Analysis of inclusive semileptonic B meson decays with τ lepton final states at the Belle experiment

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Jan Hasenbusch
aus
Herdecke

Bonn, 2017

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Jochen Dingfelder
2. Gutachter:	Prof. Dr. Klaus Desch
Tag der Promotion:	19.12.2017
Erscheinungsjahr:	2018

Für Reinhard
— Sigrid & Sophia

Abstract

In this thesis, the measurement of the inclusive branching fraction of $B \to X\tau\nu$ decays at the Belle experiment is presented. Previous measurements of the inclusive semileptonic b-hadron branching fraction $\mathcal{B}(b \to X\tau\nu)$ in $Z^0 \to b\bar{b}$ decays at LEP agree well with the prediction of the Standard Model of particle physics. Conversely, the B-factory experiments *BABAR*, Belle and LHCb consistently measured the exclusive branching fractions of the decays $B \to D\tau\nu$ and $B \to D^*\tau\nu$ larger than the theoretical prediction. The combined exclusive branching fraction measurements together exceed the Standard Model by 4σ and are even larger than the average of inclusive LEP measurements.

This work is the first measurement of the branching fraction of $B \to X\tau\nu$ decays at a B-factory. It is therefore a valuable cross check to the exclusive analyses and also probes the inclusive LEP measurements in a different experimental environment.

At Belle, a large sample of 770×10^6 $B\overline{B}$ pairs was recorded in e^+e^- collisions at the $\Upsilon(4S)$ resonance. One of the two B mesons in the $B\overline{B}$ event is fully reconstructed in a hadronic decay mode (B_{tag}) and the other one (B_{signal}) is analysed regarding the signal decay. This hadronic B tagging allows one to constrain the kinematics and the flavour of the signal B meson and enables the study of observables in the B_{signal} rest frame. The signal is selected in the leptonic decay modes of the τ and thus signal-like events are required to have a single electron or muon in the final state.

This analysis relies on Monte Carlo simulations of backgrounds, particularly on those of semileptonic B decays. The latter produce mainly charmed mesons of which $\sim$25% are the poorly known D^{**} mesons. To improve the modelling of the data, a new model of these excited charmed mesons and the related semileptonic B decays is developed and implemented in dedicated Monte Carlo samples.

Observables for the signal extraction are the lepton momentum and the squared sum of four-momenta of undetected particles (squared missing mass) in the event. The signal branching fraction is measured, normalised to the branching fraction of $B \to X\ell\nu$ ($\ell = e, \mu$) decays, to be

$$R(X) = \frac{\mathcal{B}(B \to X\tau\nu)}{\mathcal{B}(B \to X\ell\nu)} = 0.298 \pm 0.012_{\text{stat}} \pm 0.018_{\text{sys}}.$$

Assuming the isospin-average branching fraction $\mathcal{B}(B \to X\ell\nu) = (10.86 \pm 0.16)\%$, this translates to a branching fraction of

$$\mathcal{B}(B \to X\tau\nu) = (3.23 \pm 0.13_{\text{stat}} \pm 0.21_{\text{sys}})\,\%$$

and is the most precise single measurement of $B \to X\tau\nu$ decays so far.

Contents

Preface

With the discoveries of numerous new particles in the 20^{th} century, the theory to describe them evolved to today's *Standard Model of particle physics* (SM). The existence of particles predicted by the SM, e. g. the Ω baryon and the three heavy quarks, is one of the great achievements of this model.

In 1970, the *GIM-mechanism* [1] was introduced to explain the absence of flavour-changing neutral currents (FCNC) in the SM. In addition to the three light quarks which were known at that time, the mechanism required the existence of a fourth quark, which is today known as the charm quark. Based on the work of N. Cabibbo [2], M. Kobayashi and T. Maskawa in 1973 showed that expanding the Cabibbo quark-mixing matrix from two to three quark generations naturally introduces a CP violating phase, explaining the observation of CP violation in neutral Kaon decays (*CKM mechanism*) [3]. The predicted third quark generation contains two new quarks, named *bottom* and *top* (or sometimes beauty and truth) which were indeed found in 1977 [4] and 1995 [5, 6], respectively.

To confirm the time-dependent CP violation in the B sector predicted by the CKM mechanism and to study the parameters of the latter, the B-factory experiments came up in the 1990s. The Belle experiment at the KEK in Japan and the *BABAR* experiment at SLAC in the US were designed to record huge data sets of B meson decays and finished their data taking in 2010 and 2008, respectively. With the observation of time-depended CP violation in the B system in the year 2001 by both experiments [7, 8], one of the last open predictions of the SM was confirmed. Kobayashi and Maskawa were rewarded with the Nobel Prize for the CKM mechanism in 2008.

Until today, the SM has been confirmed in many studies, however, one particular particle was still missing until a few years ago: the Higgs boson. It was predicted as a consequence of electro-weak symmetry-breaking and introduced to explain the masses of the elementary particles [9, 10]. With the announcement of the Higgs boson observation in 2012, the Large Hadron Collider (LHC) experiments ATLAS and CMS at CERN completed the search for missing SM particles [11, 12] which lead to the Nobel Prize for P. Higgs and F. Englert one year later.

The SM is a very successful theory that robustly predicted new particles and interactions, nonetheless, it is known to be yet incomplete. Several aspects are not explained by the SM at all; examples are the mass of the Higgs boson and its incapability to unify all three fundamental forces. Hence, numerous models for *New Physics* (NP) beyond the SM are proposed. Supersymmetric extensions to the SM (SUSY) can solve the mentioned issues, but predict a variety of new, yet undiscovered particles. Searches for NP particles are conducted, either as *direct searches* looking for the production of those particles at high energies, or complementary, as *indirect searches* for NP signatures at lower energies which is the regime of the B-factories.

In particular, semileptonic B decays involving a τ lepton show interesting deviations from the SM prediction as reported by Belle, *BABAR* and LHCb [13–17]. The combined measurements of

the branching fractions of $B \to D\tau\nu$ and $B \to D^*\tau\nu$[1] are four standard deviations (4σ) above the SM calculations [18]. These deviations may be explained by new Higgs-like bosons and are thus of special interest.

A complementary cross check to the results of the exclusive $B \to D^{(*)}\tau\nu$ decays is the measurement of the inclusive $B \to X\tau\nu$ branching fraction which is the subject of this thesis using the large Belle data set.

After a theoretical and experimental overview in Chapter 1, the Belle experiment and its data and Monte-Carlo samples are introduced in Chapter 2 and 3, respectively. For an improved description of signal and background events, a model of excited charmed meson decays was developed and incorporated in a dedicated MC sample (Chapter 4). The analysis is described in Chapter 5 and Chapter 6 gives the final results which are discussed in Chapter 7. A summary of this work is given in Chapter 8.

[1] If no charges are given, all Isospin and CP conjugates are meant (e. g. B stands for B^+,B^-,B^0,$\bar{B}^0$). If charges are given, the CP conjugate is also meant unless stated otherwise.

CHAPTER 1

Introduction

Testing the predictions of the Standard Model of particle physics (SM) is a key aspect to understand nature at smallest scales. Possible deviations from the SM are hints to New Physics beyond the SM. The decay of a B meson involves a transition of its constituent b quark to a c or u quark by the weak force. Sec. 1.1 gives a brief introduction to the SM and the relevant interactions are discussed. The description of semileptonic B meson decays in the SM is given in Sec. 1.2. Secs. 1.3 and 1.4 discuss the experimental status of semileptonic B meson decays with τ lepton final states and their possible implications on New Physics, respectively.

1.1 The Standard Model of particle physics

The SM describes the elementary particles and their interactions (see Fig. 1.1). Spin-$\frac{1}{2}$ fermions interact via the exchange of spin-1 bosons of the electromagnetic, weak and strong forces. The fermions are divided into quarks and leptons, of which only the quarks participate in the strong force. Furthermore, the fermions are structured in three "generations" or "families", with the second and third generations being heavier copies of the first generation.

