Measurement of the Hadronic Mass Spectrum in B to Xulnu Decays and Determination of the b-Quark Mass at the BaBar Experiment

Kerstin Tackmann

SLAC-R-901

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

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

SLAC-R-901

Measurement of the Hadronic Mass Spectrum in $B \to X_u \ell \nu$ Decays and Determination of the b-Quark Mass at the *BABAR* Experiment

by

Kerstin Tackmann

Diplom (Technische Universität Dresden) 2004

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

Doctor of Philosophy
in

Physics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Marco Battaglia, Chair
Professor Hitoshi Murayama
Professor Alex Filippenko

Spring 2008

Work supported [in part] by U.S. Department of Energy contract DE-AC02-76SF00515

Abstract

Measurement of the Hadronic Mass Spectrum in $B \to X_u \ell \nu$ Decays
and Determination of the b-Quark Mass at the *BABAR* Experiment

by

Kerstin Tackmann

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Marco Battaglia, Chair

I present preliminary results of the measurement of the hadronic mass spectrum and its first three spectral moments in inclusive charmless semileptonic B-meson decays. The truncated hadronic mass moments are used for the first determination of the b-quark mass and the nonperturbative parameters μ_π^2 and ρ_D^3 in this B-meson decay channel. The study is based on 383×10^6 $B\bar{B}$ decays collected with the *BABAR* experiment at the PEP-II e^+e^- storage rings, located at the Stanford Linear Accelerator Center. The first, second central, and third central hadronic mass moment with a cut on the hadronic mass $m_X^2 < 6.4\,\mathrm{GeV}^2$ and the lepton momentum $p^* > 1\,\mathrm{GeV}$ are measured to be

$$M_1 = (1.96 \pm 0.34_{\mathrm{stat}} \pm 0.53_{\mathrm{syst}})\,\mathrm{GeV}^2$$
$$U_2 = (1.92 \pm 0.59_{\mathrm{stat}} \pm 0.87_{\mathrm{syst}})\,\mathrm{GeV}^4$$
$$U_3 = (1.79 \pm 0.62_{\mathrm{stat}} \pm 0.78_{\mathrm{syst}})\,\mathrm{GeV}^6,$$

with correlation coefficients $\rho_{12} = 0.99$, $\rho_{23} = 0.94$, and $\rho_{13} = 0.88$, respectively. Using

Heavy Quark Effective Theory-based predictions in the kinetic scheme we extract

$$m_b = (4.60 \pm 0.13_{\rm stat} \pm 0.19_{\rm syst} \pm 0.10_{\rm theo})\,{\rm GeV}$$
$$\mu_\pi^2 = (0.40 \pm 0.14_{\rm stat} \pm 0.20_{\rm syst} \pm 0.04_{\rm theo})\,{\rm GeV}^2$$
$$\rho_D^3 = (0.10 \pm 0.02_{\rm stat} \pm 0.02_{\rm syst} \pm 0.07_{\rm theo})\,{\rm GeV}^3,$$

at $\mu = 1\,{\rm GeV}$, with correlation coefficients $\rho_{m_b\mu_\pi^2} = -0.99$, $\rho_{\mu_\pi^2\rho_D^3} = 0.57$, and $\rho_{m_b\rho_D^3} = -0.59$. The results are in good agreement with earlier determinations in inclusive charmed semileptonic and radiative penguin B-meson decays and have a similar accuracy. Through the comparison of this result to those obtained in other channels, this provides a test of the Heavy Quark Effective Theory predictions that are used for the determination of $|V_{ub}|$. In addition, the measured dependence of the hadronic mass moments on the hadronic mass cut is compared to the dependence predicted by Heavy Quark Effective Theory.

Professor Marco Battaglia
Dissertation Committee Chair

Contents

Acknowledgements

I would like to thank –

Marco Battaglia, my advisor, for his mentoring, guidance and support during the past four years.

Heiko Lacker, for a very pleasant collaboration on this analysis.

The LBNL *BABAR* group, and especially Dave Brown, for his support and mentoring, and for what I learned about tracking and software design while being part of the TrkFixup team.

Bob Cahn for his physics lectures for *BABAR* students and his encouragment to not lose sight of physics outside of my analysis.

The *BABAR* collaborators from the inclusive $|V_{ub}|$ analysis with hadronic tags, Roberto Sacco, Concezio Bozzi, Antonio Petrella, Virginia Azzolini and Nicola Gagliardi.

Paolo Gambino and Paolo Giordano, for our interactions about their HQET calculations, and their explanations and advice.

The members of my qualifying exam and thesis committee, Alex Filippenko, Bob Jacobsen, and Hitoshi Murayama. My thanks also goes to those who proof-read parts of this thesis and gave valuable suggestions, Frank, Jens, Dave, Thera, and my brother Björn, and to Onsi, for his help with efficiently labeling some of my figures.

Anne and Donna, who are the most wonderful and patient grad student affairs officers I can imagine.

Frank, for his patience, encouragement, and explanations of particle physics theory, but not less for his help with typing and everyday things when RSI substantially limited the use of my hands.

Our families, for their support and for enduring us living on a different continent.

My officemates in 2177. Working here would not have been the same without the mutual help, advice, and chats on physics and non-physics issues. Thanks, Toyoko, Gena, Lindsay, Tomohiko, Aritoki, and Ben!

My friends here in Berkeley and in Germany, for their encouragment and support, and for the time we spent together, especially Tiff, Onsi, Xiaosheng, Hannah, Anna, Mark, Thera, Dan, and Charina, among many others, and Carrie, for her great humor. Their friendship is truly a blessing.

Chapter 1

Introduction

The Standard Model (SM) of particle physics describes the known elementary particles as well as their interactions through the electromagnetic, weak, and strong forces. The set of known elementary particles contains three generations of fermions, leptons, and quarks, along with 12 vector bosons, which mediate the forces. A scalar boson, the Higgs boson, thought to be responsible for the generation of mass, has not yet been discovered.

Hadrons are systems of quarks and antiquarks, bound by the strong interaction in the low-energy regime below 1 GeV. This introduces effects that cannot be treated perturbatively. The quark masses, which are fundamental parameters of the Standard Model (SM), exhibit a strong hierarchy: While u, d, and s quarks have masses of the order of a few MeV and around 100 MeV, where the strong interaction is nonperturbative, c and b quarks have masses of a few GeV and the t quark mass is around 175 GeV. For B mesons, which contain a heavy b and a light $\bar{u}$ or $\bar{d}$ quark, predictions can be made by Heavy Quark Effective Theory (HQET), which exploits the large b-quark mass m_b to expand Quantum Chromodynamics (QCD) in powers of $1/m_b$. Nonperturbative effects at $\mathcal{O}(1/m_b^2)$ and $\mathcal{O}(1/m_b^3)$ are encoded in the parameters μ_π^2, μ_G^2 and ρ_D^3, ρ_{LS}^3 (see Sec. 2.3), which cannot be computed from first principles and have to be either determined experimentally or taken from theoretical models.

The b-quark mass along with the nonperturbative parameters can be extracted from

differential decay distributions in inclusive semileptonic B meson decays, $B \to X\ell\nu$ and radiative B meson decays, $B \to X_s\gamma$. Current determinations of m_b in B meson decays have an uncertainty of $0.6-1.16\%$ [1, 2, 3, 4] and are among the most precise determinations of the b-quark mass. Apart from updates of these results with data from the *BABAR* and Belle experiments, these measurements may not be repeated in the coming years. The LHC experiments will not be able to perform the inclusive reconstruction of B meson decays with sufficient accuracy. Precise measurements of the b-quark mass will, however, be important for the study of the Higgs sector at future experiments.

The *BABAR* experiment, located at the PEP-II e^+e^- storage rings at the Stanford Linear Accelerator Center (SLAC), is designed to study the mixing and decays of B mesons. e^+e^- collisions provide a clean experimental evironment: The initial state kinematics is fully known and can be exploited to kinematically constrain the final state. In addition, there are no underlying events from soft interactions that obscure the interesting physics events. PEP-II delivers a high luminosity which provides the large datasets needed for precision measurements and studies of rare B-meson decays.

In the SM, the quark mass eigenstates and the quark eigenstates of the weak interaction are related through the Cabibbo-Kobayashi-Maskawa (CKM) matrix [5, 6], which gives rise to flavor-changing charged currents and is the source of charge-parity (CP) violation in the quark sector. Cabibbo introduced the weak quark-mixing angle θ_C in 1963 to explain the smallness of weak decay rates for particles with strangeness [5]. Kobayashi and Maskawa showed in 1973 that a quark mixing matrix for three generations of quarks, of which only two had been seen at that time, could also give rise to CP violation [6]. The three-generation CKM-matrix accounts for a large variety of observations [7]. It not only describes the CP violation observed in the K and B systems, but also explains the very different mixing frequencies observed in different neutral meson systems and gives precise predictions of rare K and B meson decays. One of the main goals of the *BABAR* experiment is the measurement of the CKM-matrix parameters in the B meson system and provide overconstraining measurements to test the CKM picture.

The determination of the different CKM matrix elements uses very different techniques,

both experimentally and theoretically [8]. The matrix element $|V_{ud}|$, which describes the strength of the coupling between the u and the d quark, is most precisely measured in super-allowed 0^+ to 0^+ nuclear β decays and is determined with a precision of 0.03%, dominated by theoretical uncertainties. Other measurements of $|V_{ud}|$ are made from the neutron lifetime and from $\pi^+ \to \pi^0 e\nu$ decays. The element $|V_{us}|$, which determines the coupling between the first-generation u and the second-generation s quark, has traditionally been extracted from semileptonic K decays, $K \to \pi\ell\nu$. The uncertainty of about 1% is dominated by the uncertainty on the decay form factor from lattice QCD. The most precise determination of $|V_{cd}|$ uses the difference of the ratio of double-muon to single-muon production by neutrino and antineutrino beams, which is proportional to the charm cross section for scattering off valence d quarks. The relative uncertainty on $|V_{cd}|$ from this method is on the order of 5%. The value of $|V_{cs}|$ can be determined at the 10% level from semileptonic $D \to K\ell\nu$, the form factors are predicted by lattice QCD. Semileptonic B-meson decays allow for a measurement of the two matrix elements $|V_{cb}|$ and $|V_{ub}|$, which describe the relative strenths of the couplings between transitions of the third-generation b quark into a second-generation c quark and a first-generation u quark. Current uncertainties are of the order of 2% and 8%, respectively. This class of decays is studied in this thesis. The determination of $|V_{td}|$ and $|V_{ts}|$ relies on loop-mediated neutral $B_{(s)}\bar{B}_{(s)}$ oscillations and loop-mediated rare B and K decays. Current uncertainties are around 10% and 7%, respectively. The ratio $|V_{td}|/|V_{ts}|$ can be determined more precisely, to about 3%, after the oscillations in the $B_s\bar{B}_s$ system have recently been observed by the CDF experiment [9]. Finally, $|V_{tb}|$ can be constrained from the ratio of branching fractions from $t \to Wb$ to $t \to Wq$ ($q = d, s, b$) decays.

One class of measurements extracts $|V_{ub}|$ from the inclusive $B \to X_u\ell\nu$ branching fraction, where the lepton ℓ is an electron or a muon and effectively all possible hadronic states X_u are summed over. In general, semileptonic decays provide a good environment to study the electroweak and strong interactions. Experimentally, electrons and muons can be identified with good accuracy. Theoretically, the computation of decays involving leptons is in general less complicated than the computation of analoguous processes involving only hadrons.

The matrix element $|V_{ub}|$ is about 10 times smaller than $|V_{cb}|$ [8]. Consequently, kinematic cuts are employed to reduce the dominant $B \to X_c \ell\nu$ background. While the dependence of $|V_{ub}|$ on the total $B \to X_u \ell\nu$ branching fraction scales as m_b^5, these kinematic cuts amplify the dependence on m_b by several powers. Hence, a good knowledge of m_b is necessary to reduce the uncertainties on $|V_{ub}|$, and presently, the uncertainty on m_b introduces a sizable part of the total uncertainty on $|V_{ub}|$. The b-quark mass as well as the nonperturbative parameters can be extracted from differential decay distributions in $B \to X_c \ell\nu$ and $B \to X_s \gamma$ decays. Yet, the kinematic cuts employed in the analysis of $B \to X_u \ell\nu$ decays not only amplify the dependence of $|V_{ub}|$ on m_b, but also lead to a break-down of the expansion in $1/m_b$ in the region of phase space that is used in the measurement, which requires the use of a modified expansion. The validity of this modified expansion is, however, not tested in $B \to X_c \ell\nu$ decays.

In this thesis, we present a preliminary result for the first determination of m_b and of nonperturbative parameters in $B \to X_u \ell\nu$ decays from the inclusive hadronic mass distribution. Inclusive $B \to X \ell\nu$ events are selected on the recoil of fully reconstructed B_{reco} meson decays into hadronic final states, $B_{\text{reco}} \to D^{(*)} Y$, and the invariant mass of the hadronic system X, m_X, is measured. The combinatorial and continuum background from the B_{reco} reconstruction is subtracted by a fit in each bin of m_X and m_X^2. To obtain the measured inclusive $B \to X_u \ell\nu$ m_X and m_X^2 spectra, veto cuts are applied to suppress $B \to X_c \ell\nu$ backgrounds, and the remaining $B \to X_c \ell\nu$ and other backgrounds are subtracted by a fit to the m_X spectrum. The measured m_X and m_X^2 spectra are then corrected for resolution and acceptance effects. We also determine the first, second central, and third central moment of the unfolded spectrum for different cuts on m_X^2. A fit is performed to the moments of the m_X^2 spectrum in the framework of the Heavy Quark Expansion (HQE) to extract the b-quark mass m_b. This allows us to perform an important test of the HQET framework by comparing the extracted value of m_b to that obtained using spectral moments in $B \to X_c \ell\nu$ decays. As an extension of this, the breakdown of the unmodified HQE can be tested by extracting m_X^2 moments with varying cuts on m_X^2 and comparing to the predicted

cut dependence. In the future, the unfolded hadronic mass spectrum can also be used to constrain the functional form of the shape function [10, 11, 12, 13].

This thesis is organized as follows: Chapter 2 gives an introduction to the physics of the CKM matrix and to the description of semileptonic B meson decays in HQET. It also gives a short historical overview about earlier determinations of the b-quark mass. Chapter 3 is an overview of the *BABAR* experiment, including brief descriptions of the different subdetectors, the reconstruction of charged particles at *BABAR*, along with the recoil method used in this analysis. The measurement of the hadronic mass spectrum is treated in Chapter 4, and Chapter 5 deals with the unfolding technique and the unfolding and extraction of the spectral moments. Chapter 6 describes the fit to the mass moments and discusses the results. Finally, Chapter 7 contains the conclusions.

Chapter 2

The CKM-Matrix Element V_{ub} and the b-Quark Mass

This chapter gives an introduction to the determination of $|V_{ub}|$ and the b-quark mass and the underlying theoretical concepts, generally following Refs. [14, 15].

Sec. 2.1 briefly introduces the Standard Model of Particle Physics, with an emphasis on those aspects that are directly relevant to this thesis. Sec. 2.2 gives an overview over common renormalization schemes in B physics and different determinations of the b-quark mass. Finally, Sec. 2.3 provides an introduction to the theoretical framework that we use to extract the b-quark mass from the measured moments.

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of Particle Physics describes all known elementary particles and their interactions. While matter is made up of spin-$\frac{1}{2}$ fermions, interactions are mediated by spin-1 bosons. We know of three generations of fermions: the charged e, μ, τ along with the corresponding neutral neutrinos, ν_e, ν_μ, ν_τ in the leptonic sector, and three up-type quarks, u, c, t, and down-type quarks, d, s, b, in the quark sector. Leptons only

carry electroweak charge and hence participate only in electroweak interactions; quarks also carry color charge and participate in electroweak and strong interactions.

The electroweak and strong forces are mediated by 12 gauge bosons. The $W^{\pm}$ and Z bosons couple to the weak charge, the photon, γ, mediates the electromagnetic force, and 8 gluons couple to the color charge.

2.1.1 Electroweak Symmetry Breaking and Quark Masses

In the SM, the electroweak interactions are described by an $SU(2) \times U(1)_Y$ gauge group. Explicit mass terms for fermions and bosons are absent, since they would violate gauge invariance. The Glashow-Weinberg-Salam theory introduces a scalar field which spontaneously breaks the $SU(2) \times U(1)_Y$ gauge symmetry in order to create mass terms for the gauge bosons and fermions by the so-called Higgs mechanism. The $SU(2) \times U(1)_Y$ symmetry is broken to $U(1)_{\rm em}$.

The scalar Higgs field $\phi = (\phi^+, \phi^0)$ is in the fundamental representation of $SU(2)$, and is assigned a charge $Y = 1/2$ under the $U(1)_Y$ symmetry. The potential of ϕ is

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda^2 (\phi^\dagger \phi)^2, \tag{2.1}$$

so that ϕ acquires a vacuum expectation value $\langle \phi \rangle = v/\sqrt{2}$ with $v = \mu/\lambda$. The gauge can be fixed such that the ground state is $\phi = (0, v/\sqrt{(2)})$. The gauge bosons are coupled to the Higgs through the covariant derivative, $D^{\rm ew}_\mu \phi$. Their mass terms arise when $D^{\rm ew}_\mu \phi$ is evaluated at vacuum expectation value of ϕ,

$$\mathcal{L}_{GM} = \frac{g^2 v^2}{8} \big(W^{+\mu} W^+_\mu + W^{-\mu} W^-_\mu \big) + \frac{v^2}{8} \big(g' B^\mu + g W^\mu_3 \big) \big(g' B_\mu + g W_{3\mu} \big). \tag{2.2}$$

Here, g and g' are the $SU(2)$ and $U(1)_Y$ coupling constants, respectively, $W^{\pm}$ the charged $SU(2)$ vector bosons, and W_3 and B the neutral vector bosons of $SU(2)$ and $U(1)_Y$. The generator of the electromagnetic $U(1)_{\rm em}$, which is a linear combination of generators of $SU(2)$ and $U(1)_Y$, leaves $\langle \phi \rangle$ invariant, and hence the photon, $\gamma = \sin(\theta_W) W_3 + \cos(\theta_W) B$, does not acquire a mass. The weak mixing angle θ_W is given by

$$\sin(\theta_W) = \frac{g'}{\sqrt{g'^2 + g^2}}, \quad \cos(\theta_W) = \frac{g}{\sqrt{g'^2 + g^2}}. \tag{2.3}$$

The mass and charge eigenstates of the remaining gauge bosons are the $W^{\pm}$ and the neutral Z. Their masses are $M_W = gv/2$ and $M_Z = M_W/\cos(\theta_W)$. The Higgs' mass is predicted to be $m_H = \sqrt{2}\mu$ at tree level, but receives non-negligible radiative corrections.

The fermions acquire masses via gauge invariant Yukawa couplings to the Higgs field,

$$\mathcal{L}_{\text{Yuk}} = \lambda_u^{ij} \bar{u}_R^i \phi^T \epsilon Q_L^j - \lambda_d^{ij} \bar{d}_R^i \phi^\dagger Q_L^j - \lambda_e^{ij} \bar{e}_R^i \phi^\dagger L_L^j + \text{h.c.}, \tag{2.4}$$

where h.c. denotes the hermitian conjugate, and the repeated generation indices i, j are summed over. The $\bar{u}_R^i$, $\bar{d}_R^i$, and $\bar{e}_R^i$ denote the conjugate right-handed fermion fields, which transform as singlets under $SU(2)$. The Q_L^j and L_L^j denote the left-handed fermion $SU(2)$ doublets. The Yukawa couplings are given by $\lambda_{u,d,e}^{ij}$, and ϵ is the antisymmetric matrix

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.5}$$

Since right-handed neutrinos do not exist in the SM, there is no Yukawa interaction for the neutrino fields.

Evaluating (2.4) at the vacuum expectation value $\langle \phi \rangle = v/\sqrt{2}$ gives rise to 3×3 mass matrices for the quarks and charged leptons,

$$M_u = \frac{v}{\sqrt{2}} \lambda_u, \quad M_d = \frac{v}{\sqrt{2}} \lambda_d, \quad M_e = \frac{v}{\sqrt{2}} \lambda_e, \tag{2.6}$$

which are in general non-diagonal. They can be diagonalized by unitary transformations, which transform the eigenstates of the weak interaction to the propagating mass eigenstates. Hence, in the SM, the masses of quarks and charged leptons depend on the expectation value of the Higgs field, as well as on the Yukawa couplings between the Higgs and the fermions.

Testing the Higgs Couplings

If a candidate for the Higgs boson will be found in future experiments, one requirement for it to be identified as the SM Higgs will be that its couplings to fermions and gauge bosons are indeed proportional to their masses. In extended models, the masses of the SM particles can receive contributions from more than one source, e.g., from couplings to more

than one Higgs doublet. Furthermore, the SM-like Higgs boson in extended models often is an admixture of states from several scalar particles, so that additional mixing angles enter the couplings to the SM particles. Precise measurements of the heavy quark, lepton, and gauge boson masses are needed for a sensitive comparison of the predicted SM couplings to the measured couplings. For example, at a $(0.5 - 1)\,\mathrm{TeV}$ e^+e^- collider, the current uncertainty on m_b will significantly reduce the volume of parameter space in which the SM can be distinguished from a supersymmetric model [16].

Experimentally, the accessable final states f in Higgs decays $H \to f\bar{f}$ depend on the mass of the Higgs boson [17]. For a smaller Higgs mass, the Higgs couplings to the heavier fermions, the b and, with larger uncertainty, c quark and the τ can be measured. For a larger Higgs mass, the couplings to the W and Z can be determined from Higgs decays to gauge bosons, while Higgs production in Higgsstrahlung and vector boson fusion can also be used for a smaller Higgs mass. For an intermediate Higgs mass, the Higgs coupling to t quarks can be determined from the Higgsstrahlung process $pp \to t\bar{t}H$, where the Higgs is radiated from the top quark. Yet, this measurement is difficult due to the small cross section and the complex topology of the $b\bar{b}b\bar{b}WW$ final state.

While the ATLAS and CMS experiments at the Large Hadron Collider (LHC) will only be able to determine ratios of Higgs couplings from $H \to f\bar{f}$ decays with an interesting accuracy due to uncertainties in the Higgs production cross sections, a future e^+e^- collider would allow for an absolute determination of these Higgs couplings.

The Cabbibo-Kobayashi-Maskawa (CKM) Matrix

The diagonalization of the mass matrices is achieved by separate unitary transformations on the left-handed and right-handed quark and lepton fields,

$$\begin{aligned} u_L &= L_u u'_L, \quad u_R = R_u u'_R \\ d_L &= L_d d'_L, \quad d_R = R_d d'_R \\ e_L &= L_e e'_L, \quad e_R = R_e e'_R, \end{aligned} \tag{2.7}$$

to yield real, diagonal non-negative mass matrices. The primed fields represent the mass eigenstates, as opposed to the unprimed flavor eigenstates, the eigenstates of the weak interaction. In general, the transformation of the u_L and d_L fields will be different, and the left-handed quark doublet Q_L can be rewritten as

$$\begin{pmatrix} u_L \\ \\ d_L \end{pmatrix} = \begin{pmatrix} L_u u'_L \\ \\ L_d d'_L \end{pmatrix} = L_u \begin{pmatrix} u'_L \\ \\ V_{\mathrm{CKM}} d'_L \end{pmatrix}. \tag{2.8}$$

The CKM matrix is given by $V_{\mathrm{CKM}} = L_u^\dagger L_d$. It enters the left-handed quark couplings to the charged gauge bosons,

$$W^+ \bar{u}_L \gamma^\mu d_L = W^+ \bar{u}'_L \gamma^\mu V_{\mathrm{CKM}} d'_L, \tag{2.9}$$

if they are written in terms of the mass eigenstates. Thus, a non-diagonal CKM matrix introduces flavor-changing charged currents. The kinetic energy terms are invariant under the transformations (2.7), and so are the couplings to the photon and Z field, as each single term involves only fields of the same handedness. Hence, there are no flavor-changing neutral currents at tree level in the SM.

Since the CKM matrix describes the tranformation between the weak eigenstates and the mass eigenstates, it is a 3×3 unitary matrix with nine real parameters. Five parameters can be eliminated by phase redefinitions of the quark fields, and the remaining four parameters are three angles and one complex phase. The latter, if nonzero, gives rise to CP violating interactions in the Standard Model. The CKM matrix can be written as a product of three complex rotation matrices [18]

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ \\ V_{cd} & V_{cs} & V_{cb} \\ \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{\mathrm{i}\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{\mathrm{i}\delta} & s_{23}c_{13} \\ \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{\mathrm{i}\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{\mathrm{i}\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.10a}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for $i, j = 1, 2, 3$.

The hierarchy found among the matrix elements is exploited in the Wolfenstein

parametrization of the CKM matrix, which uses the four parameters λ, A, ρ, and η,

$$V_{\rm CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho-\mathrm{i}\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-\mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{2.10b}$$

to expand the matrix in powers of $\lambda \approx 0.22$ [19].

Rather than expanding, one can define to all orders [20]

$$\begin{aligned} s_{12} &= \lambda \\ s_{23} &= A\lambda^2 \\ s_{13}e^{-\mathrm{i}\delta} &= A\lambda^3(\rho-\mathrm{i}\eta). \end{aligned} \tag{2.11}$$

The unitarity relations between rows and columns of the CKM matrix define triangles in the complex ρ-η plane. The different triangles have identical area, the size of which is a convention-independent measure for the size of CP violation in the SM [21]. The unitarity relation between the first and third column, which describes the B meson sector, yields a triangle with similarly-sized sides and hence predicts large observable CP violation effects in this system. After rotation and rescaling of the triangle, the apex is given by

$$\bar{\rho}+\mathrm{i}\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}. \tag{2.12}$$

Independent measurements provide constraints on the parameters of the triangle, and a fit can be used to test the consistency of the measured data with the CKM picture.

Fig. 2.1 shows a current fit in the $\bar{\rho}$-$\bar{\eta}$ plane. Shown are the constraints from semileptonic B decays ($|V_{ub}|$ and $|V_{cb}|$), $B^0\bar{B}^0$ mixing (Δm_d), $B_s^0\bar{B}_s^0$ mixing (Δm_s), CP violation in the neutral K system (ε_K), $B \to J/\Psi K^{(*)0}$ decays ($\sin(2\beta)$ and $\cos(2\beta)$), $B \to \pi\pi, \rho\rho, \rho\pi$ decays (α), and $B \to D^{(*)}K^{(*)}$ decays (γ), but a larger number of measurements enters the fit [22]. The different constraints are very consistent and the CKM picture is found to work remarkably well. The current number and precision of measurements allow for cross checks using only inputs dominated by New Physics-insensitive tree-level decays and New Physics-sensitive loop processes, and good agreement is found between the two fits.

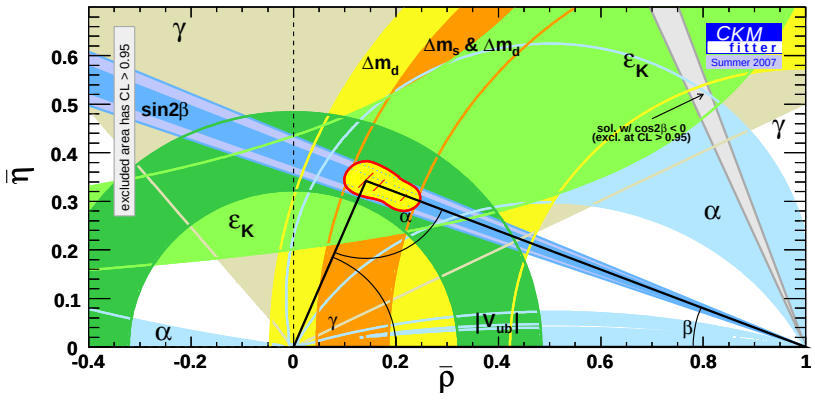

Figure 2.1. Fit to constraints in the $\bar{\rho}$-$\bar{\eta}$ plane [22].

A more precise measurement of $|V_{ub}|$ can significantly improve the current constraint on the apex of the triangle and is particularly sensitive to the consistency between the measurements of $\sin(2\beta)$ and the constraints obtained from the fit to $|V_{ub}|$ and $|V_{cb}|$.

2.1.2 Strong Interactions

In the SM, the strong interactions are described by quantum chromodynamics (QCD), which is an $SU(3)$ gauge theory.

Renormalization, that is, the removal of formal divergences arising in higher-order loop graphs, introduces a dependence on the renormalization scale μ into the renormalized coupling constant $g(\mu)$. This scale dependence is such that the QCD fine structure constant $\alpha_s(\mu) = g(\mu)^2/(4\pi)$ decreases with increasing energies,

$$\alpha_s(\mu) = \frac{12\pi}{(33 - 2n_f)\ln(\mu^2/\Lambda^2_{\rm QCD})}. \tag{2.13}$$

The number of quark flavors is given by n_f. The small energy scale where $\alpha_s(\mu)$ formally diverges defines the scale $\Lambda_{\rm QCD}$, which experimentally is of the order of a few hundred MeV. The calculation of the scale dependence is itself carried out in perturbation theory in α_s. The divergence of α_s at $\Lambda_{\rm QCD}$ is thus merely formal, since perturbation theory is no longer valid when α_s is large. Nevertheless, $\Lambda_{\rm QCD}$ can be viewed as the scale where α_s gets large, and nonperturbative effects become important. One consequence of QCD becoming strongly coupled at low energies is the non-existence of free quarks. Their properties have to

be inferred from hadronic properties with the help of effective theories or theoretical models. In the case of the b-quark mass, this introduces non-negligible theoretical uncertainties.

2.2 The b-Quark Mass

A variety of methods has been used to determine the b-quark mass both from the $b\bar{b}$-system, as well as from B meson properties. This section gives a brief overview over common renormalization schemes in B physics and some of the methods for determining m_b, following Refs. [23, 24, 25].

Each method involves its own approximations and uncertainties and in most cases, uncertainties are introduced by neglected higher-order corrections and hence are difficult to quantify. In some cases, determinations using the same method can arrive at substantially different error estimates.

2.2.1 Mass Schemes

The mass of the b quark is a fundamental parameter of the SM. Yet, due to the confinement of the b quark into hadrons, its mass is difficult to determine experimentally. Since the mass cannot be measured directly, it has to be inferred from hadron masses or other hadronic properties.

As a parameter in the Lagrangian, quark masses have to be renormalized, and the renormalized quantities are scheme- and scale-dependent. In principle, any renormalization condition can be used to define the renormalized mass, but some are better suited for B physics than others.

The Pole Scheme

The pole mass m_b^{pole} is defined as the position of the pole in the quark propagator, i.e., as the solution to[1]

$$\not{p} - m - \Sigma(\not{p}, m)|_{m=m_b^{\text{pole}}} = 0, \tag{2.14}$$

where $\Sigma(p, m)$ is the self-energy of the b quark and m is the bare mass. As a property of an unphysical on-shell quark, the pole mass is particularly sensitive to infrared physics: It suffers from a renormalon ambiguity (an ambiguity of order Λ_{QCD}), which arises from the low-momentum region of loop-integrals where QCD is strongly coupled. This results in ill-converging perturbative expansions when using the pole scheme.

The $\overline{\text{MS}}$ Scheme

The most commen short-distance mass is the $\overline{\text{MS}}$ mass, $\overline{m}_b(\mu)$. It is defined by regulating QCD with dimensional regularization and subtracting the divergences in the $\overline{\text{MS}}$ (modified minimal subtraction) scheme at the scale μ. One finds [24]

$$\frac{m_b^{pole}}{\overline{m}_b(\overline{m}_b)} = 1 + \frac{4\bar{\alpha}_s}{3\pi} + \left(\frac{\bar{\alpha}_s}{\pi}\right)^2 \left(13.44 - 1.04\, n_f\right) + \left(\frac{\bar{\alpha}_s}{\pi}\right)^3 \left(190.8 - 26.7\, n_f + 0.65\, n_f^2\right) + \dots , \tag{2.15}$$

where n_f is the number of lighter quark flavors and $\bar{\alpha}_s = \alpha_s(\overline{m}_b(\mu))$. The scale μ is typically chosen to be of the order of the characteristic energy scale Q of the process considered to resum logarithms of the form $\bar{\alpha}_s^m(\mu) \log^m(Q/\mu)$, which would otherwise be large.

The $\overline{\text{MS}}$ scheme arises naturally in processes at high energies, where the b quark is far off-shell. It treats the b quark as a fully dynamical QCD field. This is less useful in low-energy processes, where the b quark is nonrelativistic and the heavy degrees of freedom are integrated out.

The Kinetic Scheme

The shortcomings of both the pole and the $\overline{\text{MS}}$ masses can be resolved by so-called threshold masses, which have no renormalon ambiguity and have well-behaved perturbative

[1] For ease of notation, $c = 1$ will be used throughout this thesis.

expansions in the treatment of nonrelativistic heavy quarks. One of the popular threshold masses is the kinetic mass, $m_b^{\rm kin}(\mu)$. The kinetic scheme uses a Wilsonian hard cutoff at the scale μ to separate short- and long-distance contributions to the renormalization group running. Contributions from scales smaller than μ are absorbed into the definition of the mass and for $\mu \to 0$, the pole mass is regained. The nonperturbative parameters also acquire a μ dependence. The relationship between the kinetic mass and the pole mass is given by [24]

$$m_b^{\rm kin}(\mu) = m_b^{\rm pole} - \left[\bar{\Lambda}(\mu)\right]_{\rm pert} - \left[\frac{\mu_\pi^2(\mu)}{2m_b^{\rm kin}(\mu)}\right]_{\rm pert} + \ldots, \tag{2.16}$$

where $\left[\bar{\Lambda}(\mu)\right]_{\rm pert}$ and $\left[\bar{\mu_\pi^2}(\mu)\right]_{\rm pert}$ are perturbative evaluations of Heavy Quark Effective Theory (HQET, see Sec. 2.3.1) matrix elements that describe the difference between the pole and the B-meson mass. The factorization scale is commonly chosen to be $\mu = 1\,\text{GeV}$. In the following, we will drop the "kin" superscript, and all b-quark masses and nonperturbative parameters are understood to be in the kinetic scheme, unless labeled otherwise.

The $1S$ Scheme

Another commonly used threshold mass is the $1S$ mass, which, in contrast to other threshold masses, does not depend on an explicit subtraction scale. The $1S$ mass, $m_b^{\rm 1S}$, is defined as one half of the perturbative contribution to the mass of the $\Upsilon(1S)$ state in the limit that $m_b \gg m_b v \gg m_b v^2 \gg \Lambda_{\rm QCD}$, and is related to the pole mass by [24]

$$\begin{aligned}\frac{m_b^{1S}}{m_b^{pole}} =& 1 - \frac{(C_F\alpha_s(\mu))^2}{8}\left\{1 + \left(\frac{\alpha_s(\mu)}{\pi}\right)\left[\beta_0\left(L+1\right) + \frac{a_1}{2}\right]\right. \\ &+ \left(\frac{\alpha_s(\mu)}{\pi}\right)^2\left[\beta_0^2\left(\frac{3}{4}L^2 + L + \frac{\zeta_3}{2} + \frac{\pi^2}{24} + \frac{1}{4}\right) + \beta_0\frac{a_1}{2}\left(\frac{3}{2}L + 1\right)\right. \\ &\left.\left. + \frac{\beta_1}{4}\left(L+1\right) + \frac{a_1^2}{16} + \frac{a_2}{8} + \left(C_A - \frac{C_F}{48}\right)C_F\pi^2\right]\right\},\end{aligned} \tag{2.17}$$

where $L \equiv \ln(\mu/(C_F\alpha_s(\mu)\, m_b^{pole}))$ and $\zeta_3 = 1.20206$. The expression for the $1S$ mass is derived in the framework of a nonrelativistic expansion, where powers of v arise as powers of α_s. In order to achieve the renormalon cancellation in this scheme, the so-called upsilon expansion is introduced, which formally treats terms of order α_s^{n+1} in Eq. (2.17) as being of order α_s^n.

2.2.2 Determination of the *b*-Quark Mass

Spectral Moments

Moments of $e^+e^- \to b\bar{b}X$ distributions have served for a large number of m_b measurements. Their calculation starts from the correlator of two electromagnetc currents involving b quarks,

$$(-g_{\mu\nu}q^2 + q_\mu q_\nu)\Pi(q^2) \equiv \mathrm{i}\int d^4x e^{\mathrm{i}q\cdot x}\langle T j^b_\mu(x) j^b_\nu(0)\rangle, \tag{2.18}$$

with $j^b_\mu(x) \equiv \bar{b}(x)\gamma_\mu b(x)$. In the complex q^2-plane, $\Pi(q^2)$ has poles at the locations of $b\bar{b}$ bound states and a cut along the real axis, which corresponds to the continuum. Since $\Pi(q^2)$ is an analytic function of q^2, the integral along the cut is related to the nth derivative of $\Pi(q^2)$ by

$$\left.\frac{d^n}{d(q^2)^n}\Pi(q^2)\right|_{q^2=0} = \frac{\pi}{n!}\int_0^\infty \frac{\mathrm{Im}\,\Pi(s)}{s^{n+1}}, \tag{2.19}$$

and the optical theorem relates the imaginary part of $\Pi(q^2)$ to the total cross section for $e^+e^- \to \gamma^* \to b\bar{b}X$,

$$R_b(s) = 12\pi Q_b^2\,\mathrm{Im}\,\Pi(q^2 = s + \mathrm{i}\epsilon), \tag{2.20}$$

with

$$R_b(s) = \frac{\sigma(e^+e^- \to \gamma^*(s) \to \bar{b} + X)}{\sigma_{\mathrm{pt}}}, \quad \sigma_{\mathrm{pt}} = \frac{4\pi\alpha^2_{\mathrm{QED}}(m_b)}{3s}, \; Q_b = -\frac{1}{3}e. \tag{2.21}$$

The sum rules give a relationship between measurable moments of the total cross section for $e^+e^- \to b\bar{b}$ pairs and derivatives of the vacuum polarization $\Pi(q^2)$ [23],

$$\frac{12\pi^2 Q_b^2}{n!}\left.\frac{d^n}{d(q^2)^n}\Pi(q^2)\right|_{q^2=0} = \int_0^\infty ds\,\frac{R_b(s)}{s^{n+1}}. \tag{2.22}$$

For dimensional reasons, the left-hand side is proportional to m_b^{-2n}, which allows for a determination of m_b.

In practice, only a certain range of values of n can be used. Since the measurements of $R_b(s)$ have substantial uncertainties in the continuum region, n must be large enough such that the moment is dominated by the better-measured first few Υ resonances. On the other hand, nonperturbative effects become increasingly important for the calculation of

the left-hand side at larger n, for which the sum rules are dominated by low-momentum states near threshold.

Two regions can be distinguished: In the large-n region, $4 \leq n \leq 10$, the $b\bar{b}$ dynamics is dominantly nonrelativistic, and in the small-n region, $1 \leq n \leq 4$, the $b\bar{b}$ dynamics is dominantly relativistic. While the large-n region suffers from larger theoretical and perturbative uncertainties, the small-n region is more sensitive to the poorly known parts of the $b\bar{b}$ cross section above the Υ resonances.

The Υ Mass Spectrum

Some of the earliest determinations of m_b were based on the analysis of the Υ mass spectrum. In the heavy-quark limit, the $b\bar{b}$ pair forms a nonrelativistic Coulomb bound state, which can be described by the Schrödinger equation. The difficulty arises in the determination of the nonperturbative corrections, and the resulting uncertainties dominate the total uncertainty on m_b. Studies based on potential models suggest, however, that in reality, the Υ system is not well described by a Coulomb bound state and recent analyses only use the $\Upsilon(1S)$ state.

3-Jet Events at the Z Pole

In e^+e^- collisions at energies around the Z mass, exclusive observables, such as multi-jet cross sections, exhibit a sizeable dependence on the quark masses. For the 3-jet rate, calculations exist at next-to-leading order. The b-quark mass can be determined from the normalized 3-jet rate of $Z \to b\bar{b}$ to $Z \to \ell\bar{\ell}$ ($\ell = u, d, s$) events, which is defined as [25]

$$R_3^{b\ell}(y_c) = \frac{R_3^b(y_c)}{R_3^\ell(y_c)} = \frac{\Gamma_{3j}^b(y_c)/\Gamma^b}{\Gamma_{3j}^\ell(y_c)/\Gamma^\ell}, \tag{2.23}$$

where $\Gamma_{3j}^q(y_c)$ and Γ^q are the 3-jet and total decay width of the Z into $q\bar{q}$ for $q = b$ or ℓ, and y_c is the jet resolution variable.

Since these determinations are performed at a much higher energy scale than the others, they also provide quantitative evidence for the running of m_b as predicted by QCD.

Lattice QCD

Lattice QCD calculations can be used to determine the masses of light and heavy quarks from the known hadron masses, as well as from the $b\bar{b}$ spectrum. They are based on a discretization of spacetime. Since the b-quark mass is larger than the UV cutoff of the calculations, which is given by the inverse lattice spacing, the b quark cannot be simulated directly on the lattice. Rather, an effective theory, such as HQET or NRQCD (nonrelativistic QCD), is used, in which the heavy degrees of freedom are integrated out. In this framework, the mass of the B meson, m_B, is given by [24]

$$m_B = m_b^{\text{pole}} + \epsilon - \delta m, \tag{2.24}$$

where ϵ is the binding energy and δm is a mass counterterm induced by radiative corrections. The relation is valid up to order $1/m_b^2$ corrections. The binding energy is obtained from numerical simulations on the lattice. While δm depends on the underlying effective theory, it is a short-distance quantity and can be calculated in peturbation theory.

