/\epsilon_{Data}$ to estimate how well our efficiency for each cut is modeled within the MC. The results of this study can be seen in Table 8.1, and their distributions can be seen in figures 8.2 and 8.3.

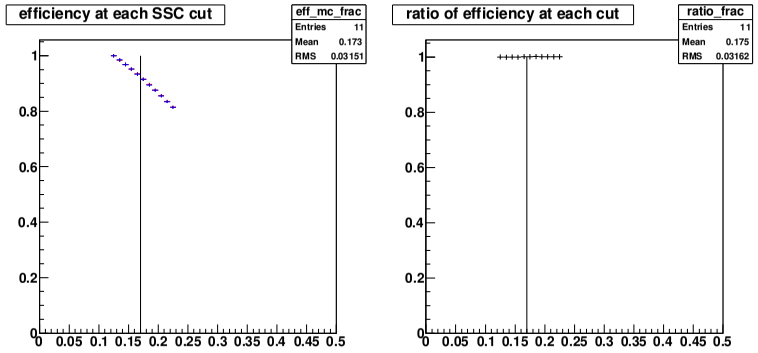
Table 8.1: The ratio between MC and data of changes in efficiencies when comparing the signal region cut value to five bins looser than the signal region cut value for each of the classifier sidebands.

BRC sideband		
mass region	SSC cut value	ratio
0.6-1.1	0.17	1.001 ± 0.009
1.1-2.0	0.21	1.002 ± 0.003
2.0-2.4	0.08	1.015 ± 0.004
2.4-2.8	0.15	0.982 ± 0.005
SSC sideband		
mass region	BRC cut value	ratio
0.6-1.1	0.29	0.99 ± 0.06
1.1-2.0	0.38	1.006 ± 0.01
2.0-2.4	0.42	1.003 ± 0.008
2.4-2.8	0.31	0.997 ± 0.003

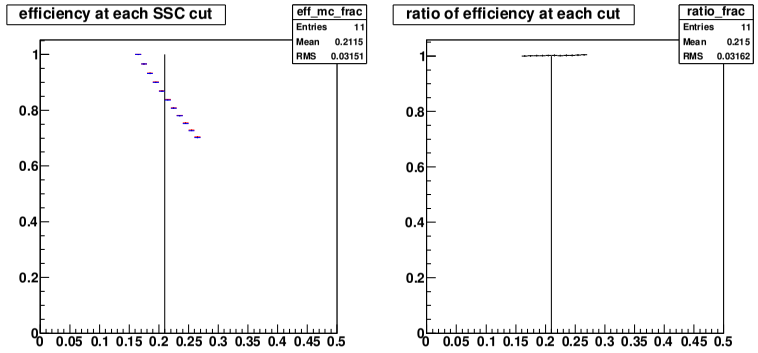
8.2 π^0 Veto Classifier Sideband

The classifier we use to veto π^0s described in Section 6.1 provides us with a different choice for a sideband in which to evaluate the SSC requirement efficiency systematic uncertainty. The classifier output is only greater than 0 if there is another photon in the event that can be combined with the high-energy transition photon giving an invariant mass consistent with a π^0. The amount of signal in this region of the classifier (response >0) is negligible compared to the amount of background. The further requirement $m_{ES} > 5.27\,\mathrm{GeV}/c^2$ substitutes our signal ($B \to X\gamma$) with events more consistent with the decay $B \to X\pi^0$, while our signal is still a relatively small component (in the K^* region, signal constitutes ∼4% of the total events, and is less in the other mass regions). We use these events as a control sample to model our signal, and evaluate our understanding of the SSC selection efficiency.

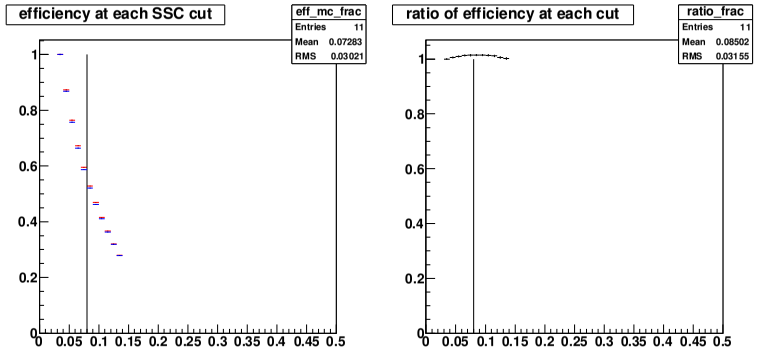
This study is highly correlated with the BRC classifier sideband mentioned above in



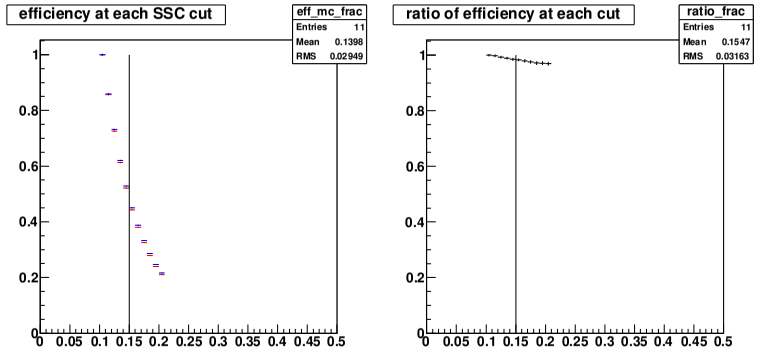
(a) mass region 0.6< m_{X_s} <1.1 GeV/c^2



(b) mass region 1.1< m_{X_s} <2.0 GeV/c^2

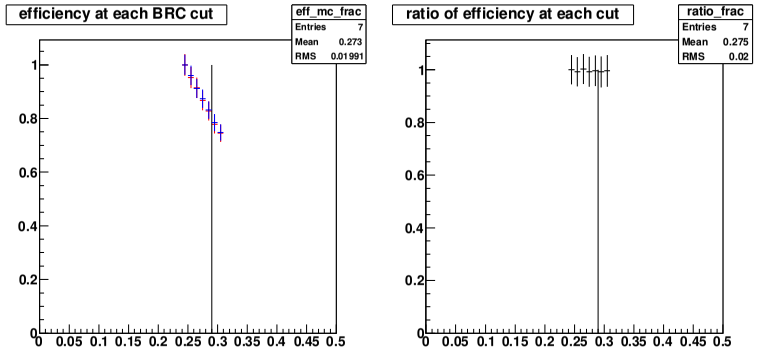


(c) mass region 2.0< m_{X_s} <2.4 GeV/c^2

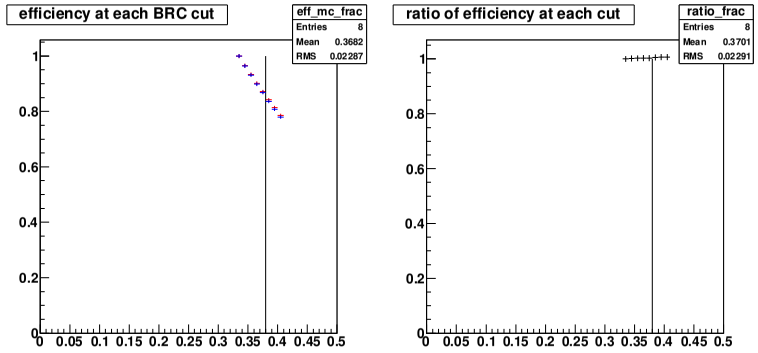


(d) mass region 2.4< m_{X_s} <2.8 GeV/c^2

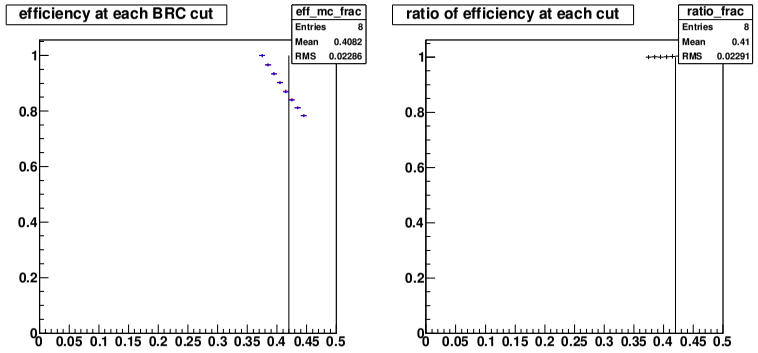
Figure 8.2: The efficiencies for MC (red) and data (blue) and their ratios when comparing the number of events that pass a given cut on the SSC response in the BRC sideband to the number that pass a cut 0.05 below the signal region cut. The x-axis reflects the cut location on the SSC response. The vertical black line indicates the location of the signal region requirement.



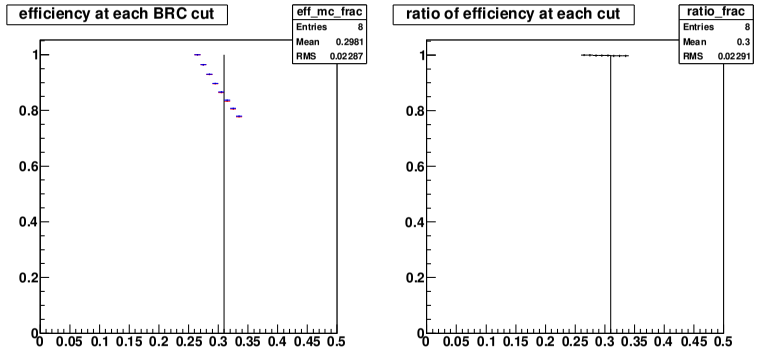
(a) mass region 0.6< m_{X_s} <1.1 GeV/c^2



(b) mass region 1.1< m_{X_s} <2.0 GeV/c^2



(c) mass region 2.0< m_{X_s} <2.4 GeV/c^2



(d) mass region 2.4< m_{X_s} <2.8 GeV/c^2

Figure 8.3: The efficiencies for MC (red) and data (blue) and their ratios when comparing the number of events that pass a given cut on the SSC response in the BRC sideband to the number that pass a cut 0.05 below the signal region cut. The x-axis reflects the cut location on the BRC response. The vertical black line indicates the location of the signal region requirement.

Section 8.1 since the output of the π^0 veto classifier is input to the BRC, and consequently would not function as an effective sideband for this classifier. The ratio of efficiencies between the MC and data at the SSC cut location as compared to five bins looser in the SSC is given in Table 8.2.

Table 8.2: The ratio between MC and data of changes in efficiencies when comparing the signal region cut value to five bins looser than the signal region cut value for the sideband with the π^0 veto > 0.

π^0-veto$>$ 0 sideband		
mass region	SSC cut value	ratio
0.6-1.1	0.17	1.001 $\pm$ 0.026
1.1-2.0	0.21	0.997 $\pm$ 0.010
2.0-2.4	0.08	1.019 $\pm$ 0.012
2.4-2.8	0.15	0.972 $\pm$ 0.014

8.3 High m_{X_s} Sideband, Peaking $B\overline{B}$

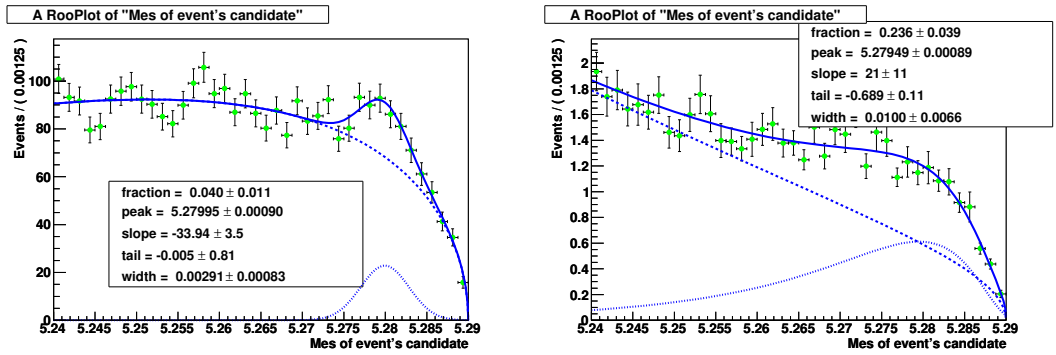
As shown in Table 7.5, we fix the total number of peaking-$B\overline{B}$ events in our final fit to the data in each mass bin. We use a high X_s mass bin as a sideband to evaluate the necessity of including a systematic correction associated with fixing this quantity and the shape of the peaking $B\overline{B}$ distribution. We choose to use a high-mass bin, specifically the bin 2.9$<$ m_{X_s} $<$3.0 GeV/c^2, as a sideband for this quantity because the signal and cross-feed contributions in this mass bin are expected to be negligible, so any peaking component in the MC and data should be almost purely peaking-$B\overline{B}$ backgrounds.

To evaluate the accuracy of our MC-based method for determining the peaking-$B\overline{B}$ amount, we duplicate much of the procedure described in Section 7.1.3 in this higher-mass bin. First, based on $B\overline{B}$ MC in this mass bin, we fit the parameters of a Nvs to describe the peaking component and extract the amount of peaking $B\overline{B}$ expected in the data. The parameters of our Nvs are shown in the inset of Figure 8.4(a). There is a total of 3311 $B\overline{B}$ MC events in this mass bin, so the fitted fraction corresponds to 132$\pm$36 peaking-$B\overline{B}$ events.

Next, we fix the shape and distribution of the cross-feed MC events according to the pa-

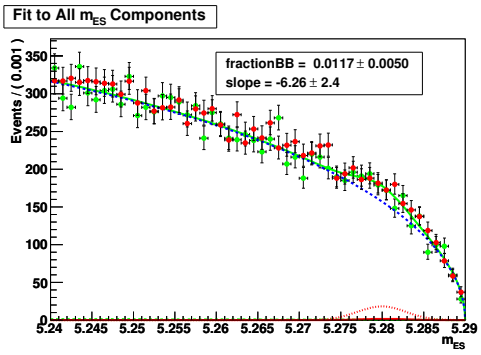
rameters given in Figure 8.4(b). We fix parameters such as fraction of signal to signal+cross-feed and fraction of signal+cross-feed to total number of events from the MC based on the number of these events given in Table 8.3. The shape of the negligible signal contribution we fix to be the same as that of the other mass bins.

In the final fit to the data in mass bin 2.9< m_{X_s} <3.0 GeV/c^2, we allow the peaking-$B\overline{B}$ contribution to float as well as the slope of the combinatoric Argus function. The final fit to the data in this mass bin is shown in Figure 8.4(c), and yields a best fit value of the number of peaking $B\overline{B}$ to be 133.4.±56.5.



(a) The peaking $B\overline{B}$ Nvs function parameters and fractional contribution of the peaking $B\overline{B}$ in the MC.

(b) The cross-feed Nvs and Argus function parameters.



(c) The total fit to data (green points) in mass bin 2.9< m_{X_s} <3.0 GeV/c^2 with the MC shown by the red points. The $B\overline{B}$ Nvs is shown by the dotted red line, the combinatoric Argus is the dashed blue line, and the total fit is the green line. Other PDF components are not visible.

Figure 8.4: The high X_s mass bin distributions.

We are too statistics-limited to make any meaningful corrections based on this sideband

Table 8.3: The different MC components and total amount of data in the 2.9< m_{X_s} <3.0 GeV/c^2 bin.

MC type	N_{events}
signal	13
cross-feed	54
B^+B^-	2111
$B^0\overline{B}^0$	1200
$c\bar{c}$	3718
uds	4534
$\tau^+\tau^-$	18
total MC	11648
data	11303

study, but the agreement between MC and data is encouraging. We take the agreement to mean our MC adequately models the $B\overline{B}$ in the data, and we will rely on MC to give us the uncertainty on the shape of the peaking $B\overline{B}$ distributions.

In order to evaluate the systematic uncertainty associated with fixing the number of peaking $B\overline{B}$, we instead decide to turn again to the π^0-veto sideband, described in Section 8.2. As shown in Figure 8.5, in which we place all of our selection requirements on the $B\overline{B}$ MC (and require $m_{ES} > 5.27$) and look at the truth matched source of the high-energy photon in the $B\overline{B}$ background, the predominant source of peaking background $B\overline{B}$ events comes from a π^0 faking our high-energy photon. We therefore conclude that this sideband can give us meaningful information about the uncertainty we should expect on the number of peaking-$B\overline{B}$ events.

In order to adequately estimate our uncertainty in the number of peaking-$B\overline{B}$ events, we decide to fit the data over three mass ranges in the π^0 sideband. We choose to fit the regions (1.1-2.0 GeV, 2.0-2.4 GeV, and 2.4-2.8 GeV) to reflect our separate selection requirement regions (the $K^*(892)$ region does not have sufficient $B\overline{B}$ statistics to make a meaningful fit, and will be described by our fit to the 1.1-2.0 GeV region). We fit the number of peaking $B\overline{B}$ in the data in this sideband, having fixed the cross-feed and signal-yield (so $B\overline{B}$ should be the only floating peaking component). We add the uncertainty from this fit to the sideband in quadrature to the uncertainty already reported from fitting the amount of peaking $B\overline{B}$

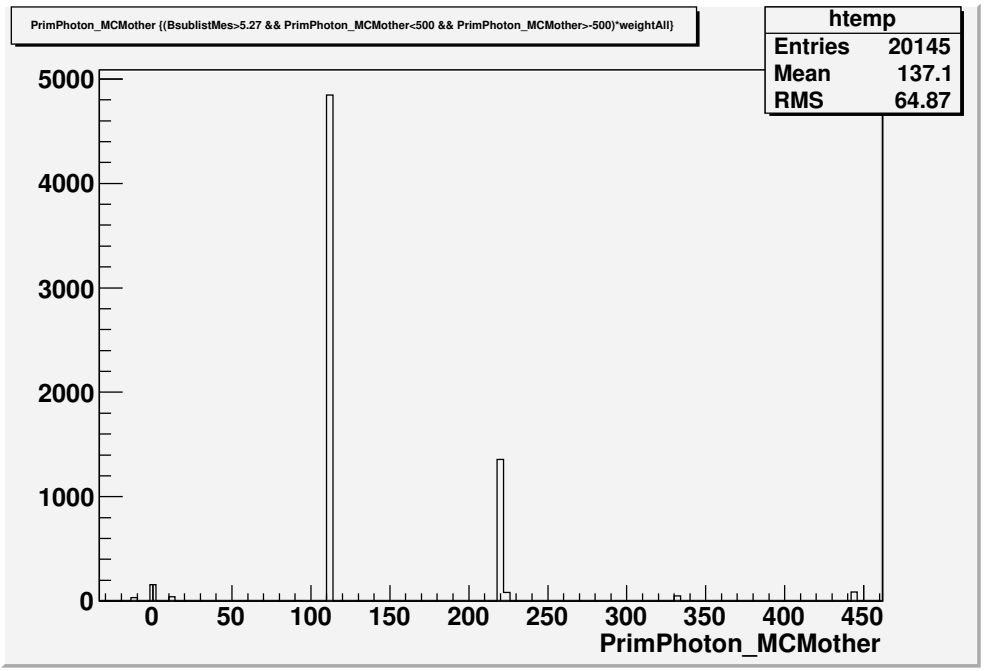


Figure 8.5: The source of the high-energy photon in the $B\overline{B}$ MC. We have plotted the LundID, which is 111 for a π^0 and 211 for an η, the two highest sources shown. We have applied all of our selection cuts in this plot and required m_{ES} >5.27 to restrict our plot to the peaking region of m_{ES}.

in the $B\overline{B}$ MC given in Table 7.5. These two uncertainties, added in quadrature, will be the amount we fluctuate the number of peaking $B\overline{B}$ in our fits to signal yield to evaluate a systematic uncertainty on our choice of fixing the amount of peaking $B\overline{B}$. We find in our sideband fits that we need to place a loose cut on the BRC (in addition to our cut on the SSC) to limit the combinatoric background. We report the number of peaking-$B\overline{B}$ events expected in the π^0 sideband of each of our three mass regions from our $B\overline{B}$ MC fit, the fit to the total MC set when we let only this peaking contribution float, and the fit to data with the resulting precision in Table 8.4.

Table 8.4: The amount of peaking $B\overline{B}$ expected from a fit to just the $B\overline{B}$ MC (N_{BB}^{BB}), a fit to the full MC set (N_{BB}^{full}), and a fit to the data (N_{BB}^{data}) in the π^0 sideband. We also explicitly report the uncertainty on this fit to the data.

Region	N_{BB}^{BB}	N_{BB}^{full}	N_{BB}^{data}	Uncertainty
$1.1 < m_{X_s} < 2.0$ SSC>0.21 BRC>0.10	311.1 ± 7.2	433.6 ± 123.6	323.3 ± 120.3	37.2%
$2.0 < m_{X_s} < 2.4$ SSC>0.08 BRC>0.10	1121.1 ± 59.9	1159.08 ± 151.5	1200.4 ± 153.5	12.8%
$2.4 < m_{X_s} < 2.8$ SSC>0.15 BRC>0.10	605.7 ± 62.7	618.7 ± 83.6	590.0 ± 85.48	14.5%

We use the uncertainty from our fits to data to supplement the uncertainty on our yield of peaking $B\overline{B}$ events. This gives us new values for the unceratinty reported already in Table 7.5, the new values are given in Table 8.5.

Table 8.5: The number of peaking $B\overline{B}$ events in each mass bin with the uncertainty from our fit to MC supplemented by the uncertainty from our fit to the data in the π^0 sideband (the uncertainty from the fit to MC is added in quadrature to the uncertainty reported in Table 8.4).

m_{X_s} range	$N_{B\overline{B}}$-peaking from fit to MC
0.6 to 0.7	0.00 ± 4.7
0.7 to 0.8	0.00 ±12.6
0.8 to 0.9	41.2 ± 15.8
0.9 to 1.0	38.8 ± 15.3
1.0 to 1.1	18.7 ± 8.8
1.1 to 1.2	12.0± 7.4
1.2 to 1.3	26.6 ± 11.5
1.3 to 1.4	17.1 ± 9.9
1.4 to 1.5	41.0 ± 17.6
1.5 to 1.6	41.1 ± 30.4
1.6 to 1.7	52.6 ± 22.5
1.7 to 1.8	43.7 ± 20.3
1.8 to 1.9	355.2 ± 133.3
1.9 to 2.0	38.8 ± 20.2
2.0 to 2.2	376.7 ± 58.5
2.2 to 2.4	501.4 ± 71.7
2.4 to 2.6	290.1 ± 48.6
2.6 to 2.8	421.3 ± 67.5

Chapter 9

Fragmentation Studies

The sum of exclusive nature of this analysis implies a unique set of systematic uncertainties associated with the hadronization of the s quark. Our signal MC uses a certain set of JETSET parameters [33] as its default setting for the X_s hadronization (or, used here interchangeably, fragmentation). The default, a phase-space hadronization model, is not the only choice, and, indeed, is not expected to model the fragmentation of the quark system perfectly.

However, our analysis is dependent on this hadronization model in two aspects. First, a change in the hadronization model will impact the relative frequency of final states among the 38 modes we reconstruct. Second, such a change will impact how often the X_s system ends up in any of our 38 modes; impacting our overall signal efficiency measurement (or estimate of the fraction of missing final states).

The first impact, the relative change within our 38 modes of how often the X_s system ends up in any one specific final state, would not have a large impact if our selection efficiency were the same for all of the 38 final states. But this is not experimentally feasible. Even with no requirements on events, such as our multivariate classifier cuts, the difference in final state multiplicity between our different modes introduces differences in detection efficiencies of the final state particles. The second impact, our overall signal efficiency based on the 38 modes we reconstruct, also needs to be investigated.

9.1 Different Fragmentation Models

The phase-space hadronization model we use as our default to model how the s quark hadronizes is not the only choice one can make. We identify two other classes of models one could use to describe the fragmentation of the system: the showering quarks model and a thermodynamics model. Two of these three different classes of models, phase-space hadronization and showering quarks models, have parameters within them that can then be easily varied as well. Often these parameters are shared between the models (such as the probability for spin-1 hadron formation). To get a sense of the differences between these models, we compare the predictions these three models make for the composition and total amount of the missing final states with parameter values set to our default in all of them.

We look at the missing fraction of final states based on generator level MC (no detector response, simply the predictions of the hadronization model at the particle level). We break down the missing final states based on s quark final state topology. The different missing final states are characterized as:

- $K_L = 1$, $K_S = 0$: This includes all final states that have 1 K_L and no K_S and are otherwise one of our 38 final states.

- FSR: This includes all 38 modes that we reconstruct in this analysis, plus the requirement of an extra photon in the final state with energy $E_\gamma > 30\,\mathrm{MeV}$.

- $\eta \to \gamma\gamma$: This includes all final states that have an $\eta \to \gamma\gamma$, plus kaon(s), and pions, but are not one of the 38 we reconstruct. These include states with $\geq 2\pi^0$ or $\geq 3\pi$ in addition to the kaon and η.

- $\leq 5 - body(\eta_{had} = 1)$: This includes final states in which some of the pions come from an $\eta \to 3\pi$ decay. This is mostly from $\eta \to 3\pi^0$, which gives a final state with one kaon and 3-4 pions, of which at least three are π^0s. Because of the large number of π^0s, they are not included among the final states investigated.

- $\leq 5 - body(\eta_{had} = 0)$: This includes events with the same multiplicity as the category immediately above, but none of the pions come from an $\eta \to 3\pi$ decay. Again, this

is mostly from $3\pi^0$s, but they are either produced inclusively or they come from the decays of other intermediate mesons.

- $\geq 6 - body(\eta_{had} \geq 1)$: final states with six or more bodies, in which at least three of the πs come from the decay of an η.
- $\geq 6 - body(\eta_{had} = 0)$: same as above, except no πs come from an η decay.
- $3Kaons(K_L = 0)$: includes final states with three kaons (no K_L) and more than one π.
- $3Kaons(K_L > 0, K_S > 0)$: This includes three kaon modes with at least one K_L; if $K_S = 0$, this would have already been included in the first type above.
- *Other* : This includes the rare meson decays and baryon decays, as well as any other modes not included above.

The default MC settings make a prediction on the missing fraction of events in each mass bin. To ascertain the breakdown of these missing fractions, we look at the generator level predictions of our signal MC by generating 500k B^+B^- and 500k $B^0\overline{B}^0$ MC events in each mass bin. The results are shown in Table 9.1.

9.1.1 Showering Quarks Fragmentation Model

As indicated above, the default settings within our JETSET modeled signal events treat the X_s system as an unstable hadron, whose mass is determined by the recoiling photon emitted in the $b \to s\gamma$ transition. This unstable hadron is then required to decay. The breakdown of final states is governed by a phase-space model of this X_s system's hadronic decay.

We can instead choose to model our X_s system as a system of showering quarks. The default model for showering quarks in JETSET uses the Lund string model [33]; for a more detailed explanation, ee the reference. To summarize, the q and $\overline{q}$ partons (in our analysis this is an $s\overline{u}$ or $s\overline{d}$, or their conjugates) move away from their common vertex, connected by a color flux tube, or, equivalently, a string. As they separate, the potential energy of

Table 9.1: The breakdown, by %, of the different missing fraction of events for the default MC settings. The events are generated with a flat photon spectrum in the mass bin, and there is no photon energy reweighting.

mass bin (GeV/c^2)	$K_L = 1$ ($K_S = 0$)	FSR	$\eta \to \gamma\gamma$	$\leq 5 - body$ ($\eta_{had} = 1$)	$\leq 5 - body$ ($\eta_{had} = 0$)	$\geq 6 - body$ ($\eta_{had} \geq 1$)
1.1-1.2	24.3	0.8	0.0	1.9	0.2	0.0
1.2-1.3	24.2	1.0	0.0	1.9	0.4	0.0
1.3-1.4	24.0	1.3	0.0	1.8	0.8	0.1
1.4-1.5	23.6	1.6	0.1	2.2	1.1	0.4
1.5-1.6	22.8	1.7	0.2	2.5	1.5	0.9
1.6-1.7	22.0	1.9	0.4	2.1	1.8	2.0
1.7-1.8	21.1	2.0	0.8	1.8	2.0	2.7
1.8-1.9	19.7	2.1	1.2	1.5	2.2	3.4
1.9-2.0	18.3	2.1	1.6	1.3	2.2	4.1
2.0-2.2	15.8	2.0	2.4	1.1	2.1	4.9
2.2-2.4	12.9	1.9	3.3	0.8	1.9	5.9
2.4-2.6	10.4	1.6	4.1	0.7	1.6	6.8
2.6-2.8	8.4	1.4	4.7	0.5	1.4	7.5

mass bin (GeV/c^2)	$\geq 6 - body$ ($\eta_{had} = 0$)	$3Kaons$ ($K_L = 0$)	$3Kaons$ ($K_L > 0, K_S > 0$)	$Other$	**Total % Missing**
1.1-1.2	0.0	0.0	0.0	0.0	**27.2**
1.2-1.3	0.0	0.0	0.0	0.0	**27.5**
1.3-1.4	0.0	0.0	0.0	0.0	**28.0**
1.4-1.5	0.2	0.0	0.0	0.1	**29.2**
1.5-1.6	1.1	0.1	0.9	0.1	**31.7**
1.6-1.7	2.7	0.1	1.1	0.1	**34.2**
1.7-1.8	5.0	0.2	1.3	0.1	**36.9**
1.8-1.9	8.8	0.4	1.5	0.1	**41.0**
1.9-2.0	12.9	0.6	2.3	0.1	**45.6**
2.0-2.2	19.3	1.5	3.3	0.6	**53.0**
2.2-2.4	26.4	2.7	4.6	1.3	**61.7**
2.4-2.6	31.8	3.9	6.1	2.0	**69.1**
2.6-2.8	35.8	5.0	7.4	2.9	**75.1**

the string increases and it may break into a new $q'\overline{q}'$ pair, thus giving two mesons ($q\overline{q}'$ and $q'\overline{q}$). If the invariant mass of either of these products is large enough, the process repeats itself until only on-shell hadrons remain.