The SM is a quantum field theory (QFT), described by the gauge symmetry group

$$\mathrm{SU(3)_{colour}} \times \mathrm{SU(2)}_T \times \mathrm{U(1)}_Y, \tag{1.1}$$

which corresponds to transformations in the space of the colour charge, the weak isopin T and the weak hypercharge Y, respectively. The strong and the weak interactions are introduced in the following sections.

1.1.1 Strong interaction

The theory of strong interactions is described by quantum chromodynamics (QCD). Quarks are strongly interacting particles bound together to mesons ($q\bar{q}$) and baryons (qqq). They carry colour charge, i. e. one of the colours red, green or blue (anti-quarks carry the corresponding anti-colour) and interact via exchange of massless gluons. The $\mathrm{SU(3)_{colour}}$ symmetry gives rise to nine massless bosons which are the eight gluons (themselves coloured, carrying colour–anti-colour combinations) and a colourless singlet which thus cannot participate in the strong interaction. As the gluons carry colour, they self-interact which leads to an effective anti-shielding and the *confinement* of quarks. This means, at large distances, or equivalently small momentum transfer q^2, the coupling becomes arbitrarily large. Due to the confinement, there are no free

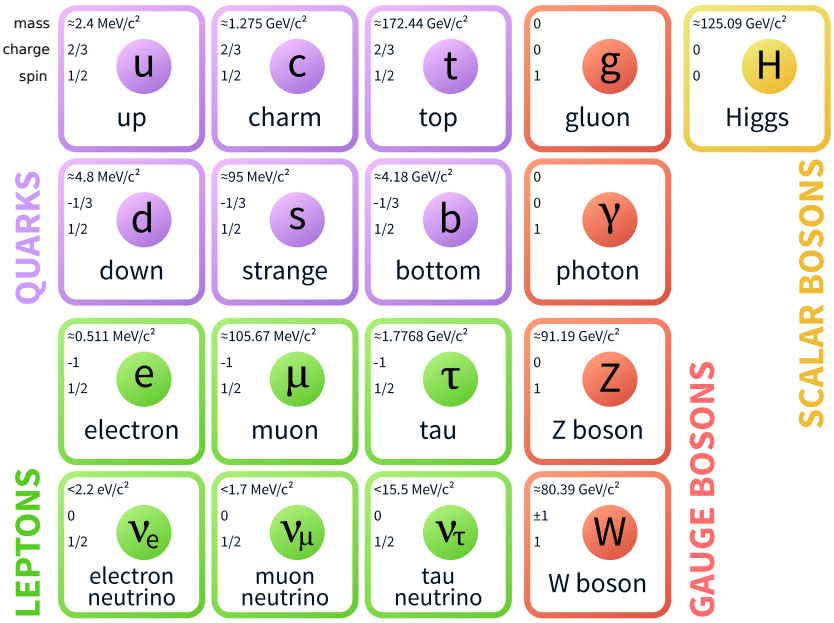


Figure 1.1: The particles of the Standard Model. Matter is made of fermions, i. e. quarks and leptons, whereas the heavier generations (second and third column) decay fast into the lighter first generation (first column). Their decays and interactions are mediated by the gauge vector bosons of the three fundamental forces (see text). The scalar Higgs boson is a consequence of the Higgs field which gives rise to the particle masses. Adapted from [19].

coloured objects, i. e. quarks are only found in colour-neutral bound states. In contrast, at small distances or large energies, the coupling is small and the bound quarks approximately behalf as free particles which is called *asymptotic freedom*. In this regime, the QCD processes can be calculated perturbatively. The coupling of the strong force, α_s, depends on the momentum transfer as

$$\alpha_s(q^2) = \frac{12\pi}{(33 - 2n_f)\ln(q^2/\Lambda^2_{\mathrm{QCD}})}, \tag{1.2}$$

with the number of quark-flavours n_f. The energy scale of the strong interaction, Λ_{QCD}, separates the short and long distance scales and is roughly set by the size of a hadron

$$R_{\mathrm{had}} \sim \frac{1}{\Lambda_{\mathrm{QCD}}} \sim 1\,\mathrm{fm} \Rightarrow \Lambda_{\mathrm{QCD}} \sim 0.2\,\mathrm{GeV}. \tag{1.3}$$

For systems of heavy quarks Q ($Q = c, b, t$) with a mass $m_Q > \Lambda_{\mathrm{QCD}}$, the momentum transfer is large and thus, for example, $Q\bar{Q}$ quarkonium states have a small α_s coupling. For this reason, they have a similar spectrum of bound states to those of the electromagnetic force.

1.1.2 Electro-weak interaction

The generators of the electro-weak symmetry group correspond to three (massless) vector bosons W_1, W_2 and W_3 from the weak isopin $\mathrm{SU}(2)_T$ group as well as a vector boson B from the hypercharge symmetry $\mathrm{U}(1)_Y$. These states mix into the physical W^+, W^-, Z^0, γ particles after the electro-weak symmetry breaking.

The electro-weak symmetry is found to be broken as high masses for the gauge bosons of the weak interaction of $m_W = (80.39 \pm 0.02)\,\mathrm{GeV}$ and $m_{Z^0} = (91.188 \pm 0.002)\,\mathrm{GeV}$ [20] were

measured and the photon is massless. The electro-weak symmetry breaking, mediated by the Higgs mechanism, is based on an additional complex scalar field doublet Φ (*Higgs field*). The Higgs field has a particularly shaped potential with a non-vanishing vacuum expectation value (vev), i. e. a ground state with reduced symmetry. At energies of $\Lambda_{\mathrm{QED}} \sim m_W$, the Higgs field settles in the minimum of the potential and the symmetry is spontaneously broken. Four possible new excitations of the Higgs field appear (so-called Goldstone bosons). These new degrees of freedom are identified as the longitudinal components of the three weak gauge bosons, which thus become massive. The remaining Goldstone boson is the spin-0 Higgs boson which was found in 2012 [12, 21] and is a further proof of the concept of symmetry breaking in the SM.

Classical physics is invariant under parity transformations $P\vec{x} = -\vec{x}$, i. e. point-mirroring the spatial components. It was thus unexpected when the experiment by Wu in 1957 [22] showed that the charged weak interaction maximally violates parity, i. e. only left-handed fermions and right-handed anti-fermions participate in the weak interactions ($V - A$ theory). This is expressed by a projector P_L in the charged weak four-current

$$j^\mu = \bar{\ell}\gamma^\mu P_L \nu, \tag{1.4}$$

which projects the fermion fields ℓ, ν onto their left-handed states $(\ell, \nu)_L = P_L(\ell, \nu)$. The projection operator P_L is given by the Dirac gamma matrices γ^μ as $P_L = \frac{1}{2}(1 + \gamma^5)$ and $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

Even though the weak interaction does not conserve parity, it was expected that the combination of charge and parity (CP) is conserved. The charge conjugation C turns a particle f into its anti-particle $\bar{f}$ by $\mathrm{C}f = \pm\bar{f}$, and thus, CP transforms a left-handed particle into its right-handed anti-particle. However, also this symmetry is broken in the weak interaction as first found in the decay of long-living neutral Kaons in 1964 [23]. The violation of CP as well as weak transitions across quark generations are implemented by the mechanism named after Cabibbo, Maskawa and Kobayashi (*CKM mechanism*) [3]. If the weak interaction acted only on particles in the same isospin doublet (as it is the case for leptons), no transitions between quark generations would occur. However, the observation of quark decays across generations suggests that the flavour- or mass-eigenstates are not the eigenstates of the weak interaction. Thus, the weak quark current of quark-fields u and v is given by

$$J^\mu = V_{qq'}\bar{u}_{q'}\gamma^\mu P_L v_q, \tag{1.5}$$

which contains the matrix element $V_{qq'}$ of the CKM quark mixing matrix as an additional factor. The CKM matrix is a complex rotation matrix which rotates the mass eigenstates of the down-type quarks (d, s and b) into eigenstates of the weak isospin, which also enables the quark transitions across generations. The elements of the CKM matrix are fundamental parameters of the SM and must be measured experimentally. Their magnitudes are found to be [20]:

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.9743 & 0.2251 & 0.0036 \\ 0.2249 & 0.9735 & 0.0411 \\ 0.0088 & 0.0403 & 0.9992 \end{pmatrix}. \tag{1.6}$$

The uncertainties are omitted for readability, but the largest error of 4% is that of the smallest matrix element $|V_{ub}|$. There is a clear hierarchy in the matrix elements, suppressing transitions between different generations (*CKM suppression*), which is represented by the

parameter $\lambda \approx 0.22$ in the Wolfenstein parametrisation [24]

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right), \tag{1.7}$$

where A, ρ and η are additional parameters of the description.