Kinematic Moments

As explained in Sec. 2.3.2, m_b, along with a number of nonperturbative parameters, can be determined from moments of kinematic decay distributions in $B \to X\ell\nu$ decays. All previous measurements rely on $B \to X_c\ell\nu$ decays. In addition, determination can be obtained from photon energy moments in radiative penguin decays, $B \to X_s\gamma$. Recently, fits to data from several experiments have been carried out in the kinetic scheme [2] and the $1S$ scheme [1].

In semileptonic B decays, $B \to X_c\ell\nu$, moments of the lepton energy E_ℓ as well as the hadronic mass m_X are used. They are measured as a function of the lower cut on E_ℓ. At parton level, the hadronic mass is fixed at m_c, so that nonzero results for moments beyond first order provide a good probe of both nonperturbative and perturbative power corrections to the leading order tree-level results.

As a two-body decay on parton level, the photon energy in $B \to X_s\gamma$ decays is very sensitive to m_b and nonperturbative parameters, which makes it an interesting observable

despite the small branching fraction of this decay. At tree level, the first and second moments are directly sensitive to m_b and μ_π^2, which describes the kinetic energy of the b quark in the B meson.

2.3 Semileptonic B-Meson Decays

The B meson[2], as the lightest particle containing a b quark, is forced to decay weakly. We consider the semileptonic decay

$$B \to X_q \ell \nu, \tag{2.25}$$

where the final state consists of a lepton-neutrino pair and a hadronic system X_q, with q being either a c or a u quark. At parton level, the b ($\bar{b}$) quark emits a W^- (W^+) gauge boson and turns into a q ($\bar{q}$) quark $b \to q\ell^-\bar{\nu}_\ell$ ($\bar{b} \to \bar{q}\ell^+\nu_\ell$), introducing the dependence of the decay amplitude on the CKM-matrix element V_{qb}. This decay is shown in the left diagram in Fig. 2.2.

To good approximation, the second valence quark q' ($\bar{q}'$) inside the B ($\bar{B}$) meson, which is a d ($\bar{d}$) quark in case of neutral B^0 ($\bar{B}^0$) mesons and a u ($\bar{u}$) quark in case of charged B^+ (B^-) mesons, acts as a spectator of the decay process, illustrated in the right diagram in Fig. 2.2. Yet, if the dynamics of the decay is to be described appropriately, the nonperturbative effects that take place inside the B meson and within the final state X_q need to be taken into account. However, since there is no strong interaction between the hadronic system X_q and the lepton-neutrino pair, semileptonic B meson decays allow both for a clean determination of the CKM-matrix elements $|V_{qb}|$ and the study of the nonperturbative effects.

2.3.1 Differential Decay Rates

The B meson, as a system containing a heavy quark b, $m_b = 4.5$ to $5\,\mathrm{GeV} \gg \Lambda_{\mathrm{QCD}}$, and a light antiquark u or d with masses much smaller than Λ_{QCD}, can be dealt with in the framework of Heavy Quark Effective Theory (HQET). The large b-quark mass has two

[2]For ease of notation, we will usually refer to "B mesons" without specifying the flavor or charge. Charge and/or flavor will be specified only when necessary.

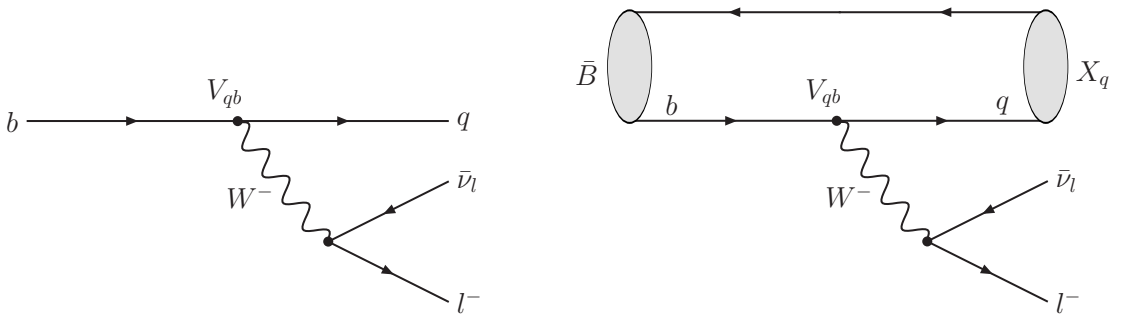

Figure 2.2. Semileptonic B-meson decay. The left Feynman diagram displays the parton level decay $b \to q\ell^-\bar{\nu}_\ell$. The right diagram displays the decay $\bar{B} \to X_q\ell^-\bar{\nu}_\ell$, where the "blobs" denote nonperturbative strong interactions.

important implications: Renormalizing QCD at a typical energy of m_b yields $\alpha_s(m_b) \approx 0.2$, and hence, perturbation theory in QCD is valid at this scale. Secondly, $\Lambda = \Lambda_{\rm QCD}/m_b \approx 0.1$[3] can be taken as a reasonable expansion parameter for nonperturbative effects. This is exploited by HQET, which performs a systematic expansion of QCD in powers of Λ and α_s.

There are two complementary approaches to semileptonic B decays. One can either be interested in exclusive decay channels, that is, in specific hadronic final states like in $B \to D^{(*)}\ell\nu$ or $B \to \pi\ell\nu$ decays. Or one can be interested in inclusive semileptonic decays. In the latter case, all possible hadronic states contribute. In particular, the X_q system can be a multi-particle state. In either case, the decay is described by the effective weak Hamiltonian density,

$$\mathcal{H}_W = \frac{4G_F}{\sqrt{2}} V_{qb}(\bar{q}\gamma^\mu P_L b)(\bar{\ell}\gamma_\mu P_L \nu). \tag{2.26}$$

Since the decay takes place at a typical scale of m_b, the W boson can be integrated out of the theory. The arising W propagator is dominated by the large W mass, $m_W = 80\,\text{GeV}$, and thus the interaction term can be replaced by the effective coupling $4G_F V_{bq}/\sqrt{2}$ and the four-fermion operator $(\bar{q}\gamma^\mu P_L b)(\bar{\ell}\gamma_\mu P_L \nu_\ell)$. The Fermi constant G_F is given by $G_F/\sqrt{2} = g^2/(8m_W^2)$, and $P_L = \frac{1}{2}(1-\gamma_5)$ is the projector on the left-handed parts of the spinors. In the following, the leptons are assumed to be massless, which is a good approximation for electrons and muons.

To obtain the differential decay rate for inclusive semileptonic decays, all possible

[3] In HQET, $\Lambda_{\rm QCD}$ is taken to be $m_B - m_b \approx 500\,\text{MeV}$.

hadronic final states are summed over, which includes the integration over the X_q phase space. The three independent kinematic variables can be taken to be the energies of the lepton and the neutrino, E_ℓ and E_ν, and the invariant mass of the lepton-neutrino system q^2 with $q = p_\ell + p_\nu$. The differential decay rate in the B rest frame is

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\mathrm{d}E_\ell\mathrm{d}E_\nu} = \frac{1}{2m_B}\int\frac{\mathrm{d}^3\vec{p}_\ell}{(2\pi)^3 2p_\ell^0}\int\frac{\mathrm{d}^3\vec{p}_\nu}{(2\pi)^3 2p_\nu^0}\delta(E_\ell - p_\ell^0)\delta(E_\nu - p_\nu^0)\delta\big(q^2 - (p_\ell + p_\nu)^2\big) \sum_{X_q}\sum_{s_\ell,s_\nu}|\big\langle X_q\ell\bar{\nu}\big|\mathcal{H}_W\big|\bar{B}\big\rangle|^2(2\pi)^4\delta^4(p_B - p_\ell - p_\nu - p_X), \tag{2.27a}$$

where s_ℓ (s_ν) denotes the spin of the lepton (neutrino). Four-momentum conservation is encoded in the four-dimensional δ function, and the other three δ functions project out the independent kinematic variables. The spins of the final state particles are summed over, which in case of the hadronic system is implicitely contained in the summation over X_q. Performing the phase space-integrations yields

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\mathrm{d}E_\ell\mathrm{d}E_\nu} &= \frac{1}{8m_B}\sum_{X_q}\sum_{s_\ell,s_\nu}|\langle X_q\ell\bar{\nu}|\mathcal{H}_W|\bar{B}\rangle|^2\delta^4(p_B - p_\ell - p_\nu - p_X) \\ &= \frac{2}{(2\pi)^3}G_F^2|V_{qb}|^2 W_{\alpha\beta}L^{\alpha\beta}.\end{aligned} \tag{2.27b}$$

In the second step, the matrix element is factored into a leptonic and a hadronic tensor, which is possible since no strong interactions occur between the two, while electroweak radiative corrections are suppressed by powers of G_F or the QED fine-structure constant α_{em}. This property makes semileptonic decays the preferred environment for the measurement of $|V_{qb}|$. The leptonic tensor is

$$L^{\alpha\beta} = \sum_{s_\ell,s_\nu}|\langle\ell\bar{\nu}|\bar{\ell}\gamma_\mu P_L\nu|0\rangle|^2 = 2(p_\ell^\alpha p_\nu^\beta + p_\ell^\beta p_\nu^\alpha - g^{\alpha\beta}p_\ell\cdot p_\nu - \mathrm{i}\varepsilon^{\alpha\beta}{}_{\kappa\lambda}\,p_\ell^\kappa p_\nu^\lambda). \tag{2.28}$$

The hadronic tensor is given by

$$W^{\alpha\beta} = \frac{1}{2m_B}\sum_{X_q}(2\pi)^3\delta^4(p_B - q - p_X)\big\langle\bar{B}(p_B)\big|J_L^{\dagger\alpha}\big|X_q(p_X)\big\rangle\big\langle X_q(p_X)\big|J_L^\beta\big|\bar{B}(p_B)\big\rangle, \tag{2.29}$$

where $J_L^\alpha = \bar{q}\gamma^\alpha P_L b$ is the left-handed quark current. The hadronic tensor contains all strong interactions relevant for the inclusive semileptonic B decay. It can be related to

the imaginary part of a time-ordered product of currents describing the forward scattering amplitude $B \to B$ via the optical theorem

$$-\frac{1}{\pi}\mathrm{Im}T^{\alpha\beta} = W^{\alpha\beta}, \tag{2.30a}$$

with

$$T^{\alpha\beta} = \frac{-i}{2m_B}\int \mathrm{d}^4x e^{-\mathrm{i}q\cdot x}\langle \bar{B}\big|T[J_L^{\dagger\alpha}(x)J_L^{\beta}(0)]\big|\bar{B}\rangle. \tag{2.30b}$$

2.3.2 Operator Product Expansion

The operator defined by Eq. (2.30b) is a non-local operator, that is, the fields in the currents are not located at a single spacetime point. For a time-ordered product of two local operatos, $T[O_a(z)O_b(0)]$, an operator product expansion (OPE) can be performed if the separation z is sufficiently small. The time-ordered product $T[O_a(z)O_b(0)]$ can be written as an expansion in local operators,

$$T[O_a(z)O_b(0)] = \sum_k C_{abk}(z)O_k(0), \tag{2.31a}$$

where the spacetime dependence is moved into the Wilson coefficients $C_{abk}(z)$. In the computation of matrix elements, $T[O_a(z)O_b(0)]$ can be replaced by $\sum_k C_{abk}(z)O_k(0)$, and the $C_{abk}(z)$ are independent of the matrix element computed, as long as the external momenta are small compared to the inverse of the separation, $1/z$. The momentum-space version of the OPE looks like

$$\int \mathrm{d}^4x e^{\mathrm{i}Q\cdot z}T[O_a(z)O_b(0)] = \sum_k C_{abk}(Q)O_k(0), \tag{2.31b}$$

and is valid for momenta Q much larger than the external momenta in the matrix elements. Since the Wilson coefficients $C_{abk}(Q)$ depend on the large momentum Q, $Q = m_b v - q \approx m_b \gg \Lambda_{\mathrm{QCD}}$ in the case of a decaying b quark, they can be computed perturbatively in QCD by taking matrix elements of both sides of Eq. (2.31b).

In the calculation of the differential decay rates, the OPE is applied to the operator in Eq. (2.30b). The coefficients $C_{abk}(Q)$ are computed by taking b quark and gluon matrix elements $\langle b|\cdot|b\rangle$ and $\langle bg|\cdot|b\rangle$. The operators in the OPE involve b-quark fields, covariant

derivatives D_μ, and the gluon field strength $G_{\mu\nu}$. The external b-quark momenta are written as $m_b v + k$, where the large b-quark mass is exploited to separate its momentum into two parts: Thinking of the b quark as residing inside a B meson, where the typical scale for momentum transfer among its components is of the order Λ_{QCD}, the b-quark velocity will only change by a small amount due to these nonperturbative interactions, $\Delta v \approx \Lambda_{\mathrm{QCD}}/m_b$. Its momentum can thus be split up into the momentum $m_b v$ arising from the B meson's motion with four-velocity v, and a residual momentum k of order Λ_{QCD}. Expanding the resulting matrix elements of the operator in Eq. (2.30b) in k/m_b thus yields an expansion in powers of the dimensionless $\Lambda = \Lambda_{\mathrm{QCD}}/m_b$. The expansion is carried out to a certain order in Λ, which requires local operators up to a certain order in k to be taken into account.

To obtain a result for the differential decay rate up to order Λ^2, the following local operators enter the OPE,

$$\begin{aligned} &\bar{b}\gamma^\alpha P_L b, \\ &\bar{b}(\mathrm{i}D - m_b v)^\mu \gamma^\alpha P_L b, \\ &\bar{b}(\mathrm{i}D - m_b v)^\mu (\mathrm{i}D - m_b v)^\nu \gamma^\alpha P_L b, \text{ and} \\ &-g\bar{b}G^{\mu\nu}\gamma^\alpha P_L b. \end{aligned} \tag{2.32}$$

From the $\langle b|\cdot|b\rangle$ and $\langle bg|\cdot|b\rangle$ matrix elements of the above local operators and of the operator in Eq. (2.30b), the coefficients $C_{abk}(Q)$ in the OPE can be inferred.

The next step in computing differential decay rates is to evaluate the local opterators in the $\langle \bar{B}|\cdot|\bar{B}\rangle$ matrix element, which, together with the coefficients, then yields $T_{\alpha\beta}$ expanded in Λ. The lowest order result reproduces the parton model result and the corrections at order Λ vanish.

2.3.3 Nonperturbative Parameters

The hadronic B-matrix elements can be parametrized in terms of two nonperturbative parameters at order Λ^2, denoted by μ_π^2 and μ_G^2, where

$$\mu_\pi^2 \equiv \frac{1}{2m_B}\langle \bar{B}|\bar{b}(\mathrm{i}\vec{D})^2 b|\bar{B}\rangle, \qquad \mu_G^2 \equiv \frac{1}{2m_B}\langle \bar{B}|\bar{b}\frac{\mathrm{i}}{2}\sigma_{jk}G^{jk}b|\bar{B}\rangle, \tag{2.33}$$

at tree level, i.e., at leading order in α_s, and $\sigma_{\mu\nu} = \frac{\mathrm{i}}{2}[\gamma_\mu, \gamma_\nu ,]$. Here, μ_π^2 describes the kinetic energy of the b quark from the residual motion inside the B meson, and μ_G^2 describes the magnetic interactions of the b quark with the light degrees of freedom in the B meson, that is, the light valence quark and possible quark-antiquark pairs as well as gluons. Both μ_π^2 and μ_G^2 are expected to be of the order Λ_{QCD}^2.

At order Λ^3, two more nonperturbative parameters arise from dimension-six operators, the expectation value of the Darwin operator ρ_D^3 and of the spin-orbit coupling ρ_{LS}^3, at tree level defined as

$$\rho_D^3 \equiv \frac{1}{2m_B}\langle \bar{B}|\bar{b}(-\frac{1}{2}\vec{D}\cdot\vec{E})b|\bar{B}\rangle, \qquad \rho_{LS}^3 \equiv \frac{1}{2m_B}\langle \bar{B}|\bar{b}(\vec{\sigma}\cdot\vec{E}\times i\vec{D})b|\bar{B}\rangle, \tag{2.34}$$

where $E_i = G_{i0}$.

The measured *decay rate* of $B \to X_q \ell\nu$ can be used to extract the CKM-matrix element $|V_{qb}|$, assuming that m_b and the nonperturbative parameters are known. It is the *shape* of differential decay distributions that carries additional information about the b-quark mass and the dynamics of the b quark in the B meson that can be used to determine m_b and the nonperturbative parameters. The shape information is usually obtained by measuring normalized moments of the decay distributions as a function of cuts on kinematic variables. In $B \to X_c \ell\nu$ decays moments of the lepton energy and of the hadronic mass are determined as a function of the lower cut on the lepton energy. Here, we determine the moments of the hadronic mass in $B \to X_u \ell\nu$ decays.

The best constraints on μ_G^2 are obtained from the mass difference between B and B^* mesons and heavy quark sum rules can be used to estimate ρ_{LS}^3. Ref. [26] finds $\mu_G^2 = (0.35^{+0.03}_{-0.02})\,\mathrm{GeV}^2$, and Ref. [27] suggests that typically, $-0.05\,\mathrm{GeV}^3 < \rho_{LS}^3 < -0.25\,\mathrm{GeV}^3$.

2.3.4 Hadronic Mass Moments

For the interpretation of the measured moments we use calculations performed in the kinetic scheme with a cutoff at $\mu = 1\,\mathrm{GeV}$.

The normalized integer moments of the squared invariant hadronic mass m_X^2 are defined

as [12]

$$\langle m_X^{2n} \rangle = \frac{\int \mathrm{d}m_X^2 \, m_X^{2n} \, \mathrm{d}\Gamma/\mathrm{d}m_X^2}{\int \mathrm{d}m_X^2 \, \mathrm{d}\Gamma/\mathrm{d}m_X^2}. \tag{2.35}$$

The calculation of the moments is performed using the parton level quantities $E_0 = 1 - \frac{v \cdot q}{m_b}$ and $s_0 = (v - q/m_b)^2$, where v_μ is the four-velocity of the B meson. In terms of E_0 and s_0, the hadronic mass can be expressed as

$$m_X^2 = \bar{\Lambda}^2 + 2 m_b \bar{\Lambda} E_0 + m_b^2 s_0. \tag{2.36}$$

The hadronic mass moments are computed from building blocks $\mathcal{M}_{(i,j)}$,

$$\begin{aligned} \mathcal{M}_{(i,j)} &= \frac{1}{\Gamma_0} \int \mathrm{d}E_0 \, \mathrm{d}s_0 \, \mathrm{d}E_\ell \, s_0^i \, E_0^j \, \frac{\mathrm{d}^3\Gamma}{\mathrm{d}E_0 \, \mathrm{d}s_0 \, \mathrm{d}E_\ell} \\ &= M_{(i,j)} + \frac{\alpha_s}{\pi} A_{(i,j)}^{(1)} + \frac{\alpha_s^2 \beta_0}{\pi^2} A_{(i,j)}^{(2)}, \end{aligned} \tag{2.37}$$

where $\Gamma_0 = G_F^2 m_b^5/(192\pi^3)$ is the tree level decay width, $\beta_0 = 11 - 2/3 n_f$ and the number of light flavors $n_f = 3$. $M_{(i,j)}$ contains the tree level contributions and the nonperturbative corrections up to order $1/m_b^3$. The values of $A_{(i,j)}^{(1,2)}$ are obtained from calculations for $B \to X_c \ell \nu$ with a small value of m_c.

The predicted moments are affected by several systematic uncertainties. Uncertainties arise from higher-order uncalculated nonperturbative corrections, of order $1/m_b^4$ and beyond, and perturbative corrections, of order α_s^2 and $\alpha_s \Lambda_{\mathrm{QCD}}^{(2,3)}/m_b^{(2,3)}$ and beyond. Further uncertainties are induced by four quark operators, which introduce additional $1/m_b^3$ corrections.

To reduce the experimental uncertainties for the determination of m_b and nonperturbative parameters from $B \to X_u \ell \nu$ decays, the moments are defined with a (variable) upper cut on the hadronic mass. However, if a tight cut is applied, too little phase space is integrated over to be able to predict the moments because the $1/m_b$ expansion breaks down. This can be remedied by using a modified expansion, usually referred to as twist expansion. Basically, the usual expansion is not justified any longer for a particular component of k, the residual momentum of the b quark in the B meson, and this component is left unexpanded. As a result, the place of the nonperturbative parameters is taken by nonperturbative functions, which, at tree level, describe the motion of the b quark in the B meson

and are commonly referred to as shape functions. Ref. [12] does not include shape function effects, which introduces an additional uncertainty on the predictions of the moments, and hence the extracted b-quark mass and nonperturbative parameters, when a cut on m_X^2 is applied. Alternatively, we can turn this argument around: By extracting mass moments with different cuts on m_X^2 and comparing the measured cut dependence with that predicted by Ref. [12], one can determine how low the cut on m_X^2 can be pushed before shape function effects become important.

2.4 New Variables for the Inclusive Determination of $|V_{ub}|$ at the B Factories

In inclusive measurements, the suppression of the dominant $B \to X_c \ell\nu$ background constitutes a major experimental challenge. Almost all extractions of $|V_{ub}|$ make use of the size of the charm quark mass, $m_c \gg m_u$, to achieve a separation of signal decays from charm backgrounds by using kinematic constraints on the reconstructed semileptonic decays. However, both theoretical uncertainties on $|V_{ub}|$ and the sensitivity to the mass of the b quark increase rapidly with tighter cuts on the available phase space for signal events. The precise dependence varies for different kinematic variables, but in all cases the theoretical and parametric uncertainties increase quickly when decreasing the fraction of signal events

$$f_u = \frac{\mathcal{B}^{\rm cut}(B \to X_u \ell\nu)}{\mathcal{B}^{\rm total}(B \to X_u \ell\nu)}\,, \tag{2.38}$$

where $\mathcal{B}^{\rm cut}$ is the signal branching fraction in the presence of kinematic cuts and $\mathcal{B}^{\rm total}$ is the total branching fraction.

Originally, $|V_{ub}|$ has been extracted using the endpoint of the lepton energy spectrum. While the experimental resolution of the lepton energy, E_ℓ, is good, the small fraction of signal decays beyond the endpoint of the lepton energy in $b \to c\ell\nu$ decays gives rise to substantial theoretical and parametric uncertainties. Lowering the cut on the lepton energy to reduce these uncertainties yields larger experimental uncertainties related to the control of the charm background. Recent analyses use $E_\ell > 2.2\,\text{GeV}$ [28], $E_\ell > 1.9\,\text{GeV}$ [29], and

$E_\ell > 2.0\,\mathrm{GeV}$ [30], resulting in a retained fraction of signal events between $f_u = 0.11$ and $f_u = 0.32$.

The use of hadronic variables such as the mass of the hadronic system, m_X, has allowed for a much larger fraction of the signal to be retained when separating the charm background. Ideally, a cut on $m_X < m_D$ would remove all charm background while keeping about 80% of the signal. However, to achieve a sufficiently good resolution on hadronic variables, the analyses need to fully reconstruct the second B meson (the "tag B") from the $\Upsilon(4S)$ decay. In addition, these "tagged" analyses allow one to perform the measurement in the rest frame of the signal B and open the possibility to separate charged and neutral Bs. The analysis strategy of tagging by a fully reconstructed second B meson in the event has become possible with the start of the B factories, where only two B mesons are produced for each B event, without soft fragmentation products or similar backgrounds. While tagged analyses allow for larger f_u and smaller systematic uncertainties from charm backgrounds, this comes at the cost of statistical uncertainty: The efficiency for fully reconstructing the tag B is only about half a percent. Furthermore, the tagged analyses have additional systematic uncertainties from the subtraction of combinatoric backgrounds on the tag side, which contribute substantially to the total systematic uncertainty [31].

An alternative method [32, 33] that retains the high efficiency of untagged analyses, but allows a larger fraction of signal events to pass the charm rejection cuts, employs the reconstructed neutrino momentum determined from the total missing momentum in the $\Upsilon(4S)$ decay. This untagged neutrino reconstruction has an efficiency of about 10%. The first measurement [33] using this technique achieves an overall uncertainty on $|V_{ub}|$ of similar size as other determinations at that time.

The method of Refs. [32, 33] uses a combined q^2–E_ℓ cut, where q^2 is the total invariant mass of the leptonic system, to discriminate against charm background. In theory, this cut has larger f_u than a pure lepton energy cut. However, the imperfect neutrino reconstruction yields a poor resolution in q^2, which again introduces uncertainties from charm background subtraction. Limiting these still necessitates an additional cut on $E_\ell > 2.0\,\mathrm{GeV}$, decreasing f_u again.

Given a specific reconstruction method, optimizing the overall uncertainty in $|V_{ub}|$ is usually achieved by optimizing the phase-space cut that discriminates against charm background and is largely a trade-off between increasing f_u, to reduce theoretical and parametric uncertainties, versus increasing the uncertainties from charm background. So far, this optimization happens by choosing one (or two) specific discriminating variable and optimizing the value of the cut on this variable.

In this section we explore alternative combinations of leptonic variables that are suited for untagged inclusive measurements. The variables we construct provide a more general way to optimize the phase-space cut, which gives the experiments an additional *continuous* degree of freedom in optimizing the total uncertainty on $|V_{ub}|$.

2.4.1 Kinematics

The experimental resolution in the discriminating kinematic variables has substantial impact on the systematic uncertainties from charm backgrounds, but is often not a part of theoretical considerations. With poor resolution one is often led to cut further away from the theoretical charm threshold, while with good resolution one can move closer to it. In general, the resolution itself depends on the region of phase space. Ideally, one would like to cut closer along the charm threshold in phase-space regions with good resolution and stay further away from it in regions with poor resolution.

Our starting point is the simple observation that the lepton energy has by far the best experimental resolution, which makes it very useful as a discriminating variable in this respect. The combined q^2–E_ℓ cut can be viewed as a q^2-dependent E_ℓ cut, which is given by the kinematic endpoint in E_ℓ for $B \to X_c \ell \nu$ for a specific q^2 value. In other words, this cut provides a way to incorporate additional kinematic information from a second variable, in this case q^2. In principle, we can use any other kinematic variable instead of q^2 as a second variable, and choosing one with good experimental resolution will result in a better charm background suppression. In the following, we will explore three different variables as second variables: q^2, the neutrino energy E_ν, and the light cone variable q_-, defined below.

In the rest frame of the B meson, the inclusive decays $B \to X\ell\nu$ can be described by three independent kinematic variables, which we can choose to be the leptonic variables q^2, E_ℓ, and E_ν or alternatively E_ℓ and $q_\pm = q^0 \pm |\vec{q}|$, where $q^0 = E_\ell + E_\nu$ and $\vec{q}$ are the total energy and three-momentum of the leptonic system. The leptonic phase space is then

$$0 \le q^2 \le 4E_\ell E_\nu \qquad \text{or} \qquad 0 \le q_- \le 2E_\ell \le q_+ \,. \tag{2.39}$$

The remaining hadronic phase-space limit comes from the restriction on the invariant mass of the hadronic system, $m_X \ge m_D$ for $B \to X_c\ell\nu$ and $m_X \ge m_\pi$ for $B \to X_u\ell\nu$. Thus, as is well known, to eliminate the charm background we need to ensure that $m_X < m_D$.

In terms of the above leptonic variables, we have

$$m_X^2 = m_B^2 + q^2 - 2m_B(E_\ell + E_\nu) = (m_B - q_+)(m_B - q_-) \,. \tag{2.40}$$

Thus, using Eq. (2.39) we can obtain an upper limit for m_X^2 given our two leptonic variables of choice:

$$s_\nu = (m_B - 2E_\ell)(m_B - 2E_\nu) \,, \tag{2.41a}$$

$$s_q = (m_B - 2E_\ell)\Big(m_B - \frac{q^2}{2E_\ell}\Big) \,, \tag{2.41b}$$

$$s_- = (m_B - 2E_\ell)(m_B - q_-) \,. \tag{2.41c}$$

The variable s_q is identical to the variable $s_h^{\max}$ defined in Ref. [32]. From Eq. (2.39) one can easily see that the s_i obey the hierarchy

$$m_X^2 \le s_\nu \le s_q \le s_- \,. \tag{2.42}$$

Assuming perfect resolution, requiring $s_i < s_i^{\text{cut}}$ with $s_i^{\text{cut}} = m_D^2$ rejects all $B \to X_c\ell\nu$ background. Since s_ν (s_-) yields a better (worse) approximation of m_X^2 than s_q, it has a higher (lower) signal efficiency than s_q.

The two-dimensional phase space in E_ℓ and E_ν, q^2, or q_-, respectively, together with the corresponding cut on s_i, is shown in Fig. 2.3. We observe that for large E_ℓ close to the endpoint, the cut on s_i becomes an almost pure E_ℓ cut and hence profits from the excellent E_ℓ resolution. With decreasing E_ℓ, the s_i cut becomes more dependent on the

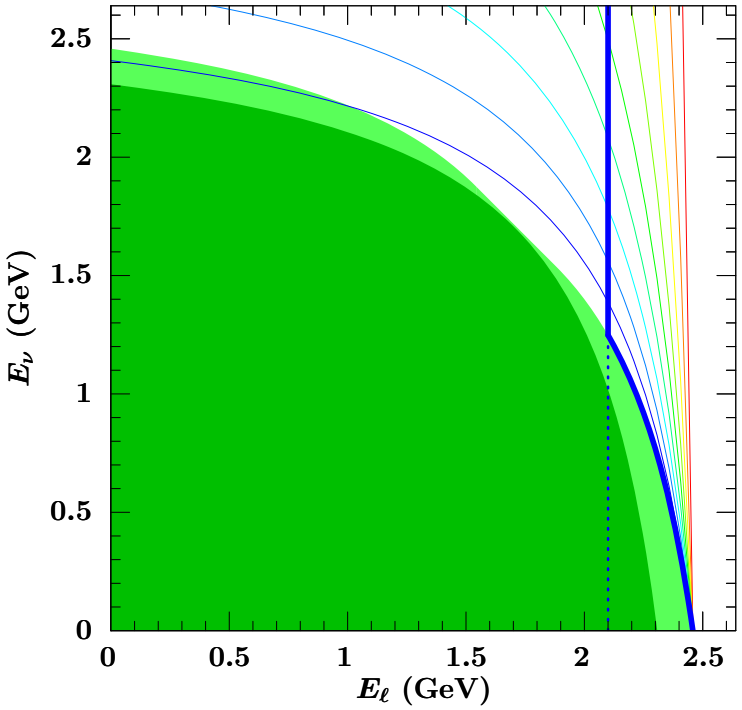

Figure 2.3. Two-dimensional phase space in E_ν–E_ℓ. The dark shaded area shows the available phase space for $B \to X_c \ell\nu$ in the B rest frame, the light shaded area shows the smearing due to the $\Upsilon(4S)$ frame. The boundaries of the $B \to X_u \ell\nu$ phase space coincide with the plot boundaries. The thick line shows the cuts $s_\nu < m_D^2$ and $E_\ell > 2.1\,\mathrm{GeV}$. The thin lines show cuts $s_\nu(x) < m_D^2$ for $x = 1$ to $x = 0$ (pure E_ℓ cut).

second leptonic variable, which will result in a poorer resolution and cause more charm background to leak into the theoretically clean phase-space region $s_i < s_i^{\rm cut}$. To limit this effect Ref. [33] imposes an additional cut $E_\ell > E_\ell^{\rm cut}$, which is also shown in Fig. 2.3. Alternatively, one could also decrease the value of $s_i^{\rm cut}$. Both possibilities reduce the charm backgrounds but also reduce the fraction of signal events f_u.

Naively, for each s_i, we have two degrees of freedom to optimize the experimental cuts: The values of $s_i^{\rm cut}$ and $E_\ell^{\rm cut}$. As mentioned before, ideally we would like the kinematic cut to be close to the kinematic boundary for $B \to X_c \ell\nu$ for large E_ℓ, taking advantage of the good E_ℓ resolution, and smoothly depart from it for smaller E_ℓ, taking into account the poorer resolution. This can be achieved by including an additional free parameter x in the definition of the s_i that controls the shape of the cut:

$$s_\nu(x) = (m_B - 2E_\ell)(m_B - x\,2E_\nu)\,, \tag{2.43a}$$

$$s_q(x) = (m_B - 2E_\ell)\Big(m_B - x\,\frac{q^2}{2E_\ell}\Big)\,, \tag{2.43b}$$

$$s_-(x) = (m_B - 2E_\ell)(m_B - x\,q_-)\,. \tag{2.43c}$$

The value of $0 \le x \le 1$ controls how much the second leptonic variable is used and thus effectively determines the resolution of $s_i(x)$. The phase space in E_ν–E_ℓ with various cuts is shown in Fig. 2.3. For $x = 1$ we recover the definitions in Eqs. (2.41), $s_i(1) \equiv s_i$, while for $x = 0$, $s_i(0) < s_i^{\rm cut}$ is equivalent to a pure E_ℓ cut

$$E_\ell > \frac{m_B^2 - s_i^{\rm cut}}{2m_B}\,. \tag{2.44}$$

Therefore, by varying the value of x one gains an independent degree of freedom in optimizing the phase space-cut, which smoothly interpolates between a pure E_ℓ cut and a given s_i cut.

So far, our discussion was specific to the rest frame of the decaying B meson. However, experimentally, untagged analyses can only be performed in the rest frame of the $\Upsilon(4S)$. There are several strategies one might consider to take into account the boost of the B meson in the $\Upsilon(4S)$ rest frame. Here, we assume that the experiments will correct for this boost and quote their results in the B rest frame, which was also the approach taken in

Ref. [33]. The reason is that the experimental reconstruction efficiencies depend on the position in phase space and also on the boost, which should be dealt with on the same footing.

Therefore, we will continue to work in the B rest frame and regard the values of the kinematic variables measured in the $\Upsilon(4S)$ rest frame as a measurement of their true values in the B rest frame. This is reasonable, since the boost essentielly causes an additional (known) smearing on top of the experimental resolution. For the neutrino energy (or q_-), the effect of the boost is negligible compared to their experimental resolution, while for the lepton energy, the smearing caused by the boost is the dominant effect. We stress that (except at the very edge of phase space) this smearing is roughly symmetric, so a measurement using an experimental cut $f(E_\ell^\Upsilon, E_\nu^\Upsilon) \le f^{\rm cut}$, where E_i^Υ are the energies measured in the $\Upsilon(4S)$ rest frame, should quote a result for $f(E_\ell, E_\nu) \le f^{\rm cut}$, with the *same* functions f and values $f^{\rm cut}$[4].

The important effect of the boost is that it effectively increases the available phase space for $b \to c$ transitions. To take this into account, we have to modify Eqs. (2.43),

$$s_\nu^\eta(x) = (m_B - \eta_- 2E_\ell)(m_B - x\,\eta_+ 2E_\nu)\,, \tag{2.45a}$$

$$s_q^\eta(x) = (m_B - \eta_- 2E_\ell)\Big(m_B - x\,\eta_+ \frac{q^2}{2E_\ell}\Big)\,, \tag{2.45b}$$

$$s_-^\eta(x) = (m_B - \eta_- 2E_\ell)(m_B - x\,\eta_+ q_-)\,, \tag{2.45c}$$

where

$$\eta_\pm = \sqrt{\frac{1\pm\beta}{1\mp\beta}}\,, \tag{2.46}$$

and β is the boost of the B in the $\Upsilon(4S)$ rest frame. As argued above, if $s_i^\eta(x)$ are used for the measurement they should also be used to quote the final results. Strictly speaking $s_i^\eta(x) \le m_D^2$ only give the theoretical $b \to c$ phase-space boundary for $E_\ell \ge \frac{1}{2}(\eta_+ m_B - m_D)$ for $i = \nu, -$ and $E_\ell \ge \frac{1}{2}\eta_+(m_B - m_D)$ for $i = q$. For practical purposes this is irrelevant, because the lepton energy cut will always be larger.

[4]This is in contrast to Ref. [33], where with a cut on $E_\ell^\Upsilon > 2.1\,\text{GeV}$ a result for $E_\ell > 2.0\,\text{GeV}$ was quoted, which implicitly extrapolates the measurement to a larger fraction of events than was actually measured. In other words, theoretically both $E_\ell^\Upsilon > 2.1\,\text{GeV}$ and $E_\ell > 2.1\,\text{GeV}$ have very similar f_u, whereas $E_\ell > 2.0\,\text{GeV}$ has a substantially larger f_u.

2.4.2 Results

We perform a study using Monte Carlo simulation (MC) to compare the expected performance of different kinematic cuts.

Monte Carlo Setup

To produce samples of $B \to X_u \ell\nu$ signal and $B \to X_c \ell\nu$ backgrounds we use the EvtGen event generator [34]. Branching fractions for $B \to X_c \ell\nu$ decays are adjusted to recent measurements [8]. $B \to X_u \ell\nu$ decays are simulated according to [35] and the hadronization of the hadronic state is performed by `JetSet` [36]. We use a pole mass of $m_b = 4.75\,\mathrm{GeV}$ and $\lambda_1 = -0.27\,\mathrm{GeV}^2$.

We use estimates for resolutions in kinematic variables that can be achieved with present analysis techniques for untagged neutrino reconstruction [37]. In particular, we take into account non-Gaussian effects and correlations between E_ℓ and the respective second leptonic variable.

Results

Taking into account experimental resolution, we obtain estimates for the ratio of signal to background for different kinematic cuts. Since we are interested in comparing the different cuts, we normalize our results to what we obtain for the cuts in Ref. [33], $s_q < 3.5\,\mathrm{GeV}^2$ and $E_\ell > 2.1\,\mathrm{GeV}$. We also obtain the value of f_u predicted by Ref. [35] for the same cuts. The results as a function of the cut on $s_i(1)$, $s_i(1)^{\mathrm{cut}}$, and x are displayed in Fig. 2.4. The size of f_u gives an indication of the theoretical uncertainties, provided that events from all regions of the associated phase space contribute significantly to the result. This would need to be tested in an actual analysis of experimental data.

We scan a wide range of values for $s_i(x)^{\mathrm{cut}}$, E_ℓ^{cut}, and x, and compare the different configurations in Fig. 2.5. Under the assumption that the uncertainty on $|V_{ub}|$ is dominated by the systematic uncertainty from charm backgrounds and theoretical uncertainties, the configuration that yields the smallest uncertainty on $|V_{ub}|$ will be found on the inner side of

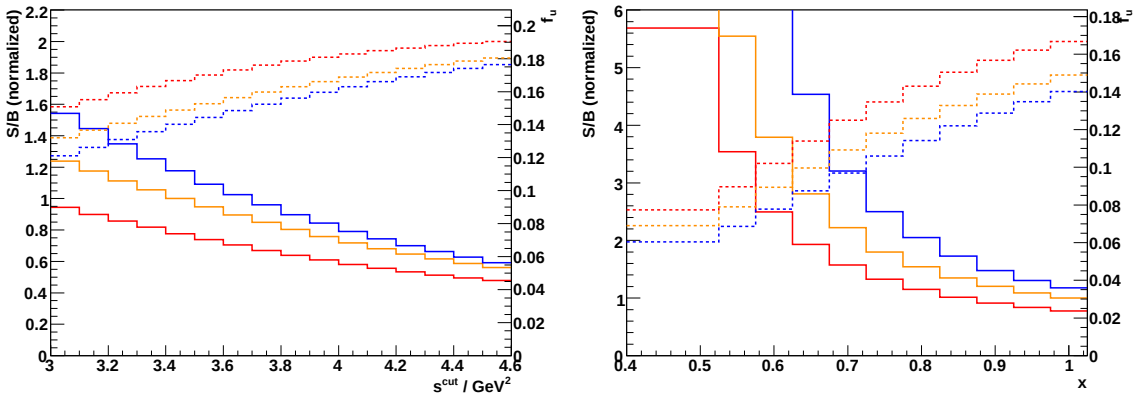

Figure 2.4. The normalized signal-to-background ratio (dashed, left axes) and f_u (solid, right axes) as a function of the cut on $s_i(1)$ (left) and as a function of x for $s_\nu(x)$ in red, $s_q(x)$ in orange, and $s_-(x)$ in blue.

the roughly hyperbolic plane. The yellow star indicates the configuration used in Ref. [33], showing that a better optimum can be found.

As a naive attempt to perform an optimization in $s_i^{\rm cut}(x)$, $E_\ell^{\rm cut}$, and x, we assume that the uncertainty induced from uncertainties in the charm background is inversely proportional to the normalized ratio of signal to background, and that the theoretical uncertainty is inversely proportional to f_u. We match the absolute size of the projected uncertainties to that observed in Ref. [33].

Under these assumptions, the configuration that minimizes the total uncertainty on $|V_{ub}|$ is $s_\nu(0.6)^{\rm cut} = 5\,{\rm GeV}^2$, $E_\ell^{\rm cut} = 1.95\,{\rm GeV}$, for which f_u exceeds 20%.

In practice, the cuts and x need to be determined taking into account details of the data analysis that are impossible to implement into this study. Yet, our study suggests strongly that a better optimum than the currently used configuration can be found and the new variables allow for additional degrees of freedom in the optimization.