At each breaking of the string, the produced quark and anti-quark are given offsetting transverse momenta, p_T (if the quark is given $+p_T$ the anti-quark is given $-p_T$); the initial $q\overline{q}$ pair has no transverse momentum excitations. The total p_T of a final hadron is made up of the p_T contributions of its constituent quarks. The amount of light-cone momentum, $z = E \pm p_T$, given to each new particle (+ for the q, − for the $\overline{q}$), is dictated by the choice of fragmentation function, $f(z)$.

There are several choices for $f(z)$ available in JETSET, but only two are applicable to our case (the others apply to charm quark and heavier jets). The default, the Lund symmetric fragmentation function, is:

$$f(z) = \frac{1}{z}(1-z)^a e^{(-bm_T^2/z)}, \tag{9.1}$$

where the parameters $a = 0.50$ and $b = 0.90\,\mathrm{GeV}^{-2}$ (the default values), and the Field-Feynman fragmentation function:

$$f(z) = 1 - a + 3a(1-z)^2, \tag{9.2}$$

where $a = 0.77$ (the default value). The impact of using this type of parametrization of the hadronization of the X_s system as compared to the default choice can be seen in Figure 9.1. As seen in the figure, we investigated several settings for the Field-Feynman parameter, with little impact on the breakdown of final states as compared to the default shower settings. The biggest difference comes from using quark showering versus phase-space hadronization. Shown in the figure is the relative breakdown between the states we reconstruct; the resulting missing final states are shown in Table 9.2.

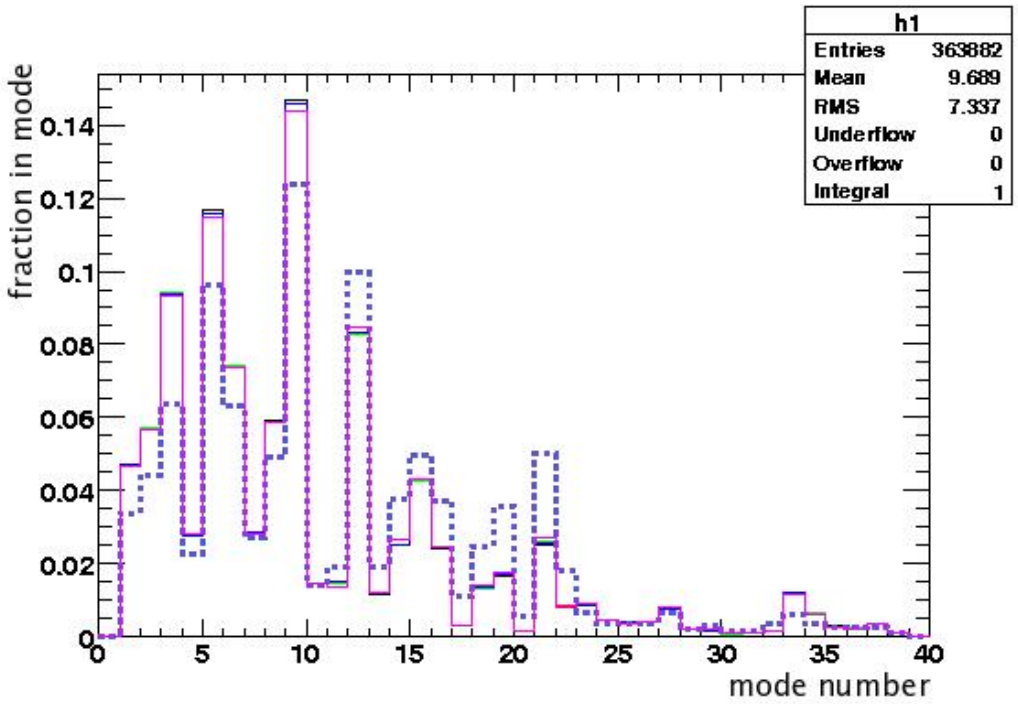


Figure 9.1: The distribution of the different 38 modes we reconstruct (the mode number is given by the x-axis), normalized to the total number of events within the 38 modes (the total area of each histogram is 1). The dashed purple line is the default setting, the pink thin line is the default shower setting, the other lines are the shower setting using the Field-Feynman fragmentation function with different values for its parameter a.

Table 9.2: The breakdown, by %, of the different missing fraction of events for the default showering quarks MC settings. The events are generated with a flat photon spectrum in the mass bin, and there is no photon energy reweighting except for the final line in which we assume our default BBU weights for the photon spectrum.

mass bin (GeV/c^2)	$K_L = 1$ ($K_S = 0$)	FSR	$\eta \to \gamma\gamma$	$\leq 5 - body$ ($\eta_{had} = 1$)	$\leq 5 - body$ ($\eta_{had} = 0$)	$\geq 6 - body$ ($\eta_{had} \geq 1$)
1.1-1.2	24.5	0.4	0.0	1.8	0.1	0.0
1.2-1.3	24.4	0.5	0.0	1.8	0.1	0.0
1.3-1.4	24.3	0.9	0.0	1.7	0.2	0.0
1.4-1.5	24.0	1.1	0.0	2.3	0.3	0.1
1.5-1.6	23.5	1.3	0.1	3.0	0.4	0.2
1.6-1.7	23.4	1.4	0.1	2.8	0.5	0.2
1.7-1.8	23.4	1.6	0.1	2.6	0.6	0.3
1.8-1.9	23.1	1.6	0.1	2.8	0.7	0.6
1.9-2.0	22.5	1.6	0.3	2.7	1.0	1.2
2.0-2.2	20.6	1.6	0.6	2.3	1.6	2.7
2.2-2.4	18.6	1.6	1.1	1.6	2.3	4.2
2.4-2.6	16.8	1.6	1.9	0.9	2.7	5.4
2.6-2.8	14.4	1.6	2.6	0.5	2.8	6.1

mass bin (GeV/c^2)	$\geq 6 - body$ ($\eta_{had} = 0$)	$3Kaons$ ($K_L = 0$)	$3Kaons$ ($K_L > 0, K_S > 0$)	$Other$	**Total % Missing**
1.1-1.2	0.0	0.0	0.0	0.0	**26.7**
1.2-1.3	0.0	0.0	0.0	0.0	**26.8**
1.3-1.4	0.0	0.0	0.0	0.0	**27.1**
1.4-1.5	0.0	0.0	0.0	0.0	**27.9**
1.5-1.6	0.0	0.0	0.9	0.1	**29.4**
1.6-1.7	0.0	0.0	1.1	0.1	**29.6**
1.7-1.8	0.0	0.0	1.0	0.1	**29.7**
1.8-1.9	0.1	0.0	1.8	0.1	**30.9**
1.9-2.0	0.5	0.1	2.8	0.1	**32.7**
2.0-2.2	2.2	0.3	4.5	1.9	**38.3**
2.2-2.4	5.8	0.8	4.6	4.0	**44.7**
2.4-2.6	10.5	2.0	4.9	3.8	**50.6**
2.6-2.8	15.7	3.4	6.4	4.5	**57.8**

9.1.2 Thermodynamics Model

In order to make sure we fully explore the effects of our default choice for final state production, we investigate a model that is very different in its approach to generating the breakdown of final states in our signal events. This model, the so called "thermodynamics model" was first proposed by Fermi [34] and further expanded on by Quigg and Rosner [35] to predict the multiplicity of final states of the F meson (now known as the D_s meson). We adapt this model to our analysis as follows:

The hadronic state (F meson in [35], X_s in this analysis) is imagined to be confined within some radius, R_0, at some temperature, T. Both the total energy and number of degrees of freedom (interpreted at average number of pions) are functions of T (for a good reference, see section 7.2, in particular equations 7.2.12 and 7.2.23 of [36] and correct by a fraction of 3/2, or the ratio of number of pion types to photon polarization states). Solving for T, one finds a relation between average decay multiplicity (interpreted as average number of pions in the final state) and the energy available for these degrees of freedom:

$$< n >= 2 + 0.528(\frac{M - M_A - M_B}{E_0/c^2})^{3/4}, \tag{9.3}$$

where $E_0 \equiv \hbar c/R_0$, M is the mass of our X_s hadronic system, M_A and M_B are the masses of the two daughter particles, and $< n >$ is the average multiplicity of extra pions in the event. For our analysis, we take E_0 to be f_B, or 0.175 GeV.

The 38 modes that we reconstruct can be broken down into three different types: i)$K\pi + n\pi$, ii)$K\eta + n\pi$, and iii)$3K + n\pi$. For the first (second) type we take $M_A = M_K$ and $M_B = M_\pi$ ($M_B = M_\eta$). For the third type of mode, we subtract a third particle, M_C, from equation 9.3 and set $M_A = M_B = M_C = M_K$ to determine the average pion multiplicity. We take M to be the central values of the different X_s mass bins (for bin 1.2-1.3 GeV/c^2 we take $M = 1.25\,\mathrm{GeV}/c^2$). The extra number of pions are then Poisson-distributed with a mean given by equation 9.3. We renormalize this Poisson distribution after imposing kinematic limits in each mass bin (for instance, decays of type $K\eta + n\pi$ in mass bin 1.2-1.3 GeV/c^2 cannot have more than one extra pion due to energy conservation). We assume

the breakdown between the different types of final states is approximately correct within the signal MC.

Based on the predictions of these three Poisson distributions within each mass bin, we can use the generator-level information to predict how much of the final states of our signal are now missing. This method obviously makes no predictions about FSR or $Other$ types of missing decays, but it can give us an indication of how large a systematic uncertainty due to missing higher multiplicity final states we are introducing. The breakdown of missing final states after reweighting based on this thermodynamics model are given in Table 9.3.

By comparing these three different classes of models, we see that their predictions for total amount of missing final states become more different as we go to higher X_s mass. Since ultimately what we want to know is the coverage of our 38 modes compared to the inclusive $b \to s\gamma$ transition rate, we compare the total missing final states (the final column in tables 9.1-9.3) between the models to evaluate our uncertainty in this quantity (see Section 9.3).

9.2 Fragmentation Study

After unblinding our data, we need to look at the breakdown of our 38 final states and correct our MC accordingly in order to correctly model our efficiency for reconstructing our 38 final states (ϵ_{38}). We performed several checks on the most effective way of doing this based on a MC study which we will present here.

9.2.1 MC-Based Fragmentation Study

The goal of this study is to better model our value of ϵ_{38}, or the efficiency of reconstructing our signal if it is generated in one of our 38 final states. The definition of ϵ_{38} is therefore:

$$\epsilon_{38} = \frac{N_{fit-38}}{N_{gen-38}}, \tag{9.4}$$

where N_{fit-38} is the number of signal events we obtain from our fits to data in a given mass bin (by definition these will be in our 38 decay modes), and N_{gen-38} is the number of

Table 9.3: The breakdown, by %, of the different missing fraction of events for reweighting the events based on the thermodynamics model detailed in the text. The events are generated with a flat photon spectrum in the mass bin, and there is no photon energy reweighting except for the final line in which we assume our default BBU weights for the photon spectrum.

mass bin (GeV/c^2)	$K_L = 1$ ($K_S = 0$)	FSR	$\eta \to \gamma\gamma$	$\leq 5 - body$ ($\eta_{had} = 1$)	$\leq 5 - body$ ($\eta_{had} = 0$)	$\geq 6 - body$ ($\eta_{had} \geq 1$)
1.1-1.2	24.1	0.8	0.0	1.9	2.7	0.0
1.2-1.3	23.2	0.9	0.0	1.9	2.1	0.0
1.3-1.4	22.2	1.1	0.0	1.6	2.2	0.0
1.4-1.5	21.7	1.4	0.0	2.0	1.7	0.0
1.5-1.6	20.6	1.5	0.4	2.1	1.9	0.8
1.6-1.7	19.8	1.7	0.7	2.0	1.9	1.3
1.7-1.8	19.0	1.8	0.9	1.8	1.9	1.7
1.8-1.9	18.0	2.0	1.2	1.7	2.1	2.1
1.9-2.0	17.0	2.0	1.5	1.6	2.1	2.8
2.0-2.2	15.5	2.1	1.9	1.5	2.1	3.6
2.2-2.4	13.7	2.0	2.6	1.3	2.0	4.7
2.4-2.6	11.9	1.9	3.2	1.1	1.8	5.7
2.6-2.8	10.3	1.7	3.7	0.9	1.7	6.6

mass bin (GeV/c^2)	$\geq 6 - body$ ($\eta_{had} = 0$)	$3Kaons$ ($K_L = 0$)	$3Kaons$ ($K_L > 0, K_S > 0$)	$Other$	**Total % Missing**
1.1-1.2	0.0	0.0	0.0	0.0	**29.6**
1.2-1.3	3.5	0.0	0.0	0.0	**31.6**
1.3-1.4	6.2	0.0	0.0	0.0	**33.3**
1.4-1.5	7.8	0.0	0.0	0.1	**34.7**
1.5-1.6	10.0	0.1	0.9	0.1	**38.3**
1.6-1.7	12.3	0.1	1.1	0.1	**41.0**
1.7-1.8	14.5	0.1	1.3	0.1	**43.2**
1.8-1.9	16.9	0.5	1.6	0.1	**46.1**
1.9-2.0	18.8	0.9	2.5	0.1	**49.3**
2.0-2.2	21.7	1.5	3.4	0.6	**53.8**
2.2-2.4	25.2	2.3	4.5	1.3	**59.5**
2.4-2.6	28.0	3.3	5.8	2.0	**64.8**
2.6-2.8	30.4	4.3	7.1	2.9	**69.7**

$b \to s\gamma$ signal events that are actually produced in one of our 38 decay modes.

One can recognize that this definition of ϵ_{38} is equivalent to:

$$\epsilon_{38} = \frac{\sum_{i=1}^{i=38} N_{fit,i}}{\sum_{i=1}^{i=38} N_{gen,i}}, \tag{9.5}$$

where we are simply explicitly referring to the 38 modes we use in our analysis (we are NOT indicating that we individually fit each of our 38 modes). Since the reconstruction efficiency for each of the modes is not the same, if our MC has the wrong breakdown of modes, the value of ϵ_{38} would not be accurate. In order to correct the breakdown of our final states in the MC, we perform fits comparing groupings of modes to data, allowing us to see which modes are over-produced and which modes are under-produced in the MC. This is equivalent to saying we reweight the MC based on weights we obtain from fits to data, and change our value of ϵ_{38} according to:

$$\epsilon_{38,new} = \frac{\sum_{i=1}^{i=38} N_{fit,i} w_i}{\sum_{i=1}^{i=38} N_{gen,i} w_i}, \tag{9.6}$$

where the w_i are the weights we obtain from comparing MC to data (if there are multiple modes grouped together, the w_i for each mode in the group is the same). Before performing this study on data, we performed a mock study on MC to ensure the efficacy of our procedure.

As a mock fragmentation study, we evaluated our procedure by trying to recover the ϵ_{38} for the "showering quarks model," based on the Lund string modeling of the quark system (described in Section 9.1.1), which we consider as "data," starting from our default phase-space hadronization model in our MC.

The procedure of the fragmentation study takes groupings of modes in our signal MC, and compares the %-contribution of each group to the %-contribution amount found in fits to the data. Since we are ultimately fitting signal+cross-feed yield in our fits to data (as described in Section 7), this is the relevant quantity we compare in our mock fragmentation study.

There is a mass dependence in the differences between the showering quarks model

(again, what we consider "data" in this study) and the phase-space hadronization model for the different modes. To account for this behavior (in particular, should there be such a mass dependence also present in the on peak data), we separate the m_{X_s} spectrum into four regions: 1.1-1.5 GeV, 1.5-2.0 GeV, 2.0-2.4 GeV, and 2.4-2.8 GeV. This set of groupings, which we also use for our fragmentation study with the on-peak data, allows us to adequately account for a mass dependence without unnecessarily decreasing the statistics in our fits to data.

The procedure we adopt to evaluate how to correct our MC to better match data is iterative in nature. We evaluate the ratio of %-yield for each of our groups of modes in data vs. MC, and use this ratio to reweight the generated MC (both the signal and cross-feed are reweighted from our findings, the cross-feed is reweighted based on which mode the event was generated in). We then re-evalaute the shape of the cross-feed PDF based on the new distribution of our 38 modes (since some modes peak in cross-feed more than others, this shape is expected to change). We use these new parameters and refit the data and compare this to our redistributed MC. This gives us a new value for the weight to be applied to each group; we use this new value and repeat the procedure. This fitting and reweighting procedure is iterated until the data and MC are found to be consistent to within $\sim$1% (generally $\leq$4 iterations). The final error on the weight is taken from the statistical error on the fits to data for the different groups. Since we are applying these weights to individual mass bins, as opposed to the mass regions given above, we scale the fitting error up by the quantity:

$$\sigma_{bin,i} = \sigma_{region,i} \times \sqrt{\frac{N_{region}}{N_{bin}}}, \tag{9.7}$$

where the $\sigma_{bin,i}$ and $\sigma_{region,i}$ represent the uncertainty on weight i, taken from the statistical error on the fit to the data of group i, N_{region} is the number of signal MC events generated in the whole region and N_{bin} is the number of signal events generated within the bin. The motivation for this scaling also comes from the potential mass dependence of the corrections: If one bin has more events than another, the correction for a given group, i, will be based more on this bin and, consequently, our uncertainty on the correction in the lower-statistics

bin should be higher.

9.2.1.1 Groupings

With the procedure in place, there still remains the choice of what groups of modes to use in evaluating our weights. We try two different sets of groupings, one based on minimizing cross-feed between groups (generally 6-8 groups, depending on mass region), the second in which we take 10 groups that are more "physics" based (they are divided more by mode-topology, and are the same groups for every mass region). Ultimately, though the final choice of groups is often quite different, we find consistent values for the corrected ϵ_{38}, and will use the 10 groups based on mode topology since it reflects a more finely grained study (the method of choosing to minimize cross-feed ultimately gives less than 10 groups for each mass range). We therefore only present the results of this study reflecting this grouping scheme given in Table 9.4.

Table 9.4: The 10 groups used in our fragmentation study based on mode topology. We only give the mode number, please see Table 3.1 for definitions of the specific modes.

Group Definition	Mode Numbers
2 bodies without π^0	1, 3
2 bodies with 1 π^0	2, 4
3 bodies without π^0	5, 8
3 bodies with 1 π^0	6, 9
4 bodies without π^0	11, 14
4 bodies with 1 π^0	12, 15
3/4 bodies with 2 π^0s	7, 10, 13, 16
5 bodies with 0-2 π^0s	17-22
$\eta \to \gamma\gamma$	23-32
$3K$ modes	33-38

9.2.1.2 Evaluating Weights

With the grouping, as well as the procedure on how to use this grouping, in place, all that remains is to iterate over the data and evaluate the weights we should use on the generated MC. We demonstrate this procedure, as pedagogically as possible, in Table 9.5, in which we show the evolution of the weights for the 10 groups in mass region 1.5-2.0 GeV.

Table 9.5: The evolution of the weights to be used to better match the MC to the data based on the 10 groups we identified, in mass region 1.5-2.0 GeV for our mock fragmentation study.

modes	% in MC	% in Data	Fragmentation Weight (%-Data/%-MC)
Initial Weights			
1,3	13.95	20.21 ± 0.29	1.448 ± 0.021
2,4	4.50	5.94 ± 0.21	1.320 ± 0.046
5,8	24.45	28.22 ± 0.52	1.540 ± 0.021
6,9	12.65	14.83 ± 0.43	1.172 ± 0.034
11,14	14.30	7.35 ± 0.73	0.514 ± 0.051
12,15	9.35	6.26 ± 0.49	0.669 ± 0.053
7,10,13,16	2.61	2.23 ± 0.34	0.856 ± 0.129
17-22	5.71	1.42 ± 0.48	0.249 ± 0.083
23-32	11.63	12.61 ± 1.06	1.084 ± 0.091
33-38	0.84	0.94 ± 0.06	1.120 ± 0.070
Apply Weight Based on Generated MC and Refit Peaking XF in MC			
1,3	19.47	19.71 ± 0.29	1.012 ± 0.015
2,4	5.80	5.84 ± 0.20	1.007 ± 0.035
5,8	27.09	27.39 ± 0.51	1.011 ± 0.019
6,9	13.95	14.13 ± 0.42	1.013 ± 0.030
11,14	9.45	8.96 ± 0.84	0.948 ± 0.089
12,15	6.74	6.70 ± 0.52	0.994 ± 0.078
7,10,13,16	2.26	2.24 ± 0.34	0.988 ± 0.148
17-22	3.03	2.52 ± 0.71	0.832 ± 0.235
23-32	11.28	11.59 ± 1.00	1.027 ± 0.088
33-38	0.93	0.93 ± 0.06	1.007 ± 0.063
Apply Weight Based on Generated MC and Refit Peaking XF in MC			
1,3	19.64	19.67 ± 0.29	1.0017 ± 0.0148
2,4	5.82	5.83 ± 0.20	1.0011 ± 0.0348
5,8	27.26	27.30 ± 0.51	1.0014 ± 0.0188
6,9	14.04	14.06 ± 0.42	1.0014 ± 0.0296
11,14	9.21	9.15 ± 0.86	0.9932 ± 0.0932
12,15	6.66	6.68 ± 0.52	1.0028 ± 0.0783
7,10,13,16	2.24	2.24 ± 0.34	0.9996 ± 0.1499
17 to 22	2.88	2.71 ± 0.74	0.9396 ± 0.2561
23 to 32	11.32	11.44 ± 0.99	1.0106 ± 0.0873
33 to 38	0.93	0.93 ± 0.06	1.0014 ± 0.0628
Apply Weight Based on Generated MC and Refit Peaking XF in MC			
1,3	19.66	19.66 ± 0.29	1.000 ± 0.015
2,4	5.82	5.83 ± 0.20	1.000 ± 0.035
5,8	27.28	27.29 ± 0.51	1.000 ± 0.019
6,9	14.05	14.05 ± 0.42	1.000 ± 0.030
11,14	9.18	9.17 ± 0.86	0.999 ± 0.094
12,15	6.66	6.66 ± 0.52	1.001 ± 0.078
7,10,13,16	2.24	2.24 ± 0.34	1.000 ± 0.150
17 to 22	2.84	2.79 ± 0.75	0.983 ± 0.264
23 to 32	11.35	11.38 ± 0.98	1.003 ± 0.087
33 to 38	0.93	0.93 ± 0.06	1.001 ± 0.063

Mode	Total Weight
1,3	1.469 ± 0.022
2,4	1.331 ± 0.046
5,8	1.169 ± 0.022
6,9	1.189 ± 0.035
11,14	0.483 ± 0.045
12,15	0.668 ± 0.052
7,10,13,16	0.846 ± 0.127
17 to 22	0.192 ± 0.051
23 to 32	1.129 ± 0.098
33 to 38	1.130 ± 0.071

We perform this procedure over all four mass regions and determine weights to apply to the signal and cross-feed MC events, based on the mode in which they were generated, to modify our value of ϵ_{38} so as to better reflect the data. The weights we obtain for the 10 groups in each mass region are given in Table 9.6. Over the course of the study, we had one instance of a consistently negative yield in our fits, that being the fit to modes 7, 10, 13, and 16 in the mass region 2.4-2.8 GeV. We elect to fix the weight of this group to 1, and take the full value of the fitting error as our uncertainty. We recognize this may be overly conservative, but we interpret this result to mean we have no information about what the weight should be (a physically meaningless negative central value in the fit), but there are many values that could be consistent with our observation (a large error) so we try them all (use the large error in our ultimate determination of the error on the corrected ϵ_{38}).

We use these weights and apply them to the events in each mass bin based on the mode in which an event is generated. Reweighting the signal MC in this way allows us to account for what our ϵ_{38} should be. The results of this study are shown in Table 9.7 in which we give the default, phase-space hadronization, ϵ_{38}, as well as the corrected value we get for each bin when using the weights evaluated with our 10 groups. We compare the amount of events we predict to have been generated in our 38 modes based on these corrected values of ϵ_{38}, as well as the actual number of events generated in our 38 modes in the showering quarks MC (again, what we used as data for this study) in Table 9.8.

Finally, as a cross check that this method works in either direction, we checked to make sure that we could recover the default efficiency if we start from the showering quarks model's ϵ_{38}. We see in Table 9.9, in which we only perform this check using the 10 groups for the mass region 1.5-2.0 GeV, that indeed we can recover the default value when starting with the showering quarks model.

9.2.2 Fragmentation Study Performed on Data

The first task after unblinding the data we perform is to implement the fragmentation study procedure we described in Section 9.2.1 above. The final weights we find in each mass region for each of the groups when comparing to unblinded data are reported in Table 9.10- 9.13,

Table 9.6: The weights determined for correcting the MC to match the showering quarks "data" to better determine ϵ_{38}, with the 10 groups we used in our study.

Mass region	Grouped Modes	Weights
$1.1 < m_{X_s} < 1.5$	1, 3	1.336 ± 0.024
	2, 4	1.172 ± 0.053
	5, 8	1.007 ± 0.024
	6, 9	1.022 ± 0.040
	11, 14	0.259 ± 0.080
	12, 15	0.563 ± 0.095
	7, 10, 13, 16	0.900 ± 0.173
	17-22	0.045 ± 0.075
	23-32	1.163 ± 0.143
	33-38	0.046 ± 0.061
$1.5 < m_{X_s} < 2.0$	1, 3	1.469 ± 0.022
	2, 4	1.331 ± 0.046
	5, 8	1.169 ± 0.022
	6, 9	1.189 ± 0.035
	11, 14	0.483 ± 0.045
	12, 15	0.668 ± 0.052
	7, 10, 13, 16	0.846 ± 0.127
	17-22	0.192 ± 0.051
	23-32	1.129 ± 0.098
	33-38	1.130 ± 0.071
$2.0 < m_{X_s} < 2.4$	1, 3	0.260 ± 0.058
	2, 4	0.249 ± 0.108
	5, 8	0.788 ± 0.089
	6, 9	0.532 ± 0.115
	11, 14	0.993 ± 0.189
	12, 15	0.822 ± 0.200
	7, 10, 13, 16	0.892 ± 0.406
	17-22	0.669 ± 0.217
	23-32	2.532 ± 0.480
	33-38	1.861 ± 0.252
$2.4 < m_{X_s} < 2.8$	1, 3	0.080 ± 0.091
	2, 4	0.087 ± 0.162
	5, 8	0.505 ± 0.196
	6, 9	0.642 ± 0.258
	11, 14	1.253 ± 0.557
	12, 15	0.300 ± 0.557
	7, 10, 13, 16	1.000 ± 9.128
	17-22	0.161 ± 0.979
	23-32	4.576 ± 1.532
	33-38	0.589 ± 0.730

Table 9.7: The default value for ϵ_{38}, as well as the values found from applying the corrections based on the weights given above, and the actual showering quarks MC value of ϵ_{38}.

Mass Bin	Default ϵ_{38} (%)	Corrected ϵ_{38}	Showering Quark ϵ_{38}
1.1-1.2	11.51	12.08 ± 0.41	12.35
1.2-1.3	11.59	12.27 ± 0.33	12.40
1.3-1.4	10.65	11.61 ± 0.31	11.96
1.4-1.5	9.47	10.73 ± 0.30	11.17
1.5-1.6	8.39	9.88 ± 0.27	10.40
1.6-1.7	7.20	8.75 ± 0.26	9.07
1.7-1.8	5.52	7.19 ± 0.27	7.30
1.8-1.9	4.54	6.17 ± 0.28	6.21
1.9-2.0	3.26	4.71 ± 0.28	4.20
2.0-2.2	4.04	3.50 ± 0.29	4.00
2.2-2.4	3.11	2.63 ± 0.35	2.26
2.4-2.6	2.29	2.47 ± 1.33	1.34
2.6-2.8	2.30	2.63 ± 2.01	1.30

Table 9.8: The number of events generated in our 38 decay modes predicted by our corrected value of ϵ_{38} (N_{gen-38}) and the actual number of events generated in our 38 modes in the showering quarks model of signal MC (Actual SQ 38).

Groupings based on minimizing cross-feed					
Mass Bin	Error - Stat.	Error - ϵ_{38}	Error Total	N_{gen-38}	Actual SQ 38
1.1-1.2	271.8	147.9	309.4	4358.0	4653.5
1.2-1.3	327.1	225.4	397.2	8381.4	9023.8
1.3-1.4	443.2	382.9	585.7	14341.7	15216.6
1.4-1.5	549.5	595.2	810.0	21287.0	21638.7
1.5-1.6	645.3	700.0	952.1	25616.4	25610.0
1.6-1.7	729.3	832.4	1106.7	28013.1	27820.4
1.7-1.8	806.9	1003.3	1287.5	26717.4	26618.0
1.8-1.9	873.6	1004.6	1331.3	22137.8	22291.5
1.9-2.0	898.9	928.0	1291.9	15609.7	17165.5
2.0-2.2	2053.3	2075.1	2919.3	25044.0	20248.5
2.2-2.4	2412.7	1712.1	2958.5	12865.0	10070.8
2.4-2.6	2256.3	4530.0	5060.8	8412.9	5817.4
2.6-2.8	2457.5	3205.0	4038.8	4193.7	3716.1

Table 9.9: A cross-check to ensure that we can recover the default reconstruction efficiency for our 38 modes if we start with the showering quarks efficiency (SQ ϵ_{38}).

Mass Bin	SQ ϵ_{38}	Corrected ϵ_{38}	Default ϵ_{38}
1.5-1.6	10.40	8.69 ± 0.34	8.40
1.6-1.7	9.07	7.06 ± 0.34	7.20
1.7-1.8	7.30	4.97 ± 0.37	5.53
1.8-1.9	6.21	4.17 ± 0.34	4.54
1.9-2.0	4.20	2.62 ± 0.30	3.26

where we have defined the "Data Subset" column for ease of plotting different quantities later. It should also be noted that this fragmentation study was performed after "detection efficiency" corrections were taken into account (described in more detail in Section 10).

Table 9.10: The reweighing factors found in mass region m_{X_s}=1.1-1.5 GeV.