The CKM matrix is a unitary matrix, i. e. $V^\dagger V = I$, due to probability conservation. A unitary complex 3×3 matrix has four free parameters, which can be interpreted as three rotation angles and a non-trivial complex phase. The latter gives rise to the CP violation in the SM. The off-diagonal unitary conditions can be interpreted as triangles in the complex plane. One of these triangles, *the unitary triangle*, is given by the condition

$$V_{ud}V_{ub}^* + V_{td}V_{tb}^* + V_{cd}V_{cb}^* = 0. \tag{1.8}$$

This condition stands out because the sides of the triangle are of comparable lengths and the apex, located at

$$\bar{\rho} + i\bar{\eta} = \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{1.9}$$

is in the experimental reach of the B factories. Various measurements have been conducted to test the CKM mechanism and determine the lengths of the sides and the angles of the unitary triangle, as shown in Fig. 1.2. The studies aim to constrain the apex of the triangle because incompatible measurements, resulting in a non-closed triangle, would be a hint for undiscovered New Physics.

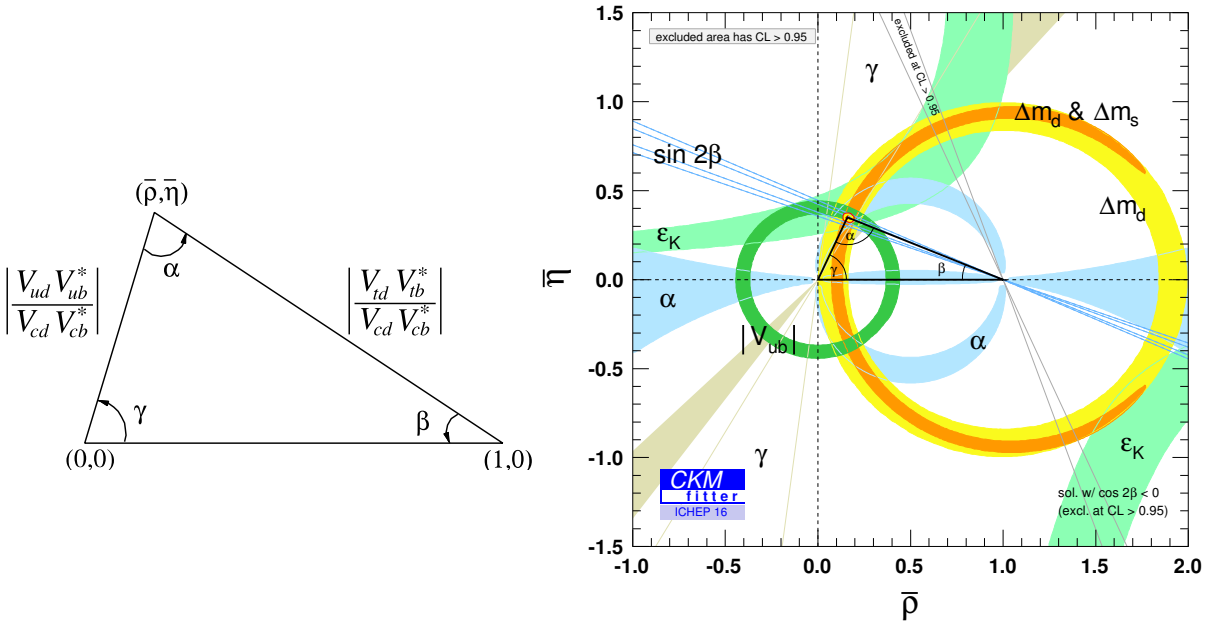


Figure 1.2: *Left:* The unitary triangle and the relevant CKM matrix elements which determine the angles an side lengths. Adapted from [20]. *Right:* The unitary triangle in the complex plane of the Wolfenstein parameters η and ρ. A global fit to various CKM measurements gives the position of the triangle's apex [25] (a similar fit is available from [26]). An inconsistent apex position would be a hint for an inconsistency in the CKM mechanism and for New Physics.

The quark mixing allows for the oscillation of neutral mesons into their anti-meson and vice-versa. The process behind meson–anti-meson mixing is depicted in Fig. 1.3 as a second-order Feynman-diagrams ("box-diagram"). The initial meson decays either in two $W^{\pm}$ or two quarks which afterwards recombine into the anti-meson. Meson–anti-meson mixing has been observed in the B, B_s, D and K systems and can have experimental consequences for analyses of these systems. In the B system, the oscillation occurs at a well observable frequency and the probability that a B^0 oscillates is found to be $\chi = (18.8 \pm 0.2)\%$ [20].

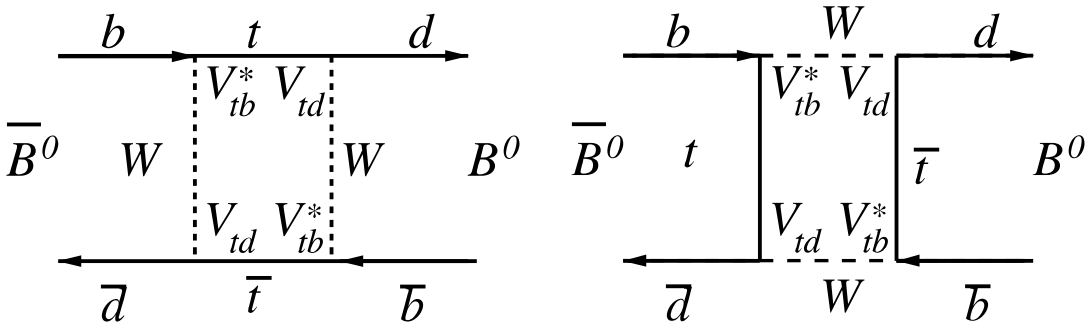


Figure 1.3: Mixing of neutral B mesons through second order Feynman "box-diagrams" [27].

1.2 Semileptonic B decays

Semileptonic $B \to X\ell\nu$ decays are ideal decay modes to measure the magnitudes of the CKM elements $|V_{ub}|$ and $|V_{cb}|$. The leptonic part makes the theoretical description much easier as there are no strong interactions between the leptonic and hadronic final states. Higher-order corrections to the tree-level process from electromagnetic and weak interactions between the final-state particles are estimated with perturbation theory. In contrast, long-distance strong interactions between hadronic final states occur, for which perturbation theory is not valid.

In general, the decay of an initial state particle i to an n-particle final state f is described by *Fermi's Golden Rule*. It relates the partial decay width $\Gamma_{i\to f}$ to the transition matrix element $\mathcal{M}_{if}$ and the phase-space volume Φ_f by

$$\mathrm{d}\Gamma_{i\to f} = \frac{(2\pi)^4}{2m_i}|\mathcal{M}_{if}|^2 \mathrm{d}\Phi_f(p_i^\mu, p_1^\mu, \ldots, p_n^\mu), \tag{1.10}$$

where m_i is the mass of the initial state particle and the p^μ are the four-momenta of the initial and final state particles. The phase space is the amount of kinematically allowed final states and is restricted by energy and momentum conservation. The transition matrix describes the quantum mechanical process as a decay operator $\mathcal{T}$ by $\mathcal{M}_{if} = \langle\psi_i|\mathcal{T}|\psi_f\rangle$. The actual SM interactions happen inside $\mathcal{T}$ and are discussed below.