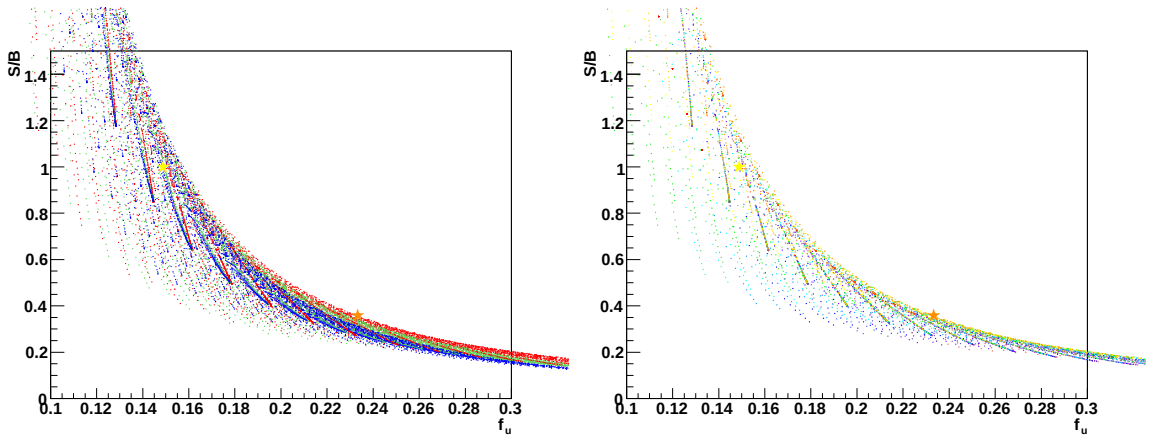

Figure 2.5. Normalized signal-to-background ratio and f_u distribution when scanning $s_i(x)^{\text{cut}}$, the cut on E_ℓ and the value of x for $s_\nu(x)$ in red, $s_q(x)$ in green, and $s_-(x)$ in blue (left) and when scanning $s_i(x)^{\text{cut}}$ and E_ℓ^{cut} for different values x in $s_\nu(x)$ in rainbow order ($x = 1$ in blue to $x = 0$ in red). The configuration from Ref. [33] is indicated by a yellow star, the optimum obtained in the toy minimization by an orange star.

Chapter 3

The *BABAR* Experiment

The *BABAR* experiment at the asymmetric PEP-II storage ring is located at the Stanford Linear Accelerator Center (SLAC). *BABAR*'s primary goal is the study of charge-parity (CP) asymmetries in B-meson decays. The large amount of recorded B-meson decays also allow for high-precision determinations of CKM-matrix parameters and the measurement of rare B decays. Other topics studied include the physics of bottom and charm mesons in general, as well as τ leptons.

The following sections give a short overview of the PEP-II B factory and the *BABAR* detector and follow the descriptions in Ref. [38] and [39]. They also give a brief overview of the reconstruction of charged particles at *BABAR*, highlighting one contribution to recent tracking improvements, and introduce the recoil method used in the analysis presented here.

3.1 The PEP-II B Factory

The PEP-II B factory is designed to deliver B mesons to the experiment and to provide a clean environment needed for the measurements. For the main physics program, PEP-II runs at a center-of-mass energy corresponding to the mass of the $\Upsilon(4S)$ resonance, $m_{\Upsilon(4S)} = 10.58\,\text{GeV}$. The $\Upsilon(4S)$ decays dominantly into $B\bar{B}$ pairs. As an e^-e^+ collider, it offers a cleaner environment than a hadron collider and a high signal-to-background ratio of

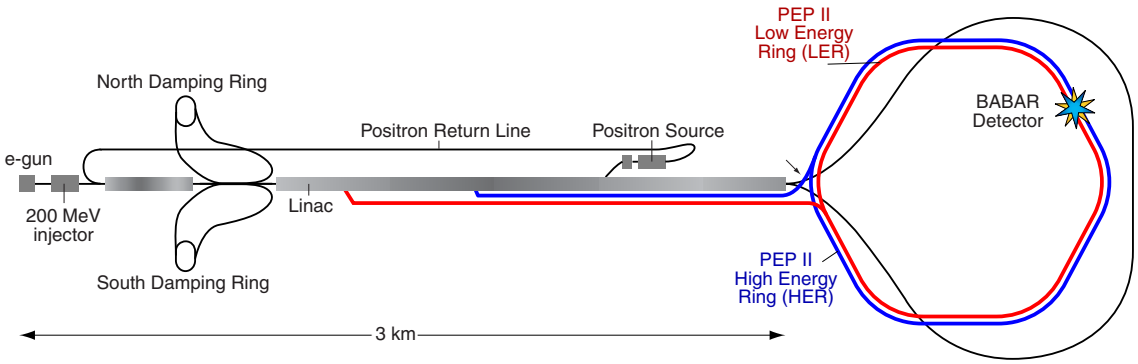

Figure 3.1. The PEP-II storage rings and the linear collider.

$\sigma_{b\bar{b}}/\sigma_{\mathrm{had}} \approx 0.28$, where σ_{had} is the total hadronic cross section. The absence of fragmentation products reduces the combinatorial background, and the knowledge of the $\Upsilon(4S)$ four momentum and thus of the $B\bar{B}$ system provides kinematic constraints, which permit a considerable suppression of combinatorial and continuum backgrounds for the reconstruction of B decays.

The B mesons are almost at rest in the center-of-mass system. Since the CP asymmetries which occur in the interference between decays with and without mixing cancel in time integrated measurements at e^+e^- colliders, the proper decay time of the B meson needs to be measured. In order to resolve the decay length of the B mesons and infer the time of their decays, which is needed for the measurement of CP violation in the B system, PEP-II works asymmetrically, that is, with different energies for the e^- and the e^+ beam. In this way, the B mesons receive significant momenta in the laboratory frame and have measureable decay lengths.

Fig. 3.1 shows the PEP-II storage rings and the linear accelerator, which serves as an injector. Electrons of 9 GeV are stored in the High Energy Ring (HER) and collide with positrons of 3.1 GeV, which are held in the Low Energy Ring (LER). The asymmetric beam energies result in a Lorentz boost of $\beta\gamma = 0.56$ of the B mesons in the laboratory frame, which allows the measurement of their decay times and thus of time-dependent CP asymmetries.

A high luminosity is needed in order to achieve enough statistics for processes with low

branching fractions. At the $\Upsilon(4S)$ resonance, the cross section for $e^-e^+ \to b\bar{b}$ is about $\sigma(e^-e^+ \to b\bar{b}) = 1.05\,\text{nb}$. PEP-II's design luminosity of $3 \cdot 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ has been surpassed by a factor of 4. Between 1999 and April 2008, PEP-II delivered an integrated luminosity of $553.48\,\text{fb}^{-1}$, of which $531.43\,\text{fb}^{-1}$ was recorded by *BABAR*. Data taking was divided into seven time periods, usually denoted by Run1 through Run7. During Run1 through Run6, data were taken mainly at the $\Upsilon(4S)$ resonance, with a small percentage of data taken 40 MeV below the $\Upsilon(4S)$ mass to study non-$B\bar{B}$ backgrounds. Run7 was devoted to data taking at the $\Upsilon(3S)$ and $\Upsilon(2S)$ resonances for the study of possible low-energy New Physics and spectroscopy. In addition, the $\Upsilon(5S)$ and $\Upsilon(6S)$ resonances were scanned. Fig. 3.2 shows the integrated luminosity delivered by PEP-II and recorded by *BABAR* between 1999 and 2008.

3.2 The *BABAR* Detector

The asymmetric design of the *BABAR* detector accounts for the asymmetric beam energies in the HER and the LER and hence garantees a large geometrical detector acceptance.

In order to achieve *BABAR*'s physics goals, an excellent decay vertex resolution of the B mesons and a good particle identification over a wide kinematical range are needed. The *BABAR* detector with its subsystems is shown in Fig. 3.3 and described in the following.

In the *BABAR* coordinate system, the direction of the B field defines the z-axis. The x-axis is the horizontal axis, the y-axis the vertical axis and x, y and z form a right-handed system. In the corresponding cylindrical coordinate system, ϕ measures the angle in the $x - y$ plane and θ measures the angle to the z-axis.

3.2.1 The Silicon Vertex Tracker (SVT)

The silicon vertex tracker (SVT) is the innermost detector component closest to the beam pipe. It provides up to ten accurate measurements of the charged tracks immediately outside the beam pipe. This allows one to precisely extrapolate charged particles to their

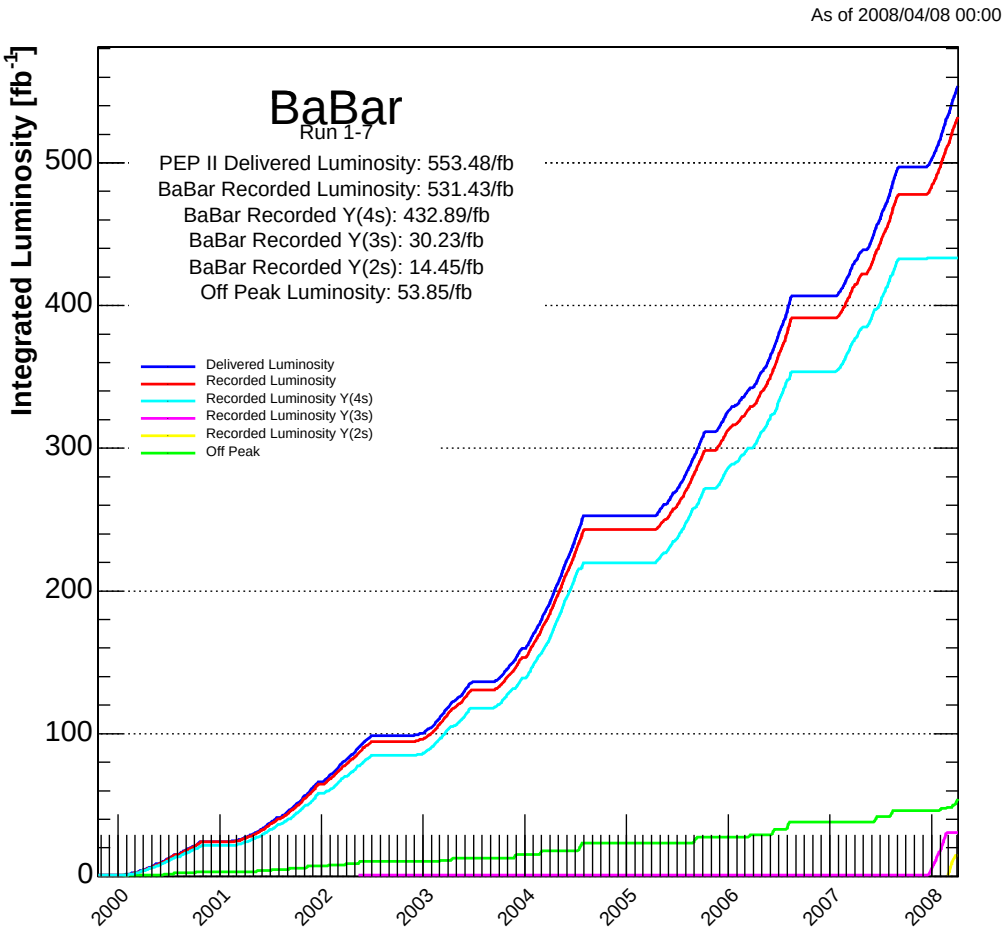

Figure 3.2. Integrated luminosity delivered by PEP-II and recorded by *BABAR*.

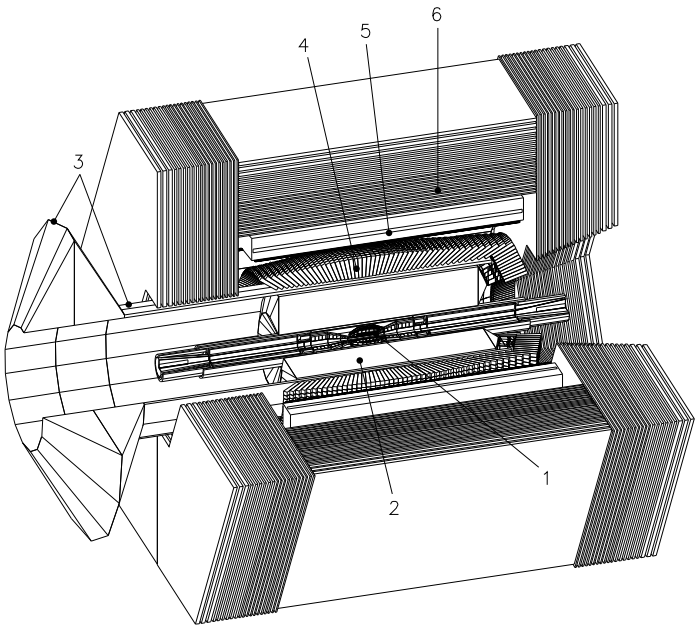

Figure 3.3. The *BABAR* detector and its subsystems: (1) Silicon Vertex Tracker (SVT), (2) Drift Chamber (DCH), (3) Detector of Internally Reflected Cherenkov Light (DIRC), (4) Electromagnetic Calorimeter (EMC), (5) Superconducting Coil, and (6) Instrumented Flux Return (IFR).

production point. The SVT serves for the reconstruction of particle trajectories and decay vertices.

Low momentum tracks with a transverse momentum below 100 MeV do not reach the drift chamber because of their curvature in the magnetic field and are measured by the SVT alone. For higher momentum tracks the SVT provides one part of the tracking information. It also yields up to ten measurements of $\mathrm{d}E/\mathrm{d}x$ per track.

The SVT consists of five double-sided, concentric, cylindrical layers of silicon detectors. The layers are composed of 6, 6, 6, 16, and 18 silicon strip detectors, respectively. The inner sides of each layer are oriented perpendicular to the beam direction and allow one to measure the z coordinate, whereas the outer sides are oriented in beam direction to determine the angle ϕ. The three inner layers reach a resolution of $(10-15)\,\mu$m, the outer layers of $(30-40)\,\mu$m, which in both cases is dominated by the uncertainty due to multiple scattering.

3.2.2 The Drift Chamber (DCH)

The drift chamber (DCH) is *BABAR*'s main tracking device for charged particles with a transverse momentum above 100 MeV. It consists of 7104 hexagonal drift cells arranged in 10 superlayers of 4 layers each. The drift cells are composed of a 20 μm gold-plated tungsten-rhenium sense wire, surrounded by six 120 μm or 80 μm gold-plated aluminum field wires. The gas mixture of 80% helium and 20% isobutane is chosen to provide good spatial and $\mathrm{d}E/\mathrm{d}x$ resolution, and a reasonably short drift time, while minimizing the material. Spatial resolution along the z direction is achieved by stereo angles between the superlayers, ranging between 40 and 76 mrad from the innermost to the outermost superlayer.

The drift chamber operates in a magnetic field of 1.5 T. The resolution for transverse momenta p_t (measured in GeV) follows a linear dependence,

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

Apart from precise measurements of the charged particles' momenta and directions, the drift chamber also provides particle identification by determination of the ionization loss $\mathrm{d}E/\mathrm{d}x$. $\mathrm{d}E/\mathrm{d}x$ is derived from the total charge deposited in each drift cell with an average resolution of typically 7.5%. Pions and Kaons can be separated for momenta lower than 700 MeV.

3.2.3 The Cherenkov Detector (DIRC)

The detector of internally reflected Cherenkov light (DIRC) is devoted to particle identification. It is designed to provide excellent kaon identification and to separate between kaons and pions at large momenta up to 4 GeV.

The DIRC consists of 144 fused silica bars with a refraction index close to $n = 1.474$ and of 4.9 m length. When a charged particle traverses the quartz bars with a velocity higher than the speed of light in the material, it radiates photons under an angle of $\cos\alpha = 1/(n\beta)$, so-called Cherenkov light. By total internal reflection and mirrors on the forward end, the photons are transported to the end of the bars on the backward end of the detector. In this

way, the light is directed into the connected Standoff Box filled with purified water. The photons are detected by photomultiplier tubes operating in the water. The reconstructed Cherenkov angle α together with the momentum measured by the tracking system provides particle identification information.

3.2.4 The Electromagnetic Calorimeter (EMC)

The electromagnetic calorimeter is designed to measure electromagnetic showers with high efficiency, and with high energy and angular resolution. It consists of a cylindrical barrel and conical foward endcap and contains 6580 thallium-doped CsI crystals. They are tilted with respect to their azimuthal direction to the interaction point to minimize losses from photons traversing though the gaps between the crystals. The calorimeter covers a solid angle of $-0.775 < \cos(\theta) < 0.962$ in the laboratory frame, corresponding to $-0.916 < \cos(\theta) < 0.895$ in the center-of-mass frame. The usable region is reduced due to shower leakage in the outmost crystals. The crystal are read out with silicon photodiodes at their rear ends.

The CSi(Tl) crystals have a radiation length of 1.85 cm and a Molière radius of 3.8 cm, compared to a typical crystal length of 30 cm and front and back face areas of $4.7 \times 4.7\,\mathrm{cm}^2$ and $6.1 \times 6.0\,\mathrm{cm}^2$, respectively.

Due to varying light yields among the crystals and damage from beam-generated radiation, frequent calibration of the individual crystals is necessary. In the low-energy region, the calibration is carried out with photons of 6.13 MeV from activated oxygen. The high-energy calibration uses Bhabha events with energies of 3 to 9 GeV.

The energy resolution is determined in several intervals of the energy range and is fitted to be

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

where $\oplus$ denotes addition in quadrature. The angular resolution is determined based on

decays of π^0 and η into photons of approximately equal energy and is found to be

$$\sigma_\theta = \sigma_\phi = \left(\frac{3.87 \pm 0.07}{\sqrt{E/\mathrm{GeV}}} + 0.00 \pm 0.04\right) \mathrm{mrad}.$$

The EMC serves for the detection of photons, for example from π^0 and η decays. It is a main device for the electron identification, which uses the shower energy and form and the track momentum. Additionally, the $\mathrm{d}E/\mathrm{d}x$ information from the drift chamber and the DIRC Cherenkov angle are used.

3.2.5 The Instrumented Flux Return (IFR)

The instrumented flux return is outside the solenoid coil providing the magnetic field for momentum measurements of charged particles. The steel serves as flux return for the magnetic field, and as muon and hadron absorber. It is segmented into iron plates varying in thickness between 2 and 10 cm, in between which resistive plate chambers (RPCs) were placed. Originally, there were 19 layers of RPCs in the barrel and 18 layers in the endcaps, covering a total active area of about $2000\,\mathrm{cm}^2$. Two layers of cylindrical RPCs are mounted between the EMC and the magnet cryostat, and detect particles exiting the EMC. The RPCs detect streamers from ionizing particles via capacitive readout strips. Their active volume is filled with a mixture of Argon (57%), Freon 134a (37%), and Isobutane (about 5%).

However, the RPCs have lost detection efficiency much faster than expected due to problems with the linseed oil used in their manufacturing [40]. The RPCs in the barrel region have therefore been replaced by Limited Streamer Tubes (LSTs) before the start of Run5 (two sextants) and Run6 (four sextants). LSTs consist of a resisitive tube, serving as cathode, with an anode wire in the center, and are operated in limited streamer mode, where a streamer discharge develops on one side of the wire.

Muon identification relies almost entirely on the IFR. Tracking information from the SVT and the drift chamber is used to extrapolate the tracks to the IFR. All detected IFR clusters within a predefined distance from the extrapolated track are associated with the track.

K_L and other neutral hadrons interact with the IFR steel and can be identified as IFR clusters not associated with a charged track. Additional information for neutral hadron detection is provided by the cylindrical RPCs and the EMC.

3.3 Reconstruction of Charged Particles

In order to correctly reconstruct B-meson decays, charged and neutral particles need to be reconstructed and their particle type needs to be identified. At *BABAR*, charged tracks are reconstructed in the SVT and the DCH. Neutral particles are seen in the EMC and identified as neutral if no charged track is found that could have caused the energy deposit in the calorimeter.

Reconstruction of charged tracks is accomplished in two steps: Track finding (pattern recognition) and track fitting.

Tracks are found independently in the SVT and the DCH. The pattern recognition algorithms start on the outside of the detectors, where the occupancy is lowest, and extends the track search inwards. Separately found tracks in the SVT and DCH are then merged into single tracks, if the χ^2 of the match between the two track pieces is good. Hit adding algorithms attempt to extend unmerged SVT and DCH tracks into the DCH and SVT by attaching unassociated DCH and SVT hits, respectively.

After the pattern recognition, the tracks are initially fit with a simple helix. Charged particles in a B field along the z axis can ideally be described by a helix around the z axis defined by the track parameters d_0, ϕ_0, ω, z_0, and $\tan\lambda$. The track parameters can be interpreted in terms of the point of closest approach between the track and the origin in the $x-y$ plane. The distance of closest approach to the origin of the $x-y$ plane is given by d_0, the sign of which is determined by the angular momentum of the track. The angle of the track in the $x-y$ plane at the point of closest approach is ϕ_0. The z projection of the distance of closest approach to the origin is given by z_0. The curvature of the track in the $x-y$ plane is given by $\omega = \frac{1}{r_t}$, where r_t is the radius of the track in the $x-y$ plane,

and the sign of ω is determined by the angular momentum of the track. Finally, $\tan\lambda$ is the tangent of the angle between the z axis and the track.

The helix fit is followed by a Kalman filter fit, where tracks are modeled as piecewise helices. The fit accounts for effects from interactions with the detector material as well as from magnetic field inhomogeneities. The Kalman filter processes the effects such as the measured hits, the expected scattering, and the deflection due to field imhomogeneities sequentially according to their flight length on the track, and updates the estimated track parameters at each step. The track fit yields the best linear unbiased estimate of the track parameters.

3.3.1 Recent Tracking Improvements

BABAR has recently introduced a set of algorithms, collected in the `TrkFixup` software package, that improve the quality of charged tracks used in physics analysis. `TrkFixup` relies on the standard reconstruction of charged particles having been performed and provides a second pass to identify and improve known tracking pathologies. The algorithms make use of the detailed hit level tracking information that is available in the *BABAR* computing model CM2.

`TrkFixup` consists of two parts. The first part deals with the reconstructed track objects, the second part creates improved track-based candidate lists to be used for physics analysis. The `TrkFixup` track algorithms can be classified as either rejecting background tracks, such as tracks arising from pattern recognition errors (referred to as ghosts tracks), non-primary branches of looping tracks, and tracks arising from material interactions or decays in flight, or as improving the resolution of the fitted track parameters by removing a few inconsistent hits or adding SVT hits to DCH-only tracks. These algorithms rely on the availability of the full Kalman fit, while normal data analysis uses a much faster parametrized version of the track fit. To make `TrkFixup` viable in terms of computing time, tracks are preselected if they show characteristics of background or being poorly measured, and the full Kalman is only restored for the preselected tracks.

Hit Filtering

In the context of `TrkFixup`, hit filtering refers to the idea of removing a few inconsistent hits from the track fit in order to improve the resolution of the fitted track parameters, and is performed by the `TrkHitFix` module. Studies and the tuning of the algorithm use simulated Monte Carlo (MC) data.

Tracks with a low track fit probability P can have a substantial number of misassociated hits. For the purpose of hit filtering, we consider a hit to be misassociated if it was not produced by the true track contributing the largest fraction of hits to the track studied, as taken from MC truth information. Thus, these misassociated hits could be background hits, i.e., not associated with any track, or could be produced by another true track. In either case, these hits do not belong on the track that is studied and are likely to degrade the track parameter resolution. The aim is to remove as many of these hits as possible from the track fit, while keeping the correctly associated hits. To remove the impact of a hit from the track fit, we disable the hit and refit the track.

`TrkHitFix` preselects tracks with a track fit probability lower than 10^{-5}, which amounts to 13% of all tracks.

In the following, we introduce the variables we use to separate misassociated and correctly associated hits. For hits in both the SVT and in the DCH, we use χ_{hit}, defined as $\chi_{\text{hit}} = r/\sigma_r$, where the residual r is the distance of closest approach between the hit and the fitted track and σ_r its error, measuring how consistent (spatially) this hit is with the track. The sign of r is determined by the angular momentum of the track with respect to the wire in the DCH or the strip in the SVT, respectively. In order to have an unbiased estimate of the residual, the effect of the respective hit is removed from the track fit. As can be seen in Fig. 3.4, χ_{hit} provides statistical separation between correctly associated and misassociated hits. Fig. 3.4 also shows that χ_{hit} for misassociated hits is peaking around 0, indicating a selection bias when hits are assigned to a track. Furthermore, we see that there is a substantially higher fraction of misassociated hits among the SVT hits than there is among the DCH hits. But, χ_{hit} of course does not provide unambiguous information

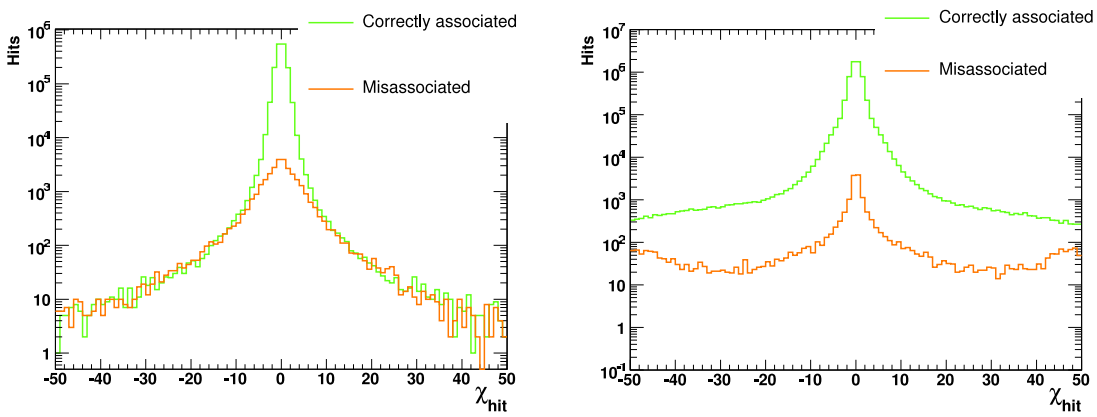

Figure 3.4. χ_{hit}, the spatial residual divided by its uncertainty, for SVT hits (left) and DCH hits (right) that are assigned to tracks. No cuts are applied to the hits or to the tracks.

whether a certain hit was produced by a given track – the track fit might be suffering from other problems and misassociated hits themselves bias the whole fit. We therefore include other quantities when deciding which hits to remove.

The second variable we consider for SVT hits is the time information from the SVT, which is independent of the spatial information from the SVT. Fig. 3.5 shows a much broader distribution in the time residual divided by its uncertainty, $\chi_t = r_t/\sigma_{r_t}$, defined analogously to the spatial χ_{hit}, for misassociated hits than for correctly associated hits. For DCH hits, we instead consider the doca of the hit, that is the distance of closest approach between the fitted track and the sense wire in the cell that was hit. As can be seen in Fig. 3.5, the distribution for correctly associated hits has a sharp drop-off at around 1 cm, which roughly corresponds to an average size of the hexagonal DCH cells. The distribution of correctly associated hits also shows a long tail, which we attribute to badly reconstructed hits or tracks. These could, for example, be hits where the wrong TDC hit out of several in that cell has been chosen and thus does not represent the hit produced by the track. These hits have been studied in more detail in [41]. Effects like this could also be due to hard scattering in the beam pipe with a poor track fit. Yet, as can be seen, a large fraction of the misassociated hits is in the tail of the doca distribution and thus can be removed by a cut, which improves the ratio of correctly associated to misassociated hits.

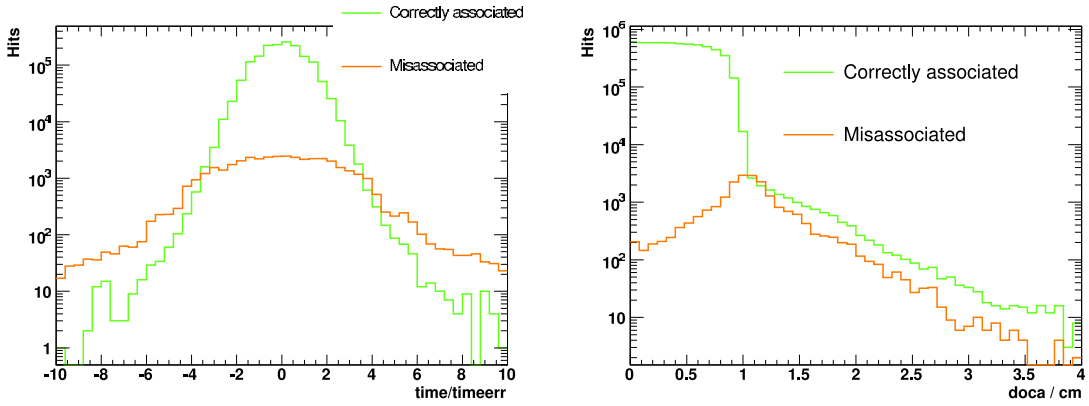

Figure 3.5. The time pull for SVT hits (left) and the doca distribution for DCH hits (right) that are assigned to tracks. No cuts are applied to the hits or to the tracks.

The hit filtering is designed to improve tracks with a few hits that should not have been added to the track. We disable up to 3 hits per track, but keep at least a $1C$ fit.

We combine the information from the different variables into an effective χ^2 and label hits with $\chi^2 > \chi^2_{\text{max}}$ as "bad". This allows us to define an absolute ranking in the quality of the hits and thus to decide which of the hits to remove in case we find more "bad" hits than we can or want to remove. Specifically, this allows us to have a common ranking and χ^2 cut for SVT and DCH hits.

For SVT hits, the χ^2 is defined as

$$\chi^2 = \frac{a^2\chi^2_{\text{hit}} + w^2\chi^2_t}{a^2 + w^2}, \tag{3.1a}$$

where w is the relative weight between the spatial and the time information and $a = (0)\ 1$ to (not) use the spatial information in the χ^2. For DCH hits, χ^2 is defined as

$$\chi^2 = \begin{cases} \infty & \text{if } \text{doca} > \text{doca}_{\text{cut}} \\ s^2\chi^2_{\text{hit}} & \text{otherwise,} \end{cases} \tag{3.1b}$$

where s^2 is the relative scaling between the χ^2 of DCH hits with respect to SVT hits.

Several parameters need to be tuned: the weight w, the scaling s, the parameter a, the cut on χ^2, χ^2_{max}, and the cut on doca, doca_{cut}. This is done using MC hit level truth information, which we use to maximize $\sqrt{\epsilon p}$, where ϵ and p are the efficiency and purity,

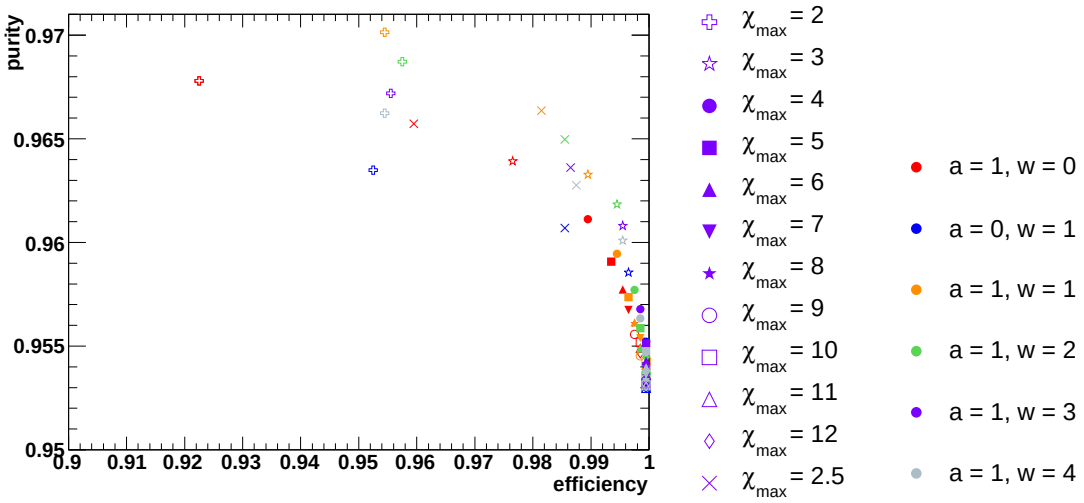

Figure 3.6. Efficiency vs. purity for SVT hits scanning $\chi_{\max}$, the weight w, and the use of the spatial information.

respectively. The efficiency is defined as the probability for a hit to be classified as good hit (i.e., a hit that passes the χ^2 cut) if it was produced by (only) that track, $\epsilon = N_{\text{good}}^{\text{cassoc}}/N^{\text{cassoc}}$. The purity is defined as the probability for a hit to have been produced by (only) that track, given that it passes the χ^2 cut, $p = N_{\text{good}}^{\text{cassoc}}/N_{\text{good}}$. For the tuning of the hit filtering we use a sample of tracks antiselected by the other `TrkFixup` modules to minimize the amount of tracks suspected to suffer from other problems.

Fig. 3.6 shows efficiency vs. purity for SVT hits, scanning $\chi_{\max} = \sqrt{\chi^2_{\max}}$ and the weight w, where points of the same marker color share the same weight w and points of the same marker style share the same cut $\chi_{\max}$. $\sqrt{\epsilon p}$ is maximized for $\chi_{\max} = 3$ and $w = 2$.

Fig. 3.7 shows efficiency vs. purity for DCH hits, scanning $\chi_{\max}$ and the cut on doca, doca_{cut}. For this scan, s is set to 1. $\sqrt{\epsilon p}$ is maximized by $\text{doca}_{\text{cut}} = 1.05\,\text{cm}$, but the scan does not give any reasonable value for $\chi^2_{\max}$, which we attribute to badly reconstructed hits which are associated to the track according to MC truth information. Lacking better information, we pick $\chi_{\max} = 9$ in order to not be sensitive to possible data and MC disagreements in the near tails of the χ_{hit} distribution. This corresponds to $s = 0.33$ if we choose to have a common $\chi_{\max}$ for SVT and DCH hits.

Using MC truth information, we can test if the track fit of a given track improved

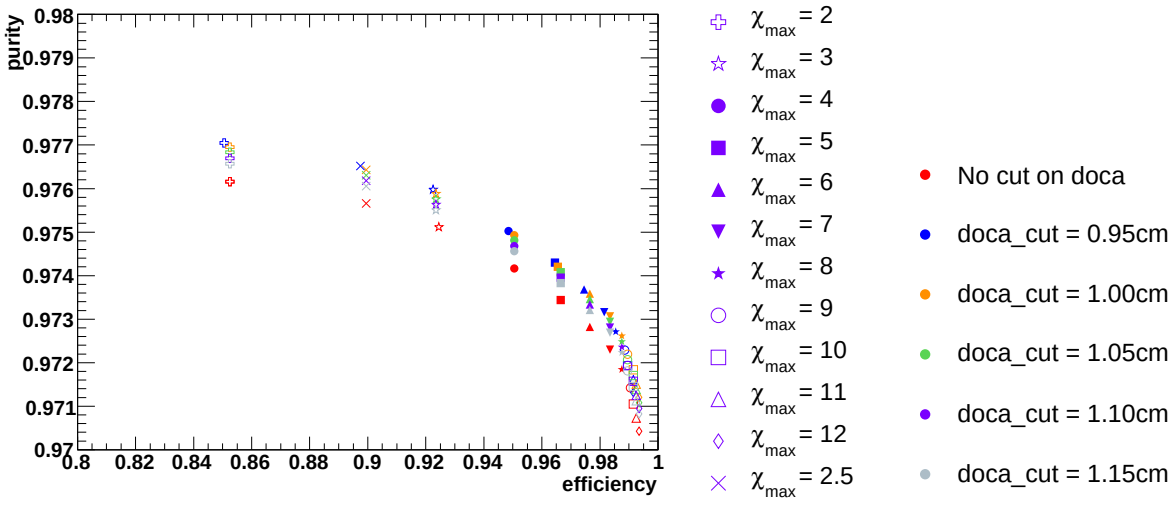

Figure 3.7. Efficiency vs. purity for DCH hits scanning $\chi_{\max}$ (setting $s = 1$ for this scan) and the cut on doca, doca_{cut}.

after disabling "bad" hits and refitting the track. For this purpose, we define the pull in a parameter p as the difference between the fitted parameter, p_{fit}, and the true parameter, p_{true}, divided by the uncertainty given by the fit, σ_p, $\frac{p_{\text{fit}} - p_{\text{true}}}{\sigma_p}$. A track fit is considered to have improved if the pull in the curvature, ω, or the impact parameter in the beam direction, z_0, improves by at least one unit. Fig. 3.8 shows the distribution of the change in the track fit χ^2 before and after the refit over the number of "bad" hits disabled before the refit, $\Delta\chi^2/n_{\text{badhits}}$. The blue distribution corresponds to those track fits where we re-enable the "bad" hits and refit since the fit did not converge after disabling those hits. Disregarding these tracks, `TrkHitFix` improves 40.5% of the track fits it changes. Statistically, we expect the track fit χ^2 to reduce by one unit when removing a random hit from the track. In order to reduce the number of tracks where we change the track fit without improving it significantly, we impose an additional cut. If the track fit does not improve by at least 10 units of χ^2 for each hit we disable, we reset the fit, that is, we re-enable the "bad" hits and refit the track. This ensures that tracks that `TrkHitFix` leaves tracks where the change is not beneficial in their original state. `TrkHitFix` then improves 44% of all track fits it changes. "Bad" hits are found on about 6% of all tracks, so `TrkHitFix` improves about 2.5% of all tracks.

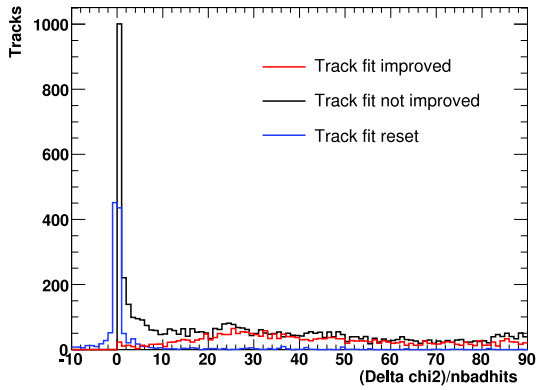

Figure 3.8. The difference in χ^2 per disabled hit before and after refitting for tracks that do improve (red) and tracks that do not improve (black) by hit filtering (see text for definition of improved). The blue distribution comes from track fits that are reset; see text for details.

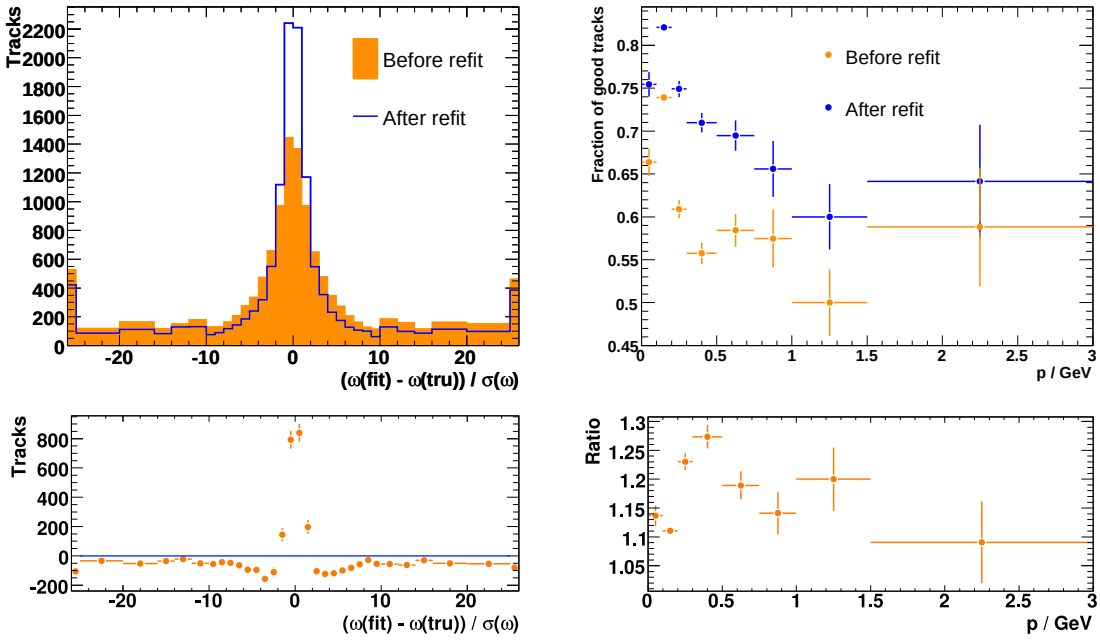

Figure 3.9. The pull in ω before (orange) and after (blue) the refit (top left) and the difference of the two distributions (bottom left). The contents of the underflow and overflow bins is added to the two outermost bins. The fraction of good tracks (for definition, see text) before (orange) and after (blue) the refit as a function of track momentum (top right) and the ratio of the two distributions (bottom right). Only tracks where at least one "bad" hit is disabled are considered for these plots.

In following, we only consider tracks with impact parameter transverse to the beam $|d_0| < 2\,\text{cm}$ which are true π, have a successful refit after disabling hits, and have not been reset because of too little improvement in the χ^2 of the track fit.

Fig. 3.9 shows the ω pull distribution before and after the refit on the left. As can be seen, tracks from the tails of the distribution are pulled into the core. On the right, we show the fraction of good tracks before and after the refit as a function of momentum. For this purpose, a track is defined as good if $|(\omega_{\text{fit}} - \omega_{\text{true}})/\sigma_\omega| < 5$, where ω_{fit} is the fitted value of ω, ω_{true} is the true value of ω, and σ_ω is the uncertainty on ω as determined by the track fit. Depending on the momentum range, the fraction of good tracks improves between a couple to more than 25%; the largest relative improvement is observed for tracks of 300 MeV to 500 MeV. The pull distributions for the other track parameters have been checked and as expected show only marginal improvement.