Data Subset	Definition	Modes Used	Frag. Weight Found in m_{X_s}=1.1-1.5 GeV
1	2 bodies without π^0	BiType = 1,3	0.650 ± 0.027
2	2 bodies with 1 π^0	BiType = 2,4	0.533 ± 0.051
3	3 bodies without π^0	BiType = 5,8	1.195 ± 0.025
4	3 bodies with 1 π^0	BiType = 6,9	1.701 ± 0.047
5	4 bodies without π^0	BiType = 11,14	0.337 ± 0.079
6	4 bodies with 1 π^0	BiType = 12,15	1.242 ± 0.125
7	3/4 bodies with 2 π^0s	BiType = 7,10,13,16	0.563 ± 0.186
8	5 bodies with 0-2 π^0s	BiType = 17-22	$1.000^{+1.048}_{-1.000}$
9	$\eta \to \gamma\gamma$	BiType = 23-32	0.938 ± 0.145
10	3K modes	BiType = 33-38	0.000 ± 0.000

In performing these fragmentation studies, there were instances where our procedure described in Section 9.2.1 was not robust enough to account for fluctuations in the data resulting in negative yields in the fit. These "exceptions," and the method we used to account for them were:

- Mass Range 1.1-1.5 GeV/c^2 - Group 8 (5-body modes). This group had a negative yield in the fits to the data. The resulting weight would have been -0.3142±0.3310. Since a negative result is meaningless, we removed the group from consideration when calculating the other group's weights, and assign a weight of $1^{+1.0536}_{-1.000}$. The value of

Table 9.11: The reweighing factors found in mass region m_{X_s}=1.5-2.0 GeV.

Data Subset	Definition	Modes Used	Frag. Weight Found in m_{X_s}=1.5-2.0 GeV
1	2 bodies without π^0	BiType = 1,3	0.376 ± 0.033
2	2 bodies with 1 π^0	BiType = 2,4	0.276 ± 0.060
3	3 bodies without π^0	BiType = 5,8	1.008 ± 0.037
4	3 bodies with 1 π^0	BiType = 6,9	1.026 ± 0.060
5	4 bodies without π^0	BiType = 11,14	1.339 ± 0.101
6	4 bodies with 1 π^0	BiType = 12,15	1.156 ± 0.108
7	3/4 bodies with 2 π^0s	BiType = 7,10,13,16	1.365 ± 0.298
8	5 bodies with 0-2 π^0s	BiType = 17-22	0.573 ± 0.159
9	$\eta \to \gamma\gamma$	BiType = 23-32	1.719 ± 0.200
10	3K modes	BiType = 33-38	0.621 ± 0.109

Table 9.12: The reweighing factors found in mass region m_{X_s}=2.0-2.4 GeV.

Data Subset	Definition	Modes Used	Frag. Weight Found in m_{X_s}=2.0-2.4 GeV
1	2 bodies without π^0	BiType = 1,3	$0.047^{+0.052}_{-0.047}$
2	2 bodies with 1 π^0	BiType = 2,4	0.323 ± 0.119
3	3 bodies without π^0	BiType = 5,8	0.723 ± 0.105
4	3 bodies with 1 π^0	BiType = 6,9	0.334 ± 0.125
5	4 bodies without π^0	BiType = 11,14	1.115 ± 0.229
6	4 bodies with 1 π^0	BiType = 12,15	1.279 ± 0.268
7	3/4 bodies with 2 π^0s	BiType = 7,10,13,16	0.828 ± 0.533
8	5 bodies with 0-2 π^0s	BiType = 17-22	0.743 ± 0.284
9	$\eta \to \gamma\gamma$	BiType = 23-32	2.470 ± 0.502
10	3K modes	BiType = 33-38	0.744 ± 0.307

1.0536 is motivated by the ratio 0.3310/0.3142, the negative error (-1.000) is motivated by the lack of information from the data (a negative fit yield) indicates a weight of 1 is just as consistent with data as 0.

- Mass Range 1.1-1.5 GeV/c^2 - Group 10 (3K modes). There are 0 events in data in this group. We fix this weight identically to 0, and since there was expected to be ∼1 event from MC (including background), we don't feel this to be a relevant source of error and simply ignore this group.

- Mass Ranges 2.0-2.4 GeV/c^2 and 2.4-2.8 GeV/c^2 - Several Groups. When the error on

Table 9.13: The reweighing factors found in mass region m_{X_s}=2.4-2.8 GeV.

Data Subset	Definition	Modes Used	Frag. Weight Found in m_{X_s}=2.4-2.8 GeV
1	2 bodies without π^0	BiType = 1,3	0.175± 0.134
2	2 bodies with 1 π^0	BiType = 2,4	$0.145\ ^{+0.246}_{-0.145}$
3	3 bodies without π^0	BiType = 5,8	$0.250^{+0.252}_{-0.250}$
4	3 bodies with 1 π^0	BiType = 6,9	$1.000^{+0.465}_{-1.000}$
5	4 bodies without π^0	BiType = 11,14	2.294 ± 0.740
6	4 bodies with 1 π^0	BiType = 12,15	$0.102^{+0.387}_{-0.102}$
7	3/4 bodies with 2 π^0s	BiType = 7,10,13,16	2.064 ± 1.643
8	5 bodies with 0-2 π^0s	BiType = 17-22	$0.294^{+1.270}_{-0.294}$
9	$\eta \to \gamma\gamma$	BiType = 23-32	$1.085^{+1.033}_{-1.085}$
10	3K modes	BiType = 33-38	$0.825^{+1.107}_{-0.825}$

the central value is greater than the central value itself, we assign the negative-side error to be the central value (indicating that the lowest weight this group could have would be 0, the lowest physically reasonable weight). This value is used to calculate the systematic uncertainty.

- Mass Range 2.4-2.8 GeV/c^2 - Group 4 (3-body, 1 π^0). This group has a negative yield in the fits to the data. As above, we fix the weight to 1, and assign the positive-side error to be the error from the fit to data, and the negative-side error to be 1.

Because we have redistributed the cross-feed based on the fragmentation weights, we need to refit the cross-feed parameters for the final fits to data. Indeed, with each iteration of the fragmentation study (as described in Section 9.2.1) we refit the cross-feed to better reflect the data. The final settings for the five cross-feed ranges (we subdivided the region 1.1-2.0 into 1.1-1.5 and 1.5-2.0 GeV/c^2 in the fragmentation study) are given in Table 9.14 and 9.15.

Finally, since the distribution of our signal and cross-feed has changed, the total efficiency for reconstructing events generated in our 38 modes, ϵ_{38}, has changed. The final values for our reconstruction efficiency are reported in Table 9.16.

Table 9.14: The final shape parameters for the cross-feed Nvs function used to model the peaking component in the different m_{X_s} regions after the fragmentation study.

m_{X_s} range	peak	width	tail
0.6 to 1.1	5.2804±0.00006	0.00350±0.00006	-0.160±0.017
1.1 to 1.5	5.2809±0.00011	0.00424 ±0.00015	-0.2449±0.025
1.5 to 2.0	5.28105±0.00012	0.00509 ±0.00020	-0.2893±0.028
2.0 to 2.4	5.28142±0.00030	0.00635±0.00068	-0.385±0.11
2.4 to 2.8	5.28127±0.00061	0.0064±0.0015	-0.438±0.18

9.3 Extracting the Missing Fraction Uncertainty

One of the largest obstacles associated with the strategy adopted for this analysis, summing a subset of the final states of the s quark system, is accurately correcting for the missing fraction of final states. Because these missing final states will be missing no matter how much luminosity one has when doing this analysis, it is the uncertainty on this correction that will be the limiting systematic uncertainty on this analysis strategy in the infinite statistics limit (and assuming everything else is "well-behaved").

That being said, we still recognize that our measurement of the transition rate of $b \to s\gamma$ based on a subset of s quark final states can give us information about the total $b \to s\gamma$ transition rate. To account for our uncertainty on the fraction of missing final states, we will adopt the strategy of being as broad as physically reasonable in the models we use for the hadronization of the s quark, and take our systematic uncertainty accordingly, based on the spread of predicted missing fractions. As mentioned above, our signal MC is generated with a phase-space hadronization model; we have presented two competing hadronization models as well: a showering quarks model and a thermodynamics model (effectively a strictly Poisson breakdown of final state multiplicity).

Our MC is generically generated, meaning above the $K^*(892)$ resonance, we do not take any specific resonances into account. While this is a practical solution to the otherwise currently intractable problem of correctly mixing higher-mass resonances (themselves not all precisely known), as well as non-resonant $b \to s\gamma$ transitions, the effects of these resonances (many of which have non-zero spin) on the observed final state multiplicity are apparent

Table 9.15: The Argus slope and peaking fraction for the cross-feed after the fragmentation study.

mass bin	Argus Slope	Peaking frac.
0.6-0.7	-64.7±6.1	0.020±0.018
0.7-0.8	-95.0±4.7	0.122±0.014
0.8-0.9	-95.5±4.9	0.629±0.007
0.9-1.0	-105.1±3.5	0.411±0.007
1.0-1.1	-103.4±2.6	0.104±0.008
1.1-1.2	-123.6±2.5	0.135±0.009
1.2-1.3	-119.8±2.2	0.180±0.008
1.3-1.4	-114.4±1.9	0.175±0.007
1.4-1.5	-107.6±1.8	0.167±0.007
1.5-1.6	-95.6±2.1	0.208±0.008
1.6-1.7	-93.3±2.0	0.179±0.008
1.7-1.8	-79.0±2.6	0.182±0.010
1.8-1.9	-73.9±2.8	0.147±0.010
1.9-2.0	-58.7±3.8	0.148±0.013
2.0-2.2	-28.7±2.6	0.149±0.009
2.2-2.4	-9.1±3.1	0.086±0.010
2.4-2.6	5.8±5.1	0.102±0.015
2.6-2.8	10.7±4.9	0.101±0.014

in tables 9.10 to 9.13. In these tables, we see we over-produce two body final states w.r.t. data, and the under/over-production of higher multiplicity final states shows a bit of a mass dependence.

Therefore, in order to thoroughly span the space of all reasonable hadronization models, we investigate generators with different probabilities to produce a spin-1 s quark-containing hadron (which itself then may decay into one of our 38 modes; this spin-1 hadron can either be interpreted as a "mock-intermediate-resonance," not otherwise accounted for, or as simply another parameter in one's generator). The two hadronization models in which this becomes relevant, the phase-space hadronization model and the showering quarks model (how to apply this to our adapted thermodynamics model is not immediately obvious), may then be investigated for dependence of missing fraction of final states on spin-1 s quark-hadron probability.

To be as thorough as possible, we also investigate different probabilities for producing a spin-1 u or d quark intermediate resonance in our generator. We investigate these parameter

Table 9.16: The value of ϵ_{38} before and after the fragmentation corrections. The uncertainty on the corrected value reflects the systematic uncertainty based on the uncertainty of the fits to the data.

mass bin	original ϵ_{38} (%)	corrected ϵ_{38} (%)
0.6-0.7	15.04	15.04
0.7-0.8	16.46	16.46
0.8-0.9	17.33	17.33
0.9-1.0	18.30	18.30
1.0-1.1	19.55	19.55
1.1-1.2	11.51	10.43±0.42
1.2-1.3	11.59	10.63±0.32
1.3-1.4	10.65	9.88±0.30
1.4-1.5	9.47	8.93±0.52
1.5-1.6	8.39	7.54±0.47
1.6-1.7	7.20	6.54±0.42
1.7-1.8	5.52	5.03±0.40
1.8-1.9	4.54	4.16±0.42
1.9-2.0	3.26	2.96±0.40
2.0-2.2	4.04	3.20±0.35
2.2-2.4	3.11	2.36±0.43
2.4-2.6	2.29	1.93±0.69
2.6-2.8	2.30	2.06±0.94

settings for both the phase space model and the showering quarks model. One can see the impact of different spin-1 resonance probability settings on our quantity $\epsilon_{incl.}$ in figures 9.2 and 9.3; it should be noted from these figures that, though changing these probabilities produces noticeable effects in both missing fraction and the breakdown of our 38 modes (shown below), the largest differences come from choice of hadronization schemes (phase-space hadronization vs. showering quarks hadronization). Our generic signal MC has the probability for an s quark to form a spin-1 hadron set at 0.60 (the JETSET default), and the u (d) quark to form a spin-1 hadron at 0.4 (the *BABAR* default) with final states generated in a phase-space hadronization model (the missing final state breakdown for the showering quarks model given in Table 9.2 has the same settings for these parameters).

Each of these settings makes different predictions for how we should reweight our generator with respect to the default settings. For mass range $1.1 < m_{X_s} < 1.5$, these would reweight our default settings by the values shown in Figure 9.4 and Figure 9.5 for the different groups given in Table 9.10. The models shown in Figure 9.4 are generated with the phase-space model with different spin-1 hadron formation probabilities, with the reweighting values found in the data from Table 9.10 included for reference (as well as the thermodynamics model). The models shown in Figure 9.5 are generated with the showering quarks model with different spin-1 hadron formation probabilities, with the reweighting values found in data included for reference. In both of these plots, the default settings are identically 1 (the default generator would not be reweighted with respect to the default generator).

In order to simplify the multitude of settings for our models now at our disposal, we focus on the extremes in order to determine a range from which we can evaluate an uncertainty on our missing fraction correction. It should be noted at this point that not all settings are stable; some cause the generator to crash. In general, and perhaps not surprisingly, as one goes to lower mass, the generator is not able to produce higher probabilities for spin-1 hadrons (we find the X_s system needs an invariant mass $>\sim$1.2-1.3 GeV/c^2 in order to ensure stability for spin-1 formation of the s quark at 100% probability, *i.e.*, it always forms a spin-1 hadron, we set the corresponding maximum photon energy cutoff for such a probablility at 2.5 GeV). Therefore we identify one set of "extreme model settings" at

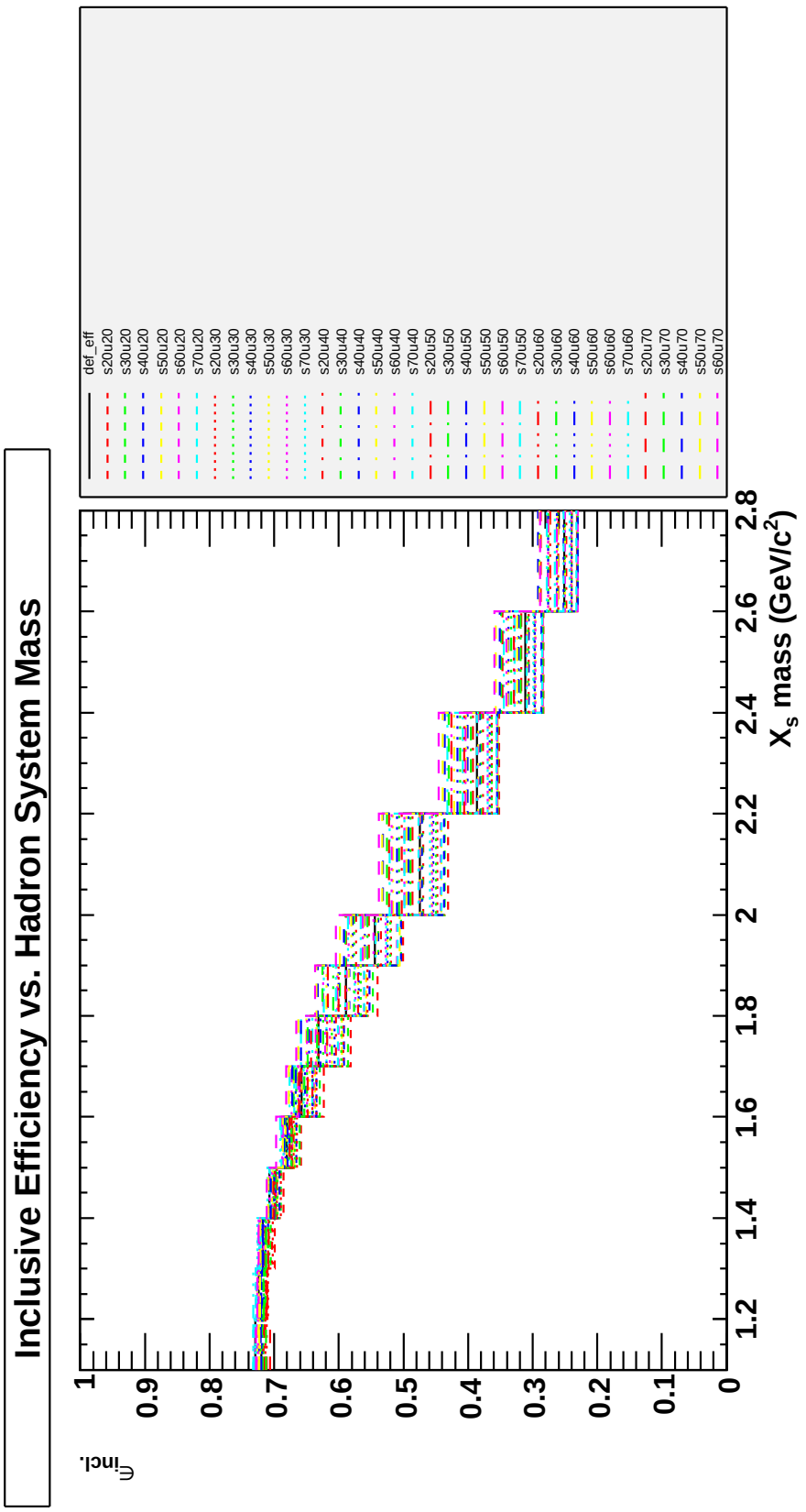


Figure 9.2: The different $\epsilon_{ind.}$ values for each mass bin for each of the different settings for spin-1 hadron formation. Each of these was generated with the phase space hadronization model with the probability of the s quark forming a spin-1 hadron given by the first number in the legend (times 10^{-2}) and the probability of the u quark (or d quark) forming a spin-1 hadron given by the second number in the legend (times 10^{-2}). The default settings are given by a black line, in the middle of the distribution.

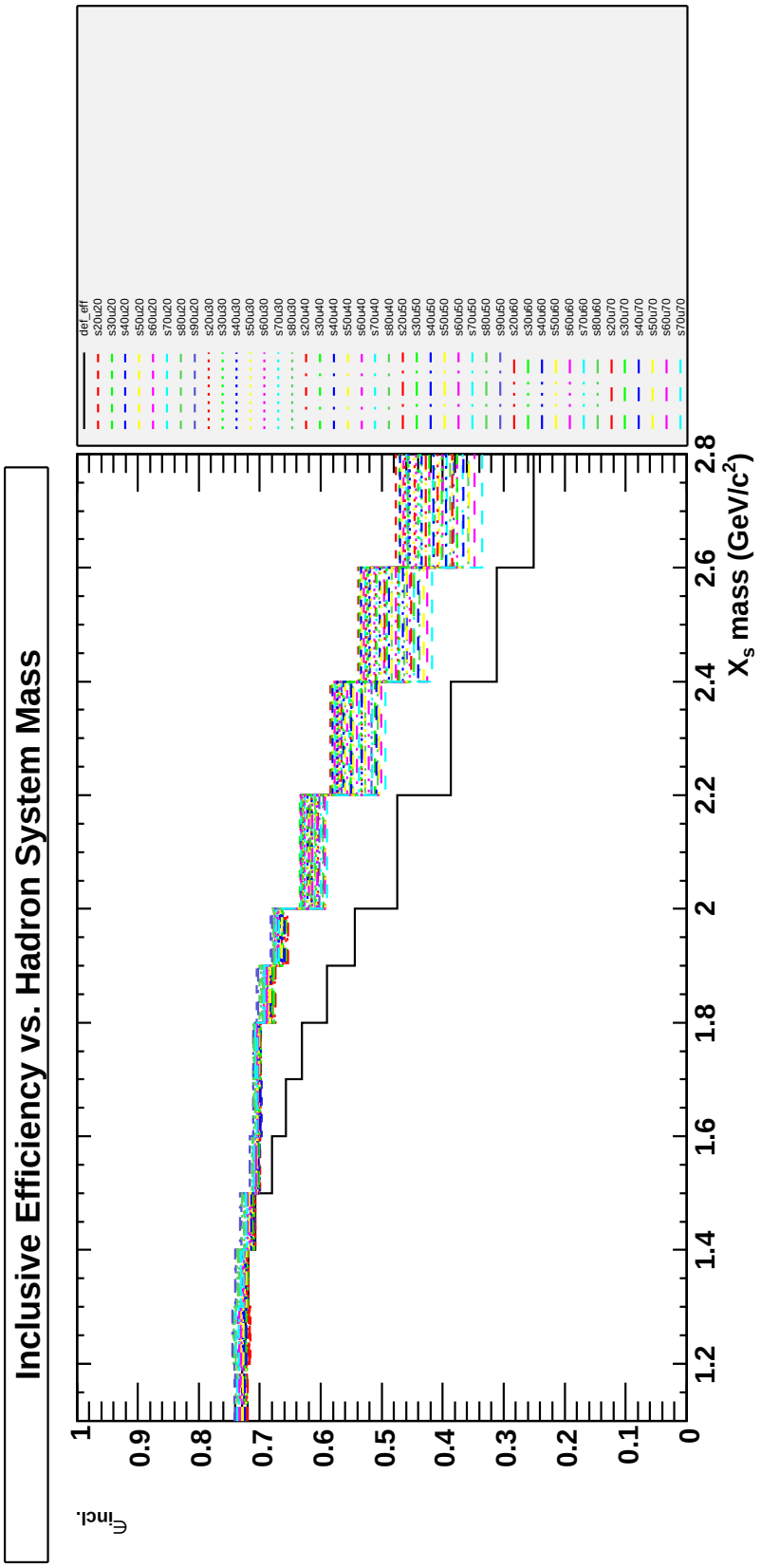


Figure 9.3: The different $\epsilon_{incl.}$ values for each mass bin for each of the different settings for spin-1 hadron formation. Each of these was generated with the showering quark hadronization model (except for the default setting, given by the black line) with the probability of the s quark forming a spin-1 hadron given by the first number in the legend (times 10^{-2}) and the probability of the u quark (or d quark) forming a spin-1 hadron given by the second number in the legend (times 10^{-2}).

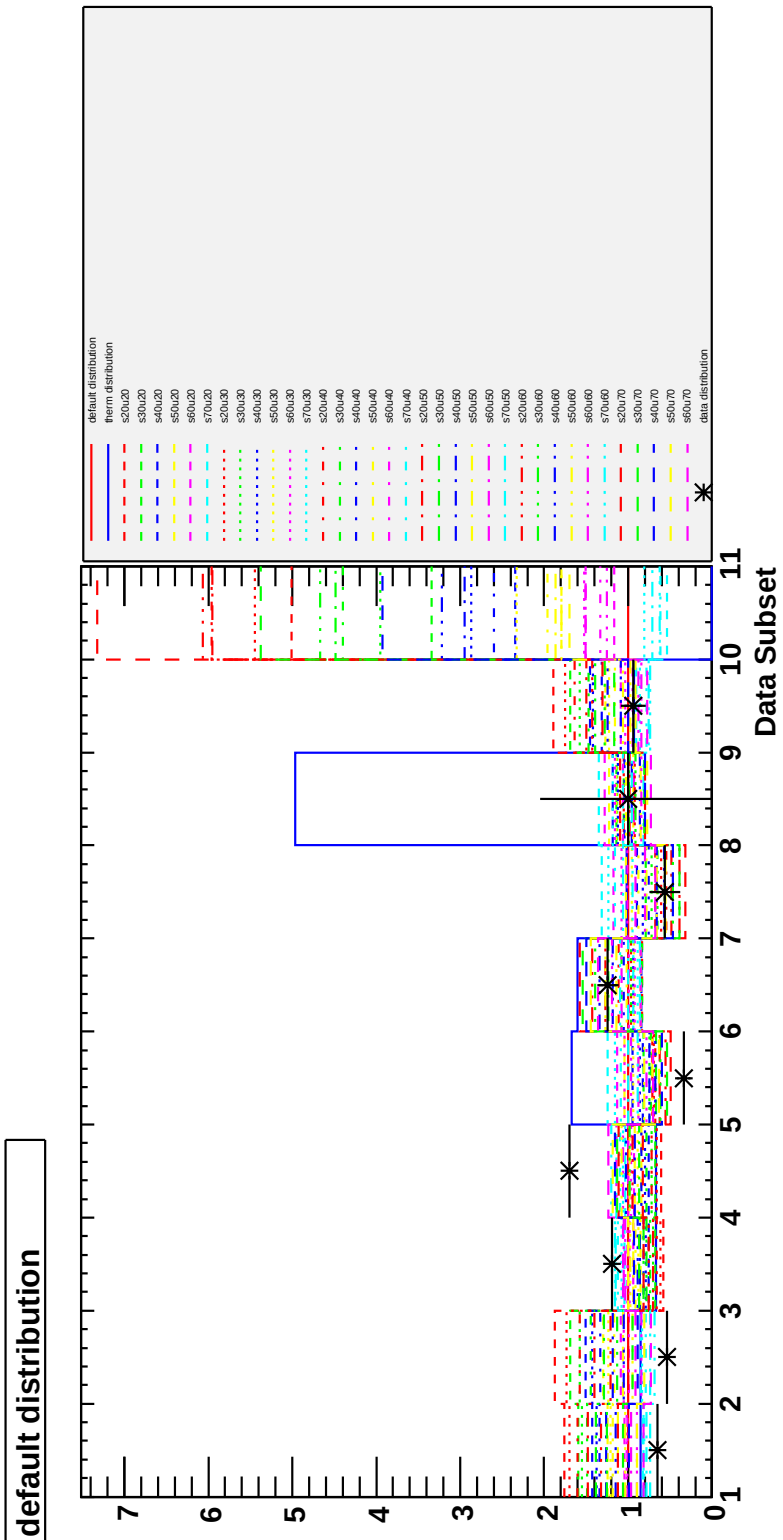


Figure 9.4: The reweighting factors predicted by different spin-1 hadron formation probalities for phase-space hadronization generated models in mass region $1.1 < m_{X_s} < 1.5$. The x-axis corresponds to the data sets given in Table 9.10; the default MC setting is a flat line of value 1 in every bin.

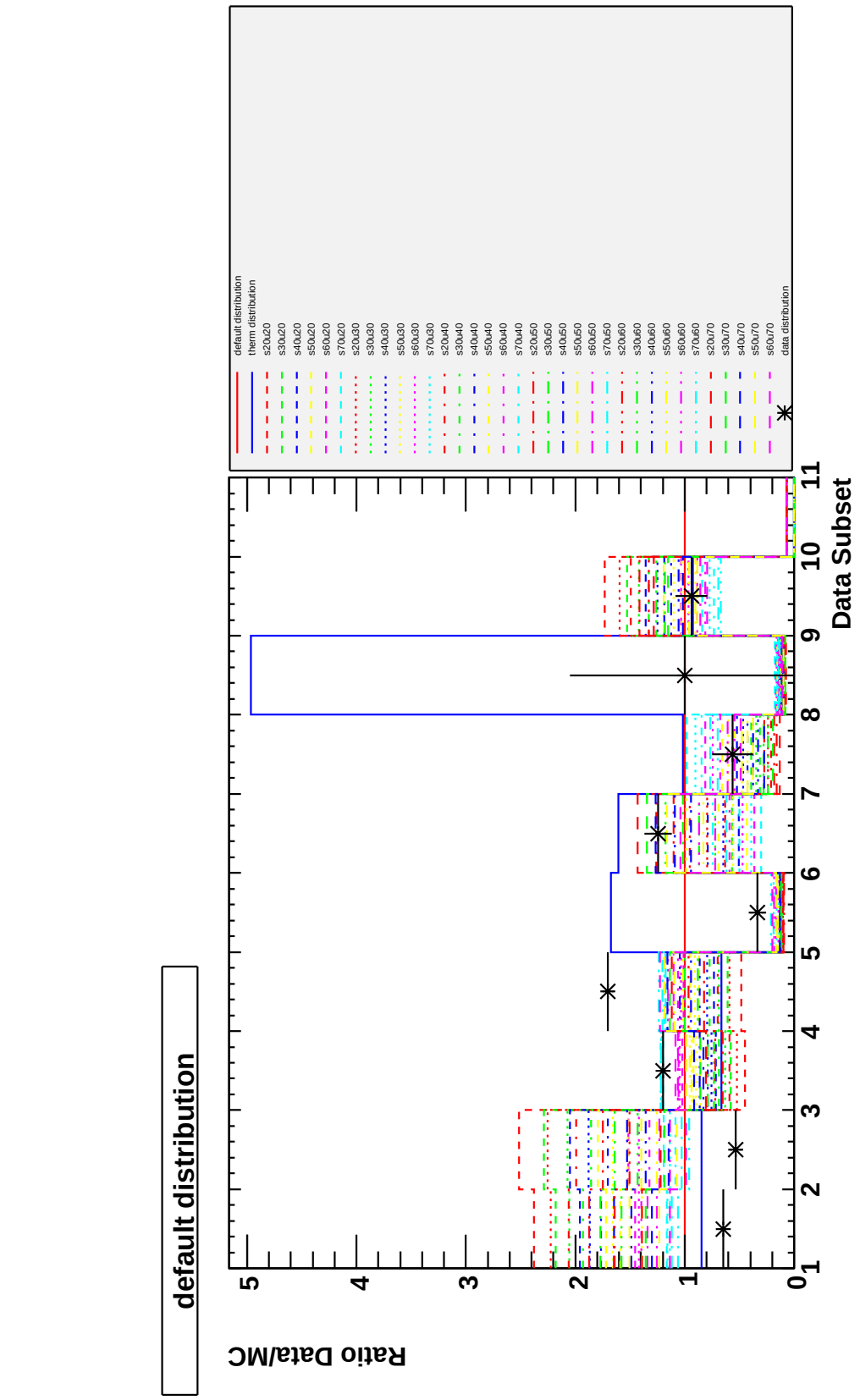


Figure 9.5: The reweighting factors predicted by different spin-1 hadron formation probalities for showering quarks hadronization generated models in mass region $1.1 < m_{X_s} < 1.5$. The x-axis corresponds to the data sets given in Table 9.10; the default MC setting is a flat line of value 1 in every bin.

lower X_s mass (1.1-1.5 GeV/c^2), and one set of "extreme model settings" for the remainder of the higher-mass bins (1.5-2.8 GeV/c^2). We also tried to generate spin-2 s hadron systems, but were unable to find a stable set of settings for our generator. We believe this could be solved if we were to restrict ourselves to a higher-mass region for our s quark system (2.0-2.8 GeV/c^2 for instance), but we see adequate coverage of our generator models when compared to the results found in data without making such a restriction, and therefore there is no motivation to pursue more intricate settings for our generator models.