In the experiments and the theory calculations, two approaches can be pursued to measure and describe (semileptonic) B decays, the inclusive and the exclusive approach. Exclusive analyses consider a concrete final state, e. g. $B \to D^*\ell\nu$. The inclusive counter-part is, for example, the decay $B \to X\ell\nu$, where no explicit final-state meson is considered. A combination are semi-exclusive approaches which investigate properties of the inclusive decay with a sum of exclusive decays. Their advantage is the correct modelling of resonant structures, however, the sum-of-exclusives approach assumes all exclusive states to be known.

1.2.1 Exclusive semileptonic B decays

In semileptonic $B \to X\ell\nu$ decays, the four-momentum transfer $q^\mu = p_\ell^\mu + p_\nu^\mu$ to the leptonic system is negligible compared to the W mass ($q^2 \leq m_B^2 + m_X^2 \ll m_W^2$) and thus are the dynamics of the W. Therefore, the tree-level interaction via the exchange of a W can effectively be expressed as a four-particle interaction (*Fermi's interaction*) with a coupling proportional to the Fermi constant G_F. The full matrix element of the semileptonic decay $B \to D^{(*)}\ell\nu$ then reads

$$\mathcal{M} = -\left[\bar{\ell}\gamma^\mu(1-\gamma^5)\nu\right]\frac{iG_F}{\sqrt{2}}\langle D^{(*)}|V_{cb}\bar{c}\gamma_\mu(1-\gamma^5)b|B\rangle, \tag{1.11}$$

whereas the left and right factors represent the leptonic and hadronic parts of the interaction, respectively.

The $|B\rangle \to |D^{(*)}\rangle$ transitions involve several form factors due to the strong interactions inside the mesons. It is convenient to define the recoil parameter $w = v^\mu \cdot v'_\mu$ as the product of the velocities of the initial and the final-state hadron (the latter is denoted with a prime). The calculations of the hadronic matrix element are split into the vector and the axial-vector component because the latter vanishes for a scalar final-state hadron. The vector-current matrix elements of the transitions $|B\rangle \to |D\rangle$ and $|B\rangle \to |D^*\rangle$ are given in terms of the two form factors $h_\pm(w)$ and a single vector form factor $h_V(w)$ [28], respectively,

$$\left\langle D(v')\left|\bar{c}\gamma^\mu b\right|B(v)\right\rangle = h_+(w)(v+v')^\mu + h_-(w)(v-v')^\mu \tag{1.12}$$

and

$$\left\langle D^*(v',\epsilon')\left|\bar{c}\gamma^\mu b\right|B(v)\right\rangle = ih_V(w)\epsilon^{\mu\nu\alpha\beta}\epsilon'_\nu v'_\alpha v_\beta\,, \tag{1.13}$$

where $\epsilon^{\mu\nu\alpha\beta}$ is the Levi-Civita symbol and ϵ' the polarisation of the D^*. The $|B\rangle \to |D^*\rangle$ axial-vector current with three form factors $h_{A_i}(w)$ is given by

$$\left\langle D^*(v',\epsilon')\left|\bar{c}\gamma^\mu\gamma^5 b\right|B(v)\right\rangle = h_{A_1}(w)(w+1)\epsilon'^\mu - \left[h_{A_2}(w)v^\mu + h_{A_3}(w)v'^\mu\right]\epsilon' v\,. \tag{1.14}$$

From these expressions, the differential decay rates are derived as

$$\frac{\mathrm{d}\Gamma(B\to D\ell\nu)}{\mathrm{d}w} = \frac{G_F^2|V_{cb}|^2 m_D^3}{48\pi^3}(m_B+m_D)^2(w^2-1)^{3/2}|\eta_\mathrm{EW}\mathcal{G}(w)|^2 \tag{1.15}$$

and

$$\frac{\mathrm{d}\Gamma(B\to D^*\ell\nu)}{\mathrm{d}w} = \frac{G_F^2|V_{cb}|^2 m_{D^*}^3}{48\pi^3}(m_B-m_{D^*})^2\sqrt{w^2-1}(w+1)^2 \times\left[1+\frac{4w}{w+1}\frac{m_B^2-2wm_Bm_{D^*}+m_{D^*}^2}{(m_B-m_{D^*})^2}\right]|\eta_\mathrm{EW}\mathcal{F}(w)|^2\,, \tag{1.16}$$

with the form factors $\mathcal{G}$ and $\mathcal{F}$, which are discussed in the next paragraph. The factor $\eta_\mathrm{EW} = 1.015 \pm 0.005$ accounts for electro-weak corrections and is estimated numerically in Ref. [29]. In contrast to the decay rate into the scalar D meson, the decay rate into the vector D^* meson depends on three angles which are already integrated out in Eq. 1.16 for simplicity. The angles further describe the decay, i. e. the distribution of final state particles due to the D^* polarisation.

1.2.2 Heavy-quark effective theory

The following section on heavy-quark effective theory (HQET) is based on the reviews in Refs. [30, 31]. The hadronic form factors for the $b \to c$ transition have to be calculated. The form factors of mesons containing a heavy b or c quark are of particular interest as they obey approximate symmetries which can be exploited. In the limit $m_Q \to \infty$, the QCD Lagrangian becomes independent of the heavy-quark flavour (i. e. the mass) and spin. The light degrees of freedom, i. e. the light quarks and gluons, decouple and the wave function approximately factorises (cf. Eq. 1.25). They strongly interact with the heavy quark, but independently of its properties such as spin or flavour. For large m_Q, the heavy quark moves with the velocity v of the hadron and its four-momentum can be written as

$$p^\mu = m_Q v^\mu + k^\mu, \tag{1.17}$$

with the residual momentum $k \sim \Lambda_{\mathrm{QCD}}$. The residual momentum k is due to the continuous interaction with the light degrees of freedom and becomes negligible for $m_Q \to \infty$. The resulting constant velocity in this limit allows one to write the QCD Lagrangian of the heavy quark as an effective Lagrangian,

$$\mathcal{L}_{\mathrm{Q}} = \mathcal{L}^{(0)}_{\mathrm{HQET}} + \mathcal{L}^{(1)}_{\mathrm{HQET}} + \mathcal{O}\left(m_Q^{-2}\right), \tag{1.18}$$

which goes under the term of *heavy-quark effective theory*. The superscript denotes the order of the $1/m_Q$ dependence, i. e. the leading order is independent of the heavy-quark mass. The heavy-quark flavour and spin symmetry is broken at $\mathcal{L}^{(1)}_{\mathrm{HQET}} \sim 1/m_Q$ due to the appearance of the kinetic and chromomagnetic energy operators. In the limit $m_Q \to \infty$, i. e. spin-flavour symmetry, the form factors in Eqs. 1.12 – 1.14 reduce to a single, so-called Isgur-Wise function which is interpreted as the overlap of the wave function of the light degrees of freedom in the initial and the final state.