3.4 The Recoil Method

As described in Sec. 3.1, the $B\bar{B}$ mesons originate from $\Upsilon(4S) \to B\bar{B}$ decays. Since the mass of the $\Upsilon(4S)$ is only about 20 MeV above the $B\bar{B}$ threshold, the momentum of the B mesons in the $\Upsilon(4S)$ frame is only of the order of 300 MeV and the two B mesons overlap completely in the detector. Inclusive analyses, i.e., analyses which do not study one or more B-meson decay channels with specific final state particle content, but rather a whole class of final states, cannot rely on spatial separation between the decay products of the two B mesons in the event.

In the case of inclusive semileptonic $B \to X\ell\nu$ decays, one commonly applies one of two strategies. First, one can measure purely leptonic quantities based on the lepton and neutrino momenta, where the neutrino momentum is inferred from the missing momentum in the event. Second, one can fully reconstruct the non-signal B meson ("tag B") decay in the event, a technique which is sometimes referred to as the recoil method and is widely used at *BABAR*. After fully reconstructing the tag B, all remaining tracks and neutral energy depositions in the event can be assigned to the signal B meson decay. The flavor, charge,

and momentum of the tag B meson are measured and thus the corresponding properties for the semileptonically decaying B meson can be inferred, since the momentum of the $\Upsilon(4S)$ in the laboratory frame is known. In the case of inclusive semileptonic analyses, the recoil method allows for a measurement of kinematic quantities related to the hadronic system X with good resolution.

The tag B, or $B_{\rm reco}$, meson is reconstructed in decays of the type $B_{\rm reco} \to D^{(*)}Y$. Here, the charmed meson $D^{(*)}$ serves as a seed for the $B_{\rm reco}$ reconstruction. Four different seeds are used, D^+ and D^{*+} candidates for the reconstruction of $B^0_{\rm reco}$ mesons and D^0 and D^{*0} candidates for $B^+_{\rm reco}$ mesons. The Y system consists of a number of charged and neutral K and π mesons with an overall charge of ± 1: $Y = n_1\pi + n_2 K + n_3 K_S + n_4\pi^0$, where $n_1 + n_2 \leq 5$, $n_3 \leq 2$ and $n_4 \leq 2$. In events with multiple $B_{\rm reco}$ candidates, the candidate with the highest a priori purity, i.e., the purity as determined on simulated events, is chosen. The reconstruction of $B_{\rm reco}$ candidates and the choice of the best $B_{\rm reco}$ candidate is provided by the centrally produced `BSemiExcl` skim.

Two types of background occur: combinatorial background, that is, where a $B_{\rm reco}$ candidate is assembled from daughters of both B mesons in a $B\bar{B}$ events, and continuum background, where a $B_{\rm reco}$ candidate is assembled from particles in $q\bar{q}$ ($q = u, d, s, c$) events.

Typically, two (mostly uncorrelated) kinematic variables are used for the separation of correctly reconstructed $B_{\rm reco}$ candidates and background: $\Delta E = E_{B_{\rm reco}} - E_b$, the difference of the measured energy of the $B_{\rm reco}$ candidate $E_{B_{\rm reco}}$ and the center-of-mass beam energy E_b, and $m_{\rm ES}$, defined as the invariant mass of the $B_{\rm reco}$ meson, with the measured $B_{\rm reco}$ energy replaced by the center-of-momentum beam energy E_b to improve the resolution.

The shortcoming of the recoil method is the low efficiency of the $B_{\rm reco}$ reconstruction, which is approximately 0.5% (0.3%) for charged (neutral) $B\bar{B}$ events, due to the large number of available B-meson decay channels.

Chapter 4

Measurement of the $B \to X_u \ell \nu$ Hadronic Mass Spectrum

This section details the analysis technique for the measurement of the hadronic mass spectrum in $B \to X_u \ell \nu$ decays. We describe the data and Monte Carlo (MC) samples used in the analysis and give details of the event selection and background subtraction. We furthermore give an overview over the estimation of the systematic uncertainties. The measurement of the hadronic mass spectrum and the estimation of its uncertainties are very similar to the inclusive recoil method $|V_{ub}|$ analysis at *BABAR* [31], with which we collaborate closely.

4.1 Analysis

4.1.1 Data and Simulation Samples

The analysis is based on a total integrated luminosity of $347\,\mathrm{fb}^{-1}$, recorded from Run1 through Run5, at a center-of-mass energy of $m_{\Upsilon(4S)} = 10.58\,\mathrm{GeV}$. This corresponds to approximately 383 Million $B\bar{B}$ pairs.

In addition, the analysis uses two types of simulated MC datasets. The simulation of B-meson decays is performed by the `EvtGen` event generator [34]. Final state radia-

tion is simulated by PHOTOS [42] and the detector simulation is performed by GEANT4 [43]. Corrections and reweightings applied to the MC events are described below.

The data and the MC samples have been reprocessed and analyzed within the *BABAR* software release series 18.

Backgrounds have been estimated on a sample of 1142×10^6 fully simulated and reconstructed $B\bar{B}$ events. In this sample both B mesons decay generically, i.e., without any preselection on their decay modes.

Properties of signal events have been studied on dedicated signal MC samples, where one B meson decays as $B \to X_u \ell \nu$ and the other one decays according to the best known B meson branching fractions. We mix a nonresonant and a resonant description of $B \to X_u \ell \nu$ decays, as both types of decays are known to be present in the data.

The nonresonant signal sample is generated according to an inclusive model based on the triple differential decay rate and the shape function parametrization of [35], using $m_b^{\rm pole} = 4.8\,{\rm GeV}$. In this sample, final state hadrons are produced with a continuous invariant hadronic mass spectrum. The hadronization of the final state is performed by JETSET's [36] parton shower algorithm, which introduces a lower bound of $2m_\pi$ on the hadronic mass.

The resonant signal sample contains charmless semileptonic decays to exclusive final states $B \to X_u \ell \nu$, where $X_u = \pi, \rho, \eta, \eta', \omega$. The decays are simulated according to the ISGW2 model [44].

The two complementary samples are mixed to produce the signal sample used in the analysis. Nonresonant $B \to X_u \ell \nu$ events are reweighted according to the generated values of the hadronic mass (m_X), lepton energy (E_ℓ), and momentum transfer (q^2) such that the three-dimensional differential branching fractions correspond as closely as possible to those obtained from the purely nonresonant sample.

MC Simulation Reweighting and Corrections

The MC samples are reweighted in order to achieve a good modeling of our data. Reweightings are motivated by known differences in particle identification, by updated

knowledge about branching fractions and by observed disagreements between data and MC simulation.

Run-by-Run weights: Running conditions, such as background levels, the DCH voltage, the muon detectors, can change from Run to Run, or even within a given Run period. This is taken into account in the production of the MC datasets. The amount of available MC data roughly reflects the amount of the available data for each Run, up to an overall factor. We use Run-dependent luminosity weights for the generic MC and the signal MC events to improve the relative MC luminosities available per Run.

Particle identification (PID): To correct data-MC differences in particle identification, we apply the PID-tweaking algorithms provided from *BABAR*'s PID group for all particle species. PID-tweaking accepts or rejects additional tracks as a particle of a given species in order to improve the model of particle identification in the MC simulation [45]. The *BABAR* PID group obtains the corrections (in bins of p, θ and ϕ, i.e. the particle's momentum vector) from the study of control samples on data and MC simulation.

***K_L*:** We correct the MC simulation for K_L detection efficiencies, energy deposition and production rate following results obtained in other *BABAR* analyses.

For each reconstructed calorimeter cluster that is matched to a simulated K_L, the energy deposition is corrected by a factor computed by the *BABAR* package `K0LTools` [46]. The K_L detection efficiency is corrected by rejecting reconstructed neutral clusters matched to simulated K_L with a probability calculated as function of the true K_L momentum using the same package. In addition, a correction due to the differences between data and simulation for the K_L production rate was applied, which is based on a study of the K_S production at *BABAR* [47]. Given that such a correction cannot be accomplished by eliminating neutral clusters, a different approach [48] has been employed. We randomly transform some reconstructed clusters matched to simulated K_L into "pseudo-photons" and in this way restore the energy and momentum balance in the event. This is achieved by rescaling the measured energy and momentum of the K_L cluster to the true K_L momentum assuming zero mass.

The probability for a K_L cluster to be transformed into a "pseudo-photon" depends on the K_L momentum, it is 22% for momenta between 0 and 0.4 GeV, 1% for momenta between 0.4 and 1.4 GeV, and 9% for momenta larger than 1.4 GeV.

$B \to X_c \ell\nu$ and D branching fractions and $B \to D^* \ell\nu$ form factors: A good knowledge of the $m_X^{(2)}$ spectrum of the remaining charm background and thus of the exclusive semileptonic branching fractions for $B \to X_c \ell\nu$ decays is needed for the background subtraction. The braching fractions of $B \to X_c \ell\nu$ and D-meson decays are scaled to agree with the latest branching fraction measurements. For the $B \to X_c \ell\nu$ branching fractions, we follow the recommendation of the Semileptonic B Decays Analysis Working Group [49]. However, the reweighting for the decays to broad D^{**} resonances (D_0 and D_1') and nonresonant charmed hadronic states ($D^{(*)}\pi$) are such that their relative ratios are the same as in the MC simulation and that the total charmed semileptonic rate ($\mathcal{B}(B^+ \to X_c \ell\nu) = (10.89 \pm 0.16)\%$ and $\mathcal{B}(B^0 \to X_c \ell\nu) = (10.15 \pm 0.16)\%$) is saturated. D-decay branching fractions are scaled to the most recent measurements [8]. In addition, we reweight $B \to D^* \ell\nu$ events according to the form factor parametrization of Caprini, Lellouch, and Neubert [50] with form factor ratios $R_1(1) = 1.417 \pm 0.061 \pm 0.044$ and $R_2(1) = 0.836 \pm 0.037 \pm 0.022$ and form factor slope $\rho^2 = 1.179 \pm 0.048 \pm 0.028$ [51].

$B \to X_u \ell\nu$ branching fractions and nonresonant model parameters: The branching fractions of exclusive $B \to X_u \ell\nu$ are scaled according to recent measurements [8]. In addition, the nonresonant events are reweighted to correspond to a nonresonant signal model with input parameters from a global fit to kinematic moments in $B \to X_c \ell\nu$ and $B \to X_s \gamma$ decays, $m_b^{\rm pole} = (4.66 \pm 0.041)\,\text{GeV}$ and $\mu_\pi^{2\rm pole} = (0.497^{+0.086}_{-0.072})\,\text{GeV}^2$ [2].

$m_{\rm miss}^2$ in $B \to X_c e\nu$: We find disagreements between the data and MC simulation which we attribute to the modeling of $B \to X_c e\nu$ decays (see App. A). We reweight simulated $B \to X_c e\nu$ events according to their missing mass squared ($m_{\rm miss}^2$) and lepton momentum in the B rest frame to correct for the observed differences.

4.1.2 Event Selection

Events are selected to increase the purity of the resulting signal sample and the resolution in the hadronic mass.

We use the recoil method, described in Sec. 3.4, to reconstruct the non-signal $B_{\rm reco}$ meson. The `BSemiExcl` skim provides the selection of the best $B_{\rm reco}$ candidate. Reconstructing signal decays on the recoil of fully reconstructed hadronic decays yields a good resolution of the hadronic mass and allows for a more efficient discrimination against cascade decays, where the lepton originates from a secondary decay of a charmed meson. We require the $B_{\rm reco}$ candidate to have $|\Delta E|$ consistent with 0 within three standard deviations, where the $|\Delta E|$ resolution is determined separately for the different modes, and we only use $B_{\rm reco}$ decay modes such that the integrated a priori purity of the total sample exceeds 0.2.

Track and Neutral Selections

After the reconstruction of the $B_{\rm reco}$ candidate, the remaining tracks and neutral energy depositions in the event are assumed to be daughters of the other B meson in the event. We select events with an electron or muon candidate and reconstruct the hadronic system X from the remaining particles. We use a set of cuts to reject background, fake and duplicate charged tracks and neutral depositions, which have been optimized by members of the *BABAR* collaboration to achieve good agreement between data and the MC simulation for a number of variables [52]. As the data processed in software release 18 do not yet profit from the `TrkFixup` tracking improvements (see Sec. 3.3.1), fake particles as well as duplicate candidates such as ghost tracks and nonprimary looper branches have to be removed by dedicated sets of cuts.

The track selection is based on the track's distance of closest approach (doca) of the track to the interaction point in the $x-y$ plane ($|d_{xy}|$) and the z direction ($|d_z|$), the track's momentum ($p_{\rm lab}$) and its component transverse to the beam axis ($p_{t,{\rm lab}}$), the number of hits in the SVT and DCH ($N_{\rm SVT}$ and $N_{\rm DCH}$) and the track's polar angle (θ). For pairs of tracks, the selection also relies on the difference of the two tracks' transverse momenta ($\Delta p_{t,{\rm lab}}$)

Select tracks with	Cut
doca in $x-y$ plane	$\lvert d_{xy}\rvert < 1.5\,\mathrm{cm}$
doca in z	$\lvert d_z\rvert < 5\,\mathrm{cm}$
Max. momentum	$p_{\mathrm{lab}} < 10\,\mathrm{GeV}$
Min. transverse momentum	$p_{t,\mathrm{lab}} > 0.06\,\mathrm{GeV}$
Max. momentum for SVT-only tracks	$p_{\mathrm{lab}} < 0.2\,\mathrm{GeV}$ if $N_{\mathrm{DCH}} = 0$
Geometrical acceptance	$0.410 < \theta_{\mathrm{lab}} < 2.54$ rad
Reject tracks if	$\Delta p_{t,\mathrm{lab}} < 0.12\,\mathrm{GeV}$ (loopers), $\Delta p_{t,\mathrm{lab}} < 0.15\,\mathrm{GeV}$ (ghosts) to other tracks and
Loopers ($p_{t,\mathrm{lab}} < 0.25\,\mathrm{GeV}$) ($\lvert\cos\theta\rvert < 0.2$, $N_{\mathrm{SVT}} > 1$)	Same sign: $\lvert\Delta\phi\rvert < 0.18$ & $\lvert\Delta\theta\rvert < 0.2$ Opp. sign: $\lvert\Delta\phi\rvert < 0.16$ & $\lvert\pi - \lvert\Delta\theta\rvert < 0.18$
Ghosts ($p_{t,\mathrm{lab}} < 0.35\,\mathrm{GeV}$) ($N_{\mathrm{DCH}} > 1$)	$\lvert\Delta\phi\rvert < 0.3$ & $\lvert\Delta\theta\rvert < 0.3$
Select neutral energy deposits with	**Cut**
Raw energy	$E_{\mathrm{raw}} > 50\,\mathrm{MeV}$
Min. number of crystals	$N_{\mathrm{crys}} > 2$
LAT shape	LAT < 0.6
Geometrical acceptance	$0.32 < \theta_{\mathrm{lab}} < 2.44$
Min. 3-d angle distance to tracks without associated clusters	$\Delta\alpha > 0.08$
List for track-neutral matching	`ChargedTracks`
Electron list for bremsstrahlung recovery	`PidLHElectrons`

Table 4.1. The selection criteria for tracks and neutral energy depositions. The *BABAR* coordinates are defined in Sec. 3.2 and variables are defined in the text.

and angles ($\Delta\theta$ and $\Delta\phi$). The selection of neutral energy deposits uses the measured energy (E_{raw}), the number of EMC crystals (N_{crys}), the lateral shower shape in the EMC (LAT), the energy deposit's polar angle (θ) and the angular distance to tracks with no associated clusters ($\Delta\alpha$).

The selection criteria for charged and neutral particles are given in Tab. 4.1. The *BABAR* `ChargedTracks` and `CalorNeutral` lists, which are lists of loosely selected charged and neutral particle candidates, serve as input lists.

Particle Identification

The *BABAR* PID group provides standard selectors for the different particles species [53]. The performance of a PID selector for a given particle species s is characterized by its efficiency to correctly identify particles of species s and by its misidentification rate for particles of other species. In this analysis, we rely on the identification of leptons to identify signal events and the identification of Kaons to reject the dominant $B \to X_c \ell \nu$ background.

Electrons We use the `PidLHElectrons` selector to identify electrons. The selector combines information from the DCH, the EMC, and the DIRC into a global likelihood. It relies on the measured energy loss in the DCH, $\mathrm{d}E/\mathrm{d}x$, the shower shape and the number of crystals in which energy was deposited in the EMC, as well as the ratio of the energy deposited in the EMC to the track momentum. Furthermore, it uses the number of Cherenkov photons and the Cherenkov angle measured in the DIRC.

Muons Muons are identified with the `muMicroTight` selector. It uses the energy deposited in the EMC along with information from the IFR, such as the number of interaction lenghts traversed, the match of the extrapolated track and the IFR cluster, and the number of IFR strips hit, and is based on cuts.

Kaons The selection of Kaons uses the `TightKaonMicro` selector. The selector is based on cuts on the measured energy loss in the DCH and the number of Cherenkov photons and the Cherenkov angle in the DIRC.

Event Selection

The semileptonic decay of the signal B meson is detected by the presence of a tagged electron or muon, identified by the `PidLHElectrons` and `muMicroTight` selector, respectively. The lepton momentum is required to exceed 0.5 GeV in the lab frame and to be contained in the fiducial region $0.36 < \theta < 2.37$ to ensure good-quality particle identification. The lepton momentum p^* in the B rest frame is required to be larger than

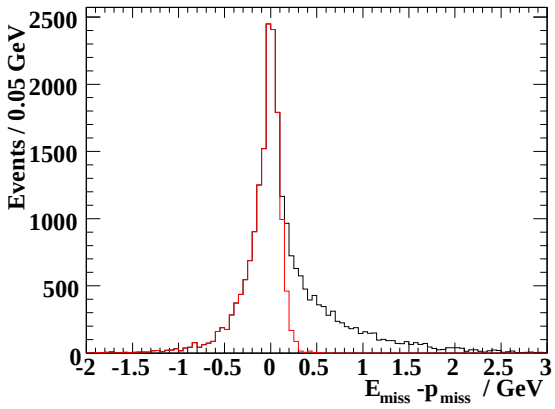

Figure 4.1. $E_{\rm miss} - p_{\rm miss}$ distribution for $B \to X_u \ell\nu$ signal MC simulation. The red distribution is obtained with $m^2_{\rm miss} < 0.5\,{\rm GeV}^2$ (see text).

1 GeV. Events with secondary leptons, i.e. leptons originating from the cascade decay chain $b \to c\ell\nu$, $c \to s\ell\nu$, are suppressed by rejecting events with any additional tagged lepton with $p^* > 1\,{\rm GeV}$. This increases the signal purity as secondary leptons from $b \to u\ell\nu$ decays are rare.

Further, the presence of missing energy ($E_{\rm miss} = E_{\Upsilon(4S)} - E_{B_{\rm reco}} - E_X - E_\ell$) and momentum ($p_{\rm miss} = |\vec{p}_{\Upsilon(4S)} - \vec{p}_{B_{\rm reco}} - \vec{p}_X - \vec{p}_\ell|$) is evidence for a neutrino in the decay. To ensure a good recontruction of this neutrino, we require $E_{\rm miss} - p_{\rm miss} > -0.3\,{\rm GeV}$ (see Fig. 4.1), where the cut value has been optimized to ensure a good resolution in the hadronic mass reconstruction, while having a high signal selection efficiency.

Semileptonic decays to charm states with unreconstructed daughters, which lead to low reconstructed values of the hadronic mass, can be reduced by cutting on the measured missing mass squared, $m^2_{\rm miss} = E^2_{\rm miss} - p^2_{\rm miss}$. Fig. 4.2 shows the $m^2_{\rm miss}$ distribution and its correlation with m_X for $B \to X_u \ell\nu$ and $B \to X_c \ell\nu$ decays. We require $m^2_{\rm miss} < 0.5\,{\rm GeV}^2$, which improves the signal-to-background ratio in the region $0 < m_X < 1.55$ GeV from 0.49 to 0.90.

To suppress events with missing or fake charged particles and improve the hadronic mass resolution, we require the total charge of the reconstructed event to be zero. Finally, the charge correlation between the lepton and the charge of the $B_{\rm reco}$ is checked. Events

with wrong sign correlation are rejected from the charged B sample, while we keep both the right and wrong sign events for the neutral B sample to correct for B meson mixing (see Sec. 4.1.5).

Events with tagged $K^{\pm}$ and reconstructed K_S in the hadronic system are also removed from the sample, since Kaons on the signal side mostly originate from $B \to X_c \ell\nu$, $D \to KY_s$, with Y_s consisting of leptons and/or pions. Charged Kaons are identified by the `TightKaonMicro` selector and K_S are reconstructed as $K_S \to \pi^+\pi^-$. Furthermore, we apply a partial $D^{*\pm}$ reconstruction based on the tag of a soft $\pi^{\pm}$. Making use of the small energy release in $D^{*\pm} \to D^0\pi^{\pm}$ decays the $D^{*\pm}$ kinematics is inferred from the soft $\pi^{\pm}$. The missing mass squared of the event is computed with the assumption that the signal side decay is $B \to D^*\ell\nu$. Neutral B events in which this missing mass squared is larger than $-3\,\mathrm{GeV}^2$ are rejected from the sample. This sample, to which the background-suppression cuts are applied, is referred to as the signal-enriched sample. The cumulative efficiencies of the different selection cuts are summarized in Tab. 4.2.

Events with an identified Kaon, a reconstructed K_S, or a partially reconstructed $D^{*\pm}$ in the hadronic system which pass all other selection cuts are used as the the signal-depleted sample, which is used for controlling the shape of the hadronic mass spectrum in background events and the background normalization (see Sec. 4.1.5).

While the reconstruction of the hadronic final state is performed inclusively, it is important to ensure that known exclusive final states are reconstructed with reasonable accuracy. In particular, $B \to \pi^{\pm}\ell\nu$ events where additional particles are mistakenly added result in a single-sidedly distorted resolution function. We consider events where the X system is accompanied by a soft neutral particle. In order to recover genuine $B \to \pi^{\pm}\ell\nu$ and other decays with low hadronic mass in the final state, we test whether the softest neutral particle is compatible with being a B_{reco} daughter if the neutral's energy is less than 250 MeV. This is performed by adding the neutral particle to the B_{reco} system and testing whether the invariant B_{reco} mass moves closer to the nominal B meson mass. If this is the case, we remove the neutral particle from the X system. The energy distribution of the excluded

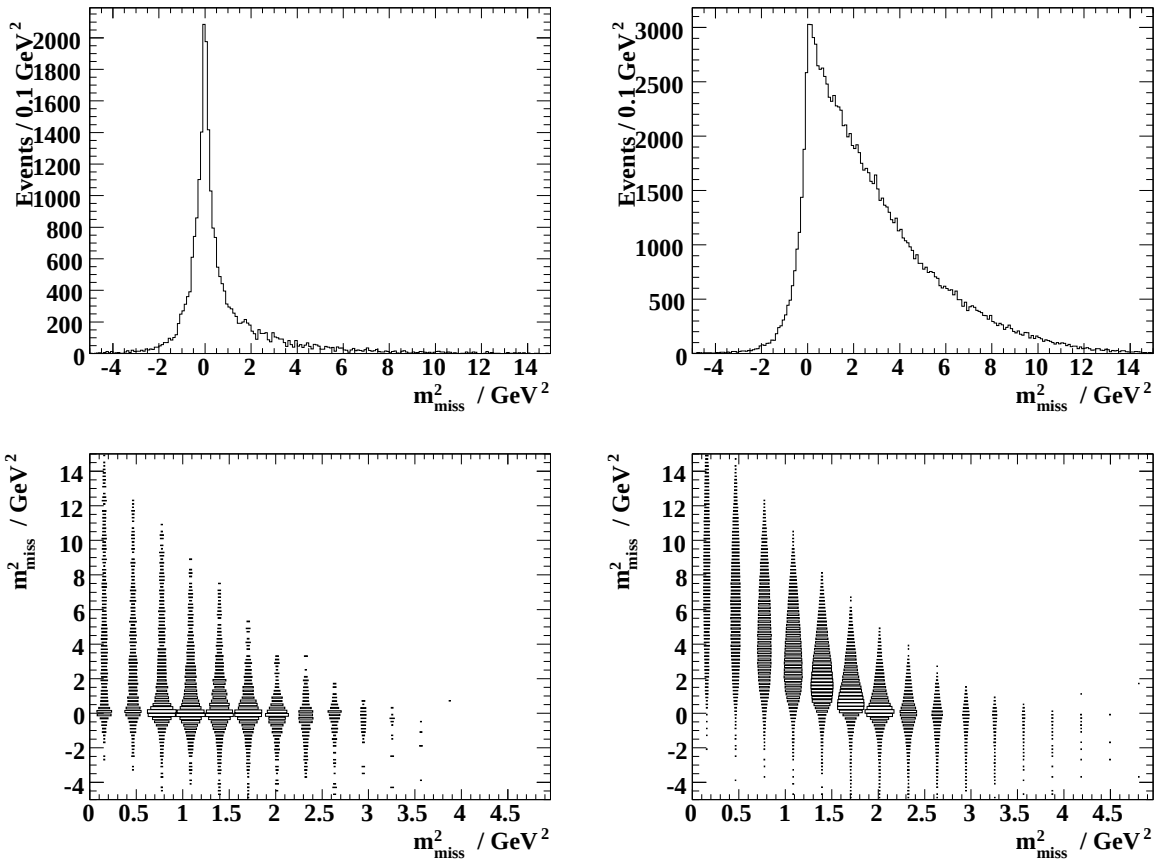

Figure 4.2. $m^2_{\rm miss}$ in the analysis: distribution for $B \to X_u \ell\nu$ (top left) and and $B \to X_c \ell\nu$ (top right), correlation of m^2_{miss} with m_X for $B \to X_u \ell\nu$ (bottom left) and $B \to X_c \ell\nu$ (bottom right).

Cut	$B \to X_u \ell\nu$ MC	$\epsilon_{\rm sig}$	$B \to X_c \ell\nu$ + other MC	$\epsilon_{\rm bkgd}$
Sl + $B_{\rm reco}$	52334	1.000	500283	1.000
$\lvert Q_{tot}\rvert = 0$	31275	0.598	263641	0.527
Exactly one lepton	31010	0.593	259438	0.519
Kaon Veto	25943	0.496	138881	0.278
D^* Veto	24943	0.477	123634	0.247
$m^2_{\rm miss}$	16270	0.311	26634	0.053
$E_{\rm miss} - p_{\rm miss}$Cut	13997	0.267	26634	0.053

Table 4.2. Cut efficiencies: given are the numbers of events (normalization given by the size of the MC sample used in this study) for $B \to X_u \ell\nu$ and generic MC events passing various cuts and the efficiencies of these cuts on the two samples. The cuts are applied consecutively. The semileptonic selection consists of the cuts on lepton momentum in the B rest frame and the lab frame, the angular acceptance cuts, and the charge-flavor correlation requirement. Charge and mixing correction (see Sec. 4.1.5) have not been applied.

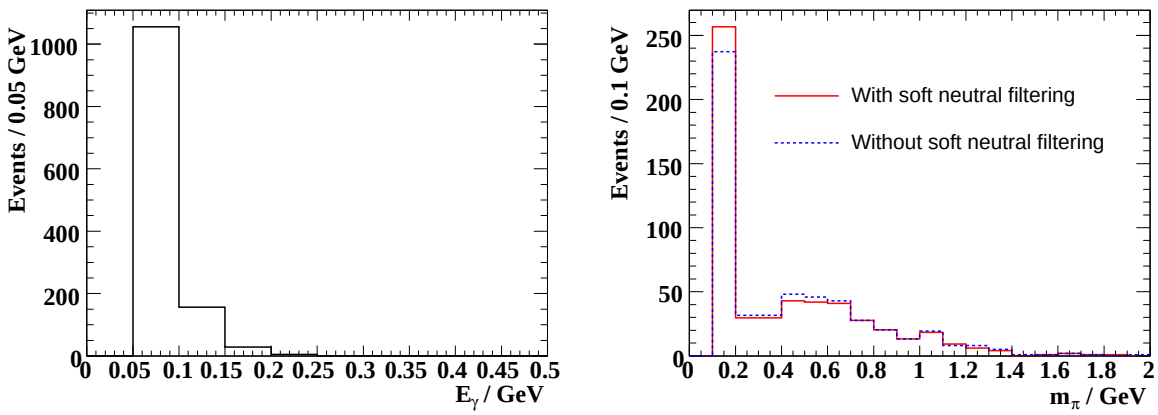

Figure 4.3. Soft neutral filtering: the energy spectrum of neutrals removed with the soft neutral filtering procedure discussed in the text (left). m_X spectrum for true $B \to \pi^{\pm}\ell\nu$ decays before (blue) and after (red) the soft neutral filtering (right).

neutral particles and the effect of this procedure on the hadronic mass spectrum in true $B \to \pi^{\pm}\ell\nu$ events is shown in Fig. 4.3.

4.1.3 Subtraction of Combinatorial $B_{\rm reco}$ and Continuum Background

After applying the selection on the signal side of the event, the combinatorial background on the tag side is subtracted in bins of m_X and m_X^2 spearately for the charged B sample and the right- and wrong-sign neutral B samples. This background mostly originates from $B_{\rm reco}$ candidates built from daughters of both B mesons, and from continuum events. The background subtraction is performed by a fit to the distribution of the energy-substituted mass $m_{\rm ES}$, defined as the invariant mass of the B meson, where the measured B energy is replaced by the center-of-momentum beam energy E_b, which is known more precisely,

$$m_{\rm ES} = \sqrt{E_b^2 - \left(\sum_i \vec{p}_i\right)^2}. \tag{4.1}$$

The $\vec{p}_i$ denote the three-momenta of the particles forming the $B_{\rm reco}$ candidate in the center-of-momentum system. For the fit, the $m_{\rm ES}$ distribution is modeled by the sum of two functions which describe the background contribution and the signal component. The background is described by an ARGUS function [54], which provides a good parametrization of the shapes of the combinatorial and continuum background. It depends on two parameters:

an upper cutoff value $m_{\rm max}$ and a shape parameter ξ,

$$\frac{{\rm d}N(m_{\rm ES})}{{\rm d}m_{\rm ES}} = N m_{\rm ES}\sqrt{1-x^2}e^{-\xi(1-x^2)}, \tag{4.2}$$

with $x = m_{\rm ES}/m_{\rm max}$. The value of $m_{\rm max}$ is kept fixed at $m_{\rm max} = 5.2785\,{\rm GeV}$, for which we observe good convergence of the fit. The $B_{\rm reco}$ signal is parametrized by a CRYSTAL BALL function (CB) [55], which has four parameters,

$$\frac{{\rm d}N(m_{\rm ES})}{{\rm d}m_{\rm ES}} = \begin{cases} N e^{-\frac{1}{2}\frac{(m_{\rm ES}-m)^2}{\sigma^2}} & \text{if } \; m_{\rm ES} > m - a\sigma \\ N\left(\frac{n}{a}\right)^n e^{-\frac{1}{2}a^2}\frac{1}{\left(-\frac{m_{\rm ES}-m}{\sigma}+\frac{n}{a}-a\right)^n} & \text{if } \; m_{\rm ES} < m - a\sigma, \end{cases} \tag{4.3}$$

where m is the peak position, σ the width of the Gaussian distribution, a determines the transition point from the main Gaussian distribution to the tail, and n describes the shape of the tail. Smaller values of n generate a longer tail, which accounts for energy loss in the EMC showers from photons used in the π^0 reconstruction.

Due to significant correlations between the fit parameters and limited statistics in the bins at large hadronic mass, the analysis uses a three-step approach for the fits to the $m_{\rm ES}$ distributions. The CB shape parameters are independently determined by binned χ^2 fits to the $m_{\rm ES}$ distribution integrated over all hadronic mass bins and B charges, and fixed for the subsequent fits (see Fig. 4.4). The procedure is performed independently for data, the $B \to X_u \ell\nu$ MC sample, the $B \to X_c \ell\nu$ MC sample, and the non-semileptonic background MC sample. In step (A), the ARGUS shape parameter ξ and the overall normalization of the ARGUS function are obtained from a fit restricted to the range $5.22\,{\rm GeV} < m_{\rm ES} < 5.255\,{\rm GeV}$, which does not contain any $B_{\rm reco}$ signal. In step (B), the peak position m and width σ of the CB peak are determined by fitting the sum of a Gaussian and an ARGUS function in the range $5.274\,{\rm GeV} < m_{\rm ES} < 5.2791\,{\rm GeV}$, with the ARGUS function fixed to the parameters determined in step (A). Finally, the fit is repeated in step (C) on the full $m_{\rm ES}$ range to the sum of a CB and an ARGUS function, where the a and n parameters and the normalization of the signal function are floated, while the peak position and width are taken from step (B) and the ARGUS shape and normalization are taken from step (A).

In the rest of the analysis, the CB shape parameters are fixed to the values extracted

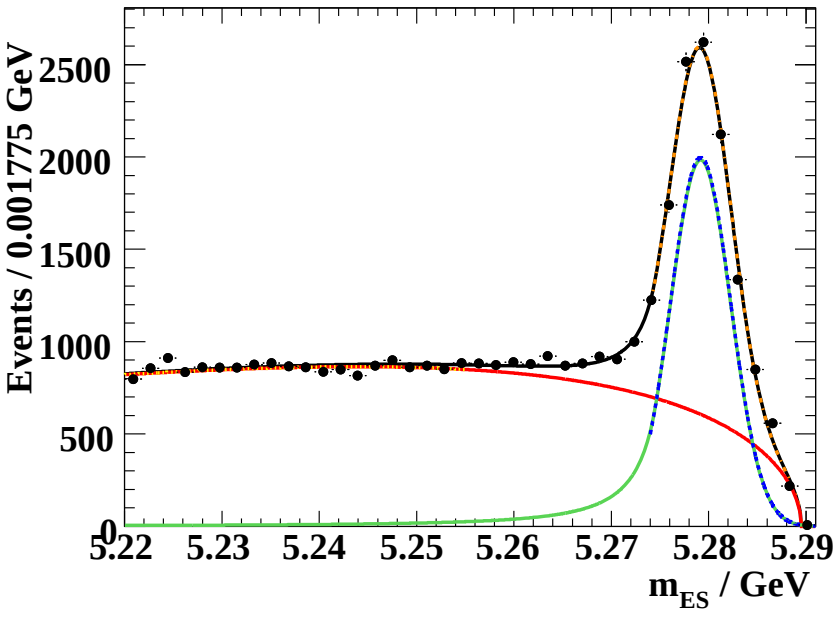

Figure 4.4. Extraction of CB shape parameters: $m_{\rm ES}$ spectrum with the various fit components. Step (A): the dotted yellow curve shows the ARGUS function in its fit interval ($5.22\,\mathrm{GeV} < m_{\rm ES} < 5.255\,\mathrm{GeV}$). Step (B): the dash-dotted orange curve shows the sum of Gaussian and ARGUS function used to determine the mean and width of the signal peak (the signal contribution is shown by the blue dashed Gaussian and the background by the red ARGUS function) in its fit interval ($5.274\,\mathrm{GeV} < m_{\rm ES} < 5.2791\,\mathrm{GeV}$). Step (C): the black curve which runs below the other functions shows the result of the fit to the sum of the CB and ARGUS function, where the CB normalization and the a and n parameters are floated. The solid red and green curves show the ARGUS and CB components of the fit to the full $m_{\rm ES}$ spectrum.

here, while the ARGUS and CB normalizations, as well as the ARGUS shape parameter[1] are floated.

To understand which $B_{\rm reco}$ candidates contribute to the peak in $m_{\rm ES}$, we develop a truth matching procedure on MC simulation as discussed in Sec. 4.1.4. The binned χ^2 fits to the $m_{\rm ES}$ distributions for the four samples integrated over hadronic mass and B charge, which we use for this purpose and which demonstrate the quality of the fits we use for the determination of the CB shape parameters, are shown in Fig. 4.5.

The final fits to the $m_{\rm ES}$ distributions are performed as unbinned extended maximum likelihood fits in bins of m_X and m_X^2, respectively, and yield spectra corrected for combinatorial and continuum background. These fits are performed for three classes of events: charged B, right- and wrong-sign neutral B candidates. As a cross-check, we perform the identical fits as binned χ^2 fits and find good agreement between the extracted mass spectra.

The number of $B_{\rm reco}$ signal events can be extracted from this fit with two different procedures: One can use the area of the fitted CB function in the $m_{\rm ES}$ signal region, $5.27\,{\rm GeV} < m_{\rm ES} < 5.2785\,{\rm GeV}$, or, alternatively, the number of signal events can be obtained by subtracting the area under the fitted background ARGUS function in the $m_{\rm ES}$ signal region from the total number of events in the $m_{\rm ES}$ signal region. In principle, both methods have some advantages and we study the statistical properties of the results from both procedures. We perform pseudo-MC experiments, which evaluate the expected combined statistical and systematic uncertainty on the yields for a high-signal-to-background sample and a low-signal-to-background sample, as well as the goodness of the error estimate. We find very similar results for the two procedures on both samples. For the subsequent analysis we use the second procedure, where we subtract the area of the ARGUS function from the total number of events in the signal region, to be less sensitive to our modeling of the $B_{\rm reco}$ signal component.

[1]The shape of the ARGUS function is correlated with the hadronic mass on the signal side, since the chances of building a $B_{\rm reco}$ candidate from daughters of both B mesons is larger and the average energy per signal B daughter is smaller in events with larger multiplicity (i.e., larger hadronic mass) on the signal side.

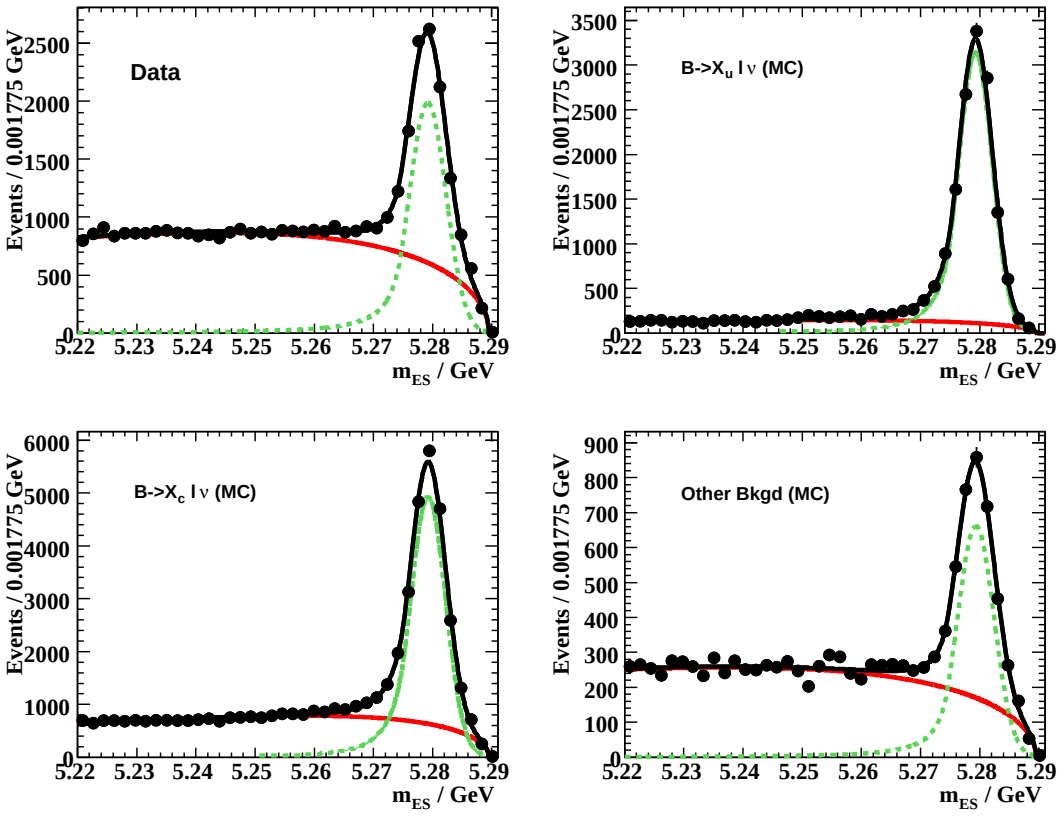

Figure 4.5. Result of the $m_{\rm ES}$ fits for data (top left), signal $B \to X_u \ell\nu$ (top right), background $B \to X_c \ell\nu$ (bottom left), and other backgrounds (bottom right). The green dashed and red solid curves show the CB and ARGUS functions, respectively.

4.1.4 $B_{\rm reco}$ Truth Matching

In inclusive analyses that use the recoil method, there is often no clear-cut definition of "signal" and "background". Misreconstructed $B_{\rm reco}$ decays affect the reconstruction of the hadronic system X when one or more daughters of one B are assigned as daughters of the other B, or when daughters of one or both B are lost during the reconstruction. The $m_{\rm ES}$ distribution of those events with "almost correctly" reconstructed $B_{\rm reco}$ candidates exhibits a broad component peaking at the B mass. These $B_{\rm reco}$ candidates make up what is usually referred to as "peaking background". A study on simulated $B \to X_u \ell \nu$ events shows that there is a class of events for which not all particles are correctly associated to their parent B meson, but which still provide good resolution of the reconstructed hadronic mass. Rejecting these peaking background events introduces a significant bias in the high-mass end of the spectrum, where the larger particle multiplicity makes it more probable for particles to be swapped between the two B mesons by the reconstruction. This motivates a loose definition of well-reconstructed $B_{\rm reco}$ decays.