For the mass range 1.5-2.8 GeV/c^2 we identify 8 total settings for our generator that define the extremities of our considered generated models. These eight settings reflect the most extreme values we can input to our generator and still have stable results; we have four different settings for the phase space hadronization model and four different settings for the showering quarks model (each setting differing in its values of spin-1 s quark or u (d) quark hadron formation probability). For the mass range 1.1-1.5 GeV/c^2, we have to restrict the parameters a bit in order to have stable generators. The extreme settings are reported in Table 9.17.

These different extreme models themselves make predictions about what the fragmentation weight should be, similar to what's been shown in Figures 9.4 and 9.5. We can compare the fragmentation weights found in each of these models to the values found in data to ensure that these models have adequate coverage of the range of physically plausable generators. For each mass region in which we performed a fragmentation study, we compare the predictions of each of our models and their settings to the values found in data in figures 9.6-9.9. We see from these figures that our models have adequate coverage to account for what has been seen in data (aside from a handful of points that generally have a large unceratainty from their fit to data). Another observation is that we cannot claim one type of hadronization model (phase space vs. showering quarks) is more correct than another, and both should be considered when correcting for the missing fraction.

Since we see that in our eight extreme models we have sufficient coverage of possible generators to account for what we see in data, we use these eight models to evaluate the range of missing fractions from each model. We show the eight extreme models (as

Table 9.17: The models we define as "extreme" for the given mass ranges, used for evaluating a reasonable range of fraction of missing final state predictions.

Model	Probility spin-1 hadron formation from s quark	Probability spin-1 hadron formation from u (d) quark
mass range 1.1< m_{X_s} <1.5		
Showering Quarks Models	0.2	0.2
	0.2	0.7
	0.9	0.2
	0.8	0.5
Phase Space Models	0.2	0.2
	0.2	0.7
	0.9	0.2
	0.8	0.5
mass range 1.5< m_{X_s} <2.8		
Showering Quarks Models	0.0	0.0
	0.0	0.7
	1.0	0.0
	1.0	0.5
Phase Space Models	0.0	0.0
	0.0	0.7
	1.0	0.0
	1.0	0.5

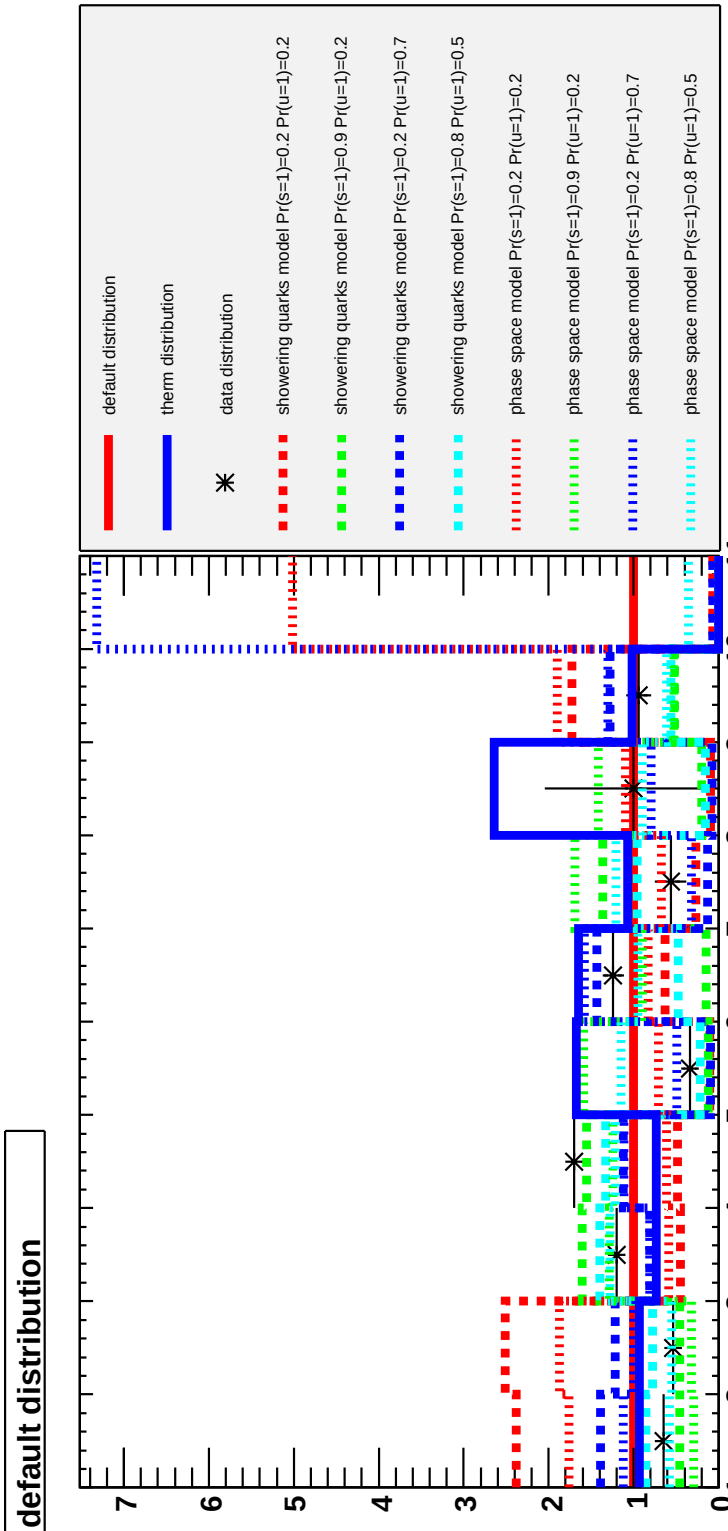


Figure 9.6: The reweighting factors in mass region $1.1 < m_{X_s} < 1.5$ predicted by the eight extreme models with different probability of spin-1 hadron formation (given as Pr(x=1), where x=*s u*), as well as the results from the measurement in the data, given in Table 9.10. We also show the thermodynamic model's predictions, and the default reweight (by definition, 1 in all bins).

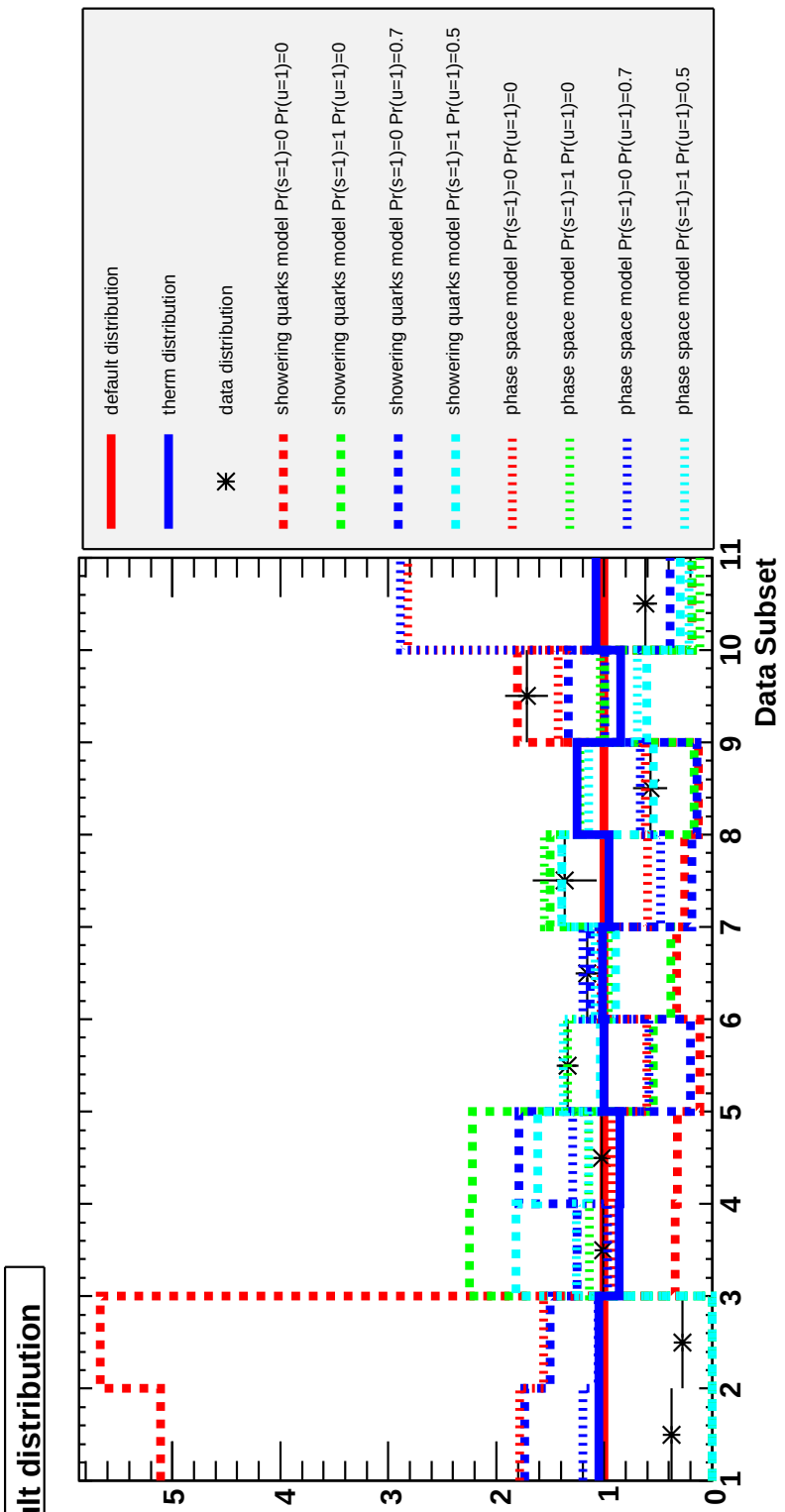


Figure 9.7: The reweighting factors in mass region $1.5 < m_{X_s} < 2.0$ predicted by the eight extreme models with different probability of spin-1 hadron formation (given as Pr(x=1), where x=*s u*), as well as the results from the measurement in the data, given in Table 9.11. We also show the thermodynamic model's predictions, and the default reweight (by definition, 1 in all bins).

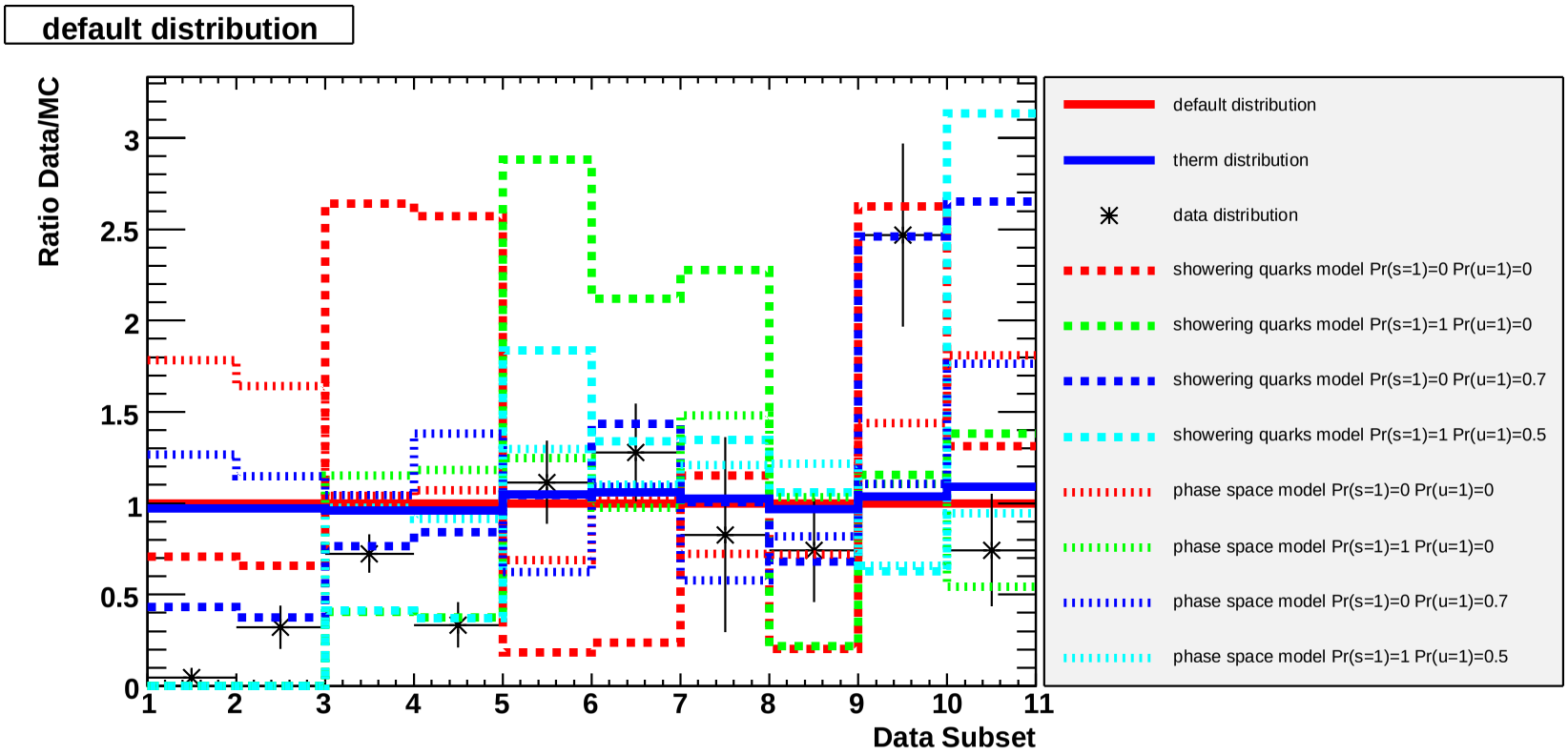


Figure 9.8: The reweighting factors in mass region $2.0 < m_{X_s} < 2.4$ predicted by the eight extreme models with different probability of spin-1 hadron formation (given as Pr(x=1), where x=s u), as well as the results from the measurement in the data, given in Table 9.12. We also show the thermodynamic model's predictions, and the default reweight (by definition, 1 in all bins).

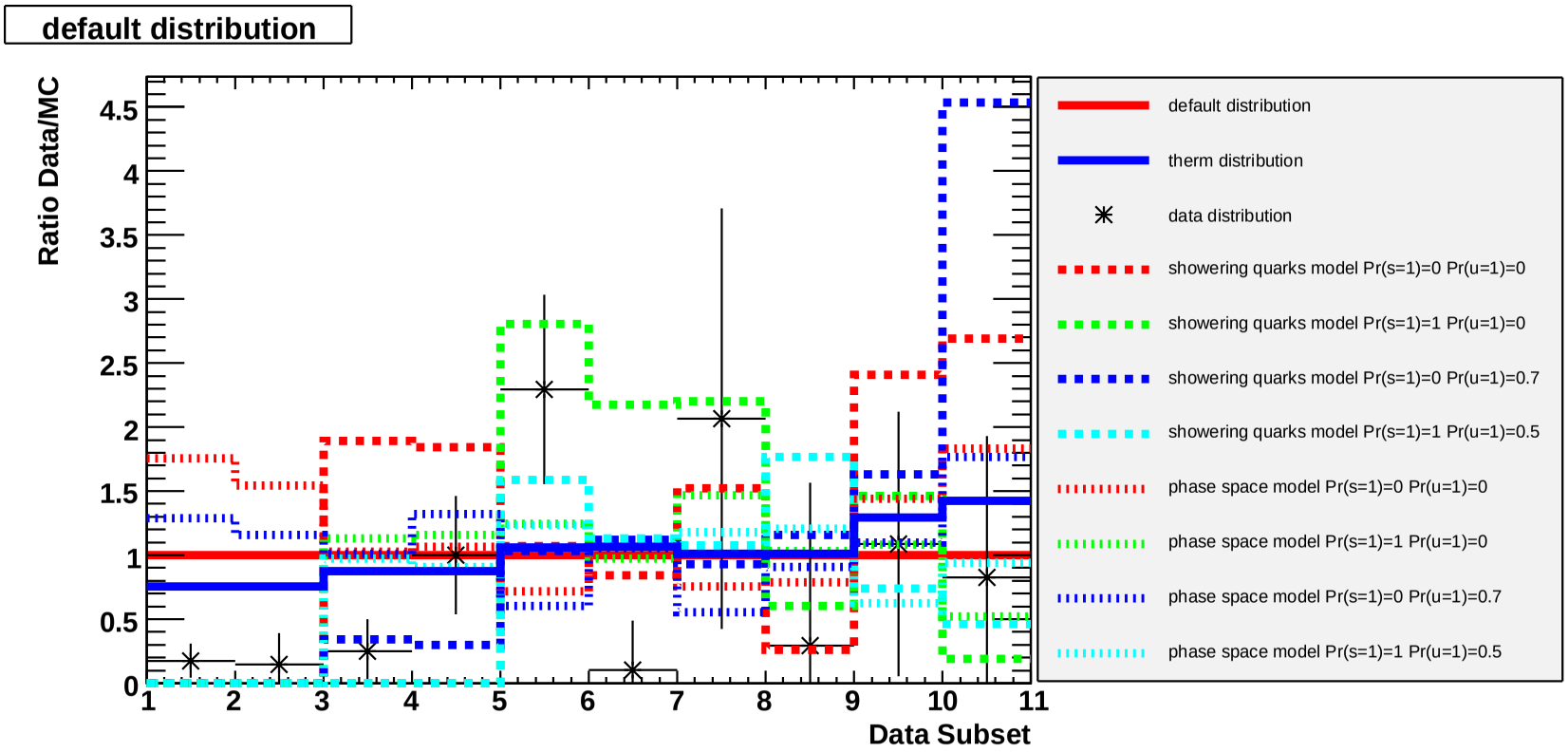


Figure 9.9: The reweighting factors in mass region $2.4 < m_{X_s} < 2.8$ predicted by the eight extreme models with different probability of spin-1 hadron formation (given as Pr(x=1), where x=s u), as well as the results from the measurement in the data, given in Table 9.13. We also show the thermodynamic model's predictions, and the default reweight (by definition, 1 in all bins).

well as the default and thermodynamics model) predictions of the quantity $\epsilon_{incl.}$ (which is $1 - (missing fraction)$) in each mass bin for the range 1.5-2.8 GeV/c^2 in Figure 9.10 (we only show this mass range since slightly different models are used from 1.1-1.5 GeV/c^2, as mentioned above). The different extreme values for $\epsilon_{incl.}$ and the default value for each mass bin is given in Table 9.18. We address the corresponding systematic uncertainty associated with using our default $\epsilon_{incl.}$ in Section 10.

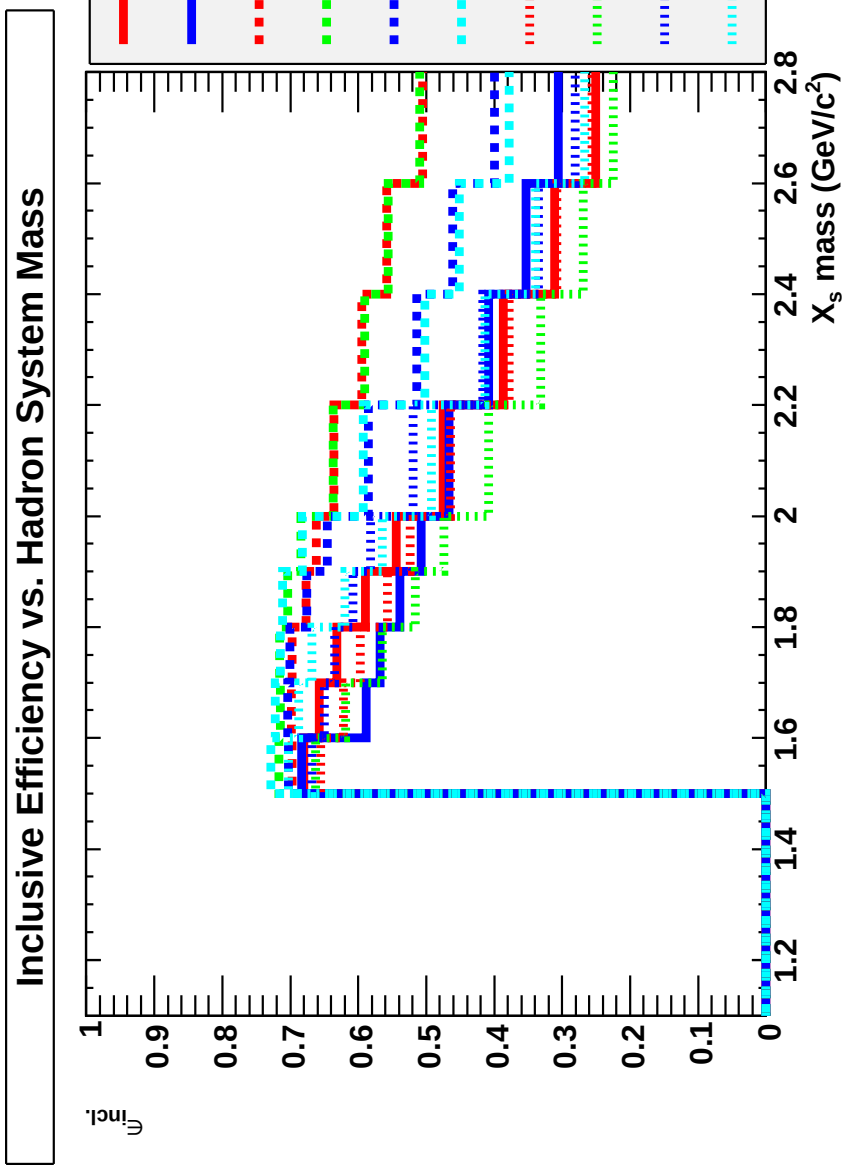


Figure 9.10: The value for $\epsilon_{incl.}$ predicted by our eight extreme models (as well as the thermodynamics model and default model) for each mass bin.

Table 9.18: The most extreme values for $\epsilon_{incl.}$ we were able to generate and the model these extremes corresponds to. The default value for each bin is also given. The notation Pr(x=1) refers to the probability for a hadron being generated in a spin-1 state from the quark x(=s, u (d)).

m_{X_x} (GeV/c^2)	Minimum $\epsilon_{incl.}$ Generated	Corresponding Model	Maximum $\epsilon_{incl.}$ Generated	Corresponding Model	Default $\epsilon_{incl.}$
1.1 to 1.2	0.707	Phase Space Pr(s =1)=0, Pr(u =1)=0.7	0.741	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.729
1.2 to 1.3	0.710	Phase Space Pr(s =1)=0, Pr(u =1)=0	0.743	Showering Quarks Pr(s =1)=1, Pr(u =1)=0.5	0.724
1.3 to 1.4	0.700	Phase Space Pr(s =1)=0, Pr(u =1)=0	0.740	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.718
1.4 to 1.5	0.686	Phase Space Pr(s =1)=0, Pr(u =1)=0	0.730	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.707
1.5 to 1.6	0.655	Phase Space Pr(s =1)=0, Pr(u =1)=0	0.729	Showering Quarks Pr(s =1)=1, Pr(u =1)=0.5	0.680
1.6 to 1.7	0.588	Thermodynamics	0.722	Showering Quarks Pr(s =1)=1, Pr(u =1)=0.5	0.657
1.7 to 1.8	0.565	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.715	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.631
1.8 to 1.9	0.516	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.712	Showering Quarks Pr(s =1)=1, Pr(u =1)=0.5	0.589
1.9 to 2.0	0.474	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.684	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.544
2.0 to 2.2	0.408	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.637	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.475
2.2 to 2.4	0.331	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.595	Showering Quarks Pr(s =1)=0, Pr(u =1)=0	0.386
2.4 to 2.6	0.269	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.558	Showering Quarks Pr(s =1)=0, Pr(u =1)=0	0.311
2.6 to 2.8	0.225	Phase Space Pr(s =1)=1, Pr(u =1)=0	0.509	Showering Quarks Pr(s =1)=1, Pr(u =1)=0	0.251

Chapter 10

Systematic Uncertainty

We identify several sources of systematic uncertainties in our analysis, and will address each of them below.

- $B\overline{B}$ Counting Uncertainty - generally taken as 1.1%, used to obtain a branching fraction.

- Classifier Selection Uncertainty - we base the efficiency of our requirements on the outputs of our classifiers on MC events. We need to account for any discrepancies in these efficiencies by including a systematic uncertainty.

- Fitting Uncertainties - there are uncertainties from fixing the different shape parameters of each component PDF. This also includes the uncertainty on the amount of peaking $B\overline{B}$ events.

- Fragmentation Uncertainties - due to the differences between the breakdown of our 38 final states in our default JETSET generator, and the breakdown of final states found in data, we need to change our reconstruction efficiency, ϵ_{38}, as described in Section 9.2. The amount by which we change this is known to within the uncertainty of the fits to data, and hence we need to incorporate the uncertainty from our fits to data into a systematic uncertainty on our new value of ϵ_{38}.

- Detection Efficiency Uncertainties - these are associated with the disagreement between data and MC efficiencies for the detection of different particles.

- Missing Fraction Uncertainties - we need to assess an uncertainty that reflects our lack

of knowledge of how much of the inclusive $b \to s\gamma$ final states are actually missing in our sum of exclusive analysis.

10.1 $B\overline{B}$ Counting Systematic Uncertainty

The uncertainty on the total sample size of $\Upsilon(4S) \to B\overline{B}$ events has been determined in *BABAR* to be 1.1%; we take this uncertainty for our analysis.

10.2 Classifier Selection Systematic Uncertainty

We base our selection systematic uncertainty on the sideband studies performed in Section 8. For the requirement on the BRC, we take our selection systematic uncertainty from the %-difference in efficiencies between data and MC at our BRC cut location in the SSC sideband. For the requirement on the SSC, we take the systematic uncertainty from the %-difference in efficiencies between data and MC at our SSC cut location in the π^0 veto sideband (which effectively acts as a $B \to X\pi^0$ control sample for our $B \to X\gamma$ signal). We assume these selection requirements to be independent of one another (there is no overlap in the variables used to build the classifiers and the classifiers are effective at rejecting different types of background), and take the total systematic on the selection efficiencies as the sum of these two values in quadrature. The total selection systematic uncertainty in each mass region is given in Table 10.1.

Table 10.1: The selection efficiency systematic uncertainty in each mass region.

X_s mass region	BRC systematic uncertainty (%)	SSC systematic uncertainty (%)	total systematic uncertainty (%)
0.6-1.1	1.0	0.1	1.00
1.1-2.0	0.6	0.3	0.67
2.0-2.4	0.3	1.9	1.92
2.4-2.8	0.3	2.8	2.81

10.3 Systematic Uncertainties in Fitting

We identify many sources of systematic uncertainties associated with our fitting procedure. We assume that the uncertainties associated with the shapes of the individual PDFs are not correlated with one another, but within a given PDF the different parameters may be correlated. We evaluate this for each PDF as described below.

10.3.1 Signal CB Shape

The CB function we use to describe the signal contribution has four parameters with which we associate a systematic uncertainty: the peak location, the width, and the alpha and tail parameters (parameterizing the powerlaw fall off of the distribution). We turn to the K^* region for guidance on how to evaluate the associated systematics, and fit the mass range $0.6 < m_{X_s} < 1.1\,\mathrm{GeV}/c^2$ in the data. Over this region, we are able to have stable fits if we allow all of the CB parameters to float in conjunction with our usual floating of the fractional yield and slope of the combinatoric Argus function. We find from this fit that the α and width parameters are slightly correlated (correlation coefficient of 0.334), these two variables are slightly anti-correlated with the peak position (correlation coefficient of -0.482 with α and -0.379 with the width), while the tail parameter is shown to be mostly uncorrelated with either of these. The optimal values quoted in the next paragraph reflect the fit with these three (anti-)correlated parameters floating along with our usual two floating parameters. The fit to the tail parameter is allowed to float individually with our two floating parameters, and we quote this value in the next paragraph.

For the peak location, the K^* data has the best fit value at 5.28020±0.00007, as compared to our default value of $5.28037\,\mathrm{GeV}/c^2$. We decide that taking the % signal yield change when varying the peak position by $0.2\,\mathrm{MeV}/c^2$ as our systematic is appropriate (the difference between the two quoted values). For the width, the K^* data has a best fit value 0.00297±0.00007, compared to our default setting of 0.002835. We take the %-yield change when varying the width by 0.00014 MeV/c^2 as our systematic. We find the optimal value for α in the K^* data is 1.238±0.129, compared with our default setting of 1.124. We take the % change in signal yield when we change our α parameter by 0.11 as our systematic.

Finally, we allow the tail parameter to float between 0 and 200, and find the optimal value to be 144.9±0.1, compared to our default value of 144.8. We therefore vary this parameter by 0.1 around the value at which it is fixed and take the %-yield change as our systematic. We do find that varying the width and tail parameters by rather large values tends to lead to non-converging fits.

In Table 10.2, we present the %-change in signal yield when each of these parameters are shifted by the above amounts. Since the alpha and width showed some correlation and are anti-correlated with the peak position, we simultaneously shift the width and alpha up and the peak down. We then shift the tail parameter up to the best fit value independently. We allow the same configuration of floating parameters in the 1.1-1.5 GeV/c^2 region and find the best fit values are consistent with the ones in the previous paragraph, so we use the changes to the parameters found in the K^* region for all regions. Since the tail parameter is uncorrelated with the others, we add the respective errors in quadrature.