The CLN parametrisation for $B \to D^{(*)}\ell\nu$ The spin-flavour symmetry breaks down at finite masses and the form factors can only be found approximately. Below, the expansion of the form factors in Eq. 1.15 and 1.16 is given in the parametrisation by Capini, Lellouch and Neubert (CLN) [32]. After introducing the conformal mapping $z = \frac{\sqrt{w+1}-\sqrt{2}}{\sqrt{w+1}+\sqrt{2}}$, the form factors are expanded around $z = 0$, i. e., vanishing meson recoil. The $B \to D\ell\nu$ form factor then reads

$$\frac{\mathcal{G}(w)}{\mathcal{G}(w=1)} \approx 1 - 8\rho_D^2 z + (51\rho_D^2 - 10)z^2 - (252\rho_D^2 - 84)z^3, \tag{1.19}$$

where the slope parameter ρ_D^2 and the normalisation $\mathcal{G}(w=1)$ are CLN parameters that can be determined experimentally. Using the mass ratio $r = m_{D^*}/m_B$, the form factor of $B \to D^*\ell\nu$ is expressed as

$$\begin{aligned}\mathcal{F}(w) = &\left[1 + \frac{4w}{1+w}\frac{1-2wr+r^2}{(1-r)^2}\right]^{-1} \\ &\times \left\{2\frac{1-2wr+r^2}{(1-r)^2}\left[1 + \frac{w-1}{w+1}R_1(w)^2\right] + \left[1 + \frac{w-1}{1-r}\left(1 - R_2(w)^2\right)\right]^2\right\} A_1(w)^2\end{aligned} \tag{1.20}$$

with the axial-vector form factor

$$\frac{A_1(w)}{A_1(w=1)} = h_{A_1}(w)\frac{w+1}{2rA_1(1)} \approx 1 - 8\rho^2_{D^*}z + (53\rho^2_{D^*} - 15)z^2 - (231\rho^2_{D^*} - 91)z^3 \tag{1.21}$$

and the expansion of the form-factor ratios

$$R_1(w) = \frac{h_V(w)}{h_{A_1}(w)} \approx R_1(w=1) - 0.12(w-1) + 0.05(w-1)^2 \quad \text{and} \tag{1.22}$$

$$R_2(w) = \frac{h_{A_3}(w) + rh_{A_2}(w)}{h_{A_1}(w)} \approx R_2(w=1) + 0.11(w-1) - 0.06(w-1)^2\,. \tag{1.23}$$

The differential decay rates of $B \to D\ell\nu$ and $B \to D^*\ell\nu$ are thus parametrised by $|\eta_{\text{EW}}\mathcal{G}(1)|$, ρ^2_D and $|\eta_{\text{EW}}\mathcal{F}_1(1)|$, $\rho^2_{D^*}$, $R_1(1) \equiv R_1$, $R_2(1) \equiv R_2$, respectively. They have been measured experimentally and are summarised in Tab. 1.1 [18] . The correlations between the $B \to D^*\ell\nu$ form-factor parameters have been measured to be [33]:

$$\rho(\rho^2_{D^*}, R_1) = 0.566 \qquad \rho(\rho^2_{D^*}, R_2) = -0.807 \qquad \rho(R_1, R_2) = -0.758. \tag{1.24}$$

$B \to D\ell\nu$	$\lvert\eta_{\text{EW}}\mathcal{G}(1)\rvert\lvert V_{cb}\rvert$	$(35.81 \pm 0.45) \cdot 10^{-3}$
	ρ^2_D	1.19 ± 0.05
$B \to D^*\ell\nu$	$\lvert\eta_{\text{EW}}\mathcal{F}_1(1)\rvert\lvert V_{cb}\rvert$	$(42.65 \pm 1.53) \cdot 10^{-3}$
	$\rho^2_{D^*}$	1.21 ± 0.03
	R_1	1.40 ± 0.03
	R_2	0.85 ± 0.02

Table 1.1: Measured parameters of the semileptonic $B \to D^{(*)}\ell\nu$ decays in the CLN parametrisation [18].

Excursus on the spectrum of charmed meson states Heavy-light $Q\bar{q}$ systems and their mass spectra are important in this analysis because semileptonic B meson decays produce almost exclusively charmed mesons ($Q = c$) of which a significant amount are higher excitations. Therefore, the focus is set on charmed mesons ($Q = c$) of which a summary of states produced in $B \to X_c\ell\nu$ decays is given in Tab. 1.2. In the case of $Q\bar{q}$ systems, the light "degrees of freedom" q decouple and the wave function factorises,

$$|Qq, s_h, s_l\rangle \approx |Q, s_h\rangle|q, s_l\rangle, \tag{1.25}$$

as depicted here for a $Q\bar{q}$ meson with heavy- and light-quark spins s_h and s_l, respectively. The factorisation of Eq. 1.25 allows one to classify charmed mesons spectroscopically; similar to an atom with a heavy nucleus and electrons. The light degrees of freedom have a spin s_l and angular momentum l which couple to their total angular momentum $j_l = s_l + l$ in the potential of the heavy quark. The total angular momentum of the charmed meson is then given by $J = s_h + j_l$.

The two lightest charmed states are the narrow 1S mesons D and D^*, with $l = 0$ and total spin 0 and 1, respectively. The next heavier mesons are four 1P states which split into a broad $j_l = s_l + l = 1/2$ and a narrow $j_l = 3/2$ doublet. The 1P states are usually referred to as D^{**} mesons. In addition, there are two radially excited 2S states which are light enough to be relevant in analyses of semileptonic B decays.

The factorisation and separation of scales is an important tool to describe the interactions inside hadrons, which is picked up again later in Sec. 1.2.1 and Chapter 4 when discussing decays of $Q\bar{q}$ systems.

	nl	$J^P_{j_l}$	Mass / MeV	Width / MeV
D	1S	$0^-_{1/2}$	1865.3 ± 0.1	negligible
D^*	1S	$1^-_{1/2}$	2008.5 ± 0.1	negligible
D_0^*	1P	$0^+_{1/2}$	2335 ± 15	249 ± 29
D_1'	1P	$1^+_{1/2}$	2427 ± 40	384 ± 120
D_1	1P	$1^+_{3/2}$	2422 ± 1	18 ± 2
D_2^*	1P	$2^+_{3/2}$	2463 ± 1	48 ± 1
D_{2S}	2S	$0^-_{1/2}$	2564 ± 20	135 ± 17
D_{2S}^*	2S	$1^-_{1/2}$	2622 ± 12	104 ± 20

Table 1.2: Properties of the relevant charmed mesons in semileptonic B meson decays. The quoted values for the masses and widths are averages over the charged and neutral state [20].

The LLSW model for $B \to D^{**}\ell\nu$ About a quarter of semileptonic B decays produces higher excitations of the charmed meson, mainly the 1P D^{**} states. These $B \to D^{**}\ell\nu$ decays are of great importance to this analysis and therefore their parametrisation is discussed below. The semileptonic B decays into D^{**} have been investigated by Leibovich, Ligeti, Stewart and Wise, usually referred to as the *LLSW model* [34]. Their calculations and findings are briefly summarised in the this section. The authors studied the semileptonic B decays into the four 1P states up to $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_Q)$, with $Q = c, b$. The differential decay width is given as a function of the recoil parameter $w = v^\mu v'_\mu$ and the cosine of the angle θ between the charmed meson and the lepton in the rest-frame of the virtual W. The narrow $j_l = 3/2$ and broad $j_l = 1/2$ states lead to a different set of functions to parametrise the form factors.

The matrix elements are not explicitly stated here, but are similarly separated into vector and axial-vector currents as previously discussed. They depend on the vector and axial-vector form factors (denoted with subscript V and A, respectively) $f_i(w)$, $k_i(w)$, $g_\pm(w)$ and $g_i(w)$ for the D_1, D_2^*, D_0^* and D_1', respectively. With the free-quark decay rate $\Gamma_b = G_F^2|V_{cb}|^2 m_B^5/(192\pi^3)$ and $r = m_{D^{**}}/m_B$, the differential decay rate into the D_1, as an example, is:

$$\frac{1}{\Gamma_b}\frac{\mathrm{d}^2\Gamma(B \to D_1\ell\nu)}{\mathrm{d}w\mathrm{d}\cos\theta} = 3r^2\sqrt{w^2-1}\Big\{\sin^2\theta\left[(w-r)f_{V_1} + (w^2-1)(f_{V_3} + r f_{V_2})\right]^2 + (1-2rw+r^2)\left[(1+\cos^2\theta)[f_{V_2}^2 + (w^2-1)f_A^2] - 4\cos\theta\sqrt{w^2-1}f_{V_1}f_A\right]\Big\}, \quad (1.26)$$

where the w dependence of the form factors is omitted for readability. The expressions for the other decay rates can be found in Appendix A.1. The expansion of the D_1 form factors $f_i(w)$ is given exemplarily below (the other form factors can also be found in Appendix A.1; Eqs. A.4 – A.6). The mass of the hadron can be expressed through the "binding energy" $\bar{\Lambda}$ as

$$m_H = m_Q + \bar{\Lambda} + \mathcal{O}\left(\frac{1}{m_Q}\right). \quad (1.27)$$