We define the $B_{\rm reco}$ signal events based on the shape of their cumulative $m_{\rm ES}$ distribution, i.e., as those events contributing to the peaking component in $m_{\rm ES}$. We study the properties of these events using $B \to X_u \ell \nu$ MC events. We define the the following quantities: the number of generated and reconstructed daughter particles of the $B_{\rm reco}$, $n_{\rm gen}$ and $n_{\rm reco}$, divided according to their charge ($n^{\rm chg}_{\rm gen}$, $n^{\rm chg}_{\rm reco}$, $n^{\rm neu}_{\rm gen}$, $n^{\rm neu}_{\rm reco}$), and the number of the reconstructed charged and neutral daughters which are truth matched, $n^{\rm chg}_{\rm tm}$ and $n^{\rm neu}_{\rm tm}$. For the truth matching of the daughters, we use the standard *BABAR* definitions, where a true particle is truth matched to a reconstructed particle if it contributed the majority of the reconstructed hits in the detector. We use $m^{\rm chg} = n^{\rm chg}_{\rm reco} - n^{\rm chg}_{\rm tm}$, $m^{\rm neu} = n^{\rm neu}_{\rm reco} - n^{\rm neu}_{\rm tm}$, $l^{\rm chg} = n^{\rm chg}_{\rm gen} - n^{\rm chg}_{\rm tm}$, and $l^{\rm neu} = n^{\rm neu}_{\rm gen} - n^{\rm neu}_{\rm tm}$ as measures of how well the $B_{\rm reco}$ candidate was reconstructed.

We only accept $B_{\rm reco}$ candidates that were generated in one of the reconstructed modes as potential signal candidates. We find that $m^{\rm neu} < 3$, $l^{\rm neu} < 3$, and $m^{\rm chg} = l^{\rm chg} = 0$ give the best agreement of the resulting $B_{\rm reco}$ signal $m_{\rm ES}$ distribution with the CB function obtained

from the fit to the full sample (see Fig. 4.5). The agreement is improved if we also accept events in the $m_{\rm ES}$-ΔE signal region, defined by $m_{\rm ES} > 5.27\,{\rm GeV}$ and $|\Delta E| < 0.026$ GeV, independent of their values of $m^{\rm neu,chg}$ and $l^{\rm neu,chg}$. The events that fail the cuts on $m^{\rm chg,neu}$ and $l^{\rm chg,neu}$, but lie in the $m_{\rm ES}$-ΔE signal region, account for approximately 8.6% of the selected $B_{\rm reco}$ signal events. They show a slightly degraded hadronic mass resolution, but their inclusion into the $B_{\rm reco}$ signal sample ensures a better estimate of the hadronic mass resolution function. Furthermore, the remaining $B_{\rm reco}$ background events are free of peaking components. The $m_{\rm ES}$ shape of the events classified as background is in good agreement with the ARGUS function obtained from the fit to the full sample.

Fig. 4.6 shows the $m_{\rm ES}$ distribution for truth matched $B_{\rm reco}$ signal events according to our truth matching definition, with the CB function with the parameters obtained from the fit to the full $m_{\rm ES}$ spectrum. The complementary $B_{\rm reco}$ background sample is compared to the ARGUS function from the same fit. The shape of the $B_{\rm reco}$ signal and background components is well reproduced by the CB and ARGUS functions.

The performance of our method (labeled as "0") is compared to possible variations of the truth-matching criteria (see Tab. 4.3). We consider the following alternative definitions:

1. $m^{\rm chg,neu} = l^{\rm chg,neu} = 0$, i.e., perfect daughter match.

2. Reconstructed $B_{\rm reco}$ decay mode identical to generated $B_{\rm reco}$ decay mode.

3. $m^{\rm neu} < 3$, $l^{\rm neu} < 3$ and $m^{\rm chg} = l^{\rm chg} = 0$, without additional events from the $m_{\rm ES} > 5.27\,{\rm GeV}$ and $|\Delta E| < 0.026$ GeV signal region.

The χ^2 measure the agreement between the CB and ARGUS shapes obtained from the fits shown in Fig. 4.5 and the truth matching method. While a large value of χ^2 does inform us about some mismatch in the $m_{\rm ES}$ shape obtained from the truth matching, the number of signal events is reproduced very well.

In addition, we study the sensitivity of our truth matching to the choice of the ΔE cut. We move the cut by $\pm 3\,{\rm MeV}$ and find that the χ^2/ν (where ν is the number of degrees

Method	$\frac{\text{Signal Evts.}-\text{CB Integral}}{\text{CB Integral}}$	$\frac{\text{Bkgd Evts.}-\text{ARGUS Integral}}{\text{ARGUS Integral}}$	Signal χ^2/ν	Bkg. χ^2/ν
0	$+0.004$	-0.099	2.73	1.02
1	-0.195	$+2.383$	62.2	19.3
2	-0.139	$+1.717$	8.95	12.4
3	-0.082	$+1.034$	4.64	6.15

Table 4.3. Performance of different possible truth matching definitions for the B_{reco}: we compare how well the different algorithms perform in selecting the correct number of events as B_{reco} signal and background in the m_{ES} signal region. We also test how well the shapes of the truth matched m_{ES} distributions match the shapes from the fit.

of freedom) for signal (background) changes from 109/40 (41/40) to 110/40 (46/40) and 110/40 (45/40).

The m_X and m_X^2 resolution is studied for truth matched B_{reco} signal events and compared to that of the fully truth matched decays according to method 1, which could be considered the standard defintion of truth matching. We first determine the width of the core of the resolution function by computing a truncated root mean square (rms) within $-0.5\,\text{GeV} < m_X^{\text{reco}} - m_X^{\text{gen}} < 0.7\,\text{GeV}$. The rms is $0.197\,\text{GeV}$ with our truth matching, to be compared with $0.180\,\text{GeV}$ for the fully truth matched events. We also study the tails by computing the fraction of events which are outside the window given above and find 4.8% and 4.3% of the events in the tails of the resolution function using our truth matching and the full truth matching, respectively. In addition, we verify that our truth matching definition provides a more stable scaling of the reconstruction efficiency as a function of the generated value of m_X. The detector response matrices extracted from $B \to X_u \ell\nu$ MC simulation that are used for the unfolding are shown in Fig. 4.7.

4.1.5 Signal-Side Background Subtraction

The measured m_X and m_X^2 spectra obtained after applying the analysis cuts and the m_{ES} fits are shown in Fig. 4.8 for the signal-enriched and the signal-depleted sample. Due to the finite resolution and the large yield of $B \to X_c \ell\nu$ decays, these still represent the

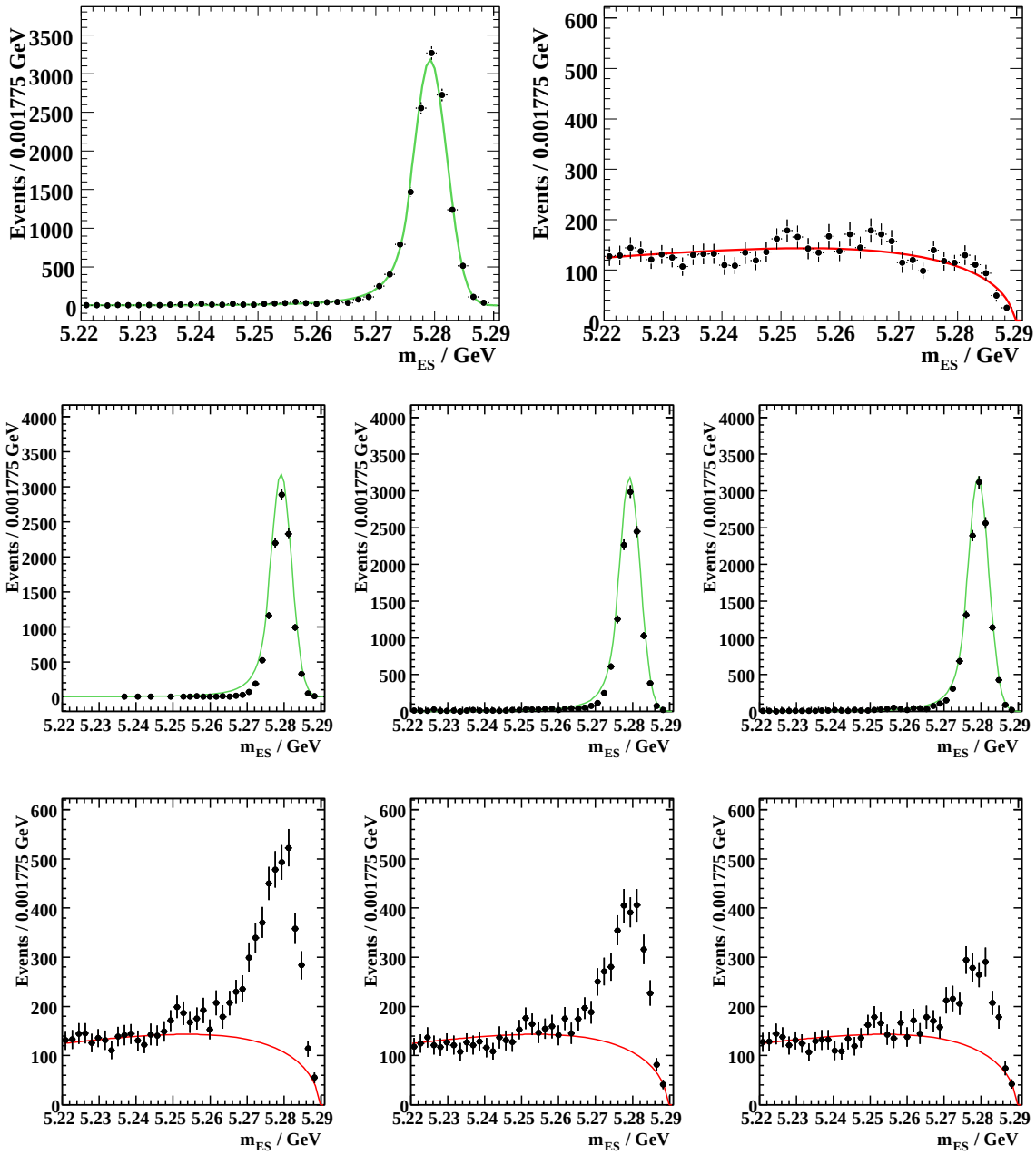

Figure 4.6. Performance of the truth match algorithms on $B \to X_u \ell \nu$ MC events: the top histograms show signal (left) and background (right) for our definition. The lower histograms show the signal and backgrounds for the methods 1 (left), 2 (center), and 3 (right) discussed in the text. The curves show the CB and ARGUS functions as obtained from a fit to the full $m_{\rm ES}$ spectrum.

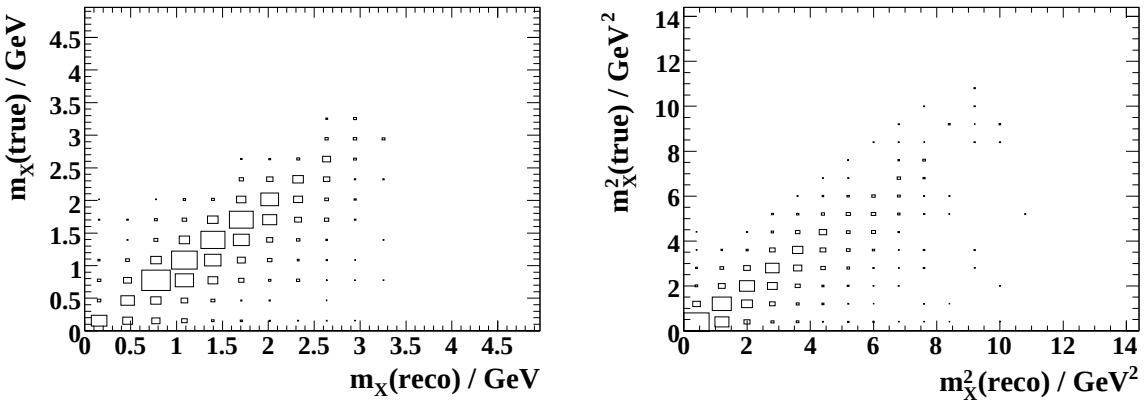

Figure 4.7. $m_X^{(2)}$ resolution: The detector response matrices for $B \to X_u \ell\nu$ events for m_X (left) and m_X^2 (right) as determined on signal MC simulation with the generated mass on the ordinate and the reconstructed mass on the abscissa. The bin content is given by the number of events in our $B \to X_u \ell\nu$ MC sample.

dominant component of the signal-enriched spectra, even in the region $m_X < m_D$. This remaining charm background, which survives the veto cuts, and the other backgrounds, including misidentified leptons, $B \to X\tau\nu \to X' l\nu\nu$ events, and secondary charm decays, which survive the lepton charge and signal B flavor correlation requirement, needs to be removed. This is achieved by a fit to the m_X spectrum for which the m_X shapes of the signal, the charm background, and the other backgrounds are taken from the MC simulation.

First, the measured spectra are corrected for $B^0\bar{B}^0$ mixing effects. The number of events (N_B) originating from direct B decays can be related to the number of right-sign (N_{rs}) and wrong-sign (N_{ws}) events, each taken after the $m_{\rm ES}$ fits, through

$$N_B = \frac{1-\chi_d}{1-2\chi_d} N_{rs} - \frac{\chi_d}{1-2\chi_d} N_{ws}, \tag{4.4}$$

where $\chi_d = 0.188$ is the B_d mixing parameter. This allows for the subtraction of background cascade decays ($b \to c\ell\nu$, $c \to s\ell\nu$ events) under the assumption that this is the only class of background events in the wrong-sign sample. This correction is applied on a bin-by-bin basis. We then correct the ratio of neutral-to-charged B events in MC simulation to reproduce the ratio observed on data to account for possible differences in the composition of the $B_{\rm reco}$ sample in MC simulation and data. The correction is obtained by performing

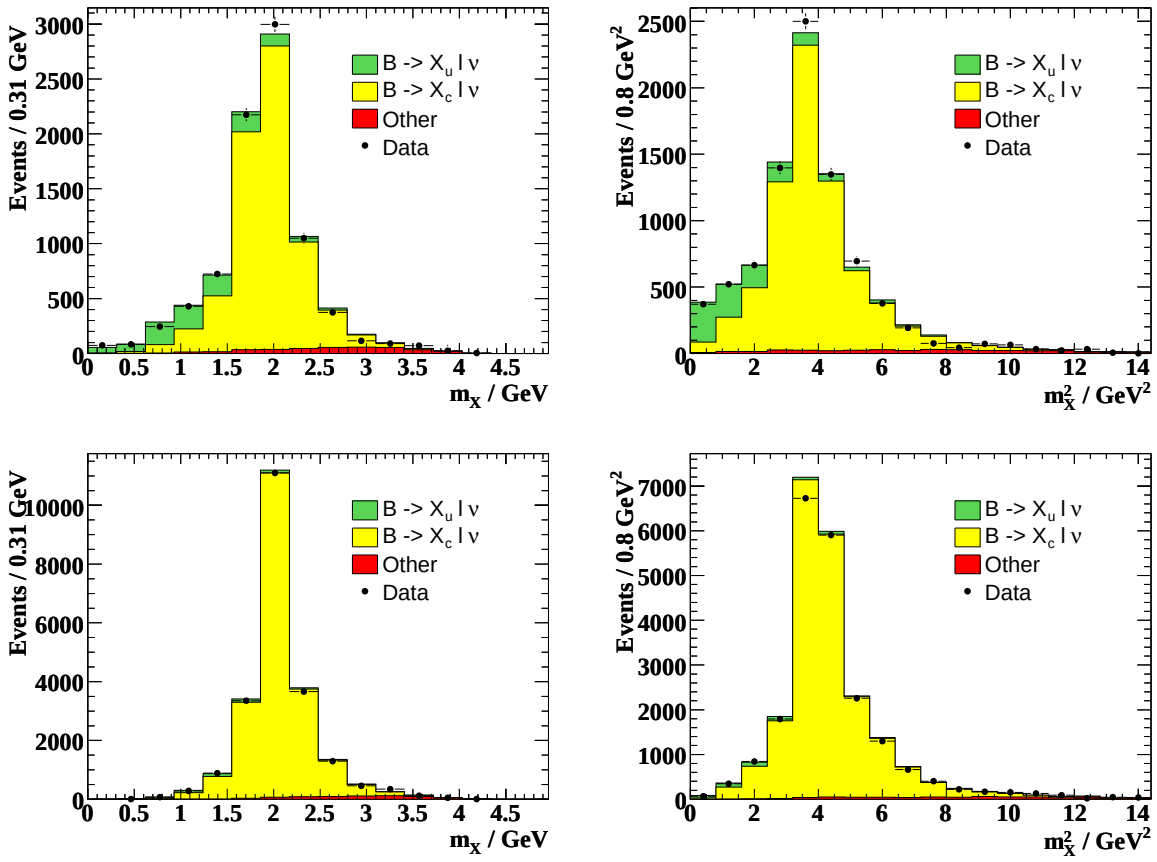

Figure 4.8. Hadronic mass spectra: m_X (left) and m_X^2 (right) distributions of the signal-enriched sample (top) and the signal-depleted sample (bottom) after analysis cuts and subtraction of combinatorial and continuum background, with the contributions of signal (green), $B \to X_c \ell \nu$ background (yellow), and other sources of background (red).

an additional $m_{\rm ES}$ fit to charged and neutral right-sign events for the full hadronic mass range. Since the data at this stage are dominated by $B \to X_c \ell \nu$ decays, the correction is obtained from the ratio of neutral-to-charged events in data and $B \to X_c \ell \nu$ MC events and applied to all MC samples.

For the fit to the m_X spectrum, we denote the content of bin i in the m_X spectra by $N_i^{\rm meas}$, $N_i^{u{\rm MC}}$, $N_i^{c{\rm MC}}$, and $N_i^{o{\rm MC}}$, for the data sample, the $B \to X_u \ell \nu$ MC sample, the $B \to X_c \ell \nu$ MC sample, and the other backgrounds MC sample, respectively. The total number of MC events in bin i is then given by

$$\mu_i = C_u N_i^{u{\rm MC}} + C_c N_i^{c{\rm MC}} + C_o N_i^{o{\rm MC}}, \tag{4.5}$$

where the C_x, $x = u, c, o$, are the relative scalings between the different MC samples, which depend on the composition of the data spectrum and the relative size of the different MC samples. The C_x are determined by a binned χ^2 fit with

$$\chi^2(C_u, C_c, C_o) = \sum_i \left(\frac{N_i^{\rm meas} - \mu_i}{\sqrt{(\delta N_i^{\rm meas})^2 + (\delta \mu_i)^2}} \right)^2. \tag{4.6a}$$

The number of measured events, $N_i^{\rm meas}$, is obtained from the fits to the $m_{\rm ES}$ distributions after applying the mixing and charge corrections per bin i, and $\delta N_i^{\rm meas}$ and $\delta \mu_i$ are the propagated uncertainties from the $m_{\rm ES}$ fits for data and the MC samples, where

$$(\delta \mu_i)^2 = (C_u \delta N_i^{u{\rm MC}})^2 + (C_c \delta N_i^{c{\rm MC}})^2 + (C_o \delta N_i^{o{\rm MC}})^2. \tag{4.6b}$$

The fit is performed using a wide first bin in m_X with $m_X < 1.55\,{\rm GeV}$, in order to reduce the sensitivity of the fit to the modeling of the hadronic mass distribution in the signal MC simulation. The fit result on the signal-enriched sample is shown in Fig. 4.9. We also perform a fit to the signal-depleted sample and compare the extracted scaling factors (see Tab. 4.4 and Fig. 4.9). We find acceptable agreement between the fits to the signal-enriched and depleted samples. Since the fit to the depleted sample yields a more precise determination of the background scaling factors, we combine the results for the subsequent analysis. The uncertainties of the C_x are propagated to the background-subtracted spectra as described in Sec. 4.2.2. In addition to the separately obtained and combined results for the C_x, Tab. 4.4 also presents scaling factors C_x^* for the combined fits, which are corrected

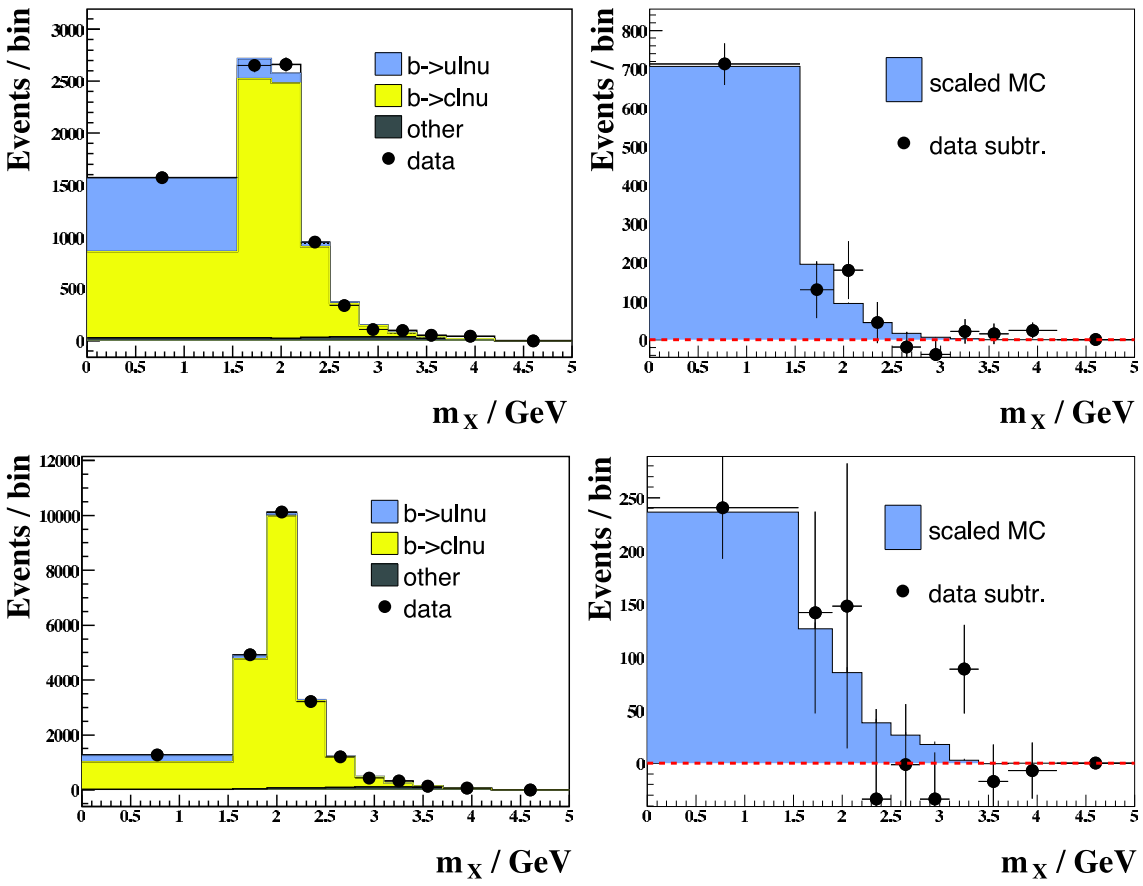

Figure 4.9. m_X spectrum with large first bin: m_X distributions of the signal-enriched (top) and signal-depleted (bottom) sample after analysis cuts and subtraction of combinatorial and continuum background (left) and all backgrounds (right).

	signal-enriched sample	signal-depleted sample	Combined
C_u	0.085 ± 0.007	0.197 ± 0.044	0.088 ± 0.007
C_c	0.372 ± 0.009	0.365 ± 0.004	0.368 ± 0.004
C_o	0.156 ± 0.072	0.267 ± 0.052	0.228 ± 0.042
C_u^*			0.860 ± 0.068
C_c^*			1.056 ± 0.011
C_o^*			0.665 ± 0.121

Table 4.4. Fitted values for the C_x coefficients on the signal-enriched and -depleted m_X spectrum and their combinations. The C_x^* are rescaled by the relative luminosities between the data and MC samples.

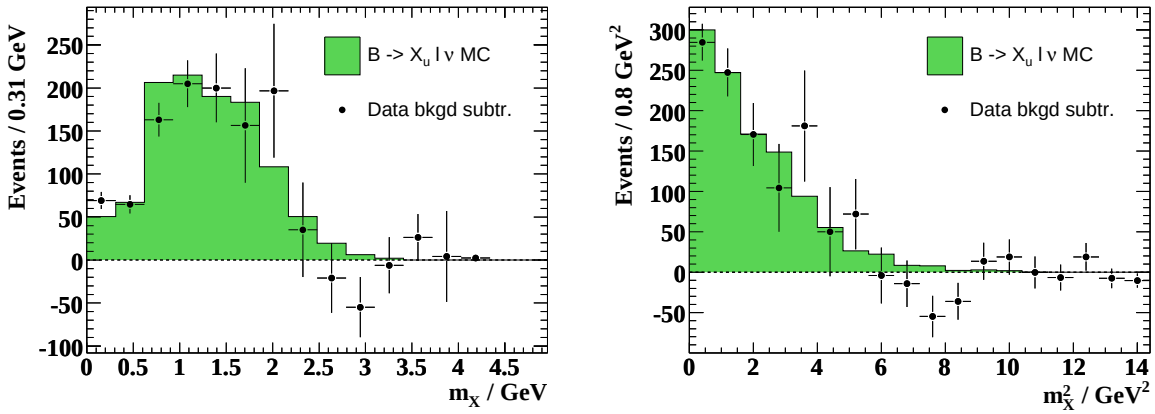

Figure 4.10. Hadronic mass spectra: m_X (left) and m_X^2 (right) distribution of the signal-enriched sample after analysis cuts and background subtraction from data (points) and MC simulation (green). The error bars show the statistical uncertainties from the $m_{\rm ES}$ fits.

for the luminosity ratios of the data and MC samples. The C_x^* are expected to be consistent with 1 and are computed as a cross-check of our fit procedure. Given the assumptions that need to be used to compute the C_x^*, such as equal efficiencies for the $B_{\rm reco}$ reconstruction in data and MC simulation, estimates of effective weights from all reweighting procedures and approximations in taking into account the enrichement in signal events, the agreement of the C_x^* with 1 is satisfactory.

The signal-side background subtraction of the equidistantly binned m_X and m_X^2 spectra uses the combined result for the C_x from the fits to the spectra with the large first bin.

After analysis cuts and $m_{\rm ES}$ fits we find 8454 (8421) events in the full data set when binning in m_X (m_X^2). The sum of the backgrounds is estimated to be 7413 (7394), which corresponds to 1041 (1027) candidate signal events. The background-subtracted spectrum is in good agreement with that predicted for $B \to X_u \ell \nu$ signal events from the simulation and is shown in Fig. 4.10.

4.2 Systematic Uncertainties

Systematic uncertainties on the measured spectrum and the detector response matrix, which is used in the unfolding of the spectrum (see Sec. 5), are introduced from uncertainties

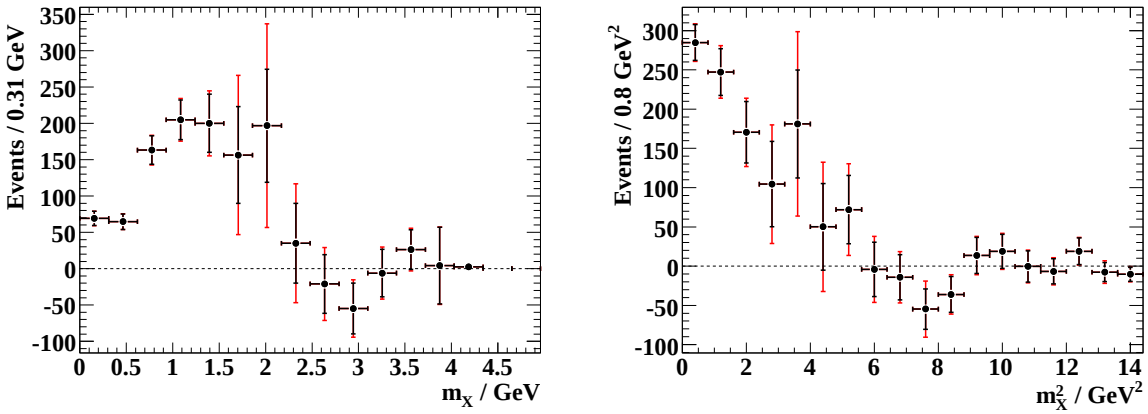

Figure 4.11. Hadronic mass spectra: m_X (left) and m_X^2 (right) spectra. The outer error bars give the total error including the systematic contribution. The statistical contribution to the uncertainty is shown by the shorter error bar.

in our modeling of the track and neutral reconstruction and particle identification, the fits used for background subtraction, the modeling of signal and backgounds, and observed disagreements between data and MC simulation.

The systematic uncertainties on the measured spectrum are evaluated and its covariance matrix C is computed as the sums of the individual covariance matrices arising from the different systematic errors, $C = \sum_k C_k$. To evaluate the C_k, we vary the respective input to the analysis and compute the covariance matrix from the default measured spectrum b and the spectra obtained with the systematic variations b^l as

$$C_{k,ij} = \begin{cases} \frac{\sum_l (b_i^l - b_i)(b_j^l - b_j)}{N-1} & \text{if } N > 2 \\ \frac{(b_i^1 - b_i) - (b_i^2 - b_i)}{2} \frac{(b_j^1 - b_j) - (b_j^2 - b_j)}{2} & \text{if } N = 2 \\ (b_i^1 - b_i)(b_j^1 - b_j) & \text{if } N = 1 \end{cases} \tag{4.7}$$

The systematic uncertainties considered in this analysis are listed in Tab. 4.5. The measured spectra with full uncertainties are shown in Fig. 4.11 and their correlation matrices are shown in Fig. 4.12.

Systematic uncertainty	Sec.	Influence on	Category
Tracking efficiency	4.2.1	b, A	Detector
Neutral efficiency	4.2.1	b, A	Detector
Lepton identification	4.2.1	b, A	Detector
Charged Kaon identification	4.2.1	b, A	Detector
K_L reconstruction	4.2.1	b,A	Detector
Fit to the $m_{\rm ES}$ distribution	4.2.2	b	Fit
$B_{\rm reco}$ truth matching	4.2.2	A	Fit
m_X fit uncertainty	4.2.2	b	Fit
$B \to D^{(*,**)} l\nu$ branching fractions	4.2.3	b	Bkgd modeling
D branching fractions	4.2.3	b	Bkgd modeling
K_S veto	4.2.3	b, A	Bkgd modeling
Nonresonant signal decay model	4.2.4	b, A	Signal modeling
Resonant signal branching fractions	4.2.4	b, A	Signal modeling
$s\bar{s}$ popping	4.2.4	b, A	Signal modeling
$B \to X_c e\nu$ $p^* - m^2_{\rm miss}$ reweighting	4.2.5	b	Data-MC
Neutral multiplicity	4.2.5	b,A	Data-MC

Table 4.5. Systematic uncertainties on the measured m_X^2 spectrum and detector response matrix. We note whether a particular uncertainty affects the measured spectrum ("b") or the detector response matrix ("A"), and group the uncertainties into categories.

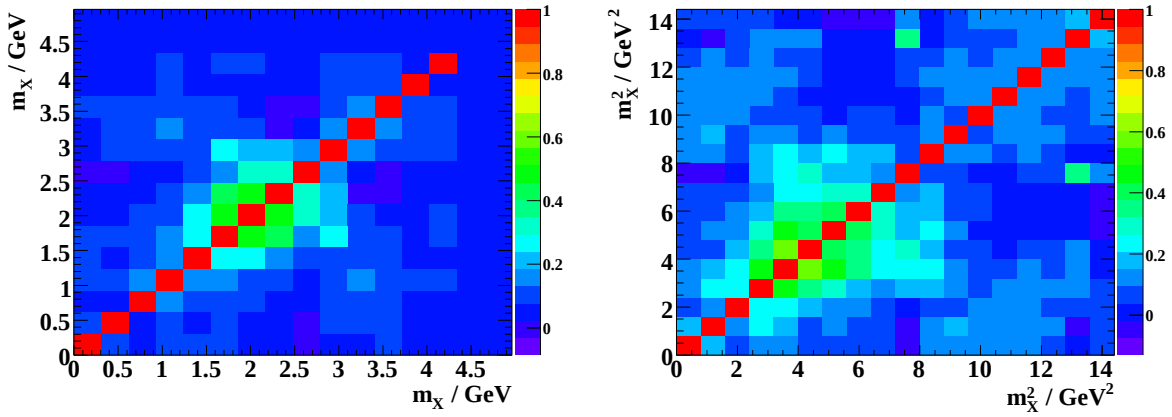

Figure 4.12. The correlation matrices for the m_X (left) and m_X^2 (right) spectra.

4.2.1 Detector-Related Uncertainties

Tracking Reconstruction Efficiency

Studies carried out on $e^+e^- \to \tau^+\tau^-$ events have shown that tracking efficiencies are well reproduced by the simulation. The *BABAR* Charged Particles Reconstruction Group [56] assigns a systematic uncertainty on the tracking efficiency as a function of the Run period. We estimate the systematic uncertainty from the tracking efficiency modeling in MC simulation by removing tracks[2] in the simulation with a Run-dependent probability. The corresponding covariance matrix is computed with $N = 1$.

Neutral Reconstruction Efficiency

The Neutral Reconstruction Group [57] studies the modeling of the neutral reconstruction using control samples. They recommend no corrections to be applied to single photons in the MC simulation. We estimate the effect of the modeling uncertainty in the neutral reconstruction efficiency on our spectrum by randomly killing neutral clusters in the simulation with a probability of 1.8% and compute the covarianz matrix with $N = 1$.

Lepton and Charged Kaon Identification

To evaluate uncertainties from particle identification, we conservatively vary the electron and charged Kaon efficiencies in MC simulation by 2% and the muon efficiencies in MC simulation by 3% based on earlier results with control samples. The variations are implemented by randomly removing identified particles with the given probabilities. The misidentification rates in MC simulation are varied by 15% for all three species. We have $N = 1$ to compute the covariance matrix.

[2]Tracks can only be removed, not added. To compute the covariance matrix, we symmetrize the effect we see from removing tracks. The analoguous is true for the estimation of systematics associated with the neutral reconstruction efficiency and the charged particle identification.

K_L Reconstruction

We apply corrections to the modeling of the K_L detection efficieny and energy deposition in MC simulation as well as to the MC K_L production rate as described in Sec. 4.1.1. We take half the difference from turning all corrections off as estimate of the systematic uncertainty and compute the covariance matrix with $N = 1$.

4.2.2 Background Subtraction Uncertainties

Fit to the m_{ES} Distribution

When performing the fits to the m_{ES} distribution in each single m_X^2 bin, the CB shape parameters are fixed to the values obtained in the dedicated m_{ES} fits to the full m_X^2 range, which are described in Sec. 4.1.3. To estimate the systematic error related to this choice of the CB shape parameters we vary their values separately for the different parameters and separately for the data and three MC samples ($B \to X_u \ell\nu$, $B \to X_c \ell\nu$, and other backgrounds) within the uncertainty quoted by the dedicated fits to the entire m_X^2 region. We have $N = 2$ for each shape parameter and sample to extract the separate covariance matrices.

B_{reco} Truth Matching

The truth matching strategy for the B_{reco}, which is used for the determination of the detector response matrix, is described in Sec. 4.1.4. We assess the uncertainty of the truth matching procedure on the detector response matrix by varying the ΔE cut by $\pm 10\,\mathrm{MeV}$, which corresponds to about a third of the resolution in ΔE.

m_X Fit Uncertainty

To propagate the uncertainty on scaling factors for the charm and other backgrounds (C_c and C_o) to the m_X^2 distribution, we vary C_c and C_o according to their uncertainties obtained from the fit to the m_X spectrum (see Tab. 4.4). While these uncertainties are

statistical, we propagate them like the systematic uncertainties to take the corresponding bin-by-bin correlations into account (with $N = 2$).

4.2.3 Background Modeling Uncertainties

$B \to D^{(*,**)}(\pi)\ell\nu$ Branching Fractions and $B \to D^*\ell\nu$ Form Factors

To estimate the uncertainty on the measured spectrum induced by the uncertainty on $B \to X_c\ell\nu$ branching fractions through the background subtraction, we perform $N = 100$ toy experiments in which we independently vary the different exclusive $B \to X_c\ell\nu$ branching fractions and the inclusive $B \to X_c\ell\nu$ branching fraction randomly within their uncertainties [49].

We find that that removing the $B \to D^*\ell\nu$ form factor reweighting has a negligible effect on the m_X^2 spectrum. We do not assign a systematic uncertainty associated with the choice of the form factor model and parameters.

D Branching Fractions

The measured spectrum depends on the exclusive D branching fractions assumed in the MC production through the background subtraction. To evaluate the uncertainties originating from this source, we perform $N = 200$ toy experiments in which we randomly and independently vary the D branching fractions within their respective uncertainties [8].

K_S Reconstruction

We depend on the simulation of the K_S reconstruction through the veto on reconstructed K_S. Effects due to the difference in K_S production rate are estimated by randomly removing K_S candidates from the K_S list with probability of 10% for K_S momenta between 0 and 10 GeV [47] ($N = 1$). We take this estimate to cover the uncertainties in the K_S efficiency.

4.2.4 Signal Modeling Uncertainties

Nonresonant Signal Decay Model

The input parameters to the nonresonant de Fazio-Neubert [35] model are $m_b^{\rm pole}$ and $\mu_\pi^{\rm 2pole}$. To assess the uncertainties introduced into the mass spectrum and detector response matrix, we take those two points on the $m_b^{\rm pole} - \mu_\pi^{\rm 2pole}$ ellipse ($m_b^{\rm pole} = (4.66 \pm 0.041)\,{\rm GeV}$ and $\mu_\pi^{\rm 2pole} = (0.497^{+0.086}_{-0.072})\,{\rm GeV}^2$ with $\rho = 0.17$ [2]) that yield the largest uncertainties on the partial branching fraction in the $|V_{ub}|$ analysis using the recoil method [31].

We also use two alternative parametrizations of the shape function: instead of the default exponential form, we use a Gaussian parametrization and the so-called Roman parametrization, while keeping the moments of the shape function fixed.

We have $N = 2$ for the variation of each parameter and of the shape function parametrization for the extraction of the separate covariance matrices.

Resonant Signal Branching Fractions

We vary the resonant signal branching fractions within their current uncertainties [8] and use hence have $N = 2$ to obtain the corresponding covariance matrix. The $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ branching fractions are treated as fully correlated and the $B \to \eta\ell\nu$ and $B \to \eta'\ell\nu$ branching fractions are also treated as fully correlated.

$s\bar{s}$ Popping

In signal event decays, a $s\bar{s}$ pair can be created from the vacuum, a process which is referred to as $s\bar{s}$ popping in this context. Since we veto on events in which we identify a charged Kaon or reconstruct a K_S to suppress charm background, we introduce a systematic uncertainty into our signal modeling due to the uncertainty in the $s\bar{s}$ popping. To evaluate this systematic uncertainty, signal events where a gluon splits into a $s\bar{s}$ pair are varied by $\pm 30\%$ in nonresonant signal events ($N = 2$).

4.2.5 Uncertainties from Data-MC Differences

$B \to X_c e\nu$ $p^* - m^2_{\mathrm{miss}}$ Reweighting

As a conservative estimate of the systematic effect related to the disagreement we see between data and MC simulation in the electron sample in m^2_{miss} (see App. A), we take the difference between the m^2_X spectra with and without applying the $p^* - m^2_{\mathrm{miss}}$ reweighting as systematic uncertainty ($N = 1$).

As a cross check, we open the m^2_{miss} cut to $m^2_{\mathrm{miss}} < 1\,\mathrm{GeV}^2$ and take the difference to the default cut ($m^2_{\mathrm{miss}} < 0.5\,\mathrm{GeV}^2$) as an estimate of the systematic uncertainty (correcting for the different selection efficiency by only taking into account variations in the shape of the measured spectrum). The systematic uncertainties on the moments of the unfolded spectrum (see Sec. 5.4) we obtain from this test are comparable to the default method (6% on M_1, 16% on U_2, 46% on U_3).

Neutral Multiplicity

The neutral particle multiplicity distribution exhibits a slightly larger mean value compared to the estimate from MC simulation (see Fig. A.2). We find $\chi^2/\nu = 2$. As the shift seems to be systematic, we estimate the associated uncertainty by rescaling the signal MC events according to their neutral particle multiplicity to get $\chi^2/\nu = 1$ and take the difference in the result ($N = 1$).

Chapter 5

Spectral Unfolding and Measurement of Moments

5.1 The Unfolding Procedure

The measured spectrum of a physical observable, like the invariant mass or the lepton energy, is usually distorted by detector effects, such as finite resolution and limited acceptance. A comparison of the measured spectrum with that predicted by theory requires a removal of these effects to obtain the true, underlying physical spectrum. This can be achieved by applying an unfolding procedure. There are several ways to achieve the unfolding of detector effects on measured spectra, and examples can be found in [58] and [59] and references therein. In this section we describe the unfolding method adopted in this analysis, which largely follows the technique proposed in [58].