10.3.2 Cross-feed Shape

In each mass bin, we completely fix the shape of the cross-feed component of the background. We describe the cross-feed by an Argus plus a Nvs, so we identify five parameters with which we need to associate a systematic uncertainty: the peak of the Nvs, the width of the Nvs, the tail of the Nvs, the slope of the Argus, and the fractional contribution of the peaking component to the total cross-feed. Similar to the signal CB section above, we turn to the K^* region for guidance on systematics. First, in this region, we allow all parameters to float in MC, and find the fractional peaking contribution to be correlated with the slope of the Argus (correlation coefficient of 0.793) and the width of the Nvs (correlation coefficient of 0.829). We find the slope of the Argus and the width themselves to also be correlated (correlation coefficient of 0.654). The peak and tail parameters are not correlated with any of these (largest correlation with the other three parameters is 0.098), but are themselves anti-correlated with each other (correlation coefficient of -0.744).

To evaluate the best values for these parameters, we allow them to float in the K^* region in the fits to the data. We find that allowing both the peak and tail parameters to float

Table 10.2: The signal CB systematics associated with fixing each parameter in our final fit. The reported values are the %-change in signal yield when changing each parameter by the uncertainty given in the text to reflect the best fit value found in the K^* region. The Δ rest column reflects shifting the three CB parameters, alpha, width, and peak position, simultaneously.

X_s mass bin	Δtail	Δ rest	Total
0.6 to 0.7	0.21	4.69	4.69
0.7 to 0.8	0.01	0.92	0.92
0.8 to 0.9	-0.04	0.99	0.99
0.9 to 1.0	-0.04	1.02	1.02
1.0 to 1.1	-0.03	1.67	1.67
1.1 to 1.2	-0.02	2.51	2.51
1.2 to 1.3	-0.02	1.25	1.25
1.3 to 1.4	-0.02	1.45	1.45
1.4 to 1.5	-0.03	1.74	1.74
1.5 to 1.6	-0.01	1.42	1.42
1.6 to 1.7	-0.02	1.15	1.15
1.7 to 1.8	-0.03	2.77	2.77
1.8 to 1.9	-0.03	1.17	1.17
1.9 to 2.0	0.01	2.69	2.69
2.0 to 2.2	-0.01	2.91	2.91
2.2 to 2.4	-0.02	2.30	2.30
2.4 to 2.6	0.00	3.10	3.10
2.6 to 2.8	-0.04	12.39	12.39

gives unreasonable values (the fitter focuses on describing a small feature with a Gaussian), so we allow the tail parameter alone to float and find the best fit value of -0.229±0.20, as compared to our default value of -0.16. We therefore will shift our value for the tail parameter down by -43% to evaluate a systematic in all mass regions (we use % shifts for the cross-feed shape parameters since we individually parameterize five separate regions). We then allowed the peak position to float, and find the best fit location at 5.278±0.0011, compared to our default value of 5.2804. We therefore shift this parameter by -0.046% in each bin to evaluate the systematic.

We allow the peaking XF fraction, slope of the Argus, and width of the Nvs to float simultaneously, since these parameters are correlated, but fix the total yield and combinatoric Argus slope. We find the best fit value of the peaking fraction is 20% higher than the default value. We therefore use this shift up in each bin to evaluate the systematic. We find the best fit slope value is shifted to -55 ± 58 from the default of -100.21. We therefore shift the slope by 45% to evaluate a systematic for fixing this shape parameter in each mass bin. Finally, we find the best width to be 0.00426±0.00079 compared to our default value of 0.0035, we therefore evaluate a systematic by shifting the width up 22% in each mass bin.

To evaluate a systematic, we evaluate the change in signal yield when shifting the width, fraction, and slope up at the same time (since these are correlated). Similarly, we evaluate a systematic by shifting the peak and tail parameters simultaneously. These two systematics we add in quadrature to get a total systematic on the cross-feed shape. We give the value of each of these shifts, and the total uncertainty for the XF shape in Table 10.3.

We combine the XF with the signal, and fix the fractional contribution of signal to signal+cross-feed; we need to associate an uncertainty with this. From fits to signal and XF MC with all signal and XF parameters floating, we find that the fractional contribution of signal is in general anti-correlated with the shape parameters of the XF distribution, and directly correlated with many of the signal shape parameters. It is not clear how the shape parameters are correlated between the signal distribution and the XF distribution. We will simplify this situation, and set the fractional contribution of signal to be completely

Table 10.3: The average %-change in signal yield when each parameter is shifted by the values given in the text. The total uncertainty is the sum in quadrature of each of the columns.

X_s mass bin	shift width, fraction, slope	shift tail and peak	Total uncertainty
0.6 to 0.7	-15.31	-4.35	15.92
0.7 to 0.8	2.30	1.86	2.96
0.8 to 0.9	-0.09	1.48	1.48
0.9 to 1.0	0.21	1.41	1.43
1.0 to 1.1	3.60	3.45	4.99
1.1 to 1.2	4.21	3.73	5.62
1.2 to 1.3	3.90	2.80	4.81
1.3 to 1.4	3.33	2.89	4.41
1.4 to 1.5	3.22	3.09	4.46
1.5 to 1.6	3.12	0.52	3.17
1.6 to 1.7	3.43	0.83	3.53
1.7 to 1.8	2.73	1.46	3.10
1.8 to 1.9	2.87	0.25	2.88
1.9 to 2.0	2.15	2.81	3.54
2.0 to 2.2	3.39	1.02	3.54
2.2 to 2.4	2.02	3.83	4.32
2.4 to 2.6	2.71	-2.33	3.57
2.6 to 2.8	-5.58	56.58	56.86

correlated with the signal shape (uncertainties add linearly), completely anti-correlated with the cross-feed shape parameters (uncertainties add in quadrature and two times the cross term is subtracted off), and the signal shape parameters are uncorrelated with the cross-feed parameters (uncertainties add in quadrature).

To obtain the uncertainty on the fractional signal contribution, we allow this value to float, fixing the fractional yield of signal+cross-feed and the Argus slope (so only one parameter is floating in this fit, the fraction of signal to signal+cross-feed). We take the difference between the fitted value and our default as our uncertainty on this parameter, and we then refit the fractional signal yield with this parameter shifted up and down by our uncertainty on this parameter. We quote the average change in the signal yield as our systematic uncertainty. The systematic uncertainty for this parameter, and the overall systematic uncertainty from the signal+cross-feed distribution in each mass bin is given in Table 10.4.

10.3.3 Peaking $B\overline{B}$

In our final fits to data, we fix both the shape of the peaking $B\overline{B}$ (consisting of three parameters similar to the peaking cross-feed) and the absolute yield. In Section 8.3, we found that the number of peaking-$B\overline{B}$ events is consistent with our MC expectations in both the high m_{X_s} sideband and the π^0 sideband, so to evaluate a systematic for this, we fluctuate the number of peaking $B\overline{B}$ by the uncertainty reported in Table 8.5. In the high m_{X_s} sideband, we find that the best fit values to the three Nvs parameters (when fixing the yield of $B\overline{B}$ and slope of the combinatoric Argus) is consistent with our MC determined values, and that the parameters show negligible correlation. Therefore, we evaluate uncertainties for all of the $B\overline{B}$ Nvs distributions by fluctuating the parameters by our statistical uncertainty from fits to MC. We also find (from fits to the high-mass sideband) that there are negligible correlations between these parameters and the number of $B\overline{B}$ events. We therefore take the total systematic uncertainty as the sum in quadrature of the change in signal yield when shifting the four parameters by their uncertainty from MC independently. Since, unlike above fits in which we were able to use the K^* region

Table 10.4: The %-uncertainty in signal yield from fixing the shape parameters in the signal+cross-feed PDF. The ρ given reflects the correlation between the parameter and the signal fraction when obtaining the total uncertainty (the correlation between columns 3 and 4 is assumed to be 0).

X_s mass bin	signal fraction	cross-feed parameters ($\rho = -1$)	signal shape ($\rho = 1$)	Total uncertainty (%)
0.6 to 0.7	2.70	15.92	4.69	14.90
0.7 to 0.8	0.78	2.96	0.92	2.65
0.8 to 0.9	-0.60	1.48	0.99	1.71
0.9 to 1.0	-0.49	1.43	1.02	1.71
1.0 to 1.1	0.32	4.99	1.67	5.06
1.1 to 1.2	-0.98	5.62	2.51	5.73
1.2 to 1.3	-0.43	4.81	1.25	4.67
1.3 to 1.4	-0.14	4.41	1.45	4.55
1.4 to 1.5	-0.23	4.46	1.74	4.66
1.5 to 1.6	-0.62	3.81	1.42	3.74
1.6 to 1.7	0.20	4.24	1.15	4.26
1.7 to 1.8	-0.04	4.03	2.77	4.88
1.8 to 1.9	0.38	3.43	1.17	3.41
1.9 to 2.0	0.04	4.52	2.69	5.25
2.0 to 2.2	-0.65	3.54	2.91	4.54
2.2 to 2.4	-0.11	4.32	2.30	4.85
2.4 to 2.6	-0.34	3.57	3.10	4.71
2.6 to 2.8	12.58	56.86	12.39	49.25

to give us information about the direction of the shift, we have no such information for the $B\overline{B}$ parameters, so in Table 10.5 we quote the average change in signal yield when the parameter is shifted up or down by the appropriate value.

Table 10.5: The %-change in signal yield when each parameter of the $B\overline{B}$ PDF is varied. The total uncertainty is the sum in quadrature of the components

X_s mass bin	peak	width	tail	number of $B\overline{B}$	total uncertainty (%)
0.6 to 0.7	0.00	0.00	0.00	21.34	21.34
0.7 to 0.8	0.00	0.00	0.00	3.05	3.05
0.8 to 0.9	0.00	0.59	0.02	0.13	0.60
0.9 to 1.0	0.00	0.71	0.03	0.19	0.73
1.0 to 1.1	0.01	2.22	0.05	1.17	2.51
1.1 to 1.2	0.00	0.14	0.04	0.91	0.92
1.2 to 1.3	0.01	0.08	0.03	0.36	0.37
1.3 to 1.4	0.00	0.05	0.02	0.29	0.29
1.4 to 1.5	0.00	0.13	0.04	0.55	0.57
1.5 to 1.6	0.01	0.19	0.05	1.45	1.47
1.6 to 1.7	0.01	0.27	0.07	1.25	1.28
1.7 to 1.8	0.01	0.28	0.08	1.43	1.46
1.8 to 1.9	0.08	2.95	0.80	12.78	13.14
1.9 to 2.0	0.02	0.70	0.15	4.08	4.15
2.0 to 2.2	0.20	3.63	0.75	5.50	6.64
2.2 to 2.4	0.52	12.77	2.32	17.73	21.98
2.4 to 2.6	2.12	8.53	1.32	22.12	23.83
2.6 to 2.8	11.19	59.79	7.14	141.36	154.06

The total systematic uncertainty associated with fixing several parameters in our fits are given in Table 10.6. In this, we assume that the $B\overline{B}$ PDF is uncorrelated with the signal+cross-feed PDF, and sum their uncertainties in quadrature for each bin. In our total systematics table, however, we keep these bins separate so we do not lose the information about which mass bins are (un)correlated with each other for the $B\overline{B}$ fitting uncertainty.

10.4 Fragmentation Uncertainties

To evaluate a systematic uncertainty associated with our fragmentation study, we shift the central value of the reweighting factors given in tables 9.10-9.13 individually by the ranges given. These ranges reflect our statistical uncertainties on the reweighting factors from

Table 10.6: The total fitting uncertainty reflecting the sum in quadrature of the components given.

X_s mass bin	signal + XF uncertainty	$B\overline{B}$ uncertainty	total fitting uncertainty (%)
0.6 to 0.7	14.90	21.34	26.02
0.7 to 0.8	2.65	3.05	4.04
0.8 to 0.9	1.71	0.60	1.82
0.9 to 1.0	1.71	0.73	1.86
1.0 to 1.1	5.06	2.51	5.65
1.1 to 1.2	5.73	0.92	5.80
1.2 to 1.3	4.67	0.37	4.69
1.3 to 1.4	4.55	0.29	4.56
1.4 to 1.5	4.66	0.57	4.70
1.5 to 1.6	3.74	1.47	4.01
1.6 to 1.7	4.26	1.28	4.45
1.7 to 1.8	4.88	1.46	5.09
1.8 to 1.9	3.41	13.14	13.58
1.9 to 2.0	5.25	4.15	6.69
2.0 to 2.2	4.54	6.64	8.04
2.2 to 2.4	4.85	21.98	22.51
2.4 to 2.6	4.71	23.83	24.30
2.6 to 2.8	49.25	154.06	161.74

fits to data; since they are statistical in nature we take them to be uncorrelated. The total uncertainty from our fragmentation study is therefore the sum in quadrature of these values.

Since some of the ranges are asymmetric, either because shifting down by the statistical uncertainty would indicate a negative reweight (*i.e.*, generate a negative number of events in that group, physically unreasonable) or the fit did not converge to a positive value, so 0 is just as consistent with the data as the central value, we quote the total uncertainty for each group as the average change in signal efficiency when fluctuating up and down by the range given. The systematic uncertainty for each group in each mass bin is given in Table 10.7.

For the mass bin 1.0-1.1, since we don't know whether this should be modeled exclusively with K^* MC or with our generic signal MC, we take the average ϵ_{38} of the two MC types as our default, and divide the difference by $\sqrt{12}$ to get the uncertainty. The average ϵ_{38} efficiency is 0.1595, we report the uncertainty on this quantity in the fragmentation column of Table 10.8.

10.5 Detection Efficiency Uncertainties

We account for several of the residual differences between data and MC associated with our choices of PID and particle lists (such as the list of K_S) by using the evaluations of these differences based on data control samples. These slight differences result in our needing to both correct our efficiency and assign an associated uncertainty. We evaluated the efficiency correction **before** performing the fragmentation study in Section 9, so that the fragmentation weight obtained would be based on the detection efficiency corrected amount of each mode.

- **Tracking:** For each of the charged tracks in a signal event, we assign an uncertainty of 0.174% as recommended by the *BABAR* **Tracking Efficiency Task Force** [37] to account for differences in tracking efficiency between data and MC.

- K_S**:** For the K_Ss in our final states, we again turn to the **Tracking Efficiency Task Force** [37], who have performed dedicated studies on the K_S detection efficiencies.

Table 10.7: The uncertainty on ϵ_{38} from each group used in our fragmentation study.

mass bin (GeV/c^2)	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)	Group 5 (%)	Group 6 (%)	Group 7 (%)	Group 8 (%)	Group 9 (%)	Group 10 (%)	Total (%)
1.1-1.2	1.61	0.24	0.85	1.76	0.01	0.49	2.94	0.02	0.08	0.00	3.92
1.2-1.3	0.97	0.06	0.65	1.04	0.06	0.96	2.29	0.37	0.01	0.00	2.96
1.3-1.4	0.69	0.01	0.56	0.59	0.11	1.60	1.95	1.25	0.04	0.00	3.01
1.4-1.5	0.54	0.08	0.50	0.32	0.22	1.56	1.88	5.09	0.05	0.00	5.71
1.5-1.6	1.19	0.21	1.32	0.43	0.63	1.92	4.80	2.50	0.01	0.06	6.06
1.6-1.7	1.08	0.32	1.31	0.21	0.76	1.95	4.56	3.44	0.19	0.04	6.33
1.7-1.8	1.33	0.39	1.51	0.08	0.79	2.19	4.94	5.39	0.10	0.07	7.94
1.8-1.9	1.54	0.64	1.68	0.12	0.96	2.42	5.51	7.52	0.13	0.26	9.97
1.9-2.0	2.17	0.89	2.33	0.38	0.73	3.18	6.34	10.82	0.24	0.70	13.40
2.0-2.2	1.41	0.84	3.59	0.83	2.07	2.28	4.71	8.57	0.25	0.65	11.04
2.2-2.4	2.18	1.60	5.88	1.62	2.81	3.58	6.43	15.22	0.66	1.00	18.43
2.4-2.6	6.26	2.45	10.70	4.09	5.15	3.58	21.95	25.28	0.05	2.70	36.65
2.6-2.8	6.80	2.76	15.18	10.75	7.82	3.84	30.05	26.59	0.52	1.73	45.71

We take the weighted average of their efficiency corrections and uncertainty from all runs, and assign this uncertainty to any event with a K_S in the final state. The weighted average correction (Data/MC) is 98.7% with an associated uncertainty of 1.24% above the K^* region, and a correction of 98.2% with an associated uncertainty of 1.02% in the K^* region.

- **π^0s:** We take a momentum-dependent efficiency correction as recommended by the *BABAR* **Neutrals** group [38], which has performed dedicated studies of neutral particle efficiency differences between data and MC. In general, this results in an efficiency correction of about 96.8% for 1 π^0 in the final state, and 93.8% for 2 π^0s in the final state. We also take the recommended 3% systematic uncertainty per π^0.

- **γs:** For the detection efficiency of the high-energy photon, there is no correction necessary, but there is a systematic of 1.8%, again based on the studies of the *BABAR* **Neutrals** group. We take this to be uncorrelated with the π^0 correction and uncertainty above, as these are different aspects of the event topology (in a signal event, this high-energy transition photon is not used to form a π^0). We take a systematic uncertainty for detection efficiency of the $\eta \to \gamma\gamma$ to be 3.6% ($2 \times 1.8\%$), again uncorrelated with the detection of the high-energy photon. We only apply these to events that have an η in the final state.

- **Kaon PID:** For the charged K's, we use the PID KaonKMLoose, which has an associated correction and uncertainty on this correction that we take from the appropriate PID tables, themselves comparisons between MC and data of high purity final states. These corrections and uncertainties have a momentum and angular dependence that we take into account. We take tracks that fall in the same bin in the PID table as being completely correlated and add their uncertainty linearly. We then take the weighted average uncertainty of all of our signal kaons, which we find evaluates to 0.1%. We take this value for all charged kaons (for events with multiple charged kaons, we add their uncertainties linearly).

We evaluate all of these uncertainties after the fragmentation study, and combine events based on the fragmentation weight. We take the weighted average of the sum of these uncertainties (taking correlations stated above into account), and report them in Table 10.8.

10.6 Missing Fraction Uncertainties

For the $K^*(892)$ region ($m_{X_s} < 1.1\,\mathrm{GeV}/c^2$), we take the missing fraction uncertainty as the %-difference between our default missing fraction of final states, and the hypothesis that only K_L modes are missing, or a missing fraction of identically 25%.

For the mass bins above the $K^*(892)$ mass region, we turn to Section 9.3 in which we have established a range of values we could have for the missing fraction based on the most extreme bounds of the generators we could produce. Since our observed fragmentation is consistent with many of the models, generated either with a showering quarks mechanism or a phase-space hadronization, we treat all generator models as equi-probable to be the correct one. We therefore take our central value from our default signal MC, then evaluate an uncertainty by taking the difference between the extremes (given in Table 9.18) and divide by the $\sqrt{12}$ (consistent with the variance of a flat distribution, in this case a flat distribution in so-called *generator-space*). We therefore take (maximum-minimum)/(default$\times\sqrt{12}$) as our %-uncertainty.

Table 10.8: Each of these subcomponent uncertainties are assumed to be uncorrelated and the total reflects their addition in quadrature. All uncertainties are given in %. We have indicated correlated uncertainties between bins with horizontal lines. We motivate these groupings in Section 11.2.

mass bin	$B\overline{B}$ counting	Classifier Selection	Non-$B\overline{B}$ Fitting	$B\overline{B}$ Fitting	Frag.	Detection Efficiency	Missing Fraction	Total
0.6-0.7	1.1	1.0	14.9	21.3	–	2.5	0.6	26.2
0.7-0.8	1.1	1.0	2.7	3.1	–	2.6	0.9	5.1
0.8-0.9	1.1	1.0	1.7	0.6	–	2.6	1.3	3.8
0.9-1.0	1.1	1.0	1.7	0.7	–	2.7	0.0	3.6
1.0-1.1	1.1	1.0	5.1	2.5	13.1	2.7	0.9	14.6
1.1-1.2	1.1	0.7	5.7	0.9	3.9	2.7	1.3	7.7
1.2-1.3	1.1	0.7	4.7	0.4	3.0	2.7	1.3	6.4
1.3-1.4	1.1	0.7	4.6	0.3	3.0	2.7	1.6	6.4
1.4-1.5	1.1	0.7	4.7	0.6	5.7	2.7	1.8	8.2
1.5-1.6	1.1	0.7	3.7	1.5	6.1	2.7	3.1	8.5
1.6-1.7	1.1	0.7	4.3	1.3	6.3	2.7	5.9	10.2
1.7-1.8	1.1	0.7	4.9	1.5	7.9	2.7	6.9	12.1
1.8-1.9	1.1	0.7	3.4	13.1	10.0	2.7	9.6	19.6
1.9-2.0	1.1	0.7	5.3	4.2	13.4	2.7	11.1	18.9
2.0-2.2	1.1	1.9	4.5	6.6	11.0	2.9	13.9	19.8
2.2-2.4	1.1	1.9	4.9	22.0	18.4	2.9	19.7	35.3
2.4-2.6	1.1	2.8	4.7	23.8	36.7	2.8	26.8	51.7
2.6-2.8	1.1	2.8	49.3	154.1	45.7	2.8	32.7	171.3

Chapter 11

Results

With the analysis procedure in place, we then unblinded the data and found the following preliminary results.

11.1 Branching Fractions

The partial branching fractions for each mass bin, after correcting for the fragmentation study described in Section 9, are given in Table 11.1. The plot comparing the mass spectrum found to the previous *BABAR* "sum-of-exclusives" analysis is shown in Figure 11.1, and the same results but binned in photon energy is shown in Figure 11.2.

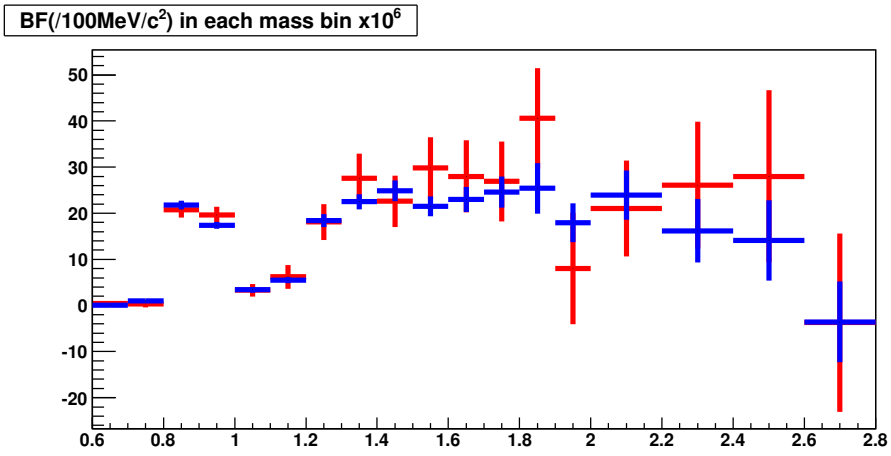


Figure 11.1: The m_{X_s} spectrum, reported in Table 11.1 (blue) as compared to the previous *BABAR* "sum of exclusives" analysis results (red). The errors for each analysis' results include the statistical and systematic uncertainties added in quadrature.

Table 11.1: The branching fractions of $\overline{B} \to X_s\gamma$ in each mass bin. These are referred to as partial branching fractions (PBF).

mass bin	signal yield (events)	ϵ_{38}	$\epsilon_{incl.}$	systematic error (%)	PBF ($\times 10^{-6}$)
0.6 to 0.7	5.9 ±12.2	0.150	0.754	26.2	0.06±0.1±0.01
0.7 to 0.8	114.7 ±24.0	0.165	0.743	5.1	1.00±0.2±0.05
0.8 to 0.9	2627.4 ±50.2	0.173	0.740	3.8	21.8±0.4±0.82
0.9 to 1.0	2249.5 ±53.1	0.183	0.750	3.6	17.4±0.4±0.63
1.0 to 1.1	380.4 ± 36.1	0.160	0.743	14.6	3.4 ± 0.3±0.50
1.1 to 1.2	393.7 ± 37.1	0.104	0.729	7.7	5.5 ± 0.5±0.43
1.2 to 1.3	1330.5 ± 47.1	0.106	0.724	6.4	18.4 ± 0.7±1.18
1.3 to 1.4	1501.0 ± 54.7	0.099	0.718	6.4	22.5 ± 0.8±1.45
1.4 to 1.5	1479.6 ± 58.3	0.089	0.707	8.2	24.9 ± 1.0±2.03
1.5 to 1.6	1039.6 ± 55.7	0.075	0.680	8.5	21.5 ± 1.2±1.82
1.6 to 1.7	929.1 ± 56.7	0.065	0.657	10.2	23.0 ± 1.4±2.34
1.7 to 1.8	736.5 ± 48.6	0.050	0.631	12.1	24.6 ± 1.6±2.97
1.8 to 1.9	586.8 ± 50.8	0.042	0.589	19.6	25.4 ± 2.2±4.98
1.9 to 2.0	272.0 ± 37.4	0.030	0.544	18.9	17.9 ± 2.5±3.39
2.0 to 2.2	684.4 ± 68.2	0.032	0.475	19.8	47.9 ± 4.8±9.49
2.2 to 2.4	277.5 ± 64.6	0.024	0.386	35.3	32.4 ± 7.6±11.40
2.4 to 2.6	159.7 ± 54.4	0.019	0.311	51.7	28.2 ± 9.6±14.60
2.6 to 2.8	-34.4 ± 62.0	0.021	0.251	171.3	-7.1 ± 12.7±12.10
				total	328.7 ±18.8±48.2

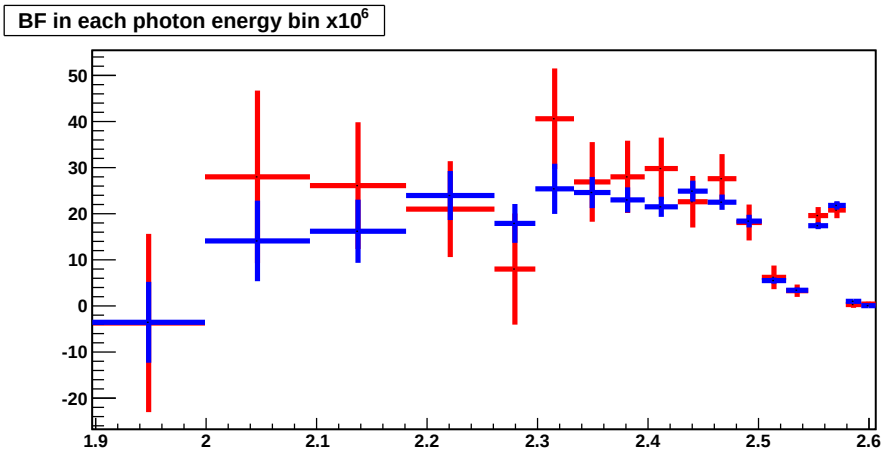


Figure 11.2: The E_γ spectrum, reported in Table 11.1 (blue) as compared to the previous *BABAR* "sum of exclusives" analysis results (red). The values shown correspond to BF/100 MeV/c^2 when binned in hadron mass (as per Figure 11.1).

The plots that show the fit to each mass bin, with the χ^2 value included (there are 48 degrees of freedom in each plot) can be seen in figures 11.3-11.7.

The total systematic uncertainties, described in Section 10, broken down by different sources are reported in Table 11.2

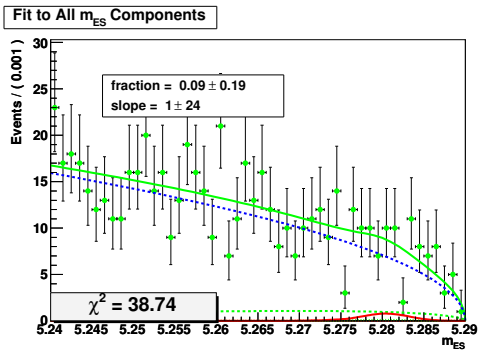
11.1.1 Comparison with Previous Analysis

We have reprinted the results from the previous *BABAR* sum-of-exclusives measurement [1] in Table 11.3.

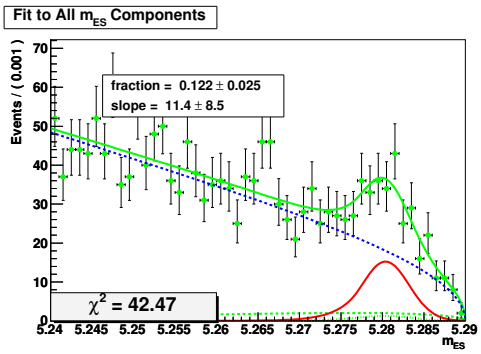
It is worth noting that the total statistical uncertainty on the previous analysis' measurement of the branching fraction is consistent with the total statistical uncertainty we have found on the branching fraction with five times more data. It is also important to note that the total statistical uncertainty from the previous analysis is smaller than the uncertainty found in four of their individual bins. The reason for their reduction in the statistical uncertainty comes from the choice of how to combine the bins to extract a branching fraction. In the previous analysis, they chose to fit the signal yield in the m_{ES} distribution for all the mass bins combined (perform one fit to m_{ES} for $0.6 < m_{ES} < 2.8\,\mathrm{GeV}/c^2$), without accounting for differences in efficiency between these bins, and extracting the statistical uncertainty from this fit. For our analysis, we quote a total branching fraction based on the linear sum of the 18 bins, and a total statistical error equal to the sum in quadrature of the statistical error on each bin.

We prefer our procedure for extracting a statistical uncertainty (and deriving a central value), as the alternative approach used by the previous analysis is not as easily defensible. Fitting all bins at once and then correcting by the statistics-weighted average efficiency (the weighted average ϵ_{38} and $\epsilon_{incl.}$) would give a reasonably consistent central value, but would under-estimate the total statistical uncertainty.