In the following, $\bar{\Lambda}$, $\bar{\Lambda}'$ and $\bar{\Lambda}^*$ are the binding energies of the (D, D^*), (D_1, D_2^*) and (D_0^*, D_1')

doublets, respectively. Their approximate values, together with those of the other parameters, are collected in Tab. A.1. When expanding the form factors in powers of $1/m_Q$ ($Q = c, b$), additional and unknown Isgur-Wise functions τ_i and ζ_1 enter the equations of the $j_l = 3/2$ and $j_l = 1/2$ states, respectively. The expansion to first order $\epsilon_Q = 1/(2m_Q)$ is given by

$$\begin{aligned}
\sqrt{6} f_A =& -(w+1)\tau - \epsilon_b\{(w-1)[(\bar{\Lambda}' + \bar{\Lambda})\tau - (2w+1)\tau_1 - \tau_2] + (w_1)\eta_b\} \\
& - \epsilon_c[4(w\bar{\Lambda}' - \bar{\Lambda})\tau - 3(w-1)(\tau_1 - \tau_2) + (w+1)(\eta_{\mathrm{ke}} - 2\eta_1 - 2\eta_3)], \\
\sqrt{6} f_{V_1} =& (1-w^2)\tau - \epsilon_b(w^2-1)[(\bar{\Lambda}' + \bar{\Lambda})\tau - (2w+1)\tau_1 - \tau_2 + \eta_b] \\
& - \epsilon_c[4(w+1)(w\bar{\Lambda}' - \bar{\Lambda})\tau - (w^2-1)(3\tau_1 - 3\tau_2 - \eta_{\mathrm{ke}} + 2\eta_1 + 3\eta_3)], \\
\sqrt{6} f_{V_2} =& -3\tau - 3\epsilon_b[(\bar{\Lambda}' + \bar{\Lambda})\tau - (2w+1)\tau_1 - \tau_2 + \eta_b \\
& - \epsilon_c[(4w-1)\tau_1 + 5\tau_2 + 3\eta_{\mathrm{ke}} + 10\eta_1 + 4(w-1)\eta_2 - 5\eta_3], \\
\sqrt{6} f_{V_3} =& (w-2)\tau + \epsilon_b\{(2+w)[(\bar{\Lambda}' + \bar{\Lambda})\tau - (2w+1)\tau_1 - \tau_2] - (2-w)\eta_b\} \\
& + \eta_c[4(w\bar{\Lambda}' - \bar{\Lambda})\tau + (2+w)\tau_1 + (2+3w)\tau_2 \\
& + (w-2)\eta_{\mathrm{ke}} - 2(6+w)\eta_1 - 4(w-1)\eta_2 - (3w-2)\eta_3].
\end{aligned} \tag{1.28}$$

The unknown functions η_{ke} and η_i ($i = 1, 2, 3$) parametrise the kinetic and chromomagnetic energy of the charm quark. Their counterparts for the bottom quark are summarised into $\eta_b = \eta_{\mathrm{ke}}^{(b)} + 6\eta_1^{(b)} - 2(w-1)\eta_2^{(b)} + \eta_3^{(b)}$. For calculating the decay rates explicitly, a linear approximation of the unknown Isgur-Wise function $\tau(w) = \tau(1)(1 + \tau'(w-1))$, with $\tau' = \mathrm{d}\tau/\mathrm{d}w\,|_{w=1}$ is chosen. LLSW also derive the leading α_s QCD-corrections of the form factors which are considered but not repeated here. For a spin-orbit independent potential of a non-relativistic quark-model, the leading order Isgur-Wise functions τ and ζ of the $\frac{3}{2}$ and $\frac{1}{2}$ doublet, respectively, are related by

$$\zeta(w) = \frac{w+1}{\sqrt{3}}\tau(w). \tag{1.29}$$

In this analysis, $B \to D^{**}\ell\nu$ decays are described by considering two of the approximations by LLSW to describe $\tau_{1,2}$ and ζ_1. Approximation "B1" assumes $\tau_1 = \tau_2 = \zeta_1 = 0$ and "B2" $\tau_1 = \bar{\Lambda}\tau$, $\tau_2 = -\bar{\Lambda}'\tau$ and $\zeta_1 = \bar{\Lambda}\zeta$. The approximations coincide for $m_Q \to \infty$ and their difference is expected to give an estimate of the unknown $\Lambda_{\mathrm{QCD}}/m_Q$ corrections.

1.2.3 Inclusive semileptonic B decays

The description of inclusive semileptonic B decays in the heavy-quark scheme is discussed briefly in this section. Representations of the inclusive triple differential decay rates have been obtained and the form factors are expanded for the b quark [35] and B meson decays [36]. The hadronic form factors of the latter, W_i, are expressed in terms of the lepton energy E_ℓ, mass m_ℓ and four-momentum transfer $q^\mu = p_\ell^\mu + p_\nu^\mu$ as dimensionless parameters

$$y = \frac{2E_\ell}{m_b}, \qquad \rho_\ell = \frac{m_\ell^2}{m_b^2}, \qquad \hat{q}^\mu = \frac{q^\mu}{m_b}. \tag{1.30}$$

The differential decay rate reads [36]

$$\frac{1}{\Gamma_b}\frac{\mathrm{d}\Gamma^3(B\to X_c\ell\nu)}{\mathrm{d}\hat{q}^2\mathrm{d}y\mathrm{d}v_\mu\hat{q}^\mu} = 24\left(2(\hat{q}^2-\rho_\ell)W_1 + \left[y(2v_\mu\hat{q}^\mu - y) - \hat{q}^2 + \rho_\ell\right]W_2 \right.$$
$$\left. +2\left[\hat{q}^2(y - v_\mu\hat{q}^\mu) - \rho_\ell v_\mu\hat{q}^\mu\right]W_3 + \rho_\ell(\hat{q}^2-\rho_\ell)W_4 + 2\rho_\ell(2v_\mu\hat{q}^\mu - y)W_5\right), \quad (1.31)$$

with the B meson four-velocity v^μ. For decays into the light leptons e and μ, the terms proportional to ρ_ℓ are negligible and thus the decays are insensitive to W_4 and W_5.

The total decay width predicts the inclusive branching fraction ratio [36]

$$R(X_c) = \frac{\mathcal{B}(B\to X_c\tau\nu)}{\mathcal{B}(B\to X_c\ell\nu)} = 0.223 \pm 0.005 \quad (1.32)$$

which will be used as the SM reference in this analysis.

1.3 Experimental status on tauonic B decays

Semileptonic B meson decays involving a τ lepton are of particular interest to test the predictions of the SM. The heavy lepton leads, on the one hand, to non-negligible terms in the SM calculations of decay rates, where usually terms of $\mathcal{O}\left(m_\ell^2/m_b^2\right)$ are neglected for $\ell = e, \mu$. On the other hand, these decays are of interest for the searches for heavy, new (e. g. Higgs-like) mediators whose coupling is proportional to the lepton mass.

In the following sections, the measurements of the inclusive decay $b\to X\tau\nu$ at LEP and the exclusive decays $B\to D^{(*)}\tau\nu$ at the B-factories are summarised.

1.3.1 Analyses of inclusive $b\to X\tau\nu$ decays

Analyses of inclusive $b\to X\tau\nu$ decays have been conducted at LEP by the ALEPH [37], DELPHI [38], L3 [39, 40] and OPAL [41] collaborations at the Z^0 resonance.