A physical quantity α, distributed according to its probability density function $f(\alpha)$, cannot be measured perfectly. Apart from statistical uncertainties, there will be effects from reconstruction efficiency and finite resolution of the detector. The reconstruction efficiency ε_{rec} is the probability to measure an occurring event, and it is in general less than one, $\varepsilon_{\text{rec}} < 1$. It may depend on the particle direction $\vec{p}$, if the detector does not cover the full solid angle. This is usually referred to as geometrical acceptance. The reconstruction

efficiency may also depend on the kinematical properties of the event, like the momentum of a particles to be measured. Limited resolution means that α can only be measured within some device-dependent accuracy. As a result, instead of the true, physical variable α, a variable β, distributed according to some distribution $g(\beta)$, is measured. The relation between $f(\alpha)$ and $g(\beta)$ can be expressed as a convolution of the true distribution $f(\alpha)$ with a kernel $\hat{A}(\alpha, \beta)$ describing the detector effects,

$$\int f(\alpha)\hat{A}(\alpha, \beta)\mathrm{d}\alpha = g(\beta).$$

Since the spectrum we intend to unfold is given as an histogram, we use vectors and matrices for the formulation of the convolution,

$$\hat{A}x = b, \tag{5.1}$$

where the i^{th} component of the n_x-dimensional vector x and the n_b-dimensional vector b contain the number of entries in bin i of the true and the measured distributions, respectively. $\hat{A}$ is a $(n_b \times n_x)$-dimensional matrix, and contains the detector effects. Here, $\hat{A}$ accounts for both effects. An event with a true value in bin j might be measured with a value in bin i (i.e., finite resolution) or might not be measured at all (i.e., $\varepsilon_{\mathrm{rec}} < 1$). The matrix element $\hat{A}_{ij}$ represents the probability for an event with a true value in bin x_j to be measured with a value in bin b_i.

Assuming that the measurement process is well simulated, $\hat{A}$ can be determined from MC events by tracking the true and reconstructed values for each event. A well-defined system of linear equations is obtained,

$$\hat{A}x^{\mathrm{ini}} = b^{\mathrm{ini}}. \tag{5.2}$$

The index "ini" denotes the use of MC spectra x^{ini}, b^{ini}. Technically, the matrix element $\hat{A}_{ij}$ is determined by taking the number of events that fall into bin j of x^{ini} and at the same time into bin i of b^{ini}, and by dividing this number by the number of events in bin j of x^{ini}. Being now in the possession of $\hat{A}$ and a measured spectrum b, one can try to solve Eq. (5.1) for the true spectrum x. However, the apparently easiest way to determine x, i.e., applying $x = \hat{A}^{-1}b$, is not adequate. Even when $\hat{A}$ can be inverted, statistical fluctuations

in the measured spectrum introduce spurious, nonphysical oscillations in the solution for x. Therefore, a more efficient method needs to be applied, and in the following we discussion the algorithm adopted here.

5.1.1 Singular Value Decomposition

One possible way to overcome these difficulties is provided by the method of Singular Value Decomposition (SVD). Any real $(n_b \times n_x)$ matrix, $\hat{A}$, can be decomposed into a diagonal $(n_b \times n_x)$ matrix S with non-negative elements and two orthogonal matrices U and V , being $(n_b \times n_b)$ and $(n_x \times n_x)$-dimensional, respectively,

$$\hat{A} = USV^T, \quad \hat{A}^{-1} = VS^{-1}U^T. \tag{5.3}$$

The diagonal elements of S, $s_i \equiv S_{ii} \geq 0$, are called singular values of $\hat{A}$. The columns of U and V are called left and right singular vectors. The rank of $\hat{A}$ is equal to the number of its nonzero singular values. The solution of Eq. (5.1) may be difficult, even if $\hat{A}$ formally has full rank, namely in the case when $\hat{A}$ and/or b are only known with some level of precision, and at the same time some singular values are significantly smaller than others. These problems can be treated with the help of the SVD.

We assume that the singular values are non-increasing for increasing i. This can always be achieved by swapping pairs of singular values and simultaneously swapping the corresponding columns of U and V. Additionally, we presume that $n_b \geq n_x$; if necessary, rows of zeroes can be added to the inital matrix $\hat{A}$.

Since U is an orthogonal matrix, its columns form an orthonormal system of vectors, which is a basis in the n_b-dimensional space. Hence, the n_b-dimensional vector b can be represented as a linear combination of these basis vectors,

$$Ud = b, \tag{5.4a}$$

and the coefficients of the decomposition are given by the vector d. Analogously, x is decomposed into the orthonormal vectors given by the columns of V, and the coefficients

of the decomposition form the vector z,

$$Vz = x. \tag{5.4b}$$

Together with Eqs. (5.1) and (5.3), this yields a diagonal system

$$z = V^T x = V^T \hat{A}^{-1} U d = S^{-1} d, \text{ and thus } \quad z_i = \frac{d_i}{s_i} \tag{5.4c}$$

for the unknown vector z. The inverse S^{-1} is obtained by inverting the singular values s_i.

At this point, the reason for largely fluctuating solutions for x becomes obvious. In most problems, $\hat{A}$ is singular or almost singular: some singular values s_i are significantly smaller than others. The non-invertability of a matrix effectively means that the induced transformation leads to a loss of information. For the detector response matrix $\hat{A}$ the two physical sources for this loss of information are limited acceptance and finite resolution. Structures in the physical distribution x cannot be resolved if they are smaller than the detector resolution. Hence, small singular values s_i (i.e., large i) are related to these fine structures, or in other words, to quickly oscillating terms in the orthogonal decompositions of x and b if we draw on the analogy to a Fourier decomposition. This means that for reasonably smooth measured distributions b we only anticipate the first k singular vectors to have statistically significant coefficients d_i. Contributions from higher oscillating – again in analogy to Fourier decomposition – basis vectors $i > k$ are expected to be compatible with zero within the statistical errors on d_i. Yet, the weight of just these d_i is enhanced by the small s_i, as shown in Eq. (5.4c). Therefore, the spurious oscillations of the solution are introduced by statistically insignificant coefficients d_i.

Rescaling

The exact solution of a well-behaved linear system does not change when the equations are multiplied by a constant. However, if the detector response matrix is (almost) singular and $n_b \geq n_x$, the linear system is (almost) overdetermined and should be treated as a least-squares problem,

$$\sum_i \Big((\hat{A}x)_i - b_i \Big)^2 \to \min.$$

Any rescaling might change its solution. By choosing a suitable rescaling, we can ensure that the significant information does not get suppressed, while insignificant information does not get enhanced. Höcker and Kartvelishvili [58] propose to rescale Eq. (5.1) by x_j^{ini}, if the detector matrix $\hat{A}$ is determined by MC simulation,

$$w_j = \frac{x_j}{x_j^{\mathrm{ini}}}, \quad A_{ij} = \hat{A}_{ij} x_j^{\mathrm{ini}} \tag{5.5a}$$

(no summation over indices implied), leading to the new system

$$Aw = b. \tag{5.5b}$$

The new vector w is a measure of how much the unknown vector x deviates from the MC truth vector x^{ini}. The matrix element A_{ij} contains the number of events that were generated in bin j and reconstructed in bin i. We use A (i.e., not $\hat{A}$) for the unfolding of the hadronic mass spectrum.

For an exact solution, Eqs. (5.1) and (5.5b) are equivalent, but for the type of problems considered here, this rescaling will improve the behavior of the system. First, if the initially generated MC distribution x^{ini} is reasonably close to the true distribution x, w will be smooth and hence requires less terms in the orthogonal decomposition. A more accurate unfolding should thus be possible, since fewer unknowns need to be determined for the solution.

The second argument given in [58] is based on formal considerations using perturbation theorems for the singular values. This is meant to account for intrinsic errors in the matrix. If the detector response matrix A is determined on MC events, some of its elements will contain only very few events. In the probability matrix $\hat{A}$ the according elements can contain very large values close to one and thus give a large weight to the respective equations, not reflecting that they actually have a comparably large error because of low statistics. Similarly, elements in $\hat{A}$ that are statistically well-determined might be significantly smaller than one. Thus, the number-of-events form of the detector response matrix A leads to a better balanced system, which gives larger weights to equations with smaller errors as far as the detector response is considered. It can be shown that if the initial MC sample statistics are at least one or two orders of magnitude higher than the data statistics, and

if the number-of-events form A of the detector matrix is used, the error on the unfolded spectrum will be dominated by the error on b.

Since in general the errors for the b_i are unequal, different equations have different significance for the solution of the problem. Hence, we should rather consider a weighted least-squares problem,

$$\sum_i \left(\frac{(Aw)_i - b_i}{\Delta b_i} \right)^2 \rightarrow \min, \tag{5.6}$$

with Δb_i being the error on b_i, and no correlations between different b_i are assumed. To obtain a balanced system with respect to the errors on b, according to Eq. (5.6), each equation is divided by the error of b_i ($\tilde{b}_i = b_i/\Delta b_i$, $\tilde{A}_{ij} = A_{ij}/\Delta b_i$), leading to

$$\tilde{A}w = \tilde{b}. \tag{5.7}$$

The covariance matrix of $\tilde{b}$ is by construction identical to the unit matrix I.

5.1.2 Regularization

With these rescalings, the problem of spurious oscillations is not solved yet, though their amplitudes might decrease. To suppress them further, we employ a priori knowledge about the solution. Technically, this is achieved by adding a regularization term to the expression being minimized,

$$\left(\tilde{A}w - \tilde{b}\right)^T \left(\tilde{A}w - \tilde{b}\right) + \tau \left(Cw\right)^T \left(Cw\right) \rightarrow \min, \tag{5.8}$$

where C is a matrix reflecting the a priori condition on w. The Lagrange multiplier τ determines the weight given to this condition in the minimization. Effectively, this expression favors solutions w that solve the linear system (5.7), and also fulfill the additional condition to have Cw small. By choosing C adequately, the small s_i causing the oscillations will be regularized. Under the assumption that the solution x does not differ too much from the simulated true distribution x^{ini}, w should be reasonably smooth, that is, w should have only small bin-to-bin variations. Spurious fluctuations, however, will introduce sharp peaks in w, since they do not exist in the initial generated distribution x^{ini}. They can thus be suppressed by requiring the solution for w to be smooth.

The "second derivative" of a discrete distribution w can be defined somewhat analogously to an ordinary second derivative,

$$w''_j = \frac{w'_{j+1} - w'_j}{\Delta w_j} = \frac{1}{(\Delta w_j)} \left(\frac{w_{j+2} - w_{j+1}}{\Delta w_{j+1}} - \frac{w_{j+1} - w_j}{\Delta w_j} \right) = \frac{1}{(\Delta w_j)^2} (w_{j+2} - 2w_{j+1} + w_j),$$

where equal binning Δw_j is assumed, and in contrast to a usual derivative, Δw_j does not approach zero.

Defining the curvature c of a discrete distribution as the sum of the squares of its second derivatives (equidistant binning assumed and the overall factor of $(\Delta w_j)^2$ omitted),

$$c = \sum_j (w_{j+1} - 2w_j + w_{j-1})^2 ,$$

the corresponding matrix C that effectively minimizes the curvature of w is

$$C = \begin{pmatrix} -1+\xi & 1 & 0 & 0 & \cdots & \\ 1 & -2+\xi & 1 & 0 & \cdots & \\ 0 & 1 & -2+\xi & 1 & & \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & 1 & -2+\xi & 1 \\ & & & 0 & 1 & -1+\xi \end{pmatrix} . \tag{5.9}$$

The minimization (5.8) corresponds to the modified linear system,

$$\begin{pmatrix} \tilde{A} \\ \sqrt{\tau} C \end{pmatrix} w = \begin{pmatrix} \tilde{A} C^{-1} \\ \sqrt{\tau} I \end{pmatrix} Cw = \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} , \tag{5.10}$$

which is equivalent to Eq. (5.7) for $\tau = 0$. For the inversion of C to be possible, a small diagonal component has been added in Eq. (5.9): $C_{ij} \to C_{ij} + \xi \delta_{ij}$. A value between 10^{-3} and 10^{-4} for ξ is large enough to make C invertable, but does not change the condition of minimum curvature significantly.

The following describes the formalism that is implemented and used for the unfolding.

In analogy to the treatment of the unregularized linear system, the SVD could be applied to the new system for each value of τ. Nevertheless, it turns out that it suffices to decompose $\tilde{A}C^{-1}$ and solve the system for $\tau = 0$. According to Eq. (5.3) the SVD yields

$$\tilde{A}C^{-1} = USV^T.$$

As before, U and V^T are orthogonal and S is a diagonal matrix with non-negative elements s_i ordered non-increasingly. Both Cw and $\tilde{b}$ can be written in terms of the orthogonal vectors forming the columns of U and V,

$$Ud = \tilde{b}, \quad Vz = Cw, \tag{5.11a}$$

to obtain the diagonal system

$$z_i = \frac{d_i}{s_i}, \quad w = C^{-1}Vz. \tag{5.11b}$$

The solution of the regularized system with $\tau \neq 0$ can then be obtained from Eqs. (5.11) by changing d_i [58],

$$d_i^{(\tau)} = d_i \frac{s_i^2}{s_i^2 + \tau}, \tag{5.12a}$$

leading to the regularized solution

$$z_i^{(\tau)} = \frac{d_i s_i}{s_i^2 + \tau} = \frac{d_i}{s_i + \frac{\tau}{s_i}}, \quad w^{(\tau)} = C^{-1}Vz^{(\tau)}. \tag{5.12b}$$

As can be seen in these equations, introducing a $\tau \neq 0$ leads to a damping of those terms with small s_i. Since small s_i (that is, large i) correspond to quickly oscillating terms, this regularization serves for the required damping of the spurious oscillations originating from statistical fluctuations. This acts as a low-pass filter with a smooth cutoff. It prevents the artificial introduction of quasi-periodic fluctuations, known as Gibbs phenomenon, which can be introduced by sharp cutoffs.

The unfolded solution $x^{(\tau)}$ is finally obtained from the regularized weights $w^{(\tau)}$,

$$x_i^{(\tau)} = x_i^{\text{ini}} w_i^{(\tau)}. \tag{5.13}$$

5.1.3 Determination of τ

The determination of the regularization parameter is of crucial importance in the unfolding process. Choosing a strong damping will bias the result toward the simulation input. As shown in Eq. (5.5a), we unfold the quotient of the end result and the true simulated spectrum. On the other hand, an insufficient damping will lead to a result substantially influenced by statistical fluctuations in the measured spectrum.

Here we choose to optimize the value of τ from a set of toy tests that are discussed in Sec. 5.2. The obtained "measured" test distribution is unfolded and the result $x^{(\tau)}$ is compared to the original distribution $x^{\mathrm{tru}} = x^{\mathrm{test}}$ for a range of a priori reasonable parameters $\tau = s_k^2$. We use the test value S^2,

$$S^2 = \sum_i \left(\frac{x_i^{(\tau)} - x_i^{\mathrm{tru}}}{\sigma(x_i^{\mathrm{tru}})} \right)^2 , \tag{5.14}$$

to describe the level of agreement between the result of the unfolding $x^{(\tau)}$ and x^{tru}. An analoguous quantity is formed for the moments,

$$S_M^2 = \left(\frac{M^{(\tau)} - M^{\mathrm{tru}}}{\sigma(M^{\mathrm{tru}})} \right)^2 , \tag{5.15}$$

where the moments M are the first (M_1) and the second central moment (U_2) defined in App. B. While this quantity can of course also be obtained for the third central moment, we expect quite large errors on this moment and optimize τ for the spectrum and the two first moments.

We consider several criteria for the determination of τ. The τ for which $x^{(\tau)}$ and x^{tru}, and $M^{(\tau)}$ and M^{tru}, respectively, agree best, is indicated by S^2 being minimal: While S^2 does not take bin-by-bin correlations in the unfolded spectrum into account, it is independent of τ and minimizes the absolute deviation of the unfolding result and the original distribution. This is a desirable property, since the uncertainties on the unfolded spectrum themselves will depend on τ. Another quantity to be taken into account is the size of the bias introduced by the unfolding relative to the statistical uncertainty – it is desirable to keep the bias, which is harder to quantify than the statistical uncertainty, small. Also, the projected bias

should not depend strongly on the true value m_b, so that it can be reliably quantified from the toy tests.

Another possibility is to use information about the d_i, which are the coefficients of the orthogonal decomposition of the rescaled measured spectrum $\tilde{b}$ (see Eq. (5.11a)). Since the covariance matrix on $\tilde{b}$ is by construction equal to the unit matrix, d also has a unit covariance matrix because of orthogonality of U in Eq. (5.11a). Hence, the d_i are expected to show some typical behavior: The first k elements d_i, which give rise to the slower oscillating terms, should be statistically significant, and have a unit error. The statistically insignificant d_i for $i > k$ should follow a standard distribution, that is, have a mean value of zero and unit variance. The average of the absolute values of the statistically insignificant d_i should thus be close to $\sqrt{2/\pi} \approx 0.8$[1]. Hence, the effective rank of the system can be determined by identifying the critical k marking the last statistically significant d_i. The regularization parameter τ should then be set to the square of the k^{th} singular value,

$$\tau = s_k^2. \tag{5.16}$$

We use S^2, S_M^2, and the size of the projected bias on the moments to determine τ.

5.2 Tests of the Unfolding Procedure

As basis for the unfolding of the m_X^2 spectrum we use the `RooUnfHistoSvd` package.

Several tests of the unfolding code, based on toy experiments, are performed. Although the method should in principle work independently of the specific distribution to be unfolded, in case of the m_X and m_X^2 spectra problems might arise because of sharp resonances in the hadronic mass. The regularization is meant to damp oscillations originating from statistical uncertainties in the measured distribution b. As can be seen in Eq. (5.8), this is achieved by adding a regularization term to the expression to be minimized. In our case this term is chosen to minimize the bin-to-bin variation of the unfolded vector. Since we actually unfold the bin-by-bin ratio w of the result $x^{(\tau)}$ and the initially generated distribu-

[1] This result differs from that of [58], which states that the average of the absolute values of the statistically insignificant d_i should be close to $1/2\sqrt{\pi} \approx 0.28$.

tion x^{ini} from the MC simulation (see Eqs. (5.5)), the regularization suppresses sharp peaks in w. Statistical fluctuations in the measured spectrum are not the only possible source for such peaks; they could also arise when sharp resonances are very different in data and MC simulation. In general, measurements of exclusive semileptonic $B \to X_u \ell \nu$ branching fractions have sizable errors or do not exist yet, and thus the existence of differences in data and MC simulation cannot be excluded[2]. A thorough test is necessary to investigate how large the differences between data and MC simulation can be for the unfolding procedure to be successful.

The toy experiments are performed using statistically independent toy MC samples for the initialization and for the distributions to be unfolded. For each toy experiment, the first MC sample serves to determine the detector response matrix A and the initial generated MC distribution x^{ini}. For b, instead of a truly measured distribution from data, the reconstructed distribution is taken from a second, statistically independent, MC sample with different values of m_b, while keeping μ_π^2 fixed. This allows us to test the performance of the unfolding dependent on the difference of the value for m_b assumed in the simulation and the actual value of m_b. Note that the appearance of the resonance structure depends on m_b, since the form of the nonresonant part of the spectrum determines how high the resonances are in comparison to the nonresonant part. For the sample serving as MC data for the unfolding, we adopt the values of m_b and μ_π^2 as determined in [60]. In the following plots, Δm_b is defined as $\Delta m_b = m_b^{\text{data}} - m_b^{\text{MC}}$, where the labels "data" and "MC" refer to the role the respective toy sample takes in the toy experiments.

For the toy experiments, we apply the same cuts to the reconstructed samples that are applied in the analysis.

The unfolding result $x^{(\tau)}$ can be compared with the true MC distribution x^{tru}, associated with the reconstructed distribution b that is unfolded. We compare the first, second central, and third central moment of $x^{(\tau)}$ and x^{tru}. For each value of Δm_b, we use 500 statistically

[2]For the final unfolding we use a bin width of 0.31 GeV and 0.8 GeV^2, respectively, which dilutes the resonant structure of the mass spectrum and hence reduces the sensitivity on different resonant structures in data and MC simulation.

independent samples to serve as "data" and we assume the resolution observed on signal MC samples.

The statistics of the second MC sample that serves as "data" for the purpose of the tests is chosen to be comparable to the data statistics in the recoil analysis, i.e., b contains roughly 750 to 800 reconstructed signal events. The statistics in the sample that serves as MC data is roughly ten times larger.

The regularization parameter τ is scanned in order to choose the optimal value for τ by comparing the different results. While by definition k can only take integer values (see Eq. (5.16)), we have defined $k = 2.5$ to denote $\tau = s_2 s_3$.

We use three different upper cuts on the hadronic mass, $m_X^2 < 6.4\,\mathrm{GeV}^2$, $m_X^2 < 7.2\,\mathrm{GeV}^2$, and no cut on m_X^2 and we study the following quantities for the moments of the m_X^2 distribution:

- The bias, $M_{\mathrm{true}} - \hat{M}^{(\tau)}$,
- The relative bias, $(M_{\mathrm{true}} - \hat{M}^{(\tau)})/M_{\mathrm{true}}$,
- The bias relative to the statistical uncertainty, $(M_{\mathrm{true}} - \hat{M}^{(\tau)})/\hat{\sigma}_{M^{(\tau)}}$,
- The relative statistical uncertainty, $\hat{\sigma}_{M^{(\tau)}}/\hat{M}^{(\tau)}$,
- The relative uncertainty including the bias, $\sqrt{\hat{\sigma}^2_{M^{(\tau)}} + (M_{\mathrm{true}} - \hat{M}^{(\tau)})^2}/\hat{M}^{(\tau)}$, and
- S^2 and S_M^2 as defined in Eqs. (5.14) and (5.15).

as obtained using the mean of 500 toys. The moments M are the first moment (M_1), the second central moment (U_2), and the third central moment (U_3). The mean value of $M_{\mathrm{true}} - \hat{M}^{(\tau)}$ is obtained from a Gaussian fit to the distribution of $M_{\mathrm{true}} - M^{(\tau)}$[3].

The color and markerstyle coding is

[3]In particular, $(M_{\mathrm{true}} - \hat{M}^{(\tau)})/\hat{\sigma}_{M^{(\tau)}}$ is not obtained on a toy-by-toy basis and averaged, but rather from the averages of $M_{\mathrm{true}} - M^{(\tau)}$ and $\sigma_{M^{(\tau)}}$.

▲ k = 2,
▼ k = 2.5,
● k = 3,
○ k = 4, and
★ k = 5.

In the case of the S_M^2 plots, the bin entry for $\Delta m_b = 0\,\text{GeV}$ is the mean value of the other bin entries.

The studies are based on the MC production version SP5/6. Changes that occured in the analysis when moving to MC production version SP8, which is used for the remainder of the analysis, are as follows:

- Increased statistics for Run1-5 ($\approx$ 1000 signal events),
- Changes in resolution due to new reconstruction, selection, and cuts, and
- Realistic estimates of statistical uncertainties on the measured spectrum (these were substantially underestimated in the SP5/6-based toys).

However, the changes in resolution and numbers of events are not very large between SP5/6 and SP8. Since we underestimated the uncertainties in the toy experiments shown here, the regularization parameter from more realistic toy studies would be smaller, if different at all, so that using the regularization parameter extracted here is conservative. Since the statistical uncertainty will be larger than predicted by the toys, the bias from the unfolding will be less significant in comparison.

We use the results of the toy studies to determine the regularization to be used in the unfolding of the measured m_X^2 spectrum. The aim is to choose k such that the bias introduced by the unfolding procedure is small compared to the statistical uncertainty, while minimizing the overall uncertainty on the moments. To use a more realistic estimate of the statistical error, however, we use the statistical error from the actual moments we obtain (see Sec. 5.4) and compare this to the bias seen in the toys, given in Fig. 5.1. The values for S_M^2 for the first and second moments are shown in Fig. 5.2. From the toy study based

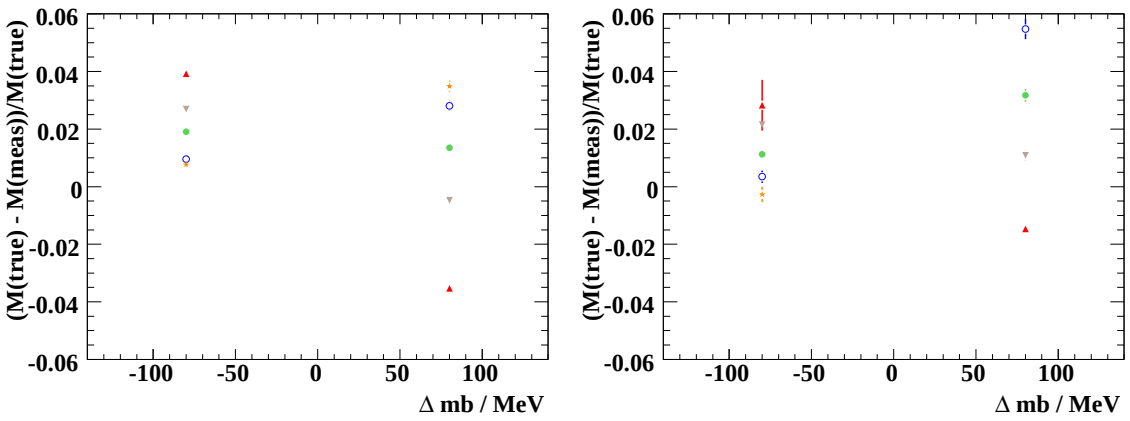

Figure 5.1. The relative bias for the first (left) and second central (right) moments from toy experiments for different values of Δm_b and τ. This toy uses the mass distribution from signal MC samples, resolution from signal MC samples, and upper cut on the hadronic mass $m_X^2 < 6.4\,\mathrm{GeV}^2$.

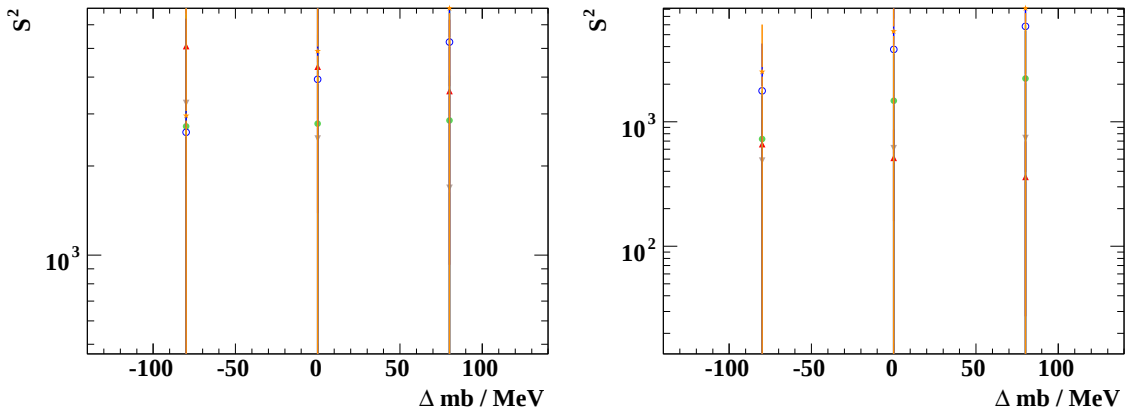

Figure 5.2. S_M^2 for the first (left) and second central (right) moments from toy experiments for different values of Δm_b and τ. This toy uses the mass distribution from signal MC samples, resolution from signal MC samples, and upper cut on the hadronic mass $m_X^2 < 6.4\,\mathrm{GeV}^2$.

on SP5/6 we choose $k = 2.5$ for the unfolding of the m_X^2 spectrum. The value $k = 2.5$ has to be understood in that we use $s_2 s_3$ instead of s_k^2 in Eq. (5.16). For the unfolding of the m_X spectrum, we use $k = 4$ as determined in an earlier analysis [61].

Rather then correcting for the bias, we use an estimate of the bias as an additional systematic uncertainty. In most cases, the bias shows a linear dependence on m_b, and we linearize and symmetrize the biases we observe at $\pm 60\,\mathrm{MeV}$ away from the mean value assumed in the toy experiments, which corresponds to 1.5σ of the uncertainty on m_b in the global fits in the kinetic scheme [2]. To summarize, we observe relative biases of 0.028 and -0.000 on M_1, 0.030 and 0.012 on U_2, and -0.165 and 0.063 on U_3, for $\pm 60\,\mathrm{MeV}$, respectively.

5.3 Unfolding the m_X and m_X^2 Spectra

This section presents the unfolded m_X and m_X^2 spectra obtained from the measured spectra. We will discuss how the initially generated spectra x^{ini} are obtained before showing the results of the unfolding.

A breakdown of the uncertainties on the moments of the unfolded m_X^2 spectrum is given in Sec. 5.4.

5.3.1 Full Signal Sample

For the initial generated spectrum x^{ini}, which enters the unfolding through Eq. (5.13), to contain the full physical information, the signal MC sample needs to contain all events, whether or not they pass any analysis cuts. In particular, this includes events for which the tag B is not correctly reconstructed. Note that for the SVD-based unfolding, which uses the number-of-events matrix A rather than the probability matrix $\hat{A}$, the matrix determination only needs the events that pass the analysis cuts.

Taking only events with a reconstructed B_{reco} tag would introduce a bias on the x^{ini} spectrum, since the efficiency of reconstructing the tag B_{reco} varies as a function of the

hadronic mass. Events with large hadronic mass have in general more particles in the X system than events with low m_X. Hence, it is easier for the reconstruction to find an acceptable B_{reco} candidate in high-m_X events, since there is a larger pool of particles from which the reconstruction can draw for building a candidate. Even if one of the actual particles from the B_{reco} candidate is lost, there is some chance that it can be replaced by one of the X system particles. Note that this B_{reco} candidate is not necessarily the correct one. After all analysis cuts are applied and the tag side background is subtracted, events with low m_X have a larger reconstruction efficiency.

As described in Sec. 4.1.2, our data samples are based on the `BSemiExcl` skim. In order to obtain unbiased initially generated spectra x^{ini}, we produced dedicated signal MC samples containing only generator level information, which uses the same input parameters and reweightings as used for the default signal MC production.

5.3.2 Final State Radiation

The unfolding can also serve to correct the m_X and m_X^2 spectra for final state radiation, assuming it is simulated well in the MC data. On the technical level, we need to guarantee that final state radiation does not introduce any bias on the generated hadronic mass in signal MC simulation. Bremsstrahlung does not need any specific attention, as it does not enter generator-level information. Provided it is simulated well in the MC data, the unfolded m_X and m_X^2 spectra will automatically be corrected for bremsstrahlung.

We define a generated hadronic mass at truth level that excludes photons likely to be radiated from the signal lepton. Photons making an angle $\theta < \pi/2$ with the signal lepton in the $\Upsilon(4S)$ rest frame are assumed to originate from final state radiation and are not included in the true X system. We find that with this procedure the generated hadronic mass m_X is not biased by final state radiation[4], and the detector matrices correct for the bias in the recontructed hadronic mass.

[4]This is tested by generating signal events with and without the simulation of final state radiation and comparing the moments of the hadronic mass spectrum between the two samples. The final state radiation veto is applied to the sample for which final state radiation is included in the simulation.

5.3.3 Determination of the Regularization Parameters

The determination of the regularization parameter τ for the unfolding of the measured spectrum is described in Sec. 5.1.3. From the results of toy experiments (Sec. 5.2), we choose $\tau = s_{2.5}^2$ for the unfolding of the m_X^2 spectrum and $\tau = s_4^2$ for the unfolding of the m_X spectrum.

5.3.4 The Unfolded Spectra and Moments

The measured m_X and m_X^2 spectra are shown in Fig. 4.10. The detector response matrices in their number-of-events form are given in Fig. 4.7. We unfold the effects of detector resolution, and of acceptance and efficiency effects, at the same time. In order to not induce an additional dependence on the modeling of the signal decays, we do not unfold the lepton momentum cut in the B rest frame. Our unfolded spectra and moments are hence determined with a cut $p^* > 1\,\mathrm{GeV}$.

Since the measured spectrum by construction contains signal events (i.e., $B \to X_u \ell \nu$ events), negative entries are not physical. Yet for the unfolding we use the spectrum as it is measured, i.e., with negative bin entries. Setting the negative bin entries to zero (while keeping the measurement error on these bins) would distort the measured spectra b and thus the unfolded spectra $x^{(\tau)}$. For the errors on the measured distributions b we take the square root of the respective diagonal element of the covariance matrices C.

The unfolded spectra $x^{(\tau)}$, taking only statistical uncertainties into account, are shown in Fig. 5.3. They are normalized to unit area. The relative statistical uncertainties on the truncated moments are given in Tab. 5.3.4 and the correlation matrices are shown in Fig. 5.4. Statistical uncertainties arise from the m_{ES} fits and the statistical uncertainties on the detector response matrix. The latter are evaluated by performing 1000 toy experiments, in which we fluctuate the detector response matrix within its statistical uncertainties according to a Poisson distributions. The covariance matrix is obtained from the spread of the results following Eq. (5.17b). The large bin-by-bin correlations are due to the bin size being very close to our resolution in the hadronic mass.

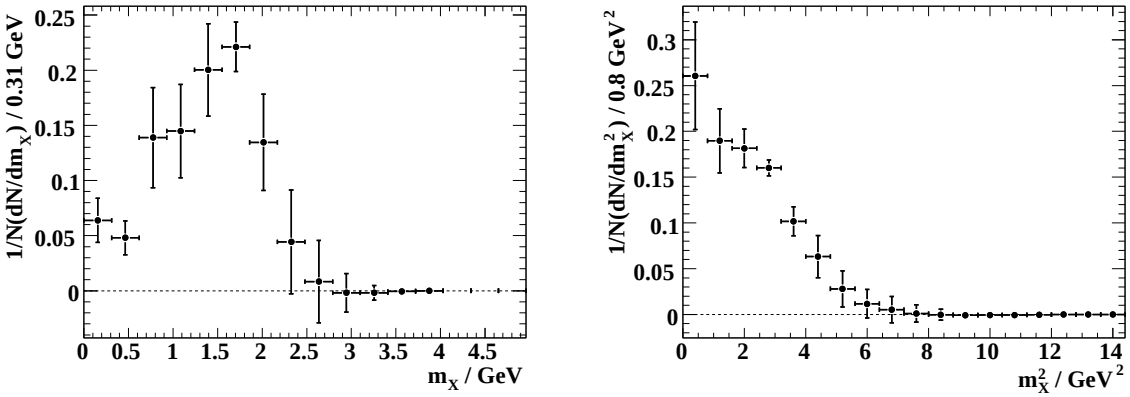

Figure 5.3. The unfolded m_X (left) and m_X^2 (right) distributions with statistical uncertainties only.

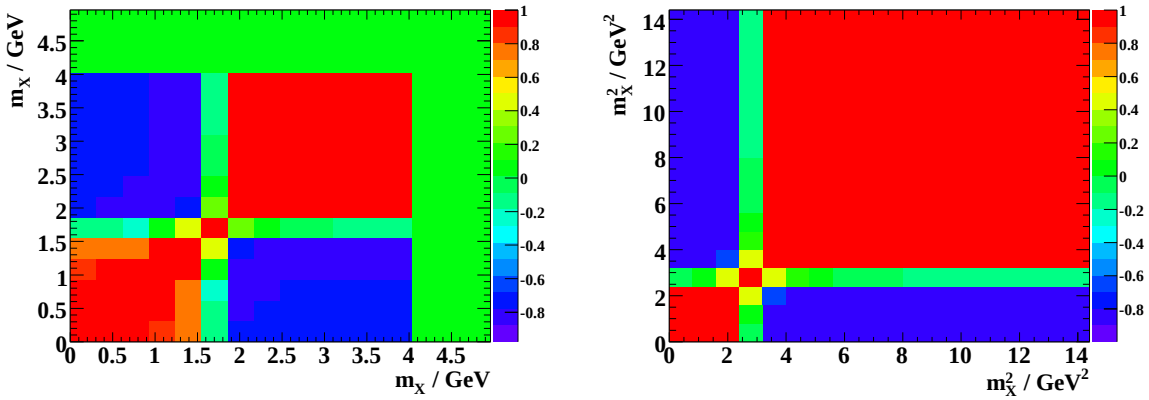

Figure 5.4. Statistical correlation matrices for the unfolded m_X (left) and m_X^2 (right) spectra.

	m_X	m_X^2
$\sigma(M_1^8)$	0.11	0.16
$\sigma(U_2^8)$	0.14	0.26
$\sigma(U_3^8)$	0.54	0.24
ρ_{12}	0.71	0.98
ρ_{23}	−0.67	0.91
ρ_{13}	−0.94	0.82

Table 5.1. Relative statistical uncertainties on the moments of the unfolded spectra and their correlations. The moments are truncated to $m_X < 2.48\,\text{GeV}$ and $m_X^2 < 6.4\,\text{GeV}^2$, respectively.

Note that both the mean values of the spectrum and moments and their statistical uncertainties are expected to change when including the effect of systematic uncertainties. The bin-by-bin rescaling of the measured spectra according to their uncertainties (see Eq. (5.7)) also takes into account systematic uncertainties for the main result, which accounts for the changes between the results presented here and the final results.

5.4 Moments and Spectra Results Including Systematic Uncertainties

The systematic uncertainties on the measured spectrum are included in the rescaling of the spectrum during the unfolding (according to Eq. (5.7)). They also are propagated through the unfolding procedure, along with systematic uncertainties on the detector response and the initially generated spectrum, to yield systematic uncertainties on the unfolded spectrum. We use two different methods for the propagation: If a given systematic uncertainty affects the detector response and/or the initially generated spectrum, they are propagated through the unfolding procedure by systematically changing all inputs to the unfolding and evaluating the systematic change in the result of the unfolding. We compute the covariance matrix from

$$C_{k,ij}^{(\tau)} = \begin{cases} \frac{\left(x_i^{1(\tau)} - x_i^{(\tau)}\right) - \left(x_i^{2(\tau)} - x_i^{(\tau)}\right)}{2} \frac{\left(x_j^{1(\tau)} - x_j^{(\tau)}\right) - \left(x_j^{2(\tau)} - x_j^{(\tau)}\right)}{2} & \text{if } N = 2 \\ (x_i^{1(\tau)} - x_i^{(\tau)})(x_j^{1(\tau)} - x_j^{(\tau)}) & \text{if } N = 1, \end{cases} \tag{5.17a}$$

where $N = 1$ or 2 depending on the nature of the systematic uncertainty.

If a given systematic uncertainty only affects the measured spectrum, we obtain the associated covariance matrix on the unfolded spectrum from toy Monte Carlo studies, in which we fluctuate the measured spectrum according to its covariance matrix (taking into account bin-by-bin correlations) and unfold those toy spectra. The covariance matrices on the unfolded spectrum associated with the different sources of systematic uncertainties $C_k^{(\tau)}$ are computed from the spread of the results,

$$C_{k,ij}^{(\tau)} = \frac{\sum_l \left(x_i^{l(\tau)} - x_i^{(\tau)}\right)\left(x_j^{l(\tau)} - x_j^{(\tau)}\right)}{N-1}, \tag{5.17b}$$

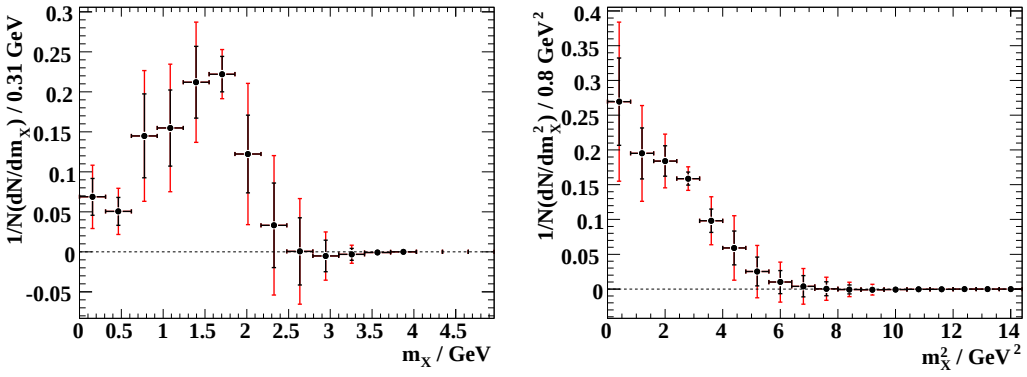

Figure 5.5. Unfolded m_X (left) and m_X^2 (right) spectra. The outer error bars give the total error including the systematic contribution. The statistical contribution to the uncertainty is shown by the shorter error bar.

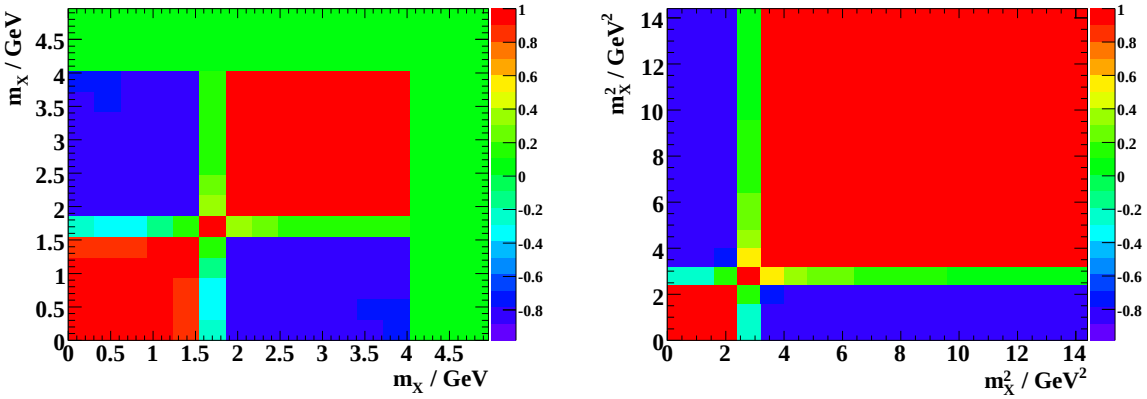

Figure 5.6. The correlation matrices for the unfolded m_X (left) and m_X^2 (right) spectra.

where l labels the unfolded, systematically varied spectrum, and $N = 1000$ is the number of toy spectra. The covariance matrices on the unfolded spectra $C^{(\tau)}$ are obtained by adding the $C_k^{(\tau)}$.