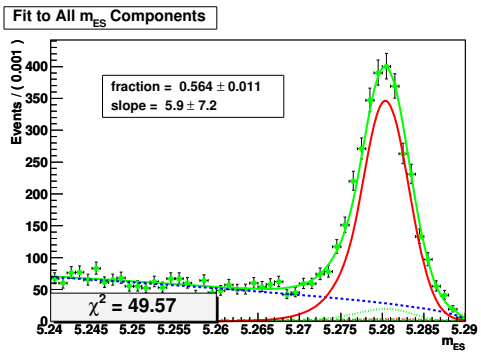
To make this point as pedagogical as possible, we reproduce the signal yield from the previous analysis in each bin in Table 11.4. In the final line of this Table, we see that the yield from fitting every mass bin at once is 1513±85.1 (statistical error is 5.6% of the central value). The linear sum of the 18 mass bins, with the statistical uncertainties added



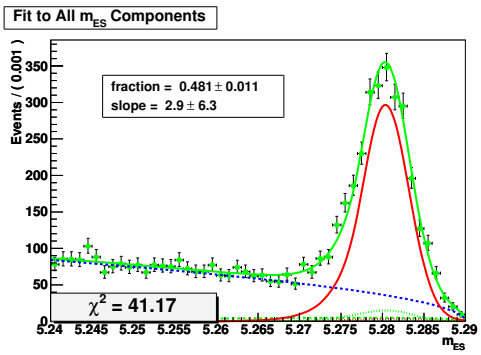
(a) bin $0.6 < m_{X_s} < 0.7\,\mathrm{GeV}/c^2$



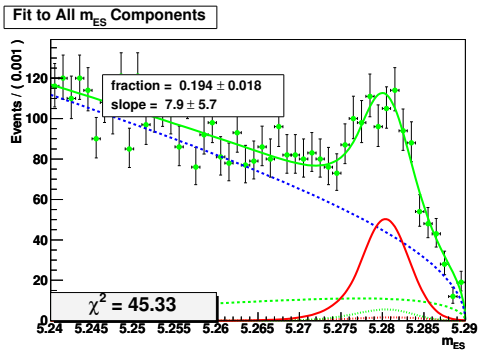
(b) bin $0.7 < m_{X_s} < 0.8\,\mathrm{GeV}/c^2$



(c) bin $0.8 < m_{X_s} < 0.9\,\mathrm{GeV}/c^2$

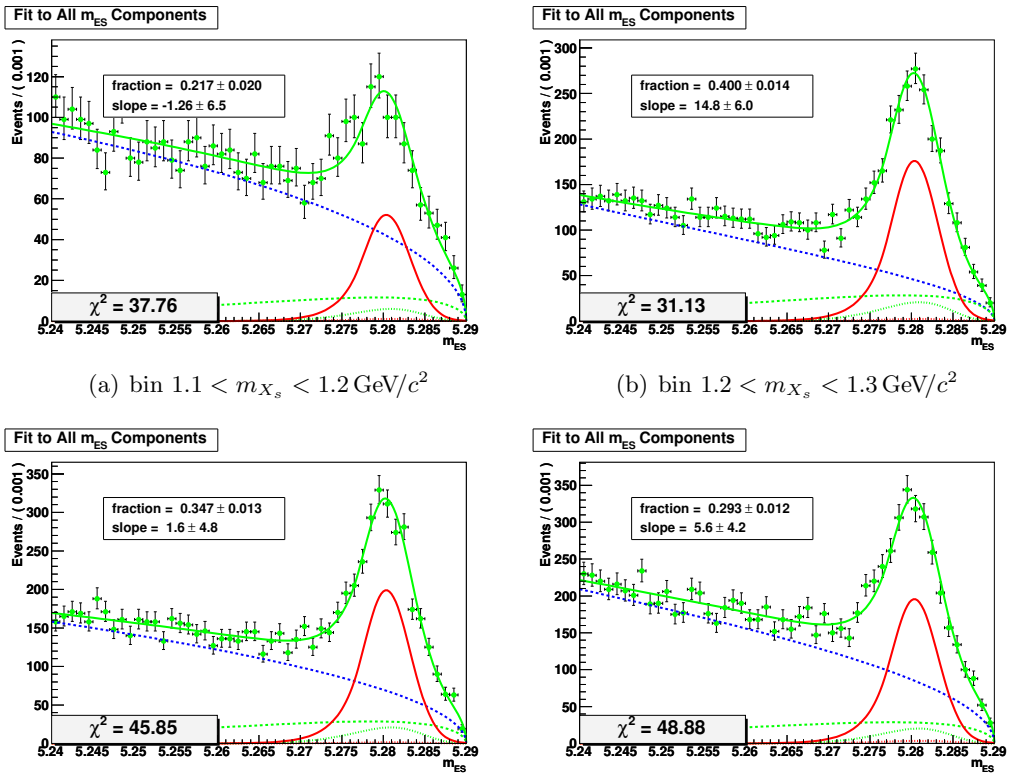


(d) bin $0.9 < m_{X_s} < 1.0\,\mathrm{GeV}/c^2$



(e) bin $1.0 < m_{X_s} < 1.1\,\mathrm{GeV}/c^2$

Figure 11.3: The fits to the mass bins in $0.6 < m_{X_s} < 1.1\,\mathrm{GeV}/c^2$.



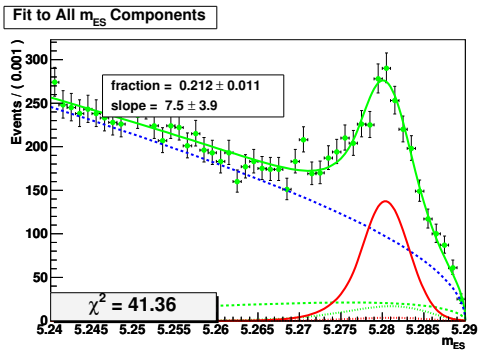
(a) bin $1.1 < m_{X_s} < 1.2\,\mathrm{GeV}/c^2$

(b) bin $1.2 < m_{X_s} < 1.3\,\mathrm{GeV}/c^2$

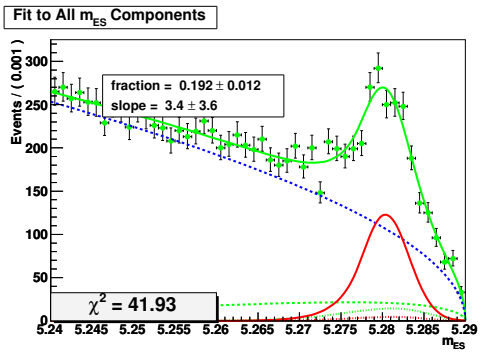
(c) bin $1.3 < m_{X_s} < 1.4\,\mathrm{GeV}/c^2$

(d) bin $1.4 < m_{X_s} < 1.5\,\mathrm{GeV}/c^2$

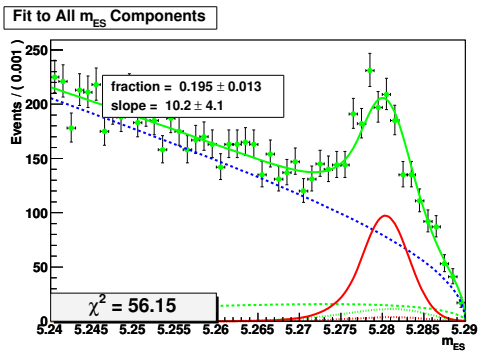
Figure 11.4: The fits to the mass bins in $1.1 < m_{X_s} < 1.5\,\mathrm{GeV}/c^2$.



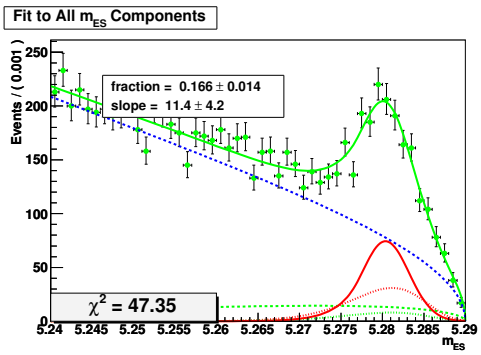
(a) bin $1.5 < m_{X_s} < 1.6\,\mathrm{GeV}/c^2$



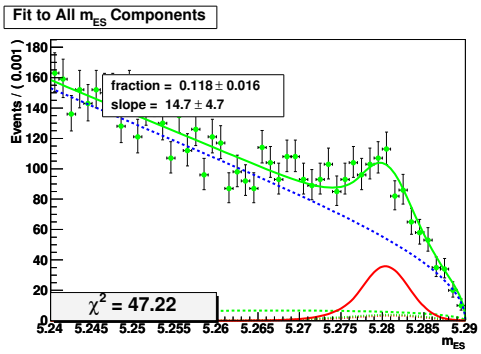
(b) bin $1.6 < m_{X_s} < 1.7\,\mathrm{GeV}/c^2$



(c) bin $1.7 < m_{X_s} < 1.8\,\mathrm{GeV}/c^2$

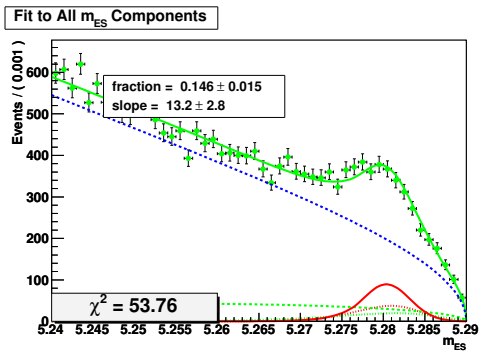


(d) bin $1.8 < m_{X_s} < 1.9\,\mathrm{GeV}/c^2$

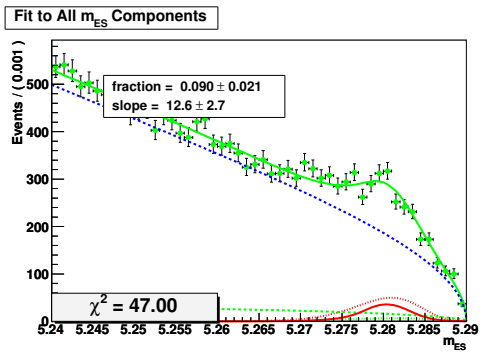


(e) bin $1.9 < m_{X_s} < 2.0 gevcc$

Figure 11.5: The fits to the mass bins in $1.5 < m_{X_s} < 2.0\,\mathrm{GeV}/c^2$.

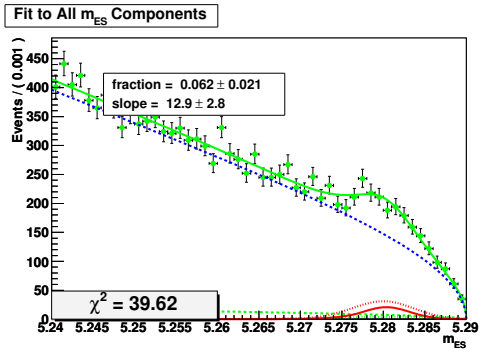


(a) bin $2.0 < m_{X_s} < 2.2\,\mathrm{GeV}/c^2$

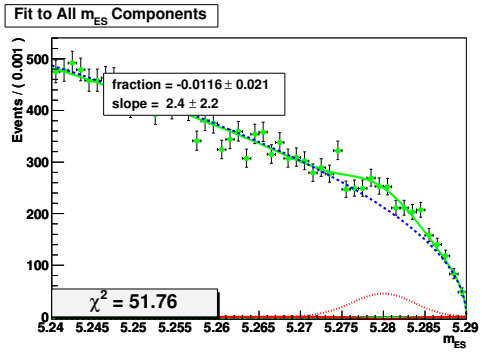


(b) bin $2.2 < m_{X_s} < 2.4\,\mathrm{GeV}/c^2$

Figure 11.6: The fits to the mass bins in $2.0 < m_{X_s} < 2.4\,\mathrm{GeV}/c^2$.



(a) bin $2.4 < m_{X_s} < 2.6\,\mathrm{GeV}/c^2$



(b) bin $2.6 < m_{X_s} < 2.8\,\mathrm{GeV}/c^2$

Figure 11.7: The fits to the mass bins in $2.4 < m_{X_s} < 2.8\,\mathrm{GeV}/c^2$.

Table 11.2: The total systematic errors, each of these subcomponent errors are assumed to be uncorrelated and the total reflects their addition in quadrature. These reflect the error on the central value reported in Table 11.1. We have indicated correlated errors between bins with horizontal lines. We motivate these groupings in Section 11.2.

Mass Bin	$B\overline{B}$ counting	Classifier Selection	Non-$B\overline{B}$ Fitting	$B\overline{B}$ Fitting	Fragmentation	Detection Efficiency	Missing Fraction	Total
0.6-0.7	6.08×10^{-10}	5.53×10^{-10}	8.24×10^{-9}	1.18×10^{-8}	0.00	1.39×10^{-9}	3.10×10^{-10}	1.45×10^{-8}
0.7-0.8	1.10×10^{-8}	9.96×10^{-9}	2.64×10^{-8}	3.03×10^{-8}	0.00	2.56×10^{-8}	9.36×10^{-9}	5.08×10^{-8}
0.8-0.9	2.39×10^{-7}	2.18×10^{-7}	3.73×10^{-7}	1.31×10^{-7}	0.00	5.73×10^{-7}	2.89×10^{-7}	8.20×10^{-7}
0.9-1.0	1.91×10^{-7}	1.74×10^{-7}	2.97×10^{-7}	1.28×10^{-7}	0.00	4.68×10^{-7}	0.00	6.25×10^{-7}
1.0-1.1	3.75×10^{-8}	3.41×10^{-8}	1.73×10^{-7}	8.55×10^{-8}	4.45×10^{-7}	9.28×10^{-8}	3.14×10^{-8}	4.97×10^{-7}
1.1-1.2	6.05×10^{-8}	3.85×10^{-8}	3.15×10^{-7}	5.06×10^{-8}	2.15×10^{-7}	1.50×10^{-7}	7.15×10^{-8}	4.25×10^{-7}
1.2-1.3	2.02×10^{-7}	1.29×10^{-7}	8.58×10^{-7}	6.78×10^{-8}	5.44×10^{-7}	4.96×10^{-7}	2.39×10^{-7}	1.18×10^{-6}
1.3-1.4	2.47×10^{-7}	1.57×10^{-7}	1.02×10^{-6}	6.54×10^{-8}	6.77×10^{-7}	6.09×10^{-7}	3.59×10^{-7}	1.45×10^{-6}
1.4-1.5	2.74×10^{-7}	1.74×10^{-7}	1.16×10^{-6}	1.41×10^{-7}	1.42×10^{-6}	6.67×10^{-7}	4.48×10^{-7}	2.03×10^{-6}
1.5-1.6	2.37×10^{-7}	1.51×10^{-7}	8.04×10^{-7}	3.16×10^{-7}	1.30×10^{-6}	5.86×10^{-7}	6.67×10^{-7}	1.82×10^{-6}
1.6-1.7	2.52×10^{-7}	1.61×10^{-7}	9.77×10^{-7}	2.94×10^{-7}	1.45×10^{-6}	6.27×10^{-7}	1.35×10^{-6}	2.34×10^{-6}
1.7-1.8	2.71×10^{-7}	1.72×10^{-7}	1.20×10^{-6}	3.60×10^{-7}	1.95×10^{-6}	6.65×10^{-7}	1.70×10^{-6}	2.97×10^{-6}
1.8-1.9	2.79×10^{-7}	1.78×10^{-7}	8.65×10^{-7}	3.34×10^{-6}	2.53×10^{-6}	6.88×10^{-7}	2.44×10^{-6}	4.98×10^{-6}
1.9-2.0	1.97×10^{-7}	1.26×10^{-7}	9.42×10^{-7}	7.44×10^{-7}	2.40×10^{-6}	4.81×10^{-7}	1.99×10^{-6}	3.39×10^{-6}
2.0-2.2	5.27×10^{-7}	9.10×10^{-7}	2.17×10^{-6}	3.18×10^{-6}	5.29×10^{-6}	1.39×10^{-6}	6.66×10^{-6}	9.49×10^{-6}
2.2-2.4	3.56×10^{-7}	6.15×10^{-7}	1.57×10^{-6}	7.12×10^{-6}	5.97×10^{-6}	9.43×10^{-7}	6.38×10^{-6}	1.14×10^{-5}
2.4-2.6	3.10×10^{-7}	7.89×10^{-7}	1.33×10^{-6}	6.72×10^{-6}	1.03×10^{-5}	7.76×10^{-7}	7.56×10^{-6}	1.46×10^{-5}
2.6-2.8	7.75×10^{-8}	1.97×10^{-7}	3.47×10^{-6}	1.09×10^{-5}	3.22×10^{-6}	1.94×10^{-7}	2.30×10^{-6}	1.21×10^{-5}

Table 11.3: The previous analysis' results for the PBF in each mass bin. The uncertainties are statistical and systematic. The total branching fraction is given in the final line.

$M(X_s)$ (GeV/c^2)	$\mathcal{B}(M(X_s))/100\,\mathrm{MeV}/c^2(10^{-6})$			
0.6 - 0.7	0.4	$\pm$	0.5	$^{+0.1}_{-0.1}$
0.7 - 0.8	0.3	$\pm$	0.7	$^{+0.1}_{-0.1}$
0.8 - 0.9	20.8	$\pm$	1.2	$^{+1.3}_{-1.3}$
0.9 - 1.0	19.6	$\pm$	1.4	$^{+1.2}_{-1.2}$
1.0 - 1.1	3.3	$\pm$	1.2	$^{+0.6}_{-0.6}$
1.1 - 1.2	6.2	$\pm$	2.4	$^{+0.9}_{-0.9}$
1.2 - 1.3	18.1	$\pm$	3.4	$^{+1.9}_{-1.9}$
1.3 - 1.4	27.6	$\pm$	4.6	$^{+2.8}_{-2.7}$
1.4 - 1.5	22.6	$\pm$	5.0	$^{+2.5}_{-2.5}$
1.5 - 1.6	29.8	$\pm$	6.0	$^{+3.1}_{-3.0}$
1.6 - 1.7	28.0	$\pm$	7.2	$^{+3.3}_{-3.1}$
1.7 - 1.8	26.9	$\pm$	8.1	$^{+3.4}_{-3.0}$
1.8 - 1.9	40.6	$\pm$	9.7	$^{+6.1}_{-5.0}$
1.9 - 2.0	8.0	$\pm$	11.7	$^{+3.1}_{-2.9}$
2.0 - 2.2	21.0	$\pm$	9.6	$^{+5.9}_{-4.0}$
2.2 - 2.4	26.1	$\pm$	12.0	$^{+10.5}_{-6.7}$
2.4 - 2.6	28.0	$\pm$	16.0	$^{+16.2}_{-9.7}$
2.6 - 2.8	-3.7	$\pm$	18.8	$^{+4.4}_{-4.5}$
	$\mathcal{B}(10^{-6})$			
0.6 - 2.8	$327.0 \pm 18.0^{+55.0}_{-40.0}$			

Table 11.4: The raw yield, before efficiency correction, found in each mass bin by the previous analysis with the statistical uncertainties from their fits shown. The final row reflects the raw yield from a fit to all mass bins at once.

$M(X_s)$ (GeV/c^2)	Data Signal yield (events)	
0.6-0.7	6.5	± 7.7
0.7-0.8	5.6	± 14.1
0.8-0.9	416.2	± 23.2
0.9-1.0	355.6	± 24.9
1.0-1.1	51.3	± 19.0
1.1-1.2	33.2	± 12.9
1.2-1.3	83.2	± 15.7
1.3-1.4	101.5	± 16.8
1.4-1.5	72.0	± 15.8
1.5-1.6	82.4	± 16.5
1.6-1.7	66.1	± 16.9
1.7-1.8	54.6	± 16.5
1.8-1.9	76.6	± 18.2
1.9-2.0	13.5	± 19.5
2.0-2.2	47.5	± 21.8
2.2-2.4	52.1	± 24.0
2.4-2.6	44.7	± 25.6
2.6-2.8	-6.2	± 31.9
0.6-2.8	1513.0	± 85.1

in quadrature, before correcting for efficiency differences, is 1556.4±83.6 (statistical error is 5.4% of the central value), consistent with the result of the fit to the entire mass range at once. The statistical error on the previous analysis' determination of the branching fraction is 5.5% of the central value.

Unfortunately, we do not believe the data used by the previous analysis supports such a small statistical error. There is almost an order of magnitude difference in total signal efficiency between the mass bins which has not been taken into account by, effectively, combining the bins together before efficiency correction. While a weighted average of the signal efficiency would give the correct central value, it would underestimate the statistical error. This is clear when one combines the PBFs in the mass bins, given in Table 11.3, and their respective statistical uncertainties, after efficiency corrections. Combining the mass bins after efficiency correction gives a statistical error that is 10.1% of the central value, or $\mathcal{B}(B \to X_s\gamma) = (395 \pm 62) \times 10^{-6}$. Because of the improved analysis procedure and the increase in data (together giving us roughly 10× the signal statistics), we have improved the statistical uncertainty to support, without having to fit all mass bins at once, what the previous analysis already reported.

11.2 Correlation Coefficients

We have identified the correlated errors in Table 10.8 with horizontal lines in different columns (the columns themselves are uncorrelated from one another). We motivate these groupings here:

- $B\overline{B}$ **Fitting:** We use independent functions to parameterize the $B\overline{B}$ distributions in the three sections indicated. We are confident of our analysis of the amount of peaking $B\overline{B}$ in the data to within three different, independently verified, uncertainties for each of these sections. We therefore assume that the errors are independent between groupings of rows, but completely correlated between the mass bins within a grouping of rows.

- **Fragmentation:** We performed four separate fragmentation studies and consider

Table 11.5: The variance matrix for two mass bins, given 7 different fully correlated systematic uncertainties.

σ_A^2	$\sigma_{A,1}\sigma_{B,1} + \sigma_{A,2}\sigma_{B,2} + \sigma_{A,3}\sigma_{B,3} + \sigma_{A,4}\sigma_{B,4}$ $+\sigma_{A,5}\sigma_{B,5} + \sigma_{A,6}\sigma_{B,6} + \sigma_{A,7}\sigma_{B,7}$
$\sigma_{A,1}\sigma_{B,1} + \sigma_{A,2}\sigma_{B,2} + \sigma_{A,3}\sigma_{B,3} + \sigma_{A,4}\sigma_{B,4}$ $+\sigma_{A,5}\sigma_{B,5} + \sigma_{A,6}\sigma_{B,6} + \sigma_{A,7}\sigma_{B,7}$	σ_B^2

their results, as well as their errors, to be independent of one another. However, the errors for a given study are completely correlated between the mass bins that were used for any given study, indicated by the groupings of rows. The bin 1.0-1.1 has been evaluated seperately from the fragmentation study as described in the text.

- **Missing Fraction:** We have used two separate methods to identify the uncertainty on the missing fraction, one for the K^* region and one for the region in which we used generically generated MC ($m_{X_s} > 1.1\,\mathrm{GeV}/c^2$); we therefore assume that the errors identified by these procedures are uncorrelated with one another.

We determine the correlation coeficients between each mass bin. To do this, we assume each component of the systematic error given in Table 10.8 is completely correlated between bins according to the groupings indicated by the horizontal lines, but the columns are uncorrelated with each other. Labeling these systematic uncertainties as 1 through 7, the total uncertainty for an arbitrary mass bin "A" would be:

$$\sigma_A^2 = \sigma_{A,stat}^2 + \sigma_{A,1}^2 + \sigma_{A,2}^2 + \sigma_{A,3}^2 + \sigma_{A,4}^2 + \sigma_{A,5}^2 + \sigma_{A,6}^2 + \sigma_{A,7}^2, \tag{11.1}$$

The variance matrix between any two mass bins, "A" and "B" for example, would then be the sum of the variance matrices of each of the uncertainty components. In tabular form, this symmetric variance matrix would look like that shown in Table 11.5.

The correlation coefficient between the two bins would then simply be the off-diagonal term, divided by the product of the total errors, $\sigma_A\sigma_B$. We report this correlation matrix, $C_{i,j}$, which is defined as

$$C_{i,j} = V_{i,j}/(\sigma_i\sigma_j). \tag{11.2}$$

We present the correlation coefficient between bins based on just the systematic uncertainty in Table 11.6. If we include the statistical error, independent between the bins and uncorrelated with any of the systematic uncertainties, the correlation coefficients for the total uncertainty are shown in Table 11.7. It should also be noted that because in the highest mass bin ($2.6 < m_{X_s} < 2.8\,\mathrm{GeV}/c^2$) we find a negative yield, we need to be explicit about the difference between additive and multiplicative errors. Most of the systematic uncertainties we report in Table 11.2 are multiplicative uncertainties. For these multiplicative uncertainties, we recognize that an increase in the first 17 mass bins would result in a decrease in the 18th (the central value would get more negative if, say, we should have divided by less B mesons when determining the PBF in the highest mass bin). We therefore incorporate this as a "-1" correlation coefficient when calculating the covariance matrices for these uncertainties (an equivalent choice would be to keep the correlation as 1, and take the uncertainty to be negative).

However, we identify the $B\overline{B}$-fitting systematic as an additive uncertainty (in most bins the uncertainty on the number of peaking $B\overline{B}$ background dominates this value), since assigning more events to $B\overline{B}$ background effectively subtracts them from signal yield. The highest mass bin, therefore, has the behaviour one would expect for this systematic uncertatinty (if we should have increased the amount of $B\overline{B}$ background in this bin, the PBF would get smaller, or more negative, just as it would for such a correction in any of the other 17 bins).

11.3 Photon Spectrum Moments

We calculate the first and second moments of the photon spectrum. We use these moments to calculate the mean and variance of the spectrum, quantities that are used as input to different models of the photon spectrum. We calculate the quantities:

$$M_n = \frac{\sum_{i=1}^{18} BF_i E_{\gamma i}^n}{\sum_{i=1}^{18} BF_i}, \tag{11.3}$$

Table 11.6: Correlation coefficient for the systematic errors between mass bins.

m_{X_S}	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0	1.0-1.1	1.1-1.2	1.2-1.3	1.3-1.4	1.4-1.5	1.5-1.6	1.6-1.7	1.7-1.8	1.8-1.9	1.9-2.0	2.0-2.2	2.2-2.4	2.4-2.6	2.6-2.8
0.6-0.7	1.000	0.851	0.485	0.532	0.363	0.561	0.511	0.490	0.421	0.432	0.373	0.356	0.661	0.355	0.150	0.089	0.060	-0.166
0.7-0.8	0.851	1.000	0.863	0.867	0.419	0.682	0.681	0.665	0.549	0.540	0.464	0.426	0.579	0.367	0.224	0.130	0.089	-0.162
0.8-0.9	0.485	0.863	1.000	0.935	0.378	0.668	0.711	0.701	0.561	0.513	0.447	0.401	0.308	0.287	0.248	0.143	0.099	-0.148
0.9-1.0	0.532	0.867	0.935	1.000	0.382	0.709	0.754	0.743	0.596	0.549	0.477	0.429	0.350	0.312	0.263	0.152	0.105	-0.155
1.0-1.1	0.363	0.419	0.378	0.382	1.000	0.360	0.361	0.352	0.287	0.259	0.230	0.214	0.208	0.168	0.118	0.069	0.047	-0.104
1.1-1.2	0.561	0.682	0.668	0.709	0.360	1.000	0.994	0.990	0.964	0.549	0.538	0.507	0.350	0.393	0.356	0.234	0.181	-0.253
1.2-1.3	0.511	0.681	0.711	0.754	0.361	0.994	1.000	0.999	0.954	0.571	0.566	0.532	0.335	0.407	0.389	0.258	0.203	-0.257
1.3-1.4	0.490	0.665	0.701	0.743	0.352	0.990	0.999	1.000	0.958	0.578	0.583	0.549	0.346	0.426	0.418	0.281	0.225	-0.260
1.4-1.5	0.421	0.549	0.561	0.596	0.287	0.964	0.954	0.958	1.000	0.475	0.483	0.456	0.309	0.361	0.349	0.237	0.191	-0.214
1.5-1.6	0.432	0.540	0.513	0.549	0.259	0.549	0.571	0.578	0.475	1.000	0.970	0.969	0.791	0.941	0.421	0.300	0.255	-0.204
1.6-1.7	0.373	0.464	0.447	0.477	0.230	0.538	0.566	0.583	0.483	0.970	1.000	0.998	0.801	0.972	0.554	0.409	0.359	-0.237
1.7-1.8	0.356	0.426	0.401	0.429	0.214	0.507	0.532	0.549	0.456	0.969	0.998	1.000	0.804	0.982	0.537	0.399	0.351	-0.231
1.8-1.9	0.661	0.579	0.308	0.350	0.208	0.350	0.335	0.346	0.309	0.791	0.801	0.804	1.000	0.868	0.410	0.312	0.280	-0.147
1.9-2.0	0.355	0.367	0.287	0.312	0.168	0.393	0.407	0.426	0.361	0.941	0.972	0.982	0.868	1.000	0.504	0.382	0.341	-0.195
2.0-2.2	0.150	0.224	0.248	0.263	0.118	0.356	0.389	0.418	0.349	0.421	0.554	0.537	0.410	0.504	1.000	0.940	0.399	-0.204
2.2-2.4	0.089	0.130	0.143	0.152	0.069	0.234	0.258	0.281	0.237	0.300	0.409	0.399	0.312	0.382	0.940	1.000	0.310	-0.148
2.4-2.6	0.060	0.089	0.099	0.105	0.047	0.181	0.203	0.225	0.191	0.255	0.359	0.351	0.280	0.341	0.399	0.310	1.000	0.099
2.6-2.8	-0.166	-0.162	-0.148	-0.155	-0.104	-0.253	-0.257	-0.260	-0.214	-0.204	-0.237	-0.231	-0.147	-0.195	-0.204	-0.148	0.099	1.000

Table 11.7: Correlation coefficient for the total errors between mass bins.

m_{X_s}	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0	1.0-1.1	1.1-1.2	1.2-1.3	1.3-1.4	1.4-1.5	1.5-1.6	1.6-1.7	1.7-1.8	1.8-1.9	1.9-2.0	2.0-2.2	2.2-2.4	2.4-2.6	2.6-2.8
0.6-0.7	1.000	0.025	0.055	0.056	0.038	0.045	0.056	0.054	0.048	0.046	0.040	0.039	0.076	0.036	0.017	0.009	0.006	-0.014
0.7-0.8	0.025	1.000	0.182	0.172	0.083	0.102	0.141	0.137	0.117	0.108	0.094	0.089	0.125	0.070	0.047	0.026	0.018	-0.026
0.8-0.9	0.055	0.182	1.000	0.697	0.283	0.378	0.556	0.544	0.451	0.387	0.342	0.314	0.251	0.207	0.198	0.107	0.074	-0.091
0.9-1.0	0.056	0.172	0.697	1.000	0.268	0.376	0.552	0.540	0.449	0.388	0.342	0.314	0.267	0.210	0.196	0.106	0.073	-0.089
1.0-1.1	0.038	0.083	0.283	0.268	1.000	0.192	0.265	0.257	0.217	0.183	0.165	0.157	0.159	0.114	0.088	0.048	0.033	-0.060
1.1-1.2	0.045	0.102	0.378	0.376	0.192	1.000	0.552	0.546	0.551	0.294	0.293	0.282	0.203	0.201	0.202	0.124	0.096	-0.111
1.2-1.3	0.056	0.141	0.556	0.552	0.265	0.552	1.000	0.762	0.753	0.423	0.425	0.409	0.268	0.288	0.305	0.189	0.148	-0.155
1.3-1.4	0.054	0.137	0.544	0.540	0.257	0.546	0.762	1.000	0.751	0.425	0.436	0.419	0.276	0.300	0.325	0.205	0.164	-0.156
1.4-1.5	0.048	0.117	0.451	0.449	0.217	0.551	0.753	0.751	1.000	0.361	0.373	0.360	0.255	0.263	0.281	0.178	0.144	-0.133
1.5-1.6	0.046	0.108	0.387	0.388	0.183	0.294	0.423	0.425	0.361	1.000	0.703	0.718	0.611	0.643	0.317	0.212	0.180	-0.119
1.6-1.7	0.040	0.094	0.342	0.342	0.165	0.293	0.425	0.436	0.373	0.703	1.000	0.751	0.629	0.674	0.424	0.293	0.257	-0.140
1.7-1.8	0.039	0.089	0.314	0.314	0.157	0.282	0.409	0.419	0.360	0.718	0.751	1.000	0.645	0.696	0.421	0.292	0.257	-0.139
1.8-1.9	0.076	0.125	0.251	0.267	0.159	0.203	0.268	0.276	0.255	0.611	0.629	0.645	1.000	0.642	0.335	0.238	0.214	-0.092
1.9-2.0	0.036	0.070	0.207	0.210	0.114	0.201	0.288	0.300	0.263	0.643	0.674	0.696	0.642	1.000	0.364	0.258	0.230	-0.109
2.0-2.2	0.017	0.047	0.198	0.196	0.088	0.202	0.305	0.325	0.281	0.317	0.424	0.421	0.335	0.364	1.000	0.701	0.297	-0.126
2.2-2.4	0.009	0.026	0.107	0.106	0.048	0.124	0.189	0.205	0.178	0.212	0.293	0.292	0.238	0.258	0.701	1.000	0.216	-0.085
2.4-2.6	0.006	0.018	0.074	0.073	0.033	0.096	0.148	0.164	0.144	0.180	0.257	0.257	0.214	0.230	0.297	0.216	1.000	0.057
2.6-2.8	-0.014	-0.026	-0.091	-0.089	-0.060	-0.111	-0.155	-0.156	-0.133	-0.119	-0.140	-0.139	-0.092	-0.109	-0.126	-0.085	0.057	1.000

where BF_i is the PBF found in bin i when evaluating these in bins of photon energy, and $E_{\gamma i}$ is the photon energy corresponding to the center of the bin (mean photon energy within the bin). We report the mean (M_1) and variance ($M_2 - M_1^2$) calculated at multiple photon energy cutoffs in Table 11.8.