The results of the LEP experiments are summarised in Fig. 1.4 and compared to the SM expectation. Their measured inclusive branching fraction matches well the SM prediction. Notably, the sum of exclusive branching fractions $\mathcal{B}(B\to D\tau\nu) + \mathcal{B}(B\to D^*\tau\nu)$ measured at the B-factories (see next section) already exceeds the inclusive LEP measurements, however, not yet significantly as the uncertainties are still large.

At LEP, the analysis strategy was different from those at the B-factories and is therefore briefly outlined below. In e^+e^- collisions, b-hadrons are produced in $e^+e^-\to Z^0\to b\bar{b}$ decays which hadronise in a two jet event shape. From the distribution of the Z^0 daughter particles, the so-called *thrust*

$$T = \max\frac{\sum_i|\vec{n}\cdot\vec{p}_i|}{\sum_i|\vec{p}_i|} \quad (1.33)$$

is obtained which maximises the projections of the final state momenta $\vec{p}_i$ on the thrust axis $\vec{n}$. The event splits into two hemispheres along a plane perpendicular to the thrust axis and intersecting the collision point as shown in Fig. 1.5. By this, the daughters of the two b quarks are separated. To avoid lost tracks in the "blind" detector region along the beam-line, the thrust axis is required to be nearly perpendicular to the beam. One of the two hemispheres is investigated for a b-hadron signature by multivariate tools to ensure a $Z^0\to b\bar{b}$ decay. The other hemisphere is analysed with respect to the signal decay. Only the hadronic decay modes $\tau\to h\nu$ are considered and charged leptons are vetoed to reduce contamination from the $b\to X\ell\nu$ background. The

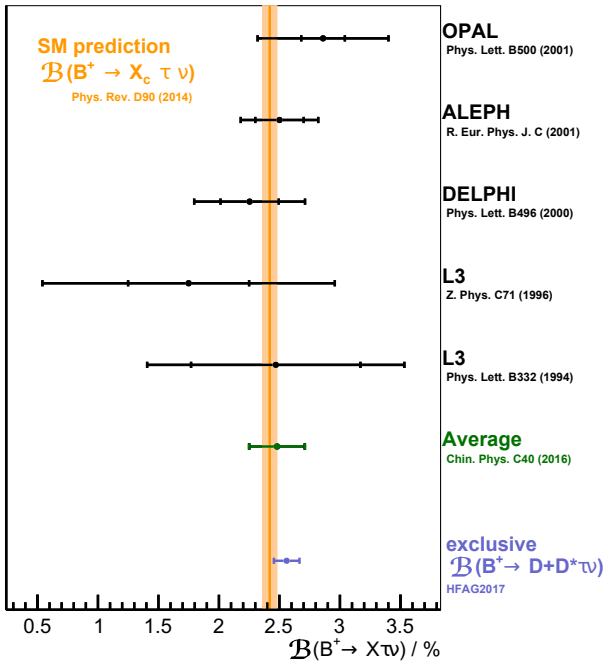


Figure 1.4: Summary of inclusive measurements of $B \to X\tau\nu$. Shown are the individual LEP measurements (black data points), the world averaged value (green) and the SM prediction (orange band). The inclusive measurements are life-time corrected to represent the decay of a charged B meson. To compare the inclusive values to the sum of exclusive $B \to D + D^*\tau\nu$ branching fractions, the latter are shown in blue, assuming the $B \to D^{(*)}\ell\nu$ branching fractions in Tab. 3.2. Without considering higher excitations of the charmed mesons, the sum of exclusive branching fractions already exceeds the inclusive branching fraction.

signal is then extracted in the missing energy distribution $E_{\text{miss}} = E_{\text{beam}} - E_{\text{visible}}$ (cf. Fig. 1.5) which is calculated from the beam energy and the measured energy in the signal hemisphere. The $b \to X\tau(\to h\nu)\nu$ signal is expected to have large missing energy because of two undetectable neutrinos in the final state.

The main systematic uncertainties are caused by the modelling of E_{miss} and are estimated from leptonic and hadronic control samples. Furthermore, the simulation of the b quark hadronisation and its parameters are important uncertainties.

In contrast to the exclusive B-factory measurements of $B \to D^{(*)}\tau\nu$ decays, which will be presented in the next section, the LEP measurements rely on a single observable, E_{miss}. The correct modelling of the visible energy is thus very important and is challenging, especially because of the numerous hadronic backgrounds which are enhanced by the lepton veto. A similar measurement in a completely different experimental environment as that of a B-factory is thus an important cross check.

1.3.2 Analyses of exclusive $B \to D^{(*)}\ell\nu$ decays

With their large data sets, the B-factory experiments *BABAR*, Belle and LHCb performed exclusive measurements of $B \to D^{(*)}\tau\nu$ decays [13–17]. The large $B\overline{B}$ samples are produced in $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ decays at Belle and *BABAR* and allow for the precise but inefficient hadronic B tagging (cf. Sec. 3.4) which gives comprehensive information on the decaying B mesons. To reduce systematic effects, the signal modes are normalised to the light-lepton modes

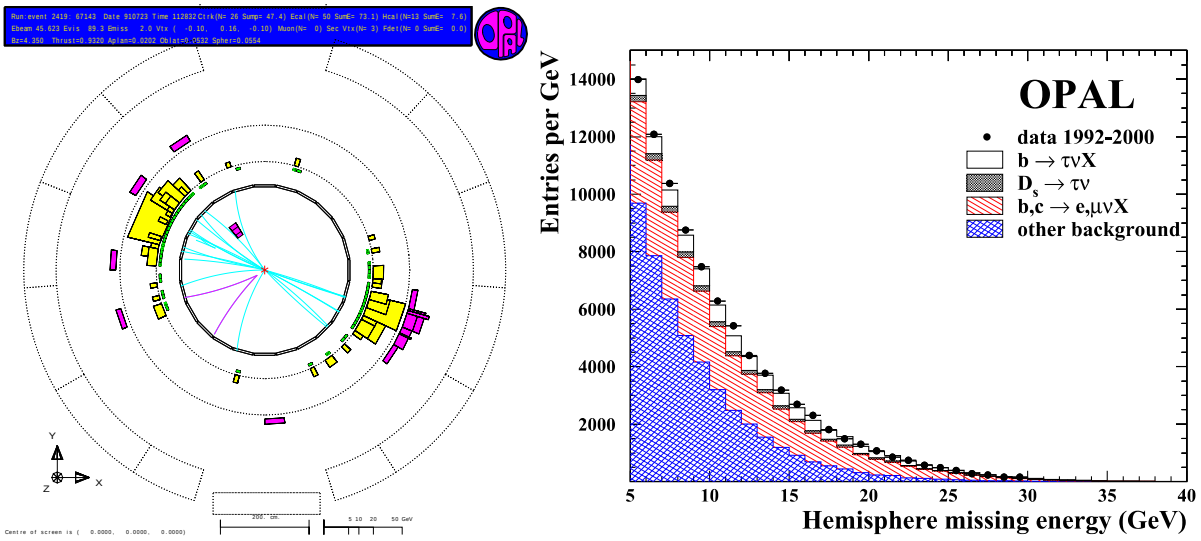


Figure 1.5: Measurement of the $b \to X\tau\nu$ branching fraction at the OPAL experiment. *Left:* Event display of a $Z \to b\bar{b}$ candidate. The view is a projection onto a plane perpendicular to the beam and the rings represent different detector components. The coloured blocks stand for energy depositions from particles in the detector and lines show reconstructed particle tracks. Clearly visible is the division of the event into two hemispheres [42]. *Right:* The distribution of missing energy in the signal hemisphere for the $b \to X\tau\nu$ signal extraction [41].

and the ratio of branching fractions

$$R(D^{(*)}) = \frac{\mathcal{B}\left(B \to D^{(*)}\tau\nu\right)}{\mathcal{B}\left(B \to D^{(*)}\ell\nu\right)} \tag{1.34}$$

is measured. *BABAR* [13] and Belle [14] reported significantly larger $B \to D^{(*)}\tau\nu$ branching fraction ratios than expected from the SM [43, 44]. The values are summarised in Tab. 1.3.