Fig. 5.5 shows the unfolded m_X and m_X^2 spectra, where the outer error bars include systematic uncertainties. The inner error bars give the statistical uncertainties. The correlation matrices for the unfolded spectra are shown in Fig. 5.6. Tab. 5.2 gives a breakdown of the uncertainties on the moments of the m_X^2 spectrum with an upper cut on the hadronic mass, $m_X^2 < 6.4\,\mathrm{GeV}^2$.

We measure the first three moments of the unfolded m_X^2 spectrum shown in Fig. 5.5

Source	Range	$\sigma_k(M_1)/M_1$	$\sigma_k(U_2)/U_2$	$\sigma_k(U_3)/U_3$
Total Statistical		0.17	0.30	0.34
Tracking efficiency	$0.2 - 0.7\%$ / track	0.04	0.06	0.03
Neutral efficiency	1.8% / photon	0.04	0.07	0.11
e efficiency	2%	0.02	0.03	0.02
μ efficiency	3%	0.05	0.07	0.05
e misid	15%	0.03	0.05	0.02
μ misid	15%	0.03	0.04	0.01
$K^{\pm}$ efficiency	2%	0.16	0.26	0.21
$K^{\pm}$ misid	15%	0.05	0.08	0.05
K_L reconstruction	Sec. 4.2.1	0.05	0.08	0.07
Detector Subtotal		0.20	0.32	0.27
$m_{\rm ES}$ fits	Sec. 4.2.2	0.06	0.12	0.13
$B_{\rm reco}$ truth matching	$\Delta E \pm 10\,{\rm MeV}$	0.01	0.01	0.00
m_X fit uncertainty	Tab. 4.4	0.08	0.14	0.15
Fit Subtotal		0.10	0.18	0.20
$B \to D^{(*,**)}(\pi)\ell\nu$ BFs	[49]	0.10	0.18	0.19
D BFs	[8]	0.06	0.09	0.07
K_S veto	10%	0.07	0.11	0.08
Bkg Subtotal		0.13	0.23	0.23
Nonresonant signal decays	[2]	0.02	0.04	0.07
Resonant signal BFs	[8]	0.01	0.02	0.01
$s\bar{s}$ popping	$\pm 30\%$	0.01	0.01	0.01
Signal Subtotal		0.03	0.05	0.07
$B \to X_c e\nu$ $m^2_{\rm miss}$ rew.	turn off	0.06	0.11	0.12
Neutral multiplicity	Sec. 4.2.5	0.06	0.06	0.03
Data-MC Subtotal		0.08	0.12	0.12
Systematics Subtotal		0.27	0.45	0.43
Unfolding bias	Sec. 5.2	0.01	0.02	0.05

Table 5.2. Relative statistical and systematic uncertainties on the measured m_X^2 moments.

$m_X^{2\text{cut}}/\,\text{GeV}^2$	$M_1/\,\text{GeV}^2$	$U_2/\,\text{GeV}^4$	$U_3/\,\text{GeV}^6$
7.2	$1.98 \pm 0.41 \pm 0.63$	$2.00 \pm 0.90 \pm 1.29$	$2.11 \pm 1.74 \pm 2.27$
	$\rho_{12} = 0.99$	$\rho_{23} = 0.99$	$\rho_{13} = 0.96$
6.4	$1.96 \pm 0.34 \pm 0.53$	$1.92 \pm 0.59 \pm 0.87$	$1.79 \pm 0.62 \pm 0.78$
	$\rho_{12} = 0.99$	$\rho_{23} = 0.94$	$\rho_{13} = 0.88$
5.6	$1.92 \pm 0.28 \pm 0.44$	$1.78 \pm 0.36 \pm 0.32$	$1.37 \pm 0.16 \pm 0.19$
	$\rho_{12} = 0.99$	$\rho_{23} = 0.08$	$\rho_{13} = -0.08$
4.8	$1.83 \pm 0.21 \pm 0.34$	$1.53 \pm 0.20 \pm 0.31$	$0.88 \pm 0.19 \pm 0.32$
	$\rho_{12} = 0.98$	$\rho_{23} = -0.86$	$\rho_{13} = -0.94$
4.0	$1.66 \pm 0.14 \pm 0.22$	$1.18 \pm 0.08 \pm 0.14$	$0.42 \pm 0.16 \pm 0.27$
	$\rho_{12} = 0.97$	$\rho_{23} = -0.94$	$\rho_{13} = -0.99$
3.2	$1.43 \pm 0.08 \pm 0.14$	$0.81 \pm 0.03 \pm 0.05$	$0.18 \pm 0.07 \pm 0.13$
	$\rho_{12} = 0.94$	$\rho_{23} = -0.93$	$\rho_{13} = -1.00$
2.4	$1.09 \pm 0.04 \pm 0.06$	$0.44 \pm 0.01 \pm 0.02$	$0.07 \pm 0.03 \pm 0.03$
	$\rho_{12} = 0.87$	$\rho_{23} = -0.84$	$\rho_{13} = -1.00$

Table 5.3. Moments and their correlations for $p^* > 1\,\text{GeV}$ and different cuts on m_X^2. The first uncertainty is statistical, the second systematic.

according to the definitions in App. B. Using an upper cut at $m_X^2 = 6.4\,\text{GeV}^2$ and $p^* > 1\,\text{GeV}$, we obtain

$$\begin{aligned} M_1 &= (1.96 \pm 0.34_{\text{stat}} \pm 0.53_{\text{syst}})\,\text{GeV}^2 \\ U_2 &= (1.92 \pm 0.59_{\text{stat}} \pm 0.87_{\text{syst}})\,\text{GeV}^4 \\ U_3 &= (1.79 \pm 0.62_{\text{stat}} \pm 0.78_{\text{syst}})\,\text{GeV}^6, \end{aligned}$$

where the correlations between the moments are $\rho_{12} = 0.99$, $\rho_{23} = 0.94$, and $\rho_{13} = 0.88$, respectively.

Moments with different upper cuts on m_X^2 are given in Tab. 5.3. Even though we are only adding (subtracting) a small fraction of phase space when moving the m_X^2 cut by $0.8\,\text{GeV}^2$, the uncertainties get substantially larger (smaller), which can be attributed to the large bin-by-bin correlations introduced by the unfolding. In addition, systematics related to the $B \to X_c \ell \nu$ background decrease for a decreasing cut on m_X^2.

Chapter 6

Determination of m_b and Nonperturbative Parameters

In this Chapter, we present the determination of the b-quark mass and the nonperturbative parameters μ_π^2 and ρ_D^3 from the first three moments of the unfolded m_X^2 spectrum.

The values of m_b and the nonperturbative parameters are usually determined from moments on the lepton energy distribution and the hadronic mass distribution in $B \to X_c \ell \nu$ decays and the photon energy distribution in $B \to X_s \gamma$ decays [1, 2], which achieve a combined uncertainty of $(30 - 39)\,\mathrm{MeV}$ on m_b from fits to data from several experiments. A recent analysis of *BABAR* data finds a combined uncertainty of 55 MeV. The branching fraction of $B \to X_c \ell \nu$ decays is about a factor of 50 larger than the branching fraction of $B \to X_u \ell \nu$. Yet, $B \to X_c \ell \nu$ decays are primarily sensitive to the difference of the b- and c-quark masses, while $B \to X_u \ell \nu$ decays are directly sensitive to the b-quark mass, which offsets the difference in the available statistics. The branching fraction of $B \to X_s \gamma$, on the other hand, is about a factor of 5 smaller than our signal branching fraction, but the parton-level two-body decay $b \to s\gamma$ is much more sensitive to m_b than the parton-level three-body decay $b \to u \ell \nu$. While the precision of our determination of m_b and the nonperturbative parameters cannot be competitive with that from $B \to X_c \ell \nu$ and $B \to X_s \gamma$, our results offer a test of the underlying theory, which is relied on for the determination of $|V_{ub}|$.

Our experimental inputs are the hadronic mass moments with an upper cut at $m_X^2 = 6.4\,\mathrm{GeV}^2$, given in Sec. 5.4. The theoretical input [12] is described in Sec. 2.3.4. The hadronic mass moments in $B \to X_u \ell\nu$ are predicted with nonperturbative and perturbative corrections up to $\mathcal{O}(1/m_b^3)$ and $\mathcal{O}(\alpha_s^2\beta_0)$, respectively, as a function of m_b and the nonperturbative parameters. They allow for a loose lower cut on the lepton energy (E_ℓ) and an upper cut on m_X^2.

The cut of m_X^2 chosen is a compromise between experimental and theoretical considerations. Experimentally, increasing the m_X^2 cut increases the uncertainties related to the $B \to X_c \ell\nu$ background, which dominates the high m_X^2 region of the spectrum. The large m_X^2 region is poorly known and the bin-by-bin correlations are large. Theoretically, a tight cut on m_X^2 introduces large uncertainties due to shape function effects (see Sec. 2.3.4). We choose $m_X^2 < 6.4\,\mathrm{GeV}^2$ as the tighest cut where theoretical uncertainties on the first moment are still thought to be under control.

The fit performs a χ^2 minimization and is based on `Minuit` [62]. The fit code uses an implementation of the moment calculation by the authors of Ref. [12]. The necessary numerical integrations can be performed by two different procedures, by a full `Vegas` integration or by using an interpolation for the $\mathcal{O}(\alpha_s^2\beta_0)$ corrections. The fit uses the interpolation of the BLM corrections since the full `Vegas` integration is too slow to be used for a fit.

The values of μ_G^2 and ρ_{LS}^3 are fixed in our fits, since our sensitivity to these parameters is small. The best constraints on μ_G^2 are obtained from the mass difference between B and B^* mesons, and heavy quark sum rules can be used to estimate ρ_{LS}^3. Ref. [26] finds $\mu_G^2 = (0.35^{+0.03}_{-0.02})\,\mathrm{GeV}^2$, and Ref. [27] suggests that typically, $-0.05\,\mathrm{GeV}^3 < \rho_{LS}^3 < -0.25\,\mathrm{GeV}^3$. We use $\alpha_s = 0.22 \pm 0.10$, where the large uncertainty is used to account for uncertainties from uncalculated higher-order perturbative corrections in addition to the parametric uncertainties in α_s.

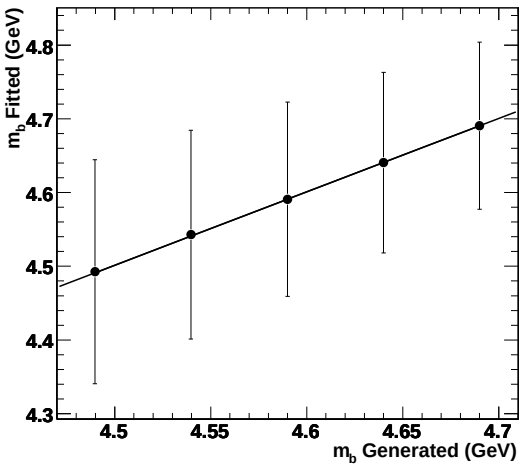

Figure 6.1. Fitted vs. generated m_b values. The fitted line has a slope consistent with unity.

6.1 Fit Validation

We validate the fitting code to verify that we obtain unbiased results and proper uncertainty estimates.

To test the fit for potential biases, we choose sets of input values for m_b, μ_π^2, and ρ_D^3, and compute the corresponding three hadronic mass moments with the full `Vegas` integration. The computed mass moments are used as input to the fit, and the fit results for m_b, μ_π^2, and ρ_D^3 are compared to the input parameters. We find that the parameters agree to better than 0.1% and hence no bias is observed. Fig. 6.1 shows the fitted values of m_b as a function of the generated values. This test validates both the fit procedure and the interpolation approximation for the BLM corrections that we use in the fit.

The uncertainties obtained from the fit are validated using a toy experiments. We choose values for m_b, μ_π^2, and ρ_D^3 that are consistent with current determinations from $B \to X_c \ell \nu$ and $B \to X_s \gamma$ decays [2]. The moments are computed with the full `Vegas` integration and are then randomly varied within the statistical uncertainties we obtain for the moments of the unfolded m_X^2 spectrum. The smeared moments serve as input for the fit, and we

Parameter	Gaussian mean	Gaussian width
m_b	0.015 ± 0.020	0.971 ± 0.015
μ_π^2	0.034 ± 0.020	0.993 ± 0.015
ρ_D^3	0.037 ± 0.020	0.987 ± 0.014

Table 6.1. Parameters of the Gaussian fits to the pull distributions from the toy experiments.

compute the pull r for the fit parameter p as

$$r = \frac{p_{\text{fit}} - p_{\text{true}}}{\sigma_p}, \tag{6.1}$$

where σ_p is the uncertainty on p_{fit} as returned by the fit. Fig. 6.2 shows the observed pull distributions for m_b, μ_π^2, and ρ_D^3. All are consistent with normal distributions, as is expected for an unbiased fit with correctly estimated uncertainties. We fit a Gaussian to the pull distributions and find the mean values and width given in Tab. 6.1, which agree with the expected values within 2σ.

6.2 Fit Results

From the hadronic moments with $m_X^2 < 6.4\,\text{GeV}^2$ and $p^* > 1\,\text{GeV}$, we determine

$$\begin{aligned} m_b &= (4.604 \pm 0.125_{\text{stat}} \pm 0.193_{\text{syst}})\,\text{GeV} \\ \mu_\pi^2 &= (0.398 \pm 0.135_{\text{stat}} \pm 0.195_{\text{syst}})\,\text{GeV}^2 \\ \rho_D^3 &= (0.102 \pm 0.017_{\text{stat}} \pm 0.021_{\text{syst}})\,\text{GeV}^3, \end{aligned}$$

where the statistical and systematic uncertainties are obtained from propagating the uncertainties on the mass moments, and the correlations are $\rho_{m_b \mu_\pi^2} = -0.99$, $\rho_{\mu_\pi^2 \rho_D^3} = 0.57$, and $\rho_{m_b \rho_D^3} = -0.59$. The large correlations on the fit results are due to the large correlations between the mass moments.

The $\Delta\chi^2 = 1$ contour in the m_b-μ_π^2 plane is shown in Fig. 6.3, along with the constraints from the different mass moments. While M_1 gives the strongest constraint on m_b, U_2 constrains mostly μ_π^2, and U_3 constrains mostly ρ_D^3. Due to the strong correlation, we

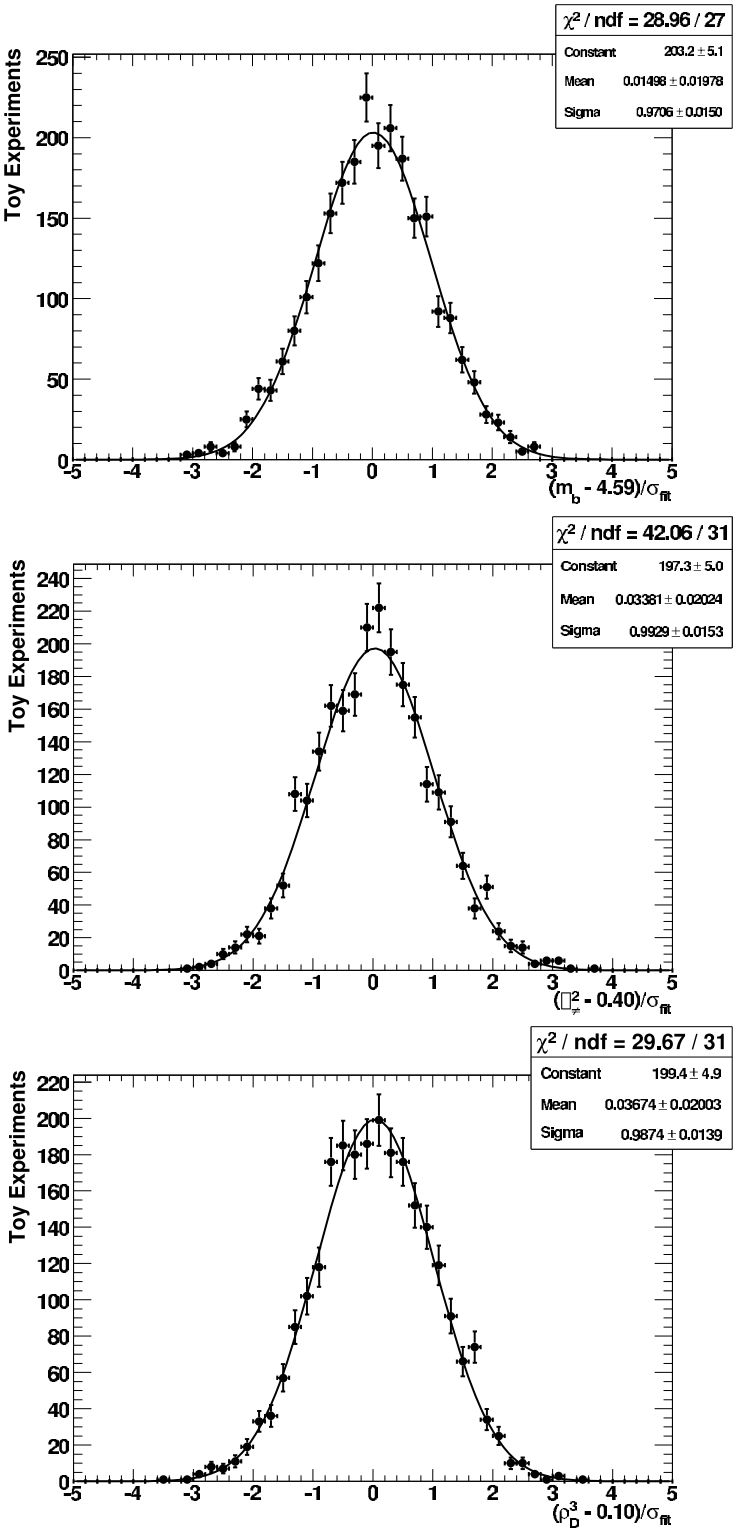

Figure 6.2. Pull distributions from 2500 toy experiments for $m_b = 4.59\,\mathrm{GeV}$ (top), $\mu_\pi^2 = 0.4\,\mathrm{GeV}^2$ (center) and $\rho_D^3 = 0.10\,\mathrm{GeV}^3$ (bottom).

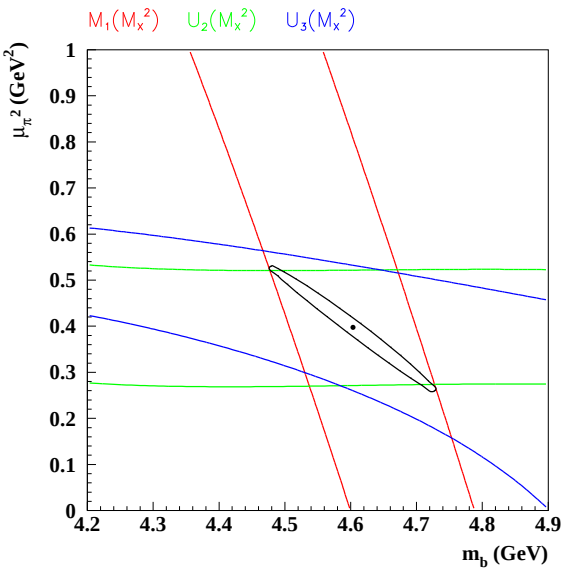

Figure 6.3. $\Delta\chi^2 = 1$ contour in the m_b-μ_π^2 plane with the constraints from the different moments: the constraint from M_1 is shown in red, the constraint from U_2 in green and the constraint from U_3 in blue.

determine a particular combination of m_b and μ_π^2 much better than the two parameters individually.

Finally, we assess the uncertainties associated with the fit itself. We vary μ_G^2 and ρ_{LS}^3, which have been fixed in the fit within the ranges given above, and estimate a systematic uncertainty from the variation of the fit results. To estimate the uncertainty due to uncalculated higher-order perturbative corrections together with the parametric uncertainty on α_s, we vary α_s by ± 0.1 and take the observed variation in the fitted parameters as the associated uncertainty. A breakdown of these uncertainties is shown in Tab. 6.2. We will label these uncertainties as theoretical uncertainties since their main contribution comes from uncalculated perturbative higher-order corrections.

To assess the stability of our results, we repeat the fit using moments with different cuts on m_X^2. We observe very stable results as shown in Tab. 6.3. It should, however, be noted that the uncertainties between results with different m_X^2 cuts are highly correlated.

Our final result from the fits to the hadronic mass moments with $m_X^2 < 6.4\,\mathrm{GeV}^2$ and

Source	Range	$\sigma_k(m_b)/\,\mathrm{GeV}$	$\sigma_k(\mu_\pi^2/\,\mathrm{GeV}^2$	$\sigma_k(\rho_D^3)/\,\mathrm{GeV}^3$
μ_G^2	$(^{+0.03}_{-0.02})\,\mathrm{GeV}^2$	0.008	0.003	0.002
ρ_{LS}^3	$\pm 0.1\,\mathrm{GeV}^3$	0.002	0.002	0.005
α_s	± 0.1	0.097	0.036	0.065
Subtotal		0.097	0.036	0.066

Table 6.2. Breakdown of the uncertainties on m_b, μ_π^2, and ρ_D^3 associated with the fit to the mass moments.

Parameter	$m_X^{2\mathrm{cut}} = 5.6\,\mathrm{GeV}^2$	$m_X^{2\mathrm{cut}} = 6.4\,\mathrm{GeV}^2$	$m_X^{2\mathrm{cut}} = 7.2\,\mathrm{GeV}^2$	No $m_X^{2\mathrm{cut}}$
$m_b/\,\mathrm{GeV}$	4.603 ± 0.099	4.604 ± 0.125	4.608 ± 0.157	4.637 ± 0.177
$\mu_\pi^2/\,\mathrm{GeV}^2$	0.399 ± 0.071	0.398 ± 0.135	0.385 ± 0.241	0.230 ± 0.339
$\rho_D^3/\,\mathrm{GeV}^3$	0.104 ± 0.011	0.102 ± 0.017	0.097 ± 0.054	0.017 ± 0.107

Table 6.3. Results of the OPE fits for several upper cuts on m_X^2 with statistical uncertainties only.

$p^* > 1\,\mathrm{GeV}$ is

$$m_b = (4.60 \pm 0.13_{\mathrm{stat}} \pm 0.19_{\mathrm{syst}} \pm 0.10_{\mathrm{theo}})\,\mathrm{GeV}$$

$$\mu_\pi^2 = (0.40 \pm 0.14_{\mathrm{stat}} \pm 0.20_{\mathrm{syst}} \pm 0.04_{\mathrm{theo}})\,\mathrm{GeV}^2$$

$$\rho_D^3 = (0.10 \pm 0.02_{\mathrm{stat}} \pm 0.02_{\mathrm{syst}} \pm 0.07_{\mathrm{theo}})\,\mathrm{GeV}^3,$$

in the kinetic scheme at $\mu = 1\,\mathrm{GeV}$, with correlation coefficients $\rho_{m_b\mu_\pi^2} = -0.99$, $\rho_{\mu_\pi^2\rho_D^3} = 0.57$, and $\rho_{m_b\rho_D^3} = -0.59$. While the systematic uncertainties constitute the largest part of the uncertainties in m_b and μ_π^2, the uncertainty in ρ_D^3 is dominated by theoretical uncertainties.

We show a comparison of our result with recent *BABAR* results in $B \to X_c\ell\nu$ and $B \to X_s\gamma$ decays in Fig. 6.4. Our results are very consistent with the parameters determined from hadronic mass moments and lepton energy moments in $B \to X_c\ell\nu$ decays. While the agreement with the results obtained in $B \to X_s\gamma$ decays is good within present uncertainties, it might be seen as curious that the small offset observed between the results from semileptonic decays and the radiative decays is also present in an earlier result from fully inclusive $B \to X_s\gamma$ decays [63].

Our measurement is the first measurement of m_b and nonperturbative parameters in

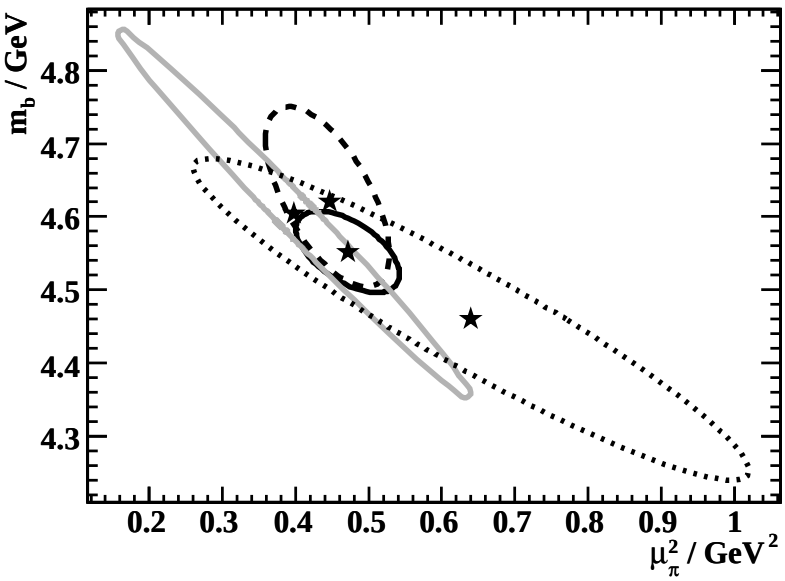

Figure 6.4. Results from recent *BABAR* analyses presented in the m_b-μ_π^2 plane in the kinetic scheme ($\mu = 1\,\text{GeV}$). The dotted ellipse shows the result of a $B \to X_s\gamma$ analysis using the recoil method [64], the dashed ellipse the result from $B \to X_c\ell\nu$ decays [4, 65], the black solid ellipse the result from its combination with earlier $B \to X_s\gamma$ measurements [66, 63], and the solid gray ellipse the result from $B \to X_u\ell\nu$ decays. The ellipses show $\Delta\chi^2 = 1$ contours.

$B \to X_u\ell\nu$ decays. With 30 times less statistics due to the small $B \to X_u\ell\nu$ branching fraction, our uncertainty in m_b is within a factor of 2 of the uncertainty from $B \to X_c\ell\nu$ decays (see Fig. 6.4) due to the much better sensitivity to the b-quark mass.

In light of the different HQEs in semileptonic and radiative penguin B decays, it would be especially interesting to combine the results from semileptonic B decays alone.

Fig. 6.5 compares the $m_X^{2\text{cut}}$ dependence of the measured and predicted dependence of the hadronic mass moments, assuming the values of m_b, μ_π^2, and ρ_D^3 extracted in the fit to the moments with $m_X^2 < 6.4\,\text{GeV}^2$ and $p^* > 1\,\text{GeV}$. We find that the $m_X^{2\text{cut}}$ dependence is well described down to $m_X^{2\text{cut}} = 4.8\,\text{GeV}^2$, with potentially a hint of a deviation starting at $m_X^{2\text{cut}} = 4\,\text{GeV}^2$. While the deviation is within the experimental error bars and theoretical uncertainties have not been considered, the uncertainties between moments with different cuts are highly correlated, which makes observed deviations more significant. It would be

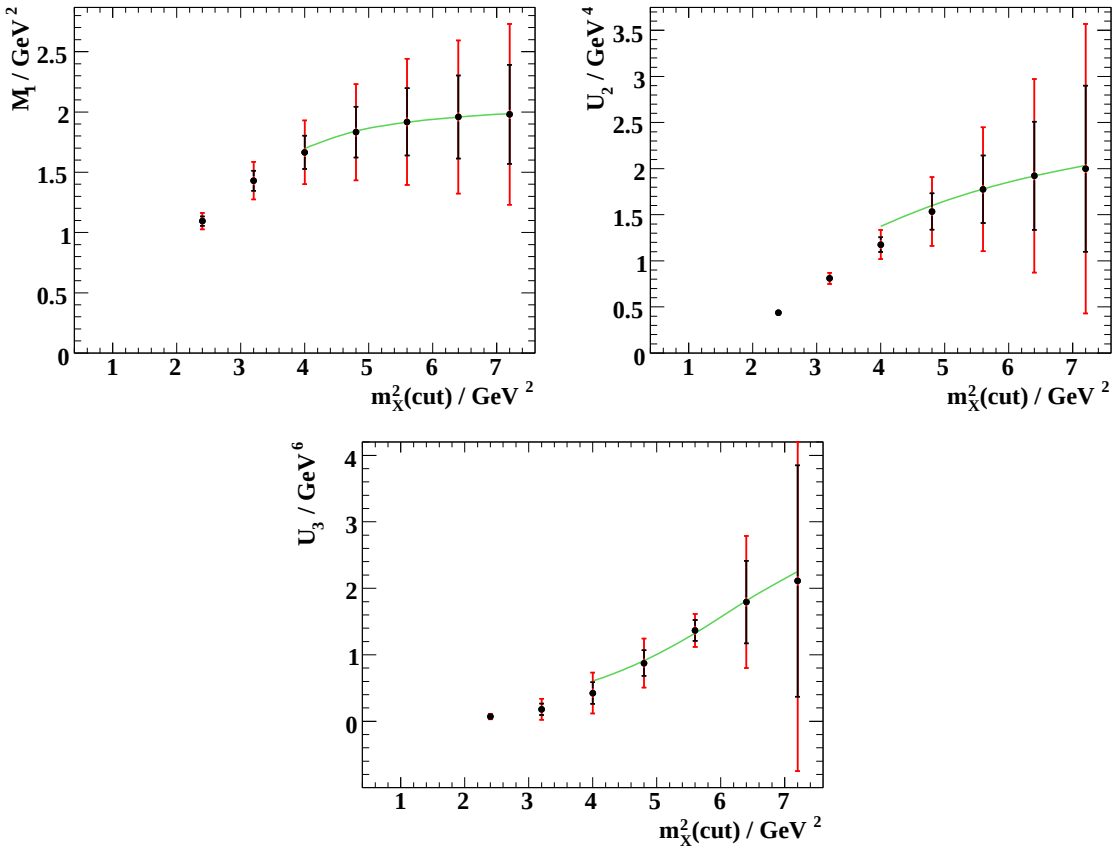

Figure 6.5. Hadronic mass moments as a function of $m_X^{2\mathrm{cut}}$. The inner error bars show the statistical uncertainties, the outer error bars the full uncertainties. The predictions for the moments [12] using our fitted values of m_b, μ_π^2, and ρ_D^3 are given by the overlaid markers.

desirable to obtain predictions with lower $m_X^{2\mathrm{cut}}$, which are not provided by the current version of the code.

Chapter 7

Conclusions

The CKM-matrix element $|V_{ub}|$ and the mass of the b quark are fundamental parameters of the SM. The precise determination of $|V_{ub}|$ is essential for testing the CKM picture, and relies on good measurements of m_b and the nonperturbative parameters, which describe the kinematics of the b quark in the B meson.

With the large datasets collected over the last nine years at the B factories, new precision measurements of $|V_{cb}|$, $|V_{ub}|$, and m_b have become possible. The $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ decays offer a clean experimental environment with low non-B backgrounds and well-known final state kinematics.

In this thesis, we present preliminary results for the measurement of the hadronic mass spectrum in inclusive $B \to X_u \ell \nu$ decays, based on 383×10^6 $B\bar{B}$ events collected by the *BABAR* experiment between 2000 and 2007. The measured spectrum is corrected for resolution, acceptance, and efficiency effects by an unfolding procedure. We also present preliminary results for the first determination of the b-quark mass and the nonperturbative parameters μ_π^2 and ρ_D^3 in $B \to X_u \ell \nu$ decays using the first three moments of the unfolded hadronic mass spectrum.

The truncated moments for $m_X^2 < 6.4\,\mathrm{GeV}^2$ and $p^* > 1\,\mathrm{GeV}$ are measured to be

$$M_1 = (1.96 \pm 0.34_{\mathrm{stat}} \pm 0.53_{\mathrm{syst}})\,\mathrm{GeV}^2$$
$$U_2 = (1.92 \pm 0.59_{\mathrm{stat}} \pm 0.87_{\mathrm{syst}})\,\mathrm{GeV}^4$$
$$U_3 = (1.79 \pm 0.62_{\mathrm{stat}} \pm 0.78_{\mathrm{syst}})\,\mathrm{GeV}^6,$$

where the correlations between the moments are $\rho_{12} = 0.99$, $\rho_{23} = 0.94$, and $\rho_{13} = 0.88$, respectively. The systematic uncertainties are somewhat larger than the statistical uncertainties and are dominated by uncertainties related to $B \to X_c \ell\nu$ background. Using HQET-based predictions in the kinetic scheme, we extract

$$m_b = (4.60 \pm 0.13_{\mathrm{stat}} \pm 0.19_{\mathrm{syst}} \pm 0.10_{\mathrm{theo}})\,\mathrm{GeV}$$
$$\mu_\pi^2 = (0.40 \pm 0.14_{\mathrm{stat}} \pm 0.20_{\mathrm{syst}} \pm 0.04_{\mathrm{theo}})\,\mathrm{GeV}^2$$
$$\rho_D^3 = (0.10 \pm 0.02_{\mathrm{stat}} \pm 0.02_{\mathrm{syst}} \pm 0.07_{\mathrm{theo}})\,\mathrm{GeV}^3,$$

at $\mu = 1\,\mathrm{GeV}$, with correlation coefficients $\rho_{m_b \mu_\pi^2} = -0.99$, $\rho_{\mu_\pi^2 \rho_D^3} = 0.57$, and $\rho_{m_b \rho_D^3} = -0.59$.

Our results for m_b, μ_π^2, and ρ_D^3 have a similar precision to the results in the $B \to X_c \ell\nu$ and $B \to X_s \gamma$ decay modes and are compatible with these within the present uncertainties. The measurement of m_b and nonperturbative parameters in $B \to X_u \ell\nu$ decays provides an independent determination of these quantities, which can be combined with results obtained in $B \to X_c \ell\nu$ and $B \to X_s \gamma$ decays.

As the precision of $|V_{ub}|$ is pushed below the 10% level, it is important to test the HQET-based predictions that are used for the extraction of $|V_{ub}|$. The HQE has been extensively tested in $B \to X_c \ell\nu$ decays with measurements by the DELPHI, CLEO, *BABAR*, Belle, and CDF experiments. The uncertainty on $|V_{cb}|$ from these measurements has reached 2% and is by now dominated by theoretical uncertainties [1, 2, 3, 4, 67, 68]. Similar tests in $B \to X_u \ell\nu$ decays are much more difficult due to the smaller $B \to X_u \ell\nu$ branching fraction and the dominant $B \to X_c \ell\nu$ background. So far, the only test of the theoretical description of $B \to X_u \ell\nu$ including shape function effects has been performed indirectly by extracting $|V_{ub}|$ with cuts on several different kinematic variables and comparing those results. The values obtained by different analyses generally agree, but a recent *BABAR* analysis [69]

found a 2.5σ difference between using kinematic cuts based on the hadronic mass and the light-cone variable $P_+ = E_X - |\vec{p}_X|$.

The observed consistency of our results for m_b, μ_π^2, and ρ_D^3 with determinations in $B \to X_c \ell\nu$ confirms the applicability of the HQE for the prediction of hadronic mass moments with moderate mass cuts within the uncertainties of our measurement. In addition to testing the consistency with earlier measurements for a fixed hadronic mass cut, we compare the measured dependence of the hadronic mass moments on the hadronic mass cut with the predicted dependence by the HQE. Within our uncertainties, we find good agreement for the mass cut dependence down to $m_X^{2\mathrm{cut}} = 4.8\,\mathrm{GeV}^2$. The small deviation at $m_X^{2\mathrm{cut}} = 4\,\mathrm{GeV}^2$ may or may not be a sign of the presence of shape function effects when lowering the mass cut.

Our result confirms that the measurement of the hadronic mass spectrum and its moments is feasible even in the charm-background dominated high-mass region, which is in agreement with previous analyses [70, 69]. It also demonstrates the feasibility of determining m_b and nonperturbative parameters using moments in $B \to X_u \ell\nu$ decays. The measurement would benefit from being repeated on a substantially larger dataset. With more statistics available, one could limit the analysis to clean B_{reco} modes and apply tighter selection criteria to improve the experimental resolution. It could also be improved by better determinations of the $B \to X_c \ell\nu$ branching fractions and of the charged K efficiency. In addition, a better hadronic mass resolution would introduce smaller correlations between the measured moments and thus between m_b and the nonperturbative parameters.

The LHC experiments will not be able to perform the inclusive reconstruction of B-meson decays with sufficient accuracy. Presently, there are proposals to build "Super$B$" factories, which would operate at the much higher luminosities of 10^{35} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 10^{36} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at KEK and Frascati, respectively. Provided that the experiments achieve a reconstruction quality comparable to that of *BABAR* and Belle despite larger backgrounds due to the high luminosity, their much larger datasets will provide an ideal basis for a repetition of our measurement.

For the time being, the determination of the b-quark mass in semileptonic and radiative B decays at the B factories is among the most precise measurements of m_b. As we enter the era of physics at the LHC experiments, precise measurements of the b-quark mass are an important ingredient for further studies of the SM as well as the Higgs sector.

Bibliography

[1] C. W. Bauer, Z. Ligeti, M. Luke, A. V. Manohar, and M. Trott, Phys. Rev. **D70**, 094017 (2004), [hep-ph/0408002].

[2] O. Buchmüller and H. Flächer, Phys. Rev. **D73**, 073008 (2006), [hep-ph/0507253].

[3] Belle Collaboration, C. Schwanda *et al.*, 0803.2158.

[4] BABAR Collaboration, B. Aubert *et al.*, 0707.2670.

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

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

[7] A. Höcker and Z. Ligeti, Ann. Rev. Nucl. Part. Sci. **56**, 501 (2006), [hep-ph/0605217].

[8] Particle Data Group Collaboration, W. M. Yao *et al.*, J. Phys. **G33**, 1 (2006 and electronic update).

[9] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 242003 (2006), [hep-ex/0609040].

[10] C. N. Burrell, M. E. Luke, and A. R. Williamson, Phys. Rev. D **69**, 074015 (2004), [hep-ph/0312366].

[11] F. J. Tackmann, Phys. Rev. D **72**, 034036 (2005), [hep-ph/0503095].

[12] P. Gambino, G. Ossola, and N. Uraltsev, JHEP **09**, 010 (2005), [hep-ph/0505091].

[13] F. J. Tackmann, Talk at the Joint Workshop on $|V_{ub}|$ and $|V_{cb}|$ at the B factories.

[14] A. V. Manohar and M. B. Wise, *Heavy Quark Physics*, Cambridge monographs on particle physics, nuclear physics, and cosmology No. 10 (Cambridge University Press, 2000).

[15] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Perseus Books, 1995).

[16] A. Droll and H. E. Logan, Phys. Rev. **D76**, 015001 (2007), [hep-ph/0612317].

[17] M. Spira and P. M. Zerwas, hep-ph/9803257.

[18] L.-L. Chau and W.-Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984).

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

[20] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, Phys. Rev. **D50**, 3433 (1994), [hep-ph/9403384].

[21] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).

[22] CKMfitter Group Collaboration, J. Charles *et al.*, Eur. Phys. J. **C41**, 1 (2005), [hep-ph/0406184], updated results and plots at: http://ckmfitter.in2p3.fr.

[23] A. X. El-Khadra and M. Luke, Ann. Rev. Nucl. Part. Sci. **52**, 201 (2002), [hep-ph/0208114].

[24] M. Battaglia *et al.*, hep-ph/0304132.

[25] DELPHI Collaboration, J. Abdallah *et al.*, Eur. Phys. J. **C46**, 569 (2006), [hep-ex/0603046].

[26] N. Uraltsev, Phys. Lett. **B545**, 337 (2002), [hep-ph/0111166].

[27] P. Gambino and N. Uraltsev, Eur. Phys. J. **C34**, 181 (2004), [hep-ph/0401063].

[28] CLEO Collaboration, A. Bornheim *et al.*, Phys. Rev. Lett. **88**, 231803 (2002), [hep-ex/0202019].

[29] Belle Collaboration, A. Limosani *et al.*, Phys. Lett. **B621**, 28 (2005), [hep-ex/0504046].

[30] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. **D73**, 012006 (2006), [hep-ex/0509040].

[31] BABAR Collaboration, B. Aubert *et al.*, 0708.3702.

[32] R. V. Kowalewski and S. Menke, Phys. Lett. **B541**, 29 (2002), [hep-ex/0205038].

[33] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **95**, 111801 (2005), [hep-ex/0506036].

[34] D. J. Lange, Nucl. Instrum. Meth. **A462**, 152 (2001).

[35] F. De Fazio and M. Neubert, JHEP **06**, 017 (1999), [hep-ph/9905351].

[36] T. Sjöstrand, hep-ph/9508391.

[37] H. Lacker and A. Volk, Private communication, 2007.

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

[39] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instrum. Meth. **A479**, 1 (2002), [hep-ex/0105044].