Table 11.8: The moments of the photon energy spectrum, calculated at multiple photon energy cutoffs. The errors are statistical and systematic, calculated as described in the text.

$E_{\gamma min}$	$\langle E \rangle$	$\langle E^2 \rangle - \langle E \rangle^2$
	Preliminary Results	
1.897	$2.346 \pm 0.018^{+0.027}_{-0.022}$	$0.0211 \pm 0.0057^{+0.0055}_{-0.0069}$
1.999	$2.338 \pm 0.010^{+0.020}_{-0.017}$	$0.0239 \pm 0.0018^{+0.0023}_{-0.0030}$
2.094	$2.365 \pm 0.006^{+0.016}_{-0.010}$	$0.0176 \pm 0.0009^{+0.0009}_{-0.0016}$
2.181	$2.391 \pm 0.003^{+0.008}_{-0.007}$	$0.0129 \pm 0.0003^{+0.0005}_{-0.0005}$
2.261	$2.427 \pm 0.002^{+0.006}_{-0.006}$	$0.0082 \pm 0.0002^{+0.0002}_{-0.0002}$
	Previous Sum of Exclusives Results	
1.897	$2.321 \pm 0.038^{+0.017}_{-0.038}$	$0.0253 \pm 0.0101^{+0.0041}_{-0.0028}$
1.999	$2.314 \pm 0.023^{+0.014}_{-0.029}$	$0.0273 \pm 0.0037^{+0.0015}_{-0.0015}$
2.094	$2.357 \pm 0.017^{+0.007}_{-0.017}$	$0.0183 \pm 0.0023^{+0.0010}_{-0.0007}$
2.181	$2.396 \pm 0.013^{+0.003}_{-0.009}$	$0.0115 \pm 0.0014^{+0.0005}_{-0.0003}$
2.261	$2.425 \pm 0.009^{+0.002}_{-0.004}$	$0.0075 \pm 0.0007^{+0.0002}_{-0.0002}$

Since the quantities themselves are not linear relations between the measured PBFs, the uncertainties on these quantities (also reported in Table 11.8) cannot be assumed to be Gaussian distributed, and we therefore simulate the uncertainties, rather than turn to more naive error propagation methods. When doing this, we also need to account for the correlations between bins in our systematic errors (though the different systematic errors themselves, given in Table 11.2, are uncorrelated with one another). One way to do this would be to input an 18-dimensional Gaussian PDF with the correlation matrix given in Table 11.7 into a program that could use it to generate values for our 18 PBFs (distributed according to this Gaussian PDF), and calculate the resulting moment a large number of times. We have not come up with a way, however, to successfully do this exact solution since such programs generally normalize the PDF, or ensure a normalized PDF, which requires a numeric integral over 18 dimensions, and is not feasable.

Our work around is to consider the uncorrelated uncertaintiess separately (there are eight

of these, seven systematic and one statistical), then combine the resulting uncertainties on the mean and variance in quadrature.

For the statistical error on each PBF, uncorrelated between the mass bins, we assume the 18 measured PBFs to be Gaussian distributed and take the measured values as the central values for each of the 18 Gaussians, and the statistical uncertainties as the widths. We then generate 18 BF values, one for each bin, and re-evaluate the mean and variance at each of our photon energy cutoff values. We repeat this procedure, generating 18 points and calculating the mean and variance, 10^5 times to obtain distributions of these quantities. We take as our uncertainty on the calculated value the difference between the central point (which we take to be the peak location of the calculated quantity, slightly different in general than the mean due to the asymmetry of the distribution) and the points that correspond to 16% of the integral of the curve (so the generally asymmetric 68% coverage region). For variance distributions that have a negative tail (an unphysical variance), we only consider > 0 values for determining our 16% coverage regions. The distributions of the mean and variance for the lowest photon energy cutoff are shown in Figure 11.8.

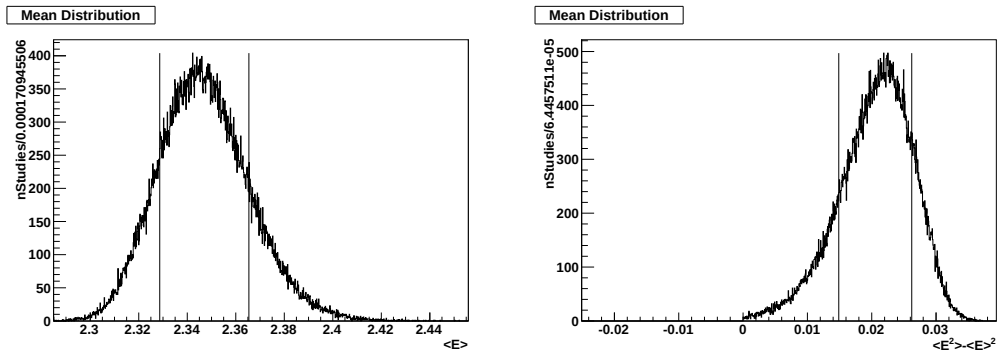


Figure 11.8: The mean (left) and variance (right) distributions used in calculating the error on these quantities. The vertical lines reflect the 16% integrals (68% coverage in the central region).

For the uncertainties that are correlated between bins, we introduce this correlation "by hand." We again interpret the uncertainties to reflect Gaussian distributions about the central, measured value, and produce 18 Gaussian PDFs with widths corresponding to the uncertainties reported in Table 11.2. However, for these correlated uncertainties, we produce a 19th Gaussian, a "reference Gaussian," which is centered on 0 and has a width of

1. We use this reference Gaussian to dictate how much of a shift we should impose on our 18 PDFs simultaneously (we take the value of the PBF in a given mass bin corresponding to a shift by the number of sigma dictated by the reference Gaussian). This way we are recovering the completely correlated error behaviour (a 1-σ shift up in one mass bin is simultaneous with a 1-σ shift up in another) in reasonable computation time.

For uncertainties which have groups of bins that are completely correlated (such as the fragmentation uncertainties), but these groups are not correlated with one another, we simply use multiple reference Gaussians (one for each correlated group of bins).

Propagating the uncertainty in this manner also allows us to evaluate the correlations in the uncertainties between different determinations of the calculated values (between mean and variance, or mean evaluated at different photon energy cutoffs, etc., an example can be seen in Figure 11.9). Each of the uncertainties produces a covariance matrix for these correlations; we add the covariance matrices (and divide by the total error) to get the total correlation coefficients. We report the correlation coefficients for the statistics errors in Table 11.9, for all of the systematic errors in Table 11.10, and for the combined statistics and systematics correlation in Table 11.11. We want to stress, however, that although we present these values as being independent of one another, there are only six independent quantities between the 10 means and variances we report (once six of the quantities are determined, the remaining four are uniquely defined). Nevertheless, we present all of the values in the table since they may be of interest to different calculations.

11.4 Spectrum Fit

Finally, we can use the mass spectrum we have measured to fit different models theorists have made to describe the photon spectrum, and use these fits to extract values for the HQET parameters (such as m_b and μ_π^2). We consider two "classes" of models, the so-called "kinetic model" [16], or BBU model we have been using as our default throughout the analysis, and the "shape-function" model [17]. Each of these models have multiple ways of parameterizing the expected shape of the photon spectrum in the non-perturbative energy region (a parton momentum distribution based on a Gaussian, exponential, and hyperbolic

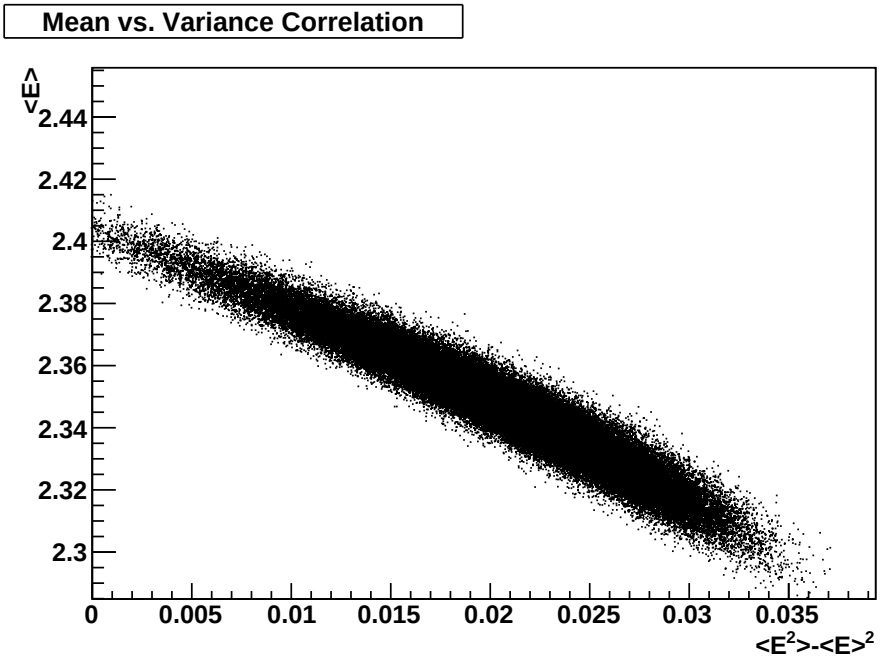


Figure 11.9: The correlation between the mean (y-axis) and variance (x-axis) when propagating the statistical error at the lowest photon energy cutoff.

Table 11.9: The correlation coefficients for the different minimum photon energies based on statistics uncertainty.

		$\langle E \rangle$					$\langle E^2 \rangle - \langle E \rangle^2$				
		1.897	1.999	2.094	2.181	2.261	1.897	1.999	2.094	2.181	2.261
$\langle E \rangle$	1.8970	1.00	0.55	0.29	0.11	0.04	-0.96	-0.50	-0.26	-0.09	-0.03
	1.9990		1.00	0.51	0.19	0.06	-0.34	-0.91	-0.46	-0.16	-0.05
	2.0940			1.00	0.41	0.14	-0.08	-0.15	-0.89	-0.33	-0.12
	2.1810				1.00	0.40	0.02	0.12	0.01	-0.76	-0.31
	2.2610					1.00	0.03	0.13	0.19	0.21	-0.70
$\langle E^2 \rangle - \langle E \rangle^2$	1.897						1.00	0.36	0.1	0.00	-0.02
	1.999							1.00	0.23	-0.02	-0.06
	2.094								1.00	0.15	-0.05
	2.181									1.00	0.11
	2.261										1.00

Table 11.10: The correlation coefficients for the different minimum photon energies based on systematic uncertainties.

		$\langle E \rangle$					$\langle E^2 \rangle - \langle E \rangle^2$				
		1.897	1.999	2.094	2.181	2.261	1.897	1.999	2.094	2.181	2.261
$\langle E \rangle$	1.897	1.00	0.80	0.53	0.51	0.25	-0.87	-0.74	-0.41	-0.36	-0.22
	1.999		1.00	0.77	0.76	0.41	-0.44	-0.06	-0.55	-0.48	-0.27
	2.094			1.00	0.93	0.45	-0.09	-0.30	-0.79	-0.65	-0.31
	2.181				1.00	0.71	-0.05	-0.25	-0.53	-0.41	-0.52
	2.261					1.00	0.06	0.01	0.17	0.34	-0.79
$\langle E^2 \rangle - \langle E \rangle^2$	1.897						1.00	0.62	0.15	0.16	0.02
	1.999							1.00	0.32	0.35	-0.03
	2.094								1.00	0.93	-0.14
	2.181									1.00	-0.21
	2.261										1.00

Table 11.11: The correlation coefficients for the different minimum photon energies based on total uncertainties (statistical and systematic).

		$\langle E \rangle$					$\langle E^2 \rangle - \langle E \rangle^2$				
		1.897	1.999	2.094	2.181	2.261	1.897	1.999	2.094	2.181	2.261
$\langle E \rangle$	1.897	1.00	0.72	0.46	0.40	0.20	-0.90	-0.66	-0.36	-0.27	-0.13
	1.999		1.00	0.71	0.65	0.35	-0.39	-0.29	-0.52	-0.40	-0.18
	2.094			1.00	0.84	0.40	-0.08	-0.25	-0.81	-0.57	-0.23
	2.181				1.00	0.67	-0.03	-0.16	-0.39	-0.48	-0.42
	2.261					1.00	0.05	0.04	0.17	0.31	-0.68
$\langle E^2 \rangle - \langle E \rangle^2$	1.897						1.00	0.51	0.13	0.10	0.00
	1.999							1.00	0.29	0.24	-0.04
	2.094								1.00	0.69	-0.10
	2.181									1.00	-0.08
	2.261										1.00

for the models in [17]; Gaussian and exponential in [16]). We do not believe that our results for the best values of the underlying HQET parameters will (nor indeed should they) depend too strongly on the choice of shape used in each model (Gaussian vs. exponential, say); though there will be differences between the central values found for each class of models as they are evaluated at different energy scales (1.0 GeV for the kinetic models and 1.5 GeV for the shape-function models), and are formulated using fundamentally different methods. For both classes of models, we choose to fit the respective exponential *ansatz*.

The kinetic models relate the perturbative and non-perturbative pieces of the first two moments directly to the HQET parameters m_b and μ_π^2 (the kinetic operator) as well as other parameters, and evaluate the bias introduced in these quantities by the experimental realities of having to place a minimum energy cut on the photon spectrum. The shape-function models instead use two parameters to tune the heavy-quark distribution functions (either in the exponential, gaussian, or hyperbolic *ansatz*), which result in shapes for the photon spectrum, and relate the values of μ_π^2 and $\overline{\Lambda}(\equiv (m_B - m_b)_{lim.\ m_b \to \infty})$ to the moments of the measured spectrum directly.

In order to fit these models, we use lookup tables provided by their authors. Though these tables are provided in bins of photon energy, we can convert them without losing information to bins of the hadron mass by inverting equation 3.1. Since these models assume quark-hadron duality, they do not take resonances, in particular the $K^*(892)$ resonance, into account in the photon nor m_{X_s} spectrum. While in general, once systematic uncertainties are included, the majority of our spectrum can be approximated by a smooth curve, the $K^*(892)$ region cannot, and we have to make special accommodations for this resonance when fitting the spectrum models.

In order to parameterize the $K^*(892)$ region, we use a P-wave relativistic Breit Wigner [39] (RBW) distribution, and fit it to the $K^*(892)$ MC. We then use the integral of the RBW over the individual mass bins in the K^* region to give us the expected values within these mass bins. When fitting the spectrum models, we require that the integral of the spectrum over the K^* region be the same as the integral of the RBW. This imposes the quark-hadron duality assumed by the theorists, but allows for the substructure of the K^* resonance. We find

from fits to the transition region that a transition point of 1.17 GeV/c^2 (the same value found by the previous analysis, independently confirmed) is a better choice than the relatively arbitrary hadron-mass bin boundaries. For this transition bin ($1.1 < m_{X_s} < 1.2\,\mathrm{GeV}/c^2$), we take the integral of the RBW up to the transition point, and sum it with the spectrum model above the transition point to get a value for the model across the bin.

We perform fits to the spectrum models by minimizing the quantity χ^2 in which we have taken correlations into account:

$$\chi^2 = (PBF_{theory} - PBF_{exp.})^T V^{-1} (PBF_{theory} - PBF_{exp}), \tag{11.4}$$

where PBF_{theory} and $PBF_{exp.}$ are the predicted and measured PBFs, respectively, and V^{-1} is the inverse of the variance matrix, generalized from Table 11.5, and related to the values in Table 11.7 through inverting equation 11.2.

For the kinetic models, we find the best fit value of the parameters, when taking statistical and systematic uncertainties as well as correlations into account, to be

$$m_b = 4.568^{+0.038}_{-0.036}\,\mathrm{GeV}/c^2$$

and

$$\mu_\pi^2 = 0.450^{+0.054}_{-0.053}\,\mathrm{GeV}^2.$$

The 1-σ error elipse, corresponding to the values of m_b and μ_π^2 over which $\Delta\chi^2 = 1$, and the spectrum distribution for the best fit model, are shown in Figure 11.10. The three-dimensional rendering of the $\Delta\chi^2 = 1$ curve is shown in Figure 11.11. The minimum χ^2 value is 17.17 for 15 degrees of freedom (18 mass bins minus 2 model parameters and the $K^*(892)$ region cutoff value). Since we used lookup tables for this model, we need a way to interpolate between the points in the μ_π^2 vs. m_b plane given to us by the theorists. The points we were given range from m_b=[4.45-4.75] GeV/c^2 and μ_π^2=[0.2-0.7] GeV ^{2} in increments of 0.05 for each parameter. To do this, we use the equation:

$$F(m_b, \mu_\pi^2) = a + b \times (m_b - 4.45) + c \times (\mu_\pi^2 - 0.2) + d \times (m_b - 4.45)(\mu_\pi^2 - 0.2), \tag{11.5}$$

and solve for $[a, b, c, d]$. The values 4.45 and 0.2 change, depending on which range we are interpolating over (for instance, we would use this equation for the range m_b=[4.45-4.50] and μ_π^2=[0.2-0.25]). This choice for an interpolation function assures continuity at the borders between regions over which we interpolate, and exact solutions at the points provided by the theorists. Analytically matching up other quantities between regions, such as derivatives, *etc.*, is beyond the scope of this procedure.

The shape-function models are determined by two parameters, b and Λ. Alone, these parameters have no real physical meaning but are related to quantities μ_π^2 and $\overline{\Lambda}$, or, equivalently, m_b (see equation (42) in [17]). We present the best value of these model parameters based on our spectrum, and show these in Figure 11.12. The three-dimensional rendering of the $\Delta\chi^2 = 1$ is shown in Figure 11.13. Again, since we use lookup tables, we interpolate between the values given by the theorists, in this case we have values for b=2.0-5.0 in increments of 0.25 and Λ=0.4-0.9 in increments of 0.05. We note that there is a slight discontinuity in the error ellipse (visible also in the 3-dimensional graphic); this is likely a result of our interpolation method. The elliptical shape of the $\Delta\chi^2 = 1$ curve is apparent however, and we take the extremes of this curve as the uncertainty on these parameter's values. We find the best parameter values at

$$b = 3.48^{+0.19}_{-0.15}$$

and

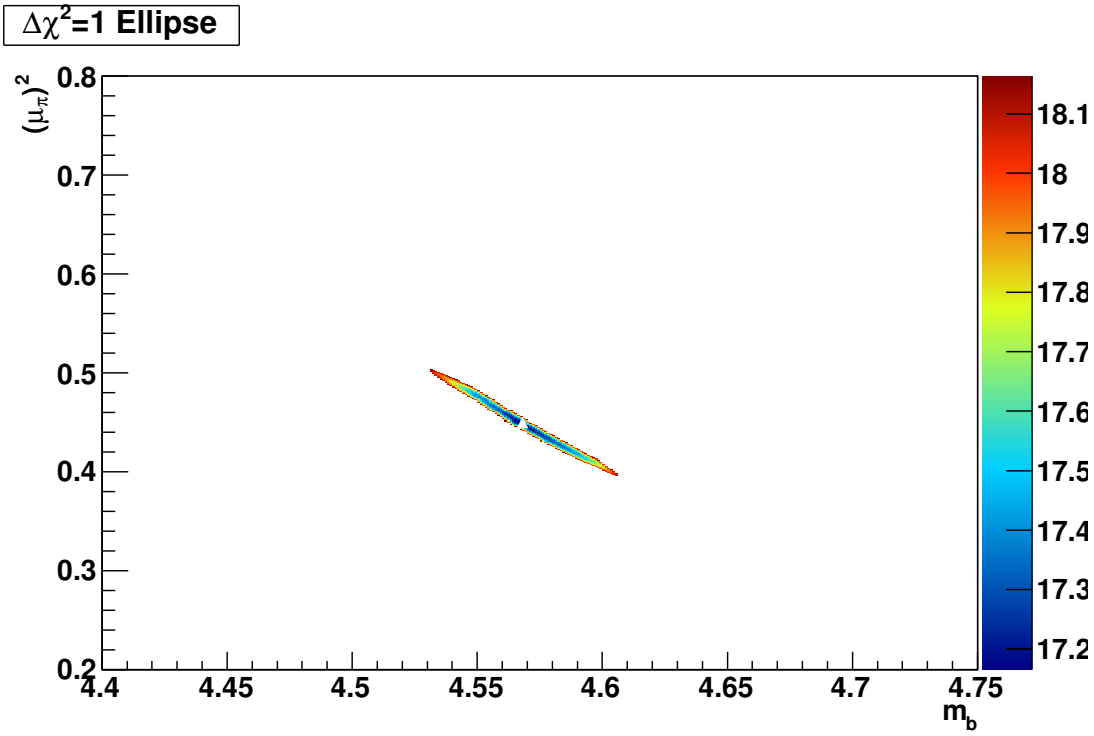
$$\Lambda = 0.77^{+0.043}_{-0.042}.$$

The central values have the minimum χ^2 value with $\chi^2 = 17.1$ with 15 d.o.f.. The corresponding values of m_b and μ_π^2 are

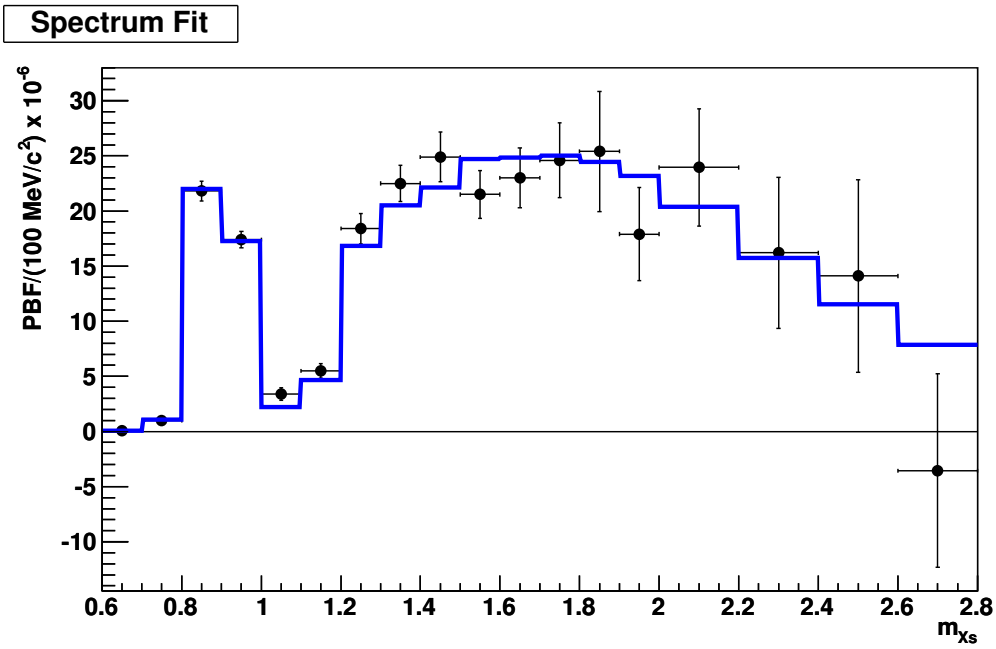
$$m_b = 4.579^{+0.032}_{-0.029}\,\mathrm{GeV}/c^2$$

and

$$\mu_\pi^2 = 0.257^{+0.034}_{-0.039}\,\mathrm{GeV}^2.$$



(a) The 1-σ elipse for the kinetic models' parameters based on our measured spectrum. The color range refers to the χ^2 value, calculated using equation 11.4.



(b) The best fit kinetic model compared to our measured PBFs in BF/100 MeV/c^2. Error bars include statistical and systematic uncertainties added in quadrature.

Figure 11.10: The kinetic model results.