[40] *BABAR* Collaboration, F. Forti *et al.*, *BABAR* Analysis Document 652v1 (2003).

[41] D. Asgeirsson, Talk at *BABAR* Collaboration Meeting, September 2006.

[42] E. Richter-Was, Phys. Lett. **B303**, 163 (1993).

[43] GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Instrum. Meth. **A506**, 250 (2003).

[44] N. Isgur and D. Scora, Phys. Rev. D **52**, 2783 (1995).

[45] *BABAR* PID group, Performing Particle ID on Monte Carlo, *BABAR* internal webpage.

[46] *BABAR* Collaboration, M. Bona, G. Cavoto, D. Cote, E. Di Marco, and M. Pelliccioni, *BABAR* Analysis Document 1191v3 (2006).

[47] *BABAR* Collaboration, J. C. Dingfelder, V. Lüth, and H. W. Wulsin, *BABAR* Analysis Document 1642v1 (2006).

[48] *BABAR* Collaboration, J. C. Dingfelder, M. H. Kelsey, V. Lüth, P. A. Walker, and H. W. Wulsin, *BABAR* Analysis Document 1520v5 (2007).

[49] D. Lopes Pegna and V. Lüth, SemiLeptonics AWG : $B \to X_c \ell \nu$ Reweighting 2007, *BABAR* internal webpage.

[50] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998), [hep-ph/9712417].

[51] *BABAR* Collaboration, M. Bomben, F. Cossutti, and G. Della Ricca, *BABAR* Analysis Document 1395v17 (2007).

[52] *BABAR* Collaboration, M. A. Mazur and J. Richman, *BABAR* Analysis Document 1111v10 (2007).

[53] *BABAR* PID Group, PID Lists, *BABAR* internal webpage.

[54] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. **C48**, 543 (1990).

[55] Crystal Ball Collaboration, T. Skwarnicki, DESY F31-86-02.

[56] *BABAR* Charged Particle Reconstruction Group, Tracking Efficiency Task Force [R18/SP8], *BABAR* internal webpage.

[57] *BABAR* Neutral Particle Reconstruction Group, Neutral Reconstruction AWG, *BABAR* interal webpage.

[58] A. Höcker and V. Kartvelishvili, Nucl. Instrum. Meth. **A372**, 469 (1996), [hep-ph/9509307].

[59] V. Blobel, DESY 84-118 (1984).

[60] A. Limosani and T. Nozaki, hep-ex/0407052.

[61] *BABAR* Collaboration, H. Lacker and K. Tackmann, *BABAR* Analysis Document 894v7 (2004).

[62] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[63] BaBar Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **97**, 171803 (2006), [hep-ex/0607071].

[64] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. **D77**, 051103 (2008), [0711.4889].

[65] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. **D69**, 111104 (2004), [hep-ex/0403030].

[66] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. **D72**, 052004 (2005), [hep-ex/0508004].

[67] M. Battaglia *et al.*, ECONF **C0304052**, WG102 (2003), [hep-ph/0210319].

[68] C. W. Bauer, Z. Ligeti, M. Luke, and A. V. Manohar, Phys. Rev. **D67**, 054012 (2003), [hep-ph/0210027].

[69] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **96**, 221801 (2006), [hep-ex/0601046].

[70] BABAR Collaboration, B. Aubert *et al.*, hep-ex/0408068.

Appendix A

Data-MC Comparisons

The analysis relies on several observables for selecting signal events as well as ensuring that the events used in the analysis have a well reconstructed hadronic system. In this section we study the agreement between the data and MC simulation for the main variables used in the event selection.

These variables are as follows:

- Number of charged particles in the X system + ℓ, (nchg);
- Number of neutral particles in the X system, (nneu);
- Number of $K^{\pm}$ in the X system, (nkp);
- Number of K_S in the X system, (nks);
- Energy of the X system, (EX);
- Momentum of the lepton in the B rest frame, (pcms);
- Total charge of the event, (Qtot);
- $E_{\rm miss} - p_{\rm miss}$, (Emiss - pmiss);
- Missing mass square, (mm2); and
- Partially reconstructed D^* tagging variable, (wdeltam).

To obtain the distributions on the various variables, we use the same procedure as for the equidistantly binned m_X and m_X^2 spectra (see Secs. 4.1.2 through 4.1.5) and apply all analysis cuts except the cuts on the variable under study.

We generally find good agreement between data and MC simulation. Differences are seen in the neutrino-related distributions ($E_{\rm miss} - p_{\rm miss}$ and $m^2_{\rm miss}$). These are addressed below. The differences in $q_{\rm tot}$ are attributed to a statistical fluctuation, while the difference in the number of neutral daughters (nneu) is accounted for in the systematic uncertainties.

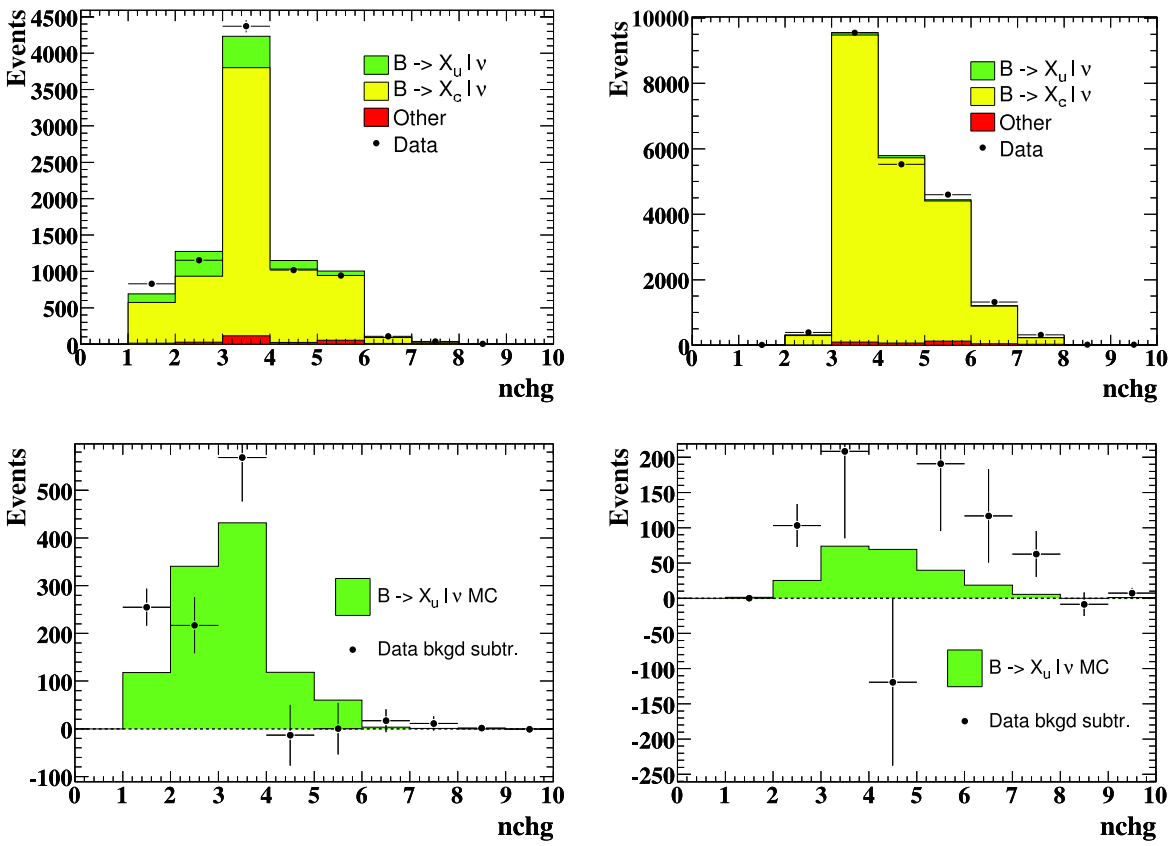

Figure A.1. Number of charged particles on the signal side: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

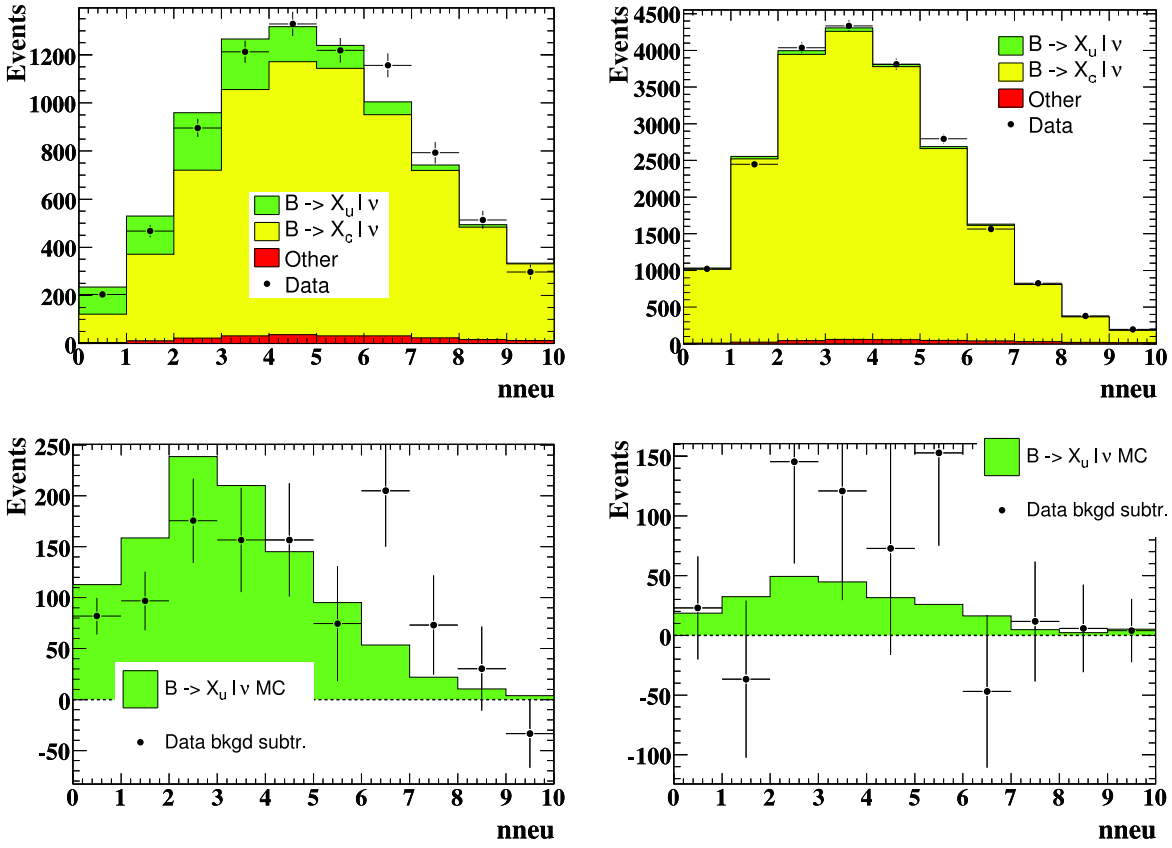

Figure A.2. Number of neutral particles on the signal side: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

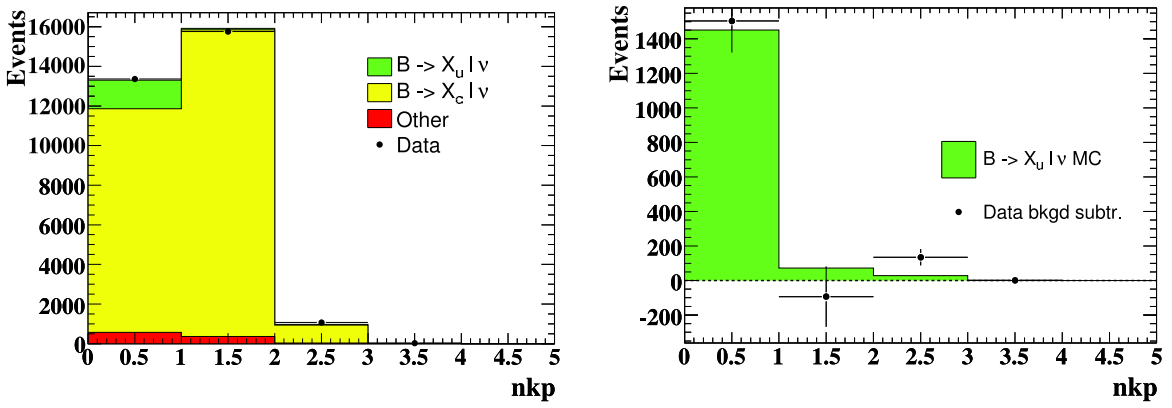

Figure A.3. Number of charged kaons on the signal side: (left) before and (right) after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

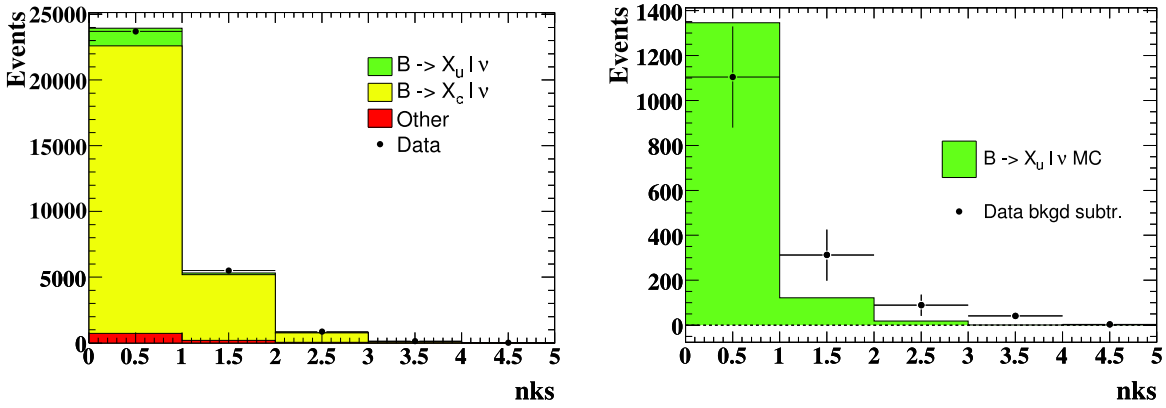

Figure A.4. Number of K_S on the signal side: (left) before and (right) after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

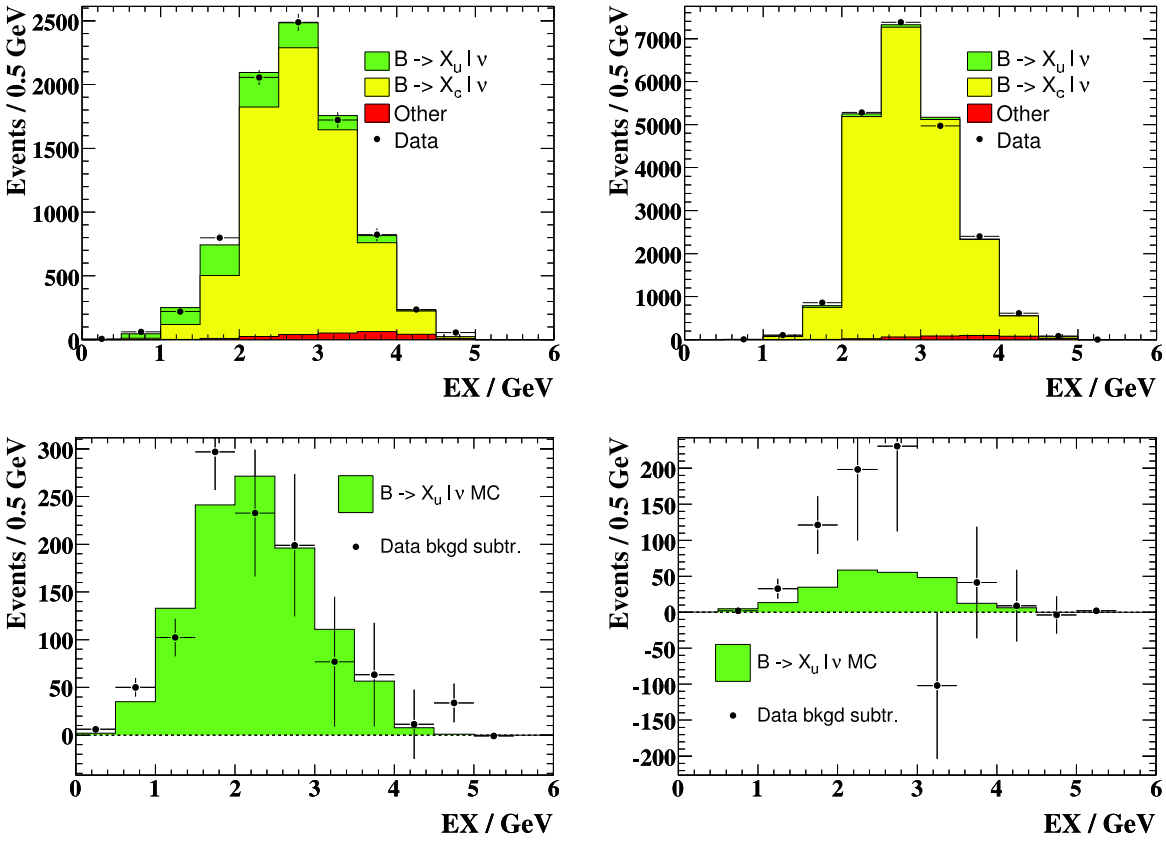

Figure A.5. Hadronic energy on the signal side: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

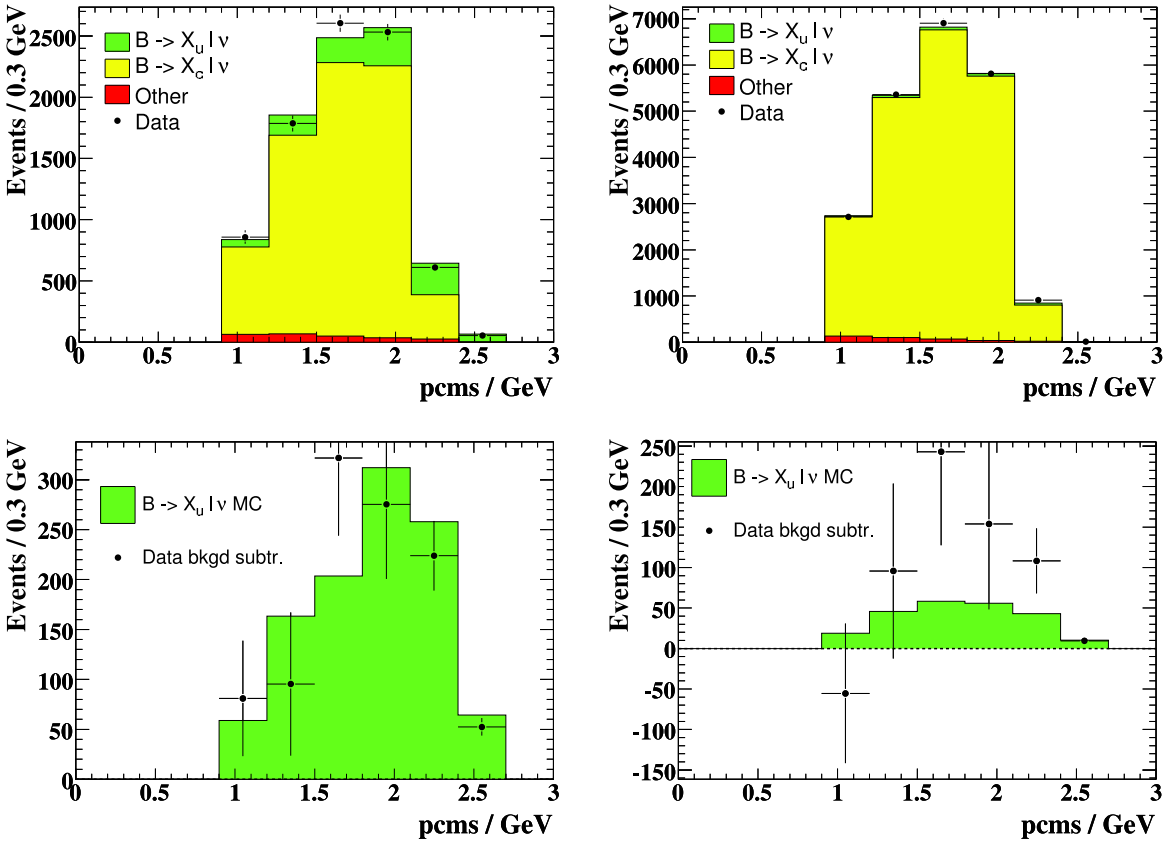

Figure A.6. Momentum of the charged lepton in the B rest frame: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

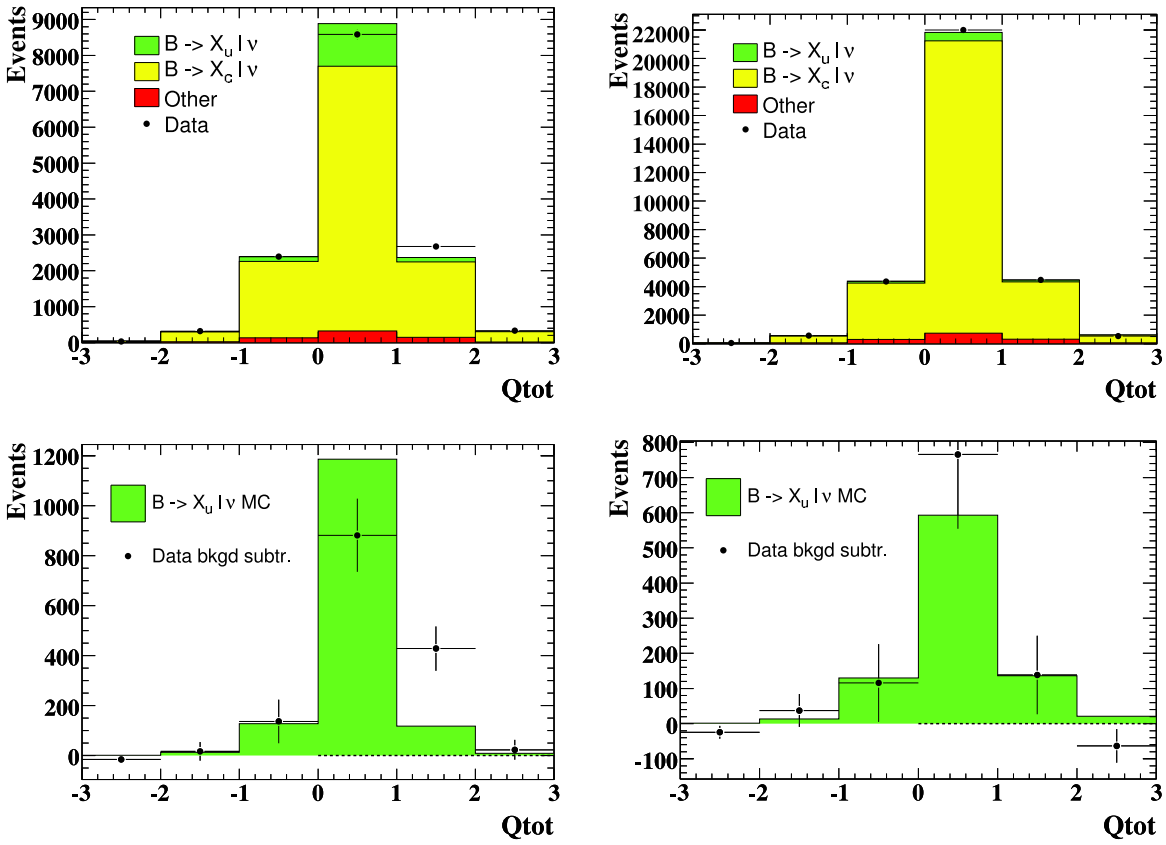

Figure A.7. Total charge of the event: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

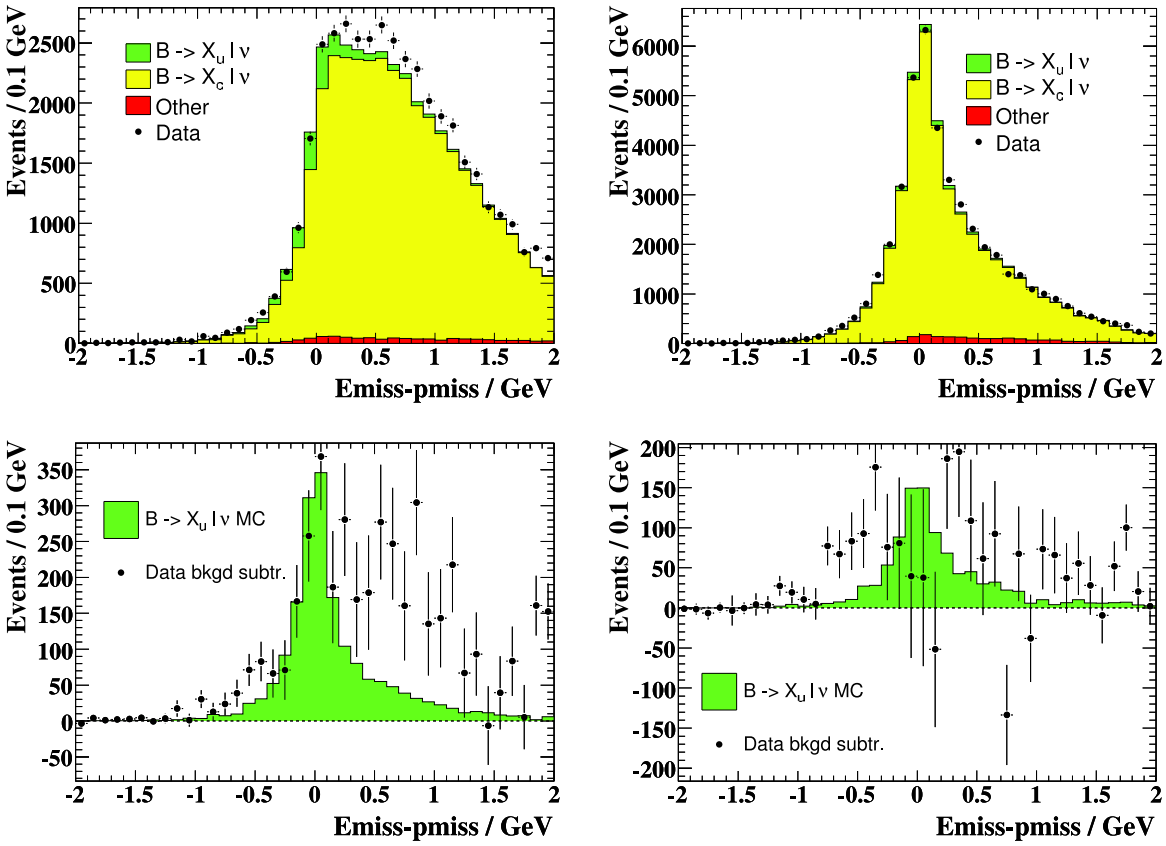

Figure A.8. $E_{\rm miss} - p_{\rm miss}$: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

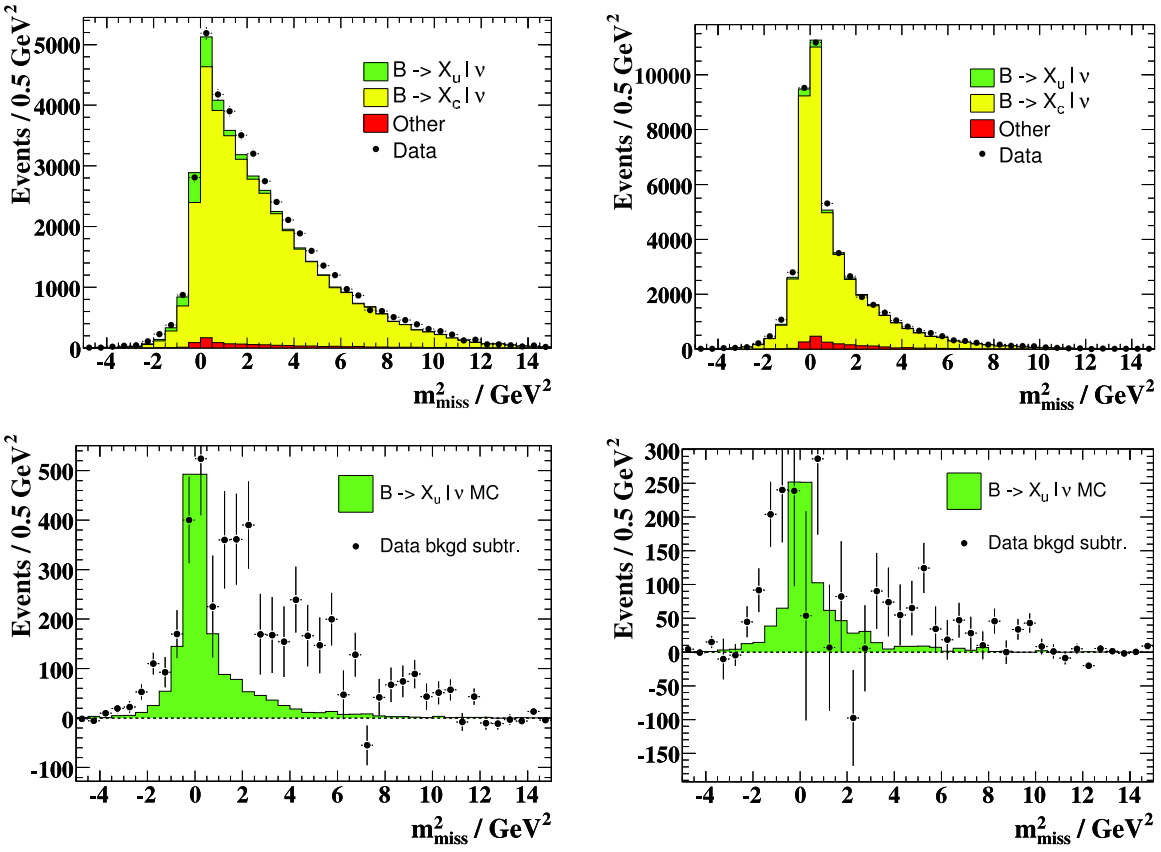

Figure A.9. $m^2_{\rm miss}$: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds.

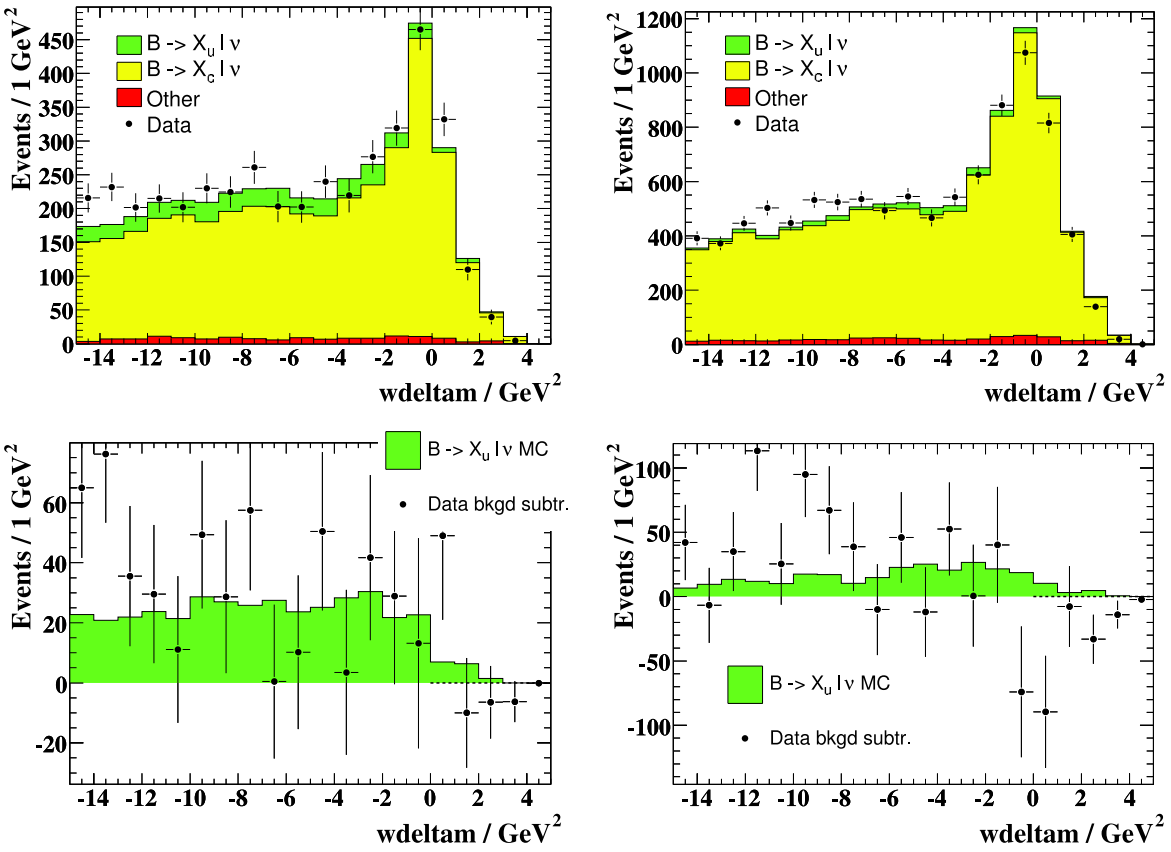

Figure A.10. Partially reconstructed D^* tagging variable: (top left) $B \to X_u \ell\nu$-enriched, (top right) $B \to X_u \ell\nu$-depleted, (bottom left) $B \to X_u \ell\nu$-enriched after subtraction of $B \to X_c \ell\nu$ and other backgrounds, (bottom right) $B \to X_u \ell\nu$-depleted after subtraction of $B \to X_c \ell\nu$ and other backgrounds. Enrichment and depletion is done only using kaons.

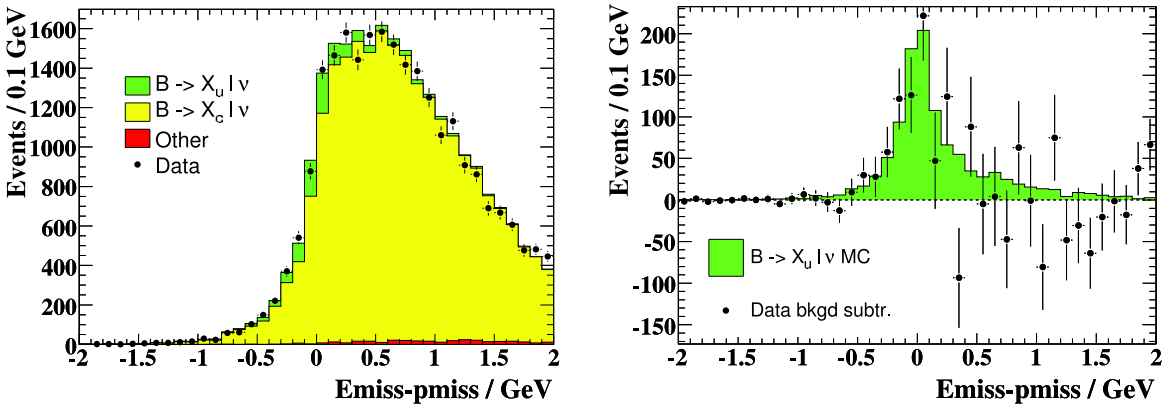

Figure A.11. $E_{\text{miss}} - p_{\text{miss}}$ with electrons as signal lepton after $p^* - m^2_{\text{miss}}$ reweighting for the signal-enriched sample: before (left) and after (right) subtraction of $B \to X_c e\nu$ and other backgrounds.

Investigation of the Neutrino Variables

As presented in Figs. A.9 and A.8, the level of agreement between data and MC simulation in the ν variables is rather poor in the signal-enriched sample, while it is fair in the signal-depleted sample. Generally, disagreements could stem from an inadequate MC simulation of the physics processes (signal and/or background), the detector response, or a combination of both.

We investigated several possible sources for the observed discrepancy: We perform comparisons in restricted regions of phase space to enrich the sample in non-charm background ($m_X > 2.5\,\text{GeV}$ and $p^* < 1.5\,\text{GeV}$) and signal ($m_X < 1.8\,\text{GeV}$ and $p^* > 2.2\,\text{GeV}$), and conclude that the MC simulation for non-charm backgrounds and the signal are adequate. We find no correlation of the discrepancy with the charge of the decaying B meson. Furthermore, we split the sample by lepton flavor. We find that the agreement between data and MC simulation for m^2_{miss} is good in the muon sample, while the discrepancy observed on the full sample is visible in the electron sample. We find that the disagreement is not located in any range of lepton momentum in the B rest frame p^* or laboratory frame p_{lab}, but takes a different shape as a function of p^*.

Reweighting the Electron Sample

Given that we observe the disagreement in the electron sample and that it takes different shapes in different lepton momentum regions, we introduce a two-dimensional binning in p^* $(1.0, 1.3, 1.6, 1.9, 2.7\,\text{GeV})$ and m^2_{miss} (40 equidistant bins between $-5\,\text{GeV}^2$ and $15\,\text{GeV}^2$). In these bins, we compute weights for the $B \to X_c \ell\nu$ events in MC simulation to correct for the observed differences. By construction, the m^2_{miss} distributions agree well after applying the reweighting. We observe that the agreement for $E_{\text{miss}} - p_{\text{miss}}$ is also good (Fig. A.11) and at the same level as for the muon sample (Fig. A.12), for which no additional reweighting is applied.

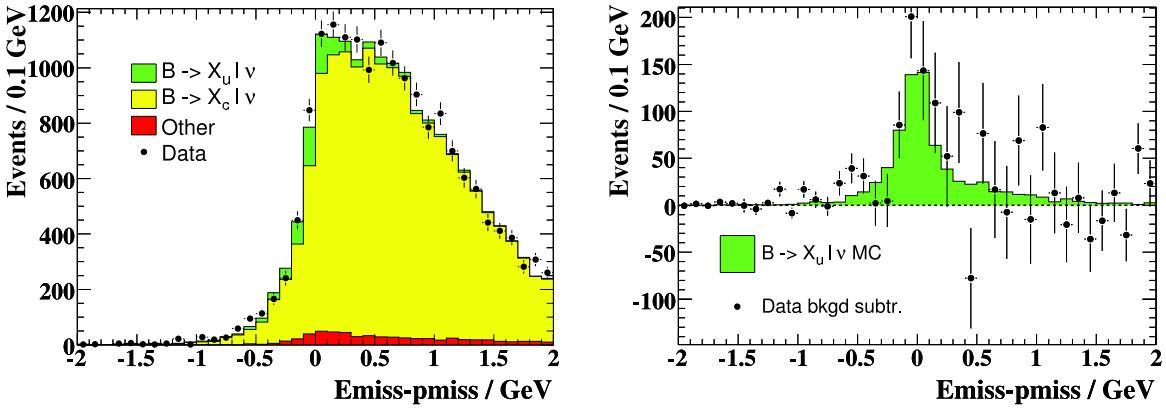

Figure A.12. $E_{\text{miss}} - p_{\text{miss}}$ with muons as signal lepton after $p^* - m^2_{\text{miss}}$ reweighting for the signal-enriched sample: before (left) and after (right) subtraction of $B \to X_c \mu\nu$ and other backgrounds.

Appendix B

Definition of the Moments

This section introduces the moments we will compute from the unfolded spectrum.

For a probability density function with $\int f(x)\mathrm{d}x = 1$, the n^{th} moment $\mathcal{M}_n$ is defined as

$$\mathcal{M}_n = \int x^n f(x)\mathrm{d}x, \tag{B.1a}$$

and the n^{th} central moment $\mathcal{U}_n$ as

$$\mathcal{U}_n = \int (x - \mathcal{M}_1)^n f(x)\mathrm{d}x. \tag{B.1b}$$

For a subrange of x, $0 < x < x_2$, $\mathcal{M}_n^{x_2}$ and $\mathcal{U}_n^{x_2}$ are defined as

$$\mathcal{M}_n^{x_2} = \frac{\int^{x_2} x^n f(x)\mathrm{d}x}{\int^{x_2} f(x)\mathrm{d}x} \quad \text{and} \quad \mathcal{U}_n^{x_2} = \frac{\int^{x_2} (x - M_1^{x_2})^n f(x)\mathrm{d}x}{\int^{x_2} f(x)\mathrm{d}x}. \tag{B.1c}$$

For a discrete spectrum, given as a histogram with N bins, normalized to unit area, the above definitions turn into

$$M_n = \sum_{i=1}^{N} m_i^n h_i \quad \text{and} \quad U_n = \sum_{i=1}^{N} (m_i - M_1)^n h_i \tag{B.2a}$$

and

$$M_n^K = \frac{\sum_{i=0}^{K} m_i^n h_i}{\sum_{i=0}^{K} h_i} \quad \text{and} \quad U_n^K = \frac{\sum_{i=0}^{K} (m_i - M_1^K)^n h_i}{\sum_{i=0}^{K} h_i}, \tag{B.2b}$$

respectively, where the bin centers are denoted by m_i and the content of the i^{th} bin is denoted by h_i.

The use of a finite bin width will in general bias the discretely computed moments M_n and U_n with respect to the moments $\mathcal{M}_n$ and $\mathcal{U}_n$, respectively, obtained from the respective continuous function. At this point we make no attempt to correct for this but give the moments M_1 and U_n ($n = 2, 3$), which we obtain from the unfolded m_X^2 spectrum, with a bin width of $0.8\,\mathrm{GeV}^2$.