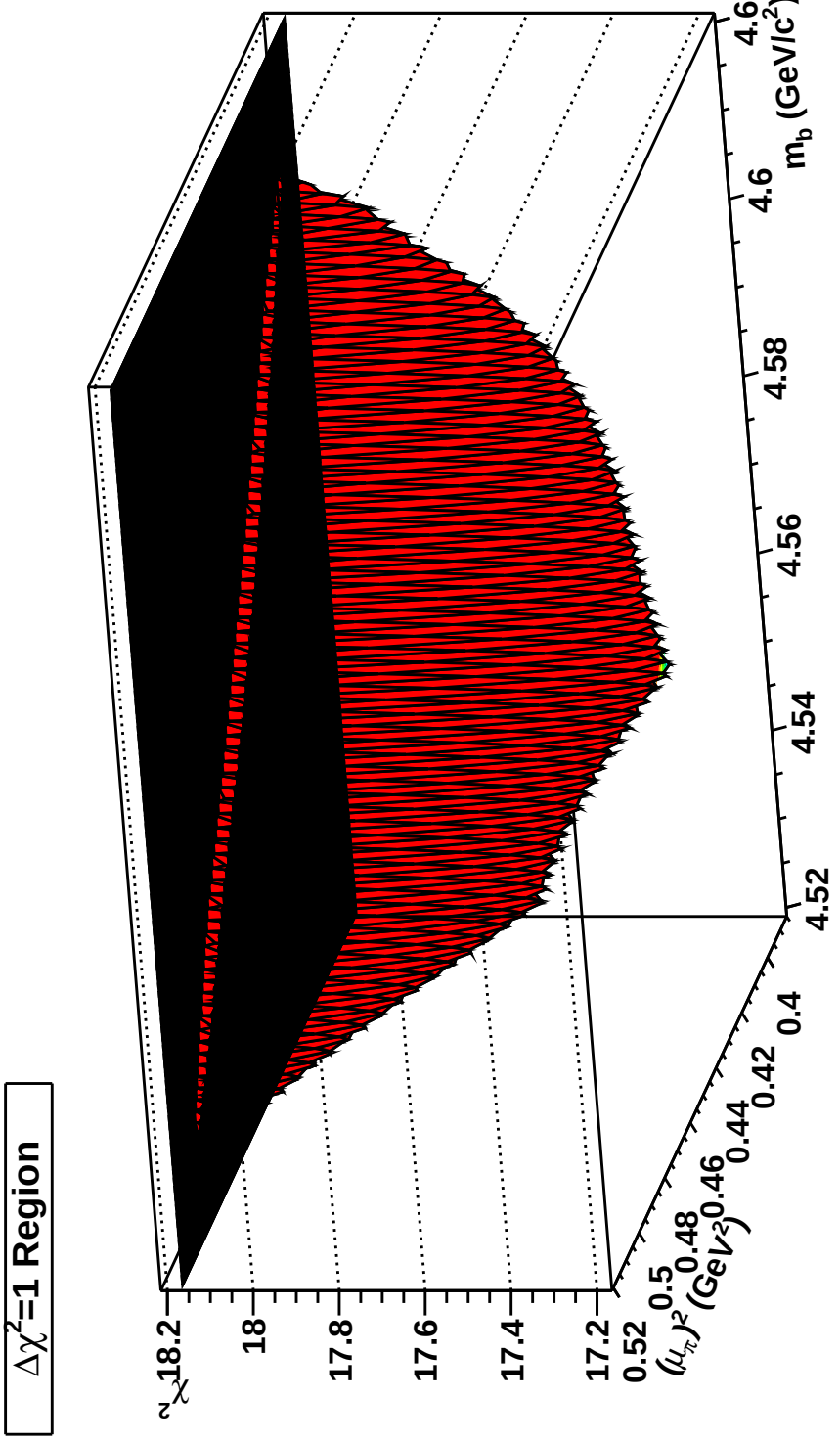
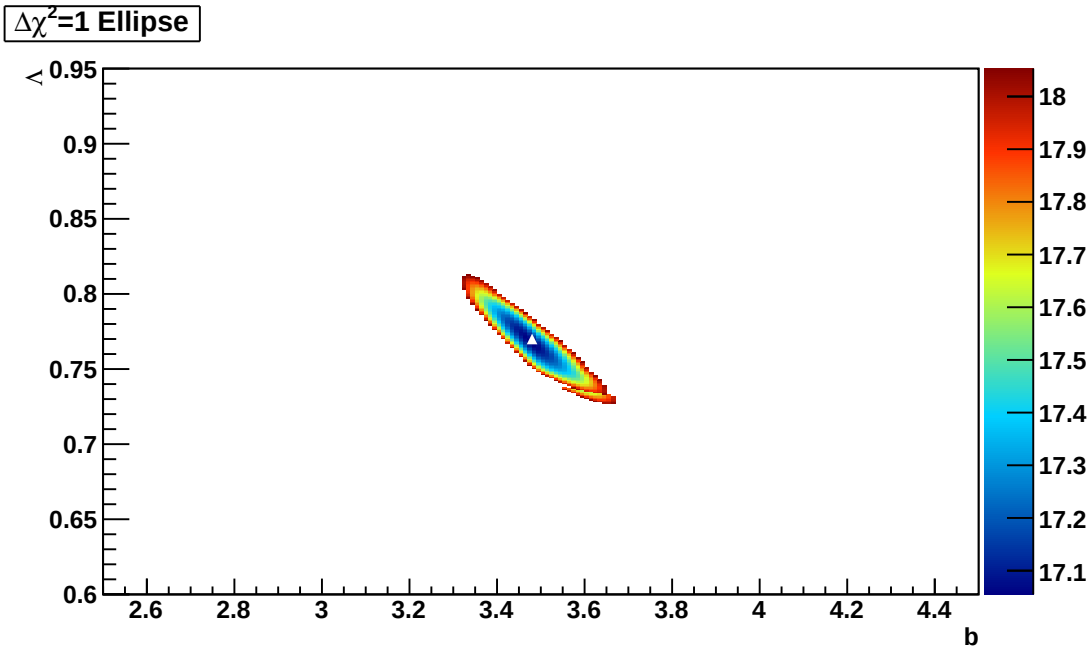
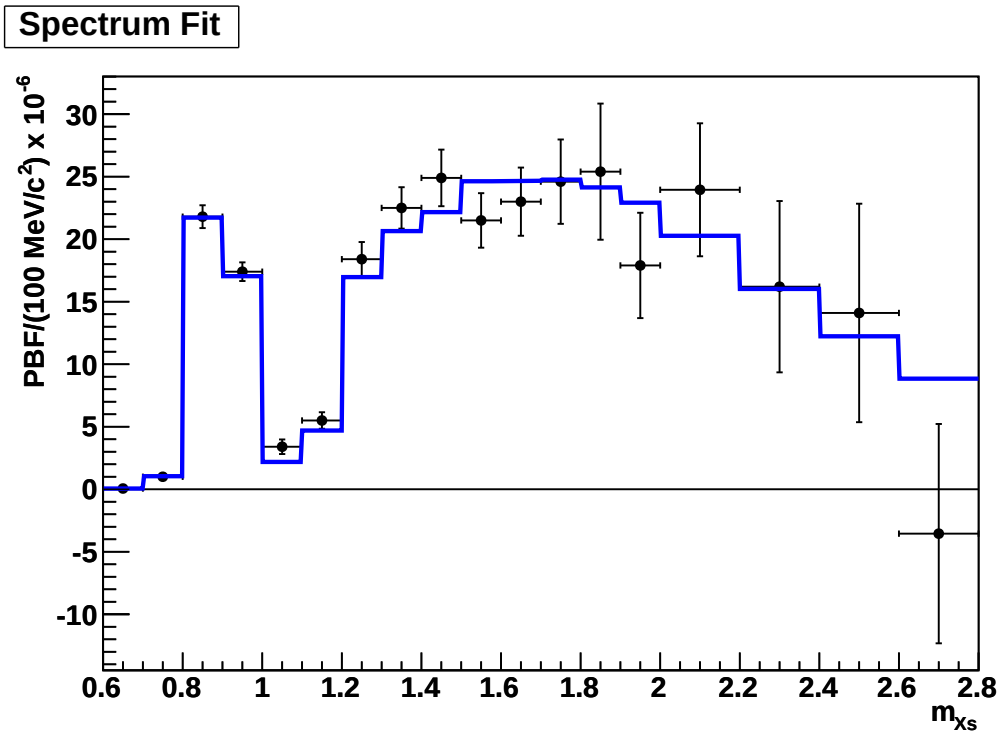


Figure 11.11: A 3-dimensional rendering of the 1-σ region for the kinetic models. Points with $\Delta\chi^2 > 1$ are fixed to a shift of 1.



(a) The 1-σ elipse for the shape-function models' parameters based on our measured spectrum. The color range refers to the χ^2 value, calculated using equation 11.4.



(b) The best fit shape-function model compared to our measured PBFs in BF/100 MeV/c^2. Error bars include statistical and systematic uncertainties added in quadrature.

Figure 11.12: The shape-function model results.

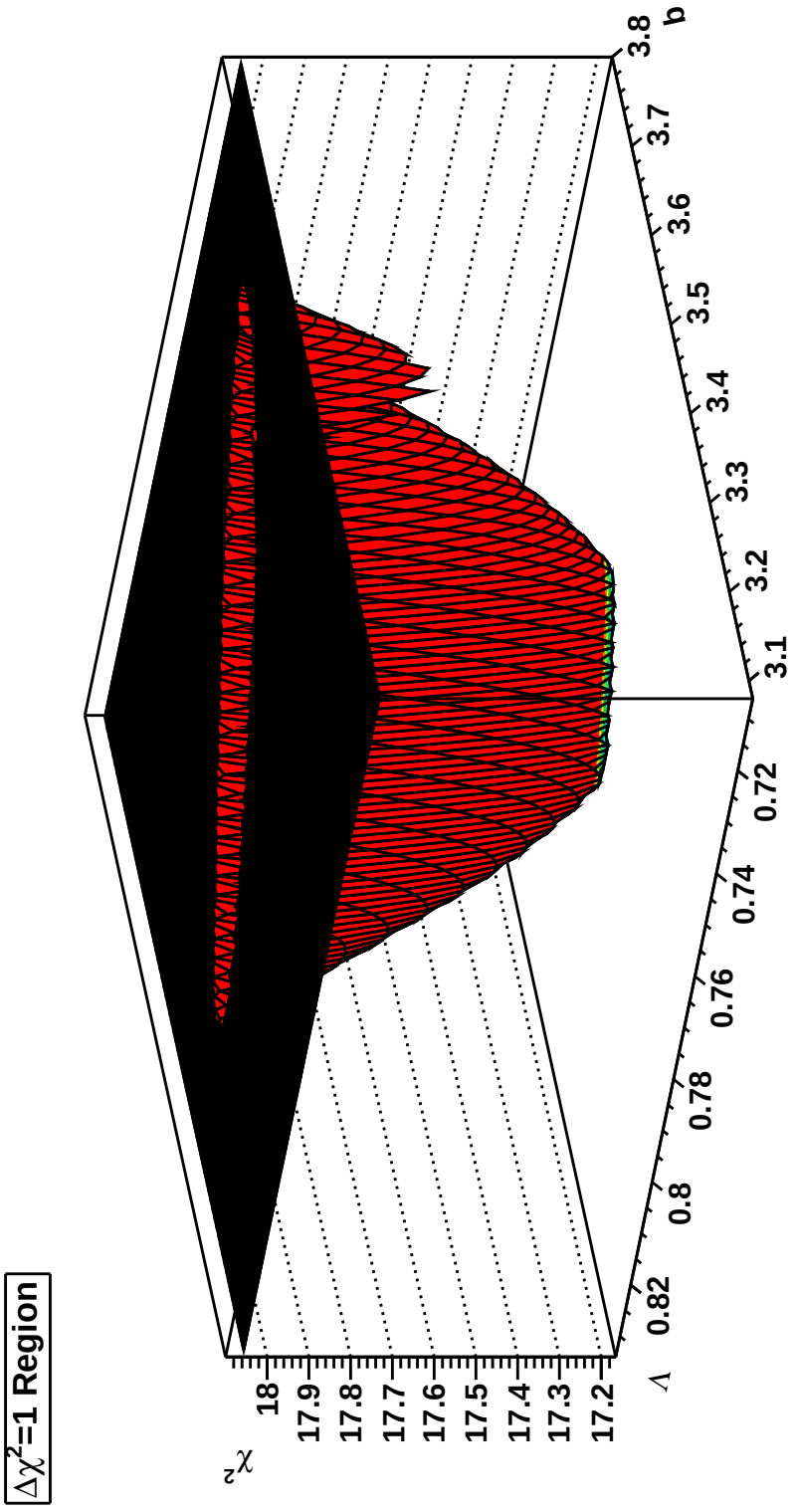


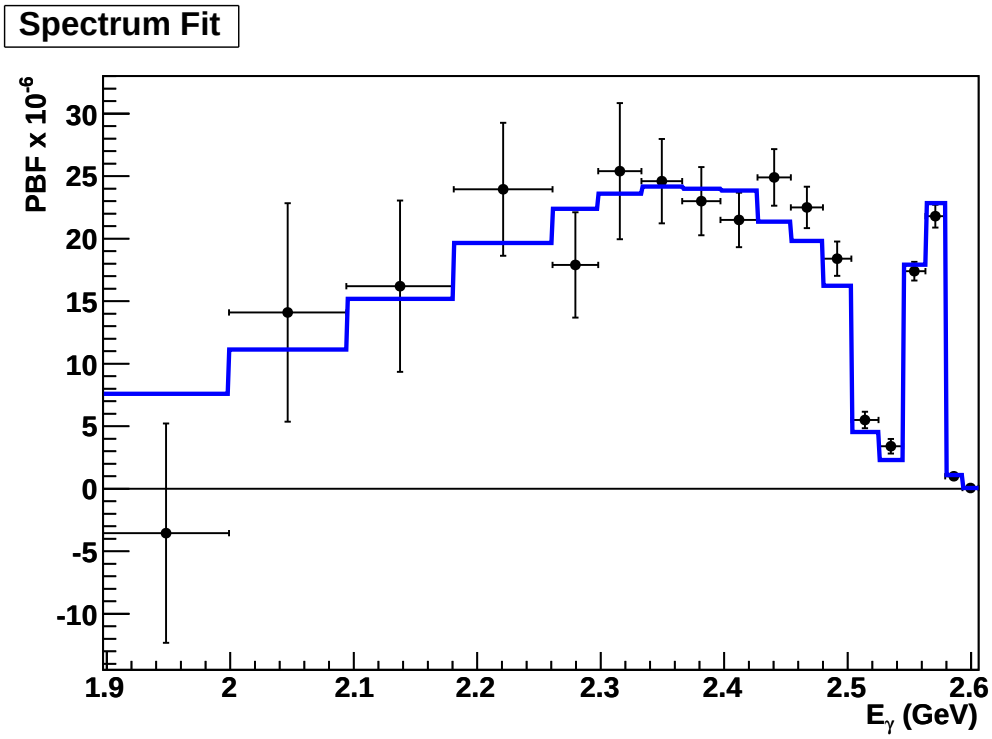
Figure 11.13: A 3-dimensional rendering of the 1-σ region. Points with $\Delta\chi^2 > 1$ are fixed to a shift of 1..

For both of these models, we also plot the expected photon spectrum compared to our measured photon spectrum (using the results measured in bins of hadronic mass and converting them to photon energy bins via equation 3.1), as well as the $\Delta\chi^2 = 3$ curve to see the long distance behavior of the uncertainties on the quantities we fit. These plots are shown in Figures 11.14 and 11.15.

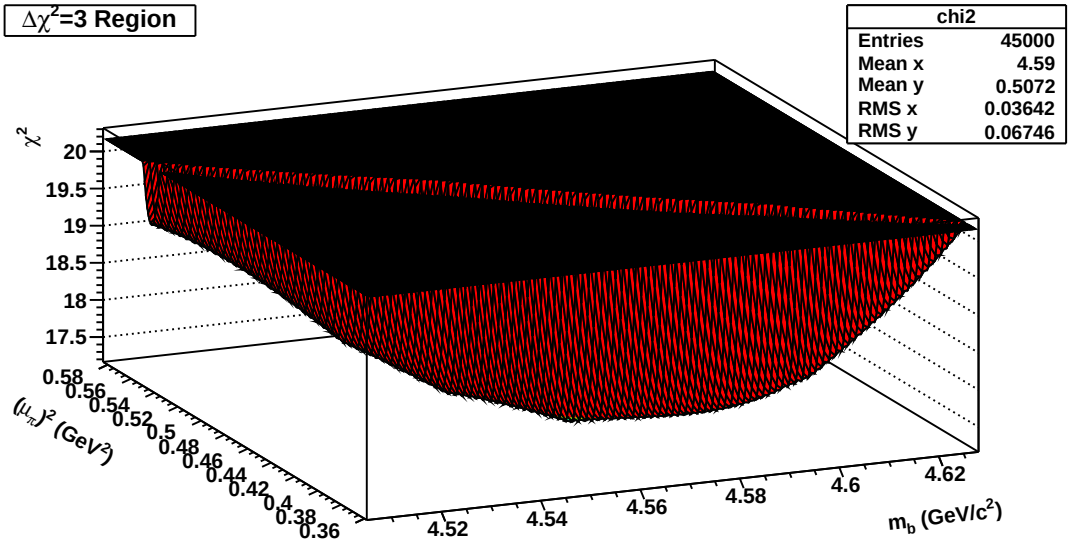
Finally, we can use the correction factors for each of these models provided by the Heavy Flavor Averaging Group (HFAG) to extrapolate from a minimum photon energy cutoff of $E_\gamma > 1.9\,\mathrm{GeV}$ to $E_\gamma > 1.6\,\mathrm{GeV}$ [5]. The resulting branching fractions at this lower photon energy cutoff are given in table 11.12.

Table 11.12: The branching fraction for $b \to s\gamma$ extrapolated to a minimum photon energy of $E_\gamma = 1.6\,\mathrm{GeV}$. The uncertainties are statistical then systematic.

Model	Branching Fraction ($E_\gamma > 1.6\,\mathrm{GeV}$)
Kinetic	$(350 \pm 20 \pm 51) \times 10^{-6}$
Shape Function	$(353 \pm 20 \pm 52) \times 10^{-6}$

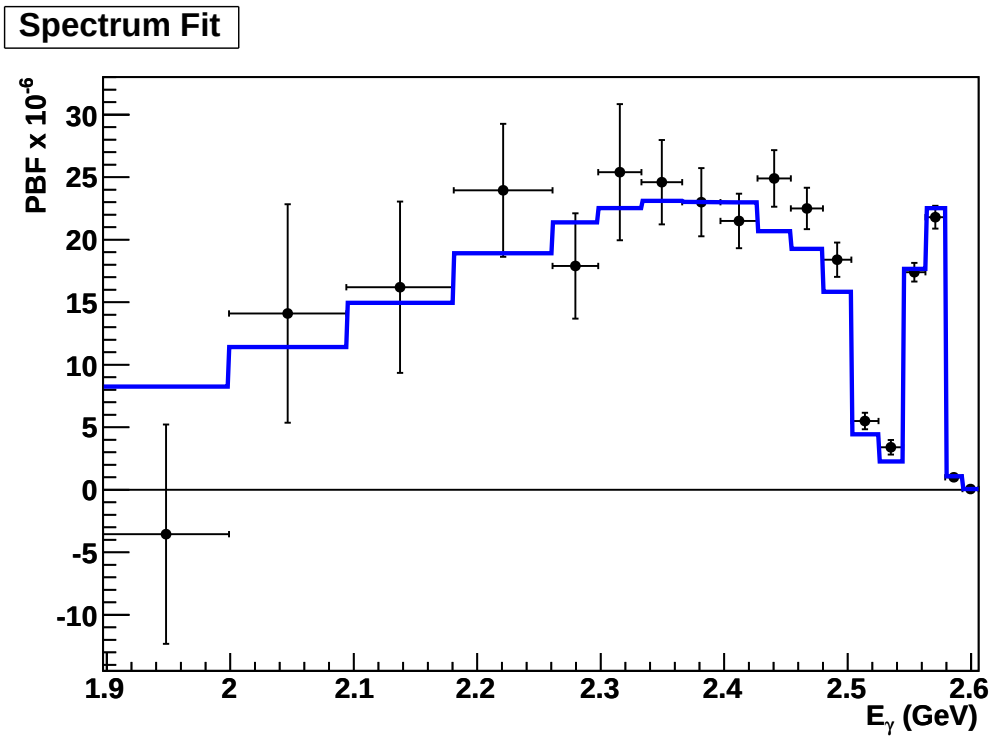


(a) The photon spectrum and the best fit kinetic model shape.

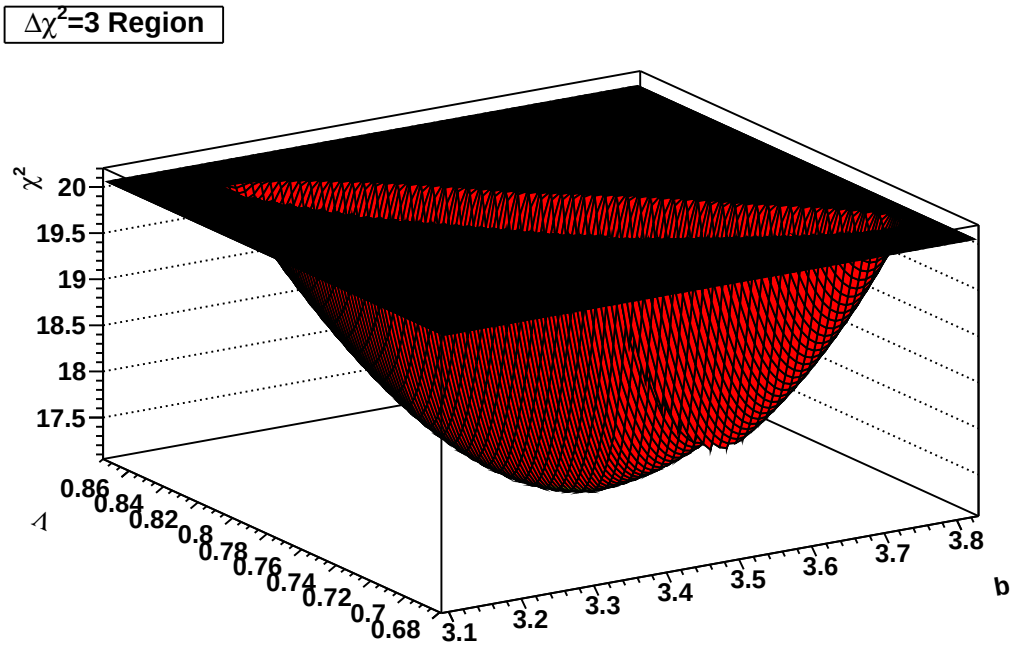


(b) The $\Delta\chi^2 = 3$ curve for the kinetic model.

Figure 11.14: The kinetic model results for the E_γ spectrum and the three-sigma curve.



(a) The photon spectrum and the best fit shape-function model shape.



(b) The $\Delta\chi^2 = 3$ curve for the shape-function model.

Figure 11.15: The shape function model results for the E_γ spectrum and the three-sigma curve.

Chapter 12

Conclusion

Using the entire *BABAR* data set, 429 fb^{-1}, we have measured the transition rate $b \to s\gamma$ using a "sum of exclusives" approach (reconstructing the s quark system in one of 38 final states). After reconstructing the B mesons and removing much of the background, we fit the m_{ES} distribution to extract our signal yield in bins of the hadronic system's mass, m_{X_s}. We find a total branching fraction of

$$\mathcal{B}(\overline{B} \to X_s\gamma) = 3.28 \pm 0.19 \pm 0.48 \times 10^{-4}$$

with a minimum photon energy cutoff of 1.9 GeV. This result is compatible with measurements from Belle, CLEO, and other independent measurements at *BABAR* (using an inclusive approach, for instance). Accounting for the missing region of the photon spectrum between $E_\gamma = 1.6 - 1.9$ GeV, we derive for the kinetic models

$$\mathcal{B}(\overline{B} \to X_s\gamma)_{E_\gamma > 1.6\,\mathrm{GeV}} = 3.50 \pm 0.20 \pm 0.51 \times 10^{-4}$$

and for the shape-function models

$$\mathcal{B}(\overline{B} \to X_s\gamma)_{E_\gamma > 1.6\,\mathrm{GeV}} = 3.53 \pm 0.20 \pm 0.52 \times 10^{-4}.$$

With the first order assumption that the branching fraction of $b \to s\gamma$ can be summarized by the Wilson coefficient C_7^{eff}, we can derive limits on the impact of NP models through [40]:

$$R_{b\to s\gamma} = \frac{BR(B \to X_s\gamma)}{BR(B \to X_s\gamma)^{SM}} = \left|1 + \frac{C_7^{NP}}{C_7^{eff,SM}}\right|^2, \tag{12.1}$$

where the branching ratio in the numerator reflects the measured value, and the branching ratio in the denominator is the value predicted by the SM. With the Wilson coefficient evaluated at m_b, using our measured central value for the shape function models, and the world average for the uncertainties (as these will likely remain unchanged with this new measurement), we find a limit on C_7^{NP}:

$$C_7^{NP} = (0.059 \pm 0.055) \times C_7^{eff,SM}. \tag{12.2}$$

Because we measure the partial branching fraction in bins of invariant mass of the hadron system (a very precise quantity), we are able to evaluate properties of the recoiling transition photon spectrum, such as mean and variance, independent of any input spectrum model. We evaluate these quantities at several minimum photon energy cutoffs, the lowest of which is $\sim 1.9\,\mathrm{GeV}$, where we find

$$\langle E \rangle = 2.346 \pm 0.018^{+0.027}_{-0.022}$$

and

$$\left\langle E^2 \right\rangle - \langle E \rangle^2 = 0.0211 \pm 0.0057^{+0.0055}_{-0.0069}$$

. The correlation between the total uncertainties on these two quantities (with the statistical and average systematic uncertainties added in quadrature) is -0.90. The precision of these moments is substantially improved over the previous *BABAR* sum of exclusives measurement, and is now limited by systematic uncertainties.

We are also able to use the results of the partial branching fraction measurements to fit specific models of the photon energy spectrum. The two particular classes of models we choose to fit are those by [16] and [17]. Each of these classes of models of the photon spectrum are based on HQET parameters, though they are derived in different ways and at different energy scales (so the fitted parameters are not immediately comparable). Each of the classes of models has multiple forms of the heavy quark distribution function, and we

choose to fit the parameters for the exponential form in each class of models, as the fitted parameter values are not expected to be too dependent on functional form. We find for the models by [16] that the best fit values are

$$m_b = 4.568^{+0.038}_{-0.036}\,\mathrm{GeV}/c^2$$

and

$$\mu_\pi^2 = 0.450^{+0.054}_{-0.053}\,\mathrm{GeV}^2$$

. For the models by [17] the best fit parameter values are

$$b = 3.48^{+0.19}_{-0.15}$$

and

$$\Lambda = 0.77^{+0.043}_{-0.042}$$

. These correspond to values of the HQET parameters

$$m_b = 4.579^{+0.032}_{-0.029}\,\mathrm{GeV}/c^2$$

and

$$\mu_\pi^2 = 0.257^{+0.034}_{-0.039}\,\mathrm{GeV}^2.$$

The uncertainty on m_b for this class of models is roughly 25% smaller than the uncertainty on this quantity used to calculate V_{ub} [21] using this class of models, at the time this document was written. As the uncertainty on m_b is one of the largest components of uncertainty in extracting V_{ub}, our measurement can reduce the theoretical error on V_{ub} by roughly 10%.

We are now limited by the systematic uncertainties in all aspects of this type of analysis. Any improvements at the next generation experiments, for instance SuperB, would have to be made by a different approach to measuring the transition $b \to s\gamma$, such as hadronically reconstructing the companion B, and then performing a more inclusive analysis on the $b \to s\gamma$ transition. This would help limit continuum backgrounds, but would require much

more data than is currently available at the B Factory.

Bibliography

[1] B. Aubert *et al.*, Phys. Rev. **D72**, 052004 (2005).

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

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

[4] M. Misiak and M. Steinhauser, Nucl. Phys. **B764**, 62 (2007).

[5] D. Asner *et al.*, "Averages of b-hadron, c-hadron, and tau-lepton Properties," arXiv:1010.1589 (2011).

[6] P. Gambino and M. Misiak, Nucl. Phys. **B611**, 338 (2001).

[7] M. Misiak *et al.*, Phys. Rev. Lett. **98**, 022002 (2007).

[8] T. van Ritbergen, Phys. Lett. **B454**, 353 (1999).

[9] U. Haisch, FPCP Conference, arXiv:0805.2141 [hep-ph] (2008).

[10] O. Deschamps *et al., the CKMfitter Group*, Phys. Rev. **D82**, 073012 (2010); *CKMfitter Group* (J. Charles *et al.*), Eur. Phys. J. **C41**, 1 (2005), [hep/ph/0406184], updated results and plots available at: http://ckmfitter.in2p3.fr

[11] M. Wick and W. Altmannshofer, AIP Conf.Proc. **1078**, 348 (2009).

[12] N. Chen *et al.*, Phys. Lett. **B685** 174 (2010).

[13] A. Ferroglia, Mod. Phys. Lett. **A23** 3123 (2008) and references therein.

[14] I.I. Bigi, M.A. Shifman, N.G. Uraltsev, and A.I. Vainshtein, Int. J. Mod. Phys. **A9** 2467 (1994).

[15] A.L. Kagan and M. Neubert, Eur. Phys. J. **C7** 5 (1999).

[16] D. Benson, I. Bigi and N. Uraltsev, Nucl. Phys. **B710** 371 (2005).

[17] B. Lange, M. Neubert and G. Paz, Phys. Rev. **D72** 073006 (2005).

[18] W. Kozanecki, Nucl. Instrum. Meth. **A446**, 59 (2000).

[19] *SLAC report 504, "The BABAR Physics Book. Physics at an Asymmetric B Factory"* (1998).

[20] B. Aubert *et al.*, Nucl. Instrum. Meth. **A479**, 1 (2002).

[21] K. Nakamura *et al.* (Particle Data Group), J. Phys. **G37**, 075021 (2010)

[22] B. Aubert *et al.*, Phys. Rev. **D79**, 072009 (2009).

[23] R. Zhu, Nucl. Instrum. Meth. **A413**, 297 (1998).

[24] F. Anulli *et al.*, Nucl. Instrum. Meth. **A539**, 155 (2005).

[25] T. G. Dietterich and G. Bakiri, Journal of Articifical Intelligence Research **2**, 263 (1995).

[26] P. Ongmongkolkul, Future Ph.D. Thesis, California Institute of Technology.

[27] G. Fox and S. Wolfram, Nucl. Phys. **B149**, 413 (1979).

[28] L. Breiman, Machine Learning **45**, 5 (2001); L. Breiman *et al.*, Classification and Regression Trees, Waldsworth International (1984); L. Breiman, Bagging Predictors, Machine Learning **26**, 123 (1996).

[29] I. Narsky, StatPatternRecognition: A C++ Package for Statistical Analysis of High Energy Physics Data, physics/0507143 (2005).

[30] J. Gaiser, Ph.D. Thesis, SLAC-R-255 (1982).

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

[32] R. Seitz, "Describing Energy Deposits in CsI Crystals", *BABAR Note* #294 (1996).

[33] T. Sjostrand, Computer Physics Commun. **82**, 74 (1994).

[34] E. Fermi, Prog. Theor. Phys., **Vol. 5** 570 (1950).

[35] C. Quigg and J. L. Rosner, Phys. Rev. **D17** 239 (1978).

[36] R.K. Pathria, *Statistical Mechanics*, (1996).

[37] *Tracking Efficiency Task Force*,
http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TauEff/R24/TauEff.html

[38] *Neutrals Group*,
http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals/validation/recipie18.html

[39] See equation 1 in B. Aubert *et al.*, Phys. Rev. **D78** 071103 (2008).

[40] R. Barbieri, P. Lodone and D. Straub, JHEP **2011** 1 (2011).