Experiment	$R(D)$	$R(D^*)$	
BABAR	0.44 ± 0.07	0.33 ± 0.3	had. tag, $\tau \to e, \mu$ [13]
Belle	0.38 ± 0.07	0.29 ± 0.04	had. tag, $\tau \to e, \mu$ [14]
Belle		0.30 ± 0.03	sl. tag, $\tau \to e, \mu$ [15]
Belle		0.27 ± 0.04	had. tag, $\tau \to had.$ [16]
LHCb		0.34 ± 0.04	pp collisions, $\tau \to \mu$ [17]
Average	0.40 ± 0.05	0.31 ± 0.02	[45]
SM	0.299 ± 0.003	0.252 ± 0.003	[43, 44]

Table 1.3: Summary of exclusive $R(D^{(*)})$ measurements at the B-factories as well as the SM predictions. The numbers are visualised in Fig. 1.6.

The analysis performed by *BABAR* extracts the signal yield in a 2D fit to the momentum of the τ daughter lepton and the sqared missing mass. The latter corresponds to the invariant mass of undetected particles in an event which can be estimated from the well known collision conditions in e^+e^- collisions and the detected particles.

The Belle measurement of $R(D^{(*)})$ relies on the squared missing mass and the event classification by an artificial neural network. The neural network is trained with Monte Carlo samples

of the signal decay and backgrounds.

An overview of current $R(D^{(*)})$ measurements is given in Fig. 1.6 and illustrates the tension between the measured and the expected $R(D^{(*)})$. Further analyses of $R(D^*)$ have been performed by Belle and LHCb. One of the Belle analyses used a semileptonic B tagging [15], whereas the other one used a hadronic tag and measured the $\pi\nu$ and $\rho\nu$ final states of the signal τ lepton [16]. These results are also depicted in Fig. 1.6. LHCb measured $R(D^*)$ untagged in the $\tau \to \mu\nu\nu$ decay channel, however, in a very different experimental environment of proton-proton collisions [17].

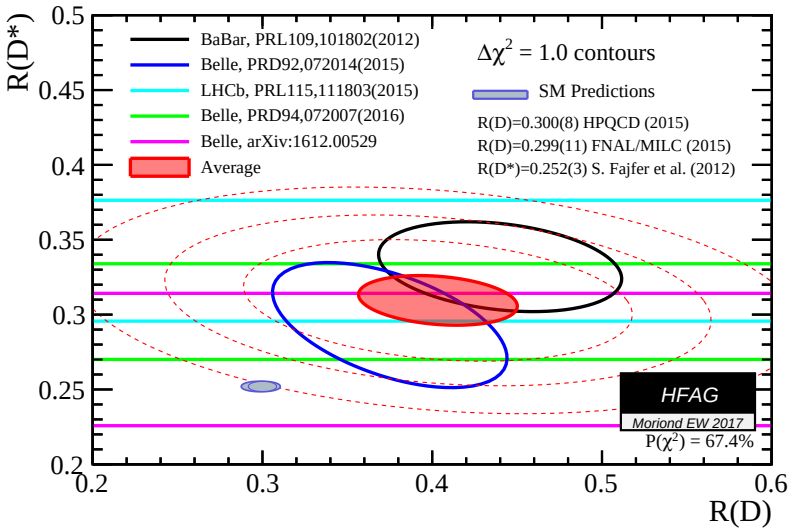


Figure 1.6: Measurements of $R(D^{(*)})$ at the B factories. The deviation to the SM prediction is at 4.0σ level when taking the experimental correlation between $R(D)$ and $R(D^*)$ into account [45].

Taking the average of all those measurements (Tab. 1.3), $R(D^{(*)})$ exceeds the SM prediction by 4.0σ when taking the experimental correlations into account. This is one of the major discrepancies in the B sector and is subject to extensive theoretical research [46–49] regarding New Physics.

1.4 Physics beyond the Standard Model in semileptonic B decays

The semileptonic B decays into light leptons are measured precisely and lie within the SM expectation. However, semileptonic decays involving a heavy τ show tensions with the SM which hints to possible NP contributions.

In general, in a model independent view, NP appear as scalar, vector or tensor-currents in particle interactions. They can have impact on observables, e. g. the total decay rate or angular distributions.

Two-Higgs-doublet models Tensions with the SM are measured in exclusive semileptonic B decays involving a heavy τ. Hence, a new mediator whose couplings are related to the masses, i. e. a Higgs-like interaction, is of particular interest. A common class of models are the two-Higgs-doublet models (2HDM) with an additional Higgs-doublet to that of the SM. In 2HDMs, the additional doublet results in three neutral Higgs-bosons (two scalars h, H and a

pseudo-scalar A) and two charged Higgs-bosons $H^{\pm}$. The latter can interfere with the charged weak SM current at tree-level in $B \to \tau$ transitions.

An important parameter of 2HDMs is the ratio of the vacuum expectation values v_1 and v_2 of the two doublets

$$\tan\beta = \frac{v_1}{v_2} \tag{1.35}$$

which can, in general, have an imaginary, CP violating part. The choice of the fermion couplings to the Higgs-doublets Φ_1 and Φ_2 results in different models [50]. A model which is often considered is the type-II 2HDM (2HDM-II) in which the down-type quarks couple to a different Higgs-doublet than charged leptons and up-type quarks. This leads to a conservation of flavour at tree-level.

In a 2HDM-II the free parameters reduce to the ratio

$$r = \frac{\tan\beta}{m_{H^+}}, \tag{1.36}$$

where m_{H^+} is the mass of the charged Higgs. The impact of 2HDM-II has been studied for exclusive and inclusive $B \to \tau$ decays.

The 2HDM-II inclusive $B \to X\ell\nu$ decay rate generally takes the form [51]

$$\frac{\Gamma(B \to X\ell\nu)}{\Gamma_b} = \Gamma_W - 2m_l^3 m_b \mathfrak{Re}(r^2)\Gamma_I + \frac{m_l^2 m_b^2}{4}|r|^4\Gamma_H, \tag{1.37}$$

whereas Γ_b is the previously defined free quark decay rate. Beside the SM decay rate Γ_W introduced in Sec. 1.2.3, the Higgs decay rate Γ_H and a (destructive) interference term Γ_I appear, which have been calculated up to $\mathcal{O}\left(m_b^{-2}\right)$ [51]. The Higgs contribution to the total decay rate scales with the lepton mass and thus its influence should be larger for the τ compared to the light leptons. The typical dependence of the inclusive decay rate on r is illustrated in Fig. 1.7 and holds for the exclusive rates [48], too. As the exclusive semileptonic B decays scale differently with r, the composition of $R(X_c)$ is also sensitive to the NP.

Experimental constraints on New Physics *BABAR* tested the 2HDM-II with their $R(D^{(*)})$ results and found them to be incompatible. The obtained branching fraction ratios $R(D)$ and $R(D^*)$ cannot be described by the same value of r as shown in Fig. 1.8. Further limits on NP contributions have been obtained by *BABAR* for the 2HDM-III which is a generalisation of the type-II model (cf. Ref. [52]) and is compatible with their result. In general, an additional vector current is preferred by their findings.

Belle tested its result on $R(D^{(*)})$ (using hadronic B tagging) for a SM and a 2HDM-II hypothesis and found both to be compatible (cf. Fig. 1.8). Belle's semileptonic tag analysis investigated generic NP contributions from new scalar, vector or tensor-currents. Only the scalar and vector currents are compatible with their measurement which can be interpreted as the 2HDM-II and SM currents, respectively. Additional tensor-currents are ruled out by their measurements.

The large deviation of the measured exclusive $R(D^{(*)})$ from the SM prediction is contrary to the compatibility of the inclusive LEP measurements with the SM. This analysis is the first $R(X)$ measurement at a B-factory and is the first analysis that directly measures the inclusive $B \to X\tau\nu$ decay rather than a b-hadron admixture and features the same experimental environment as the exclusive $R(D^{(*)})$ analyses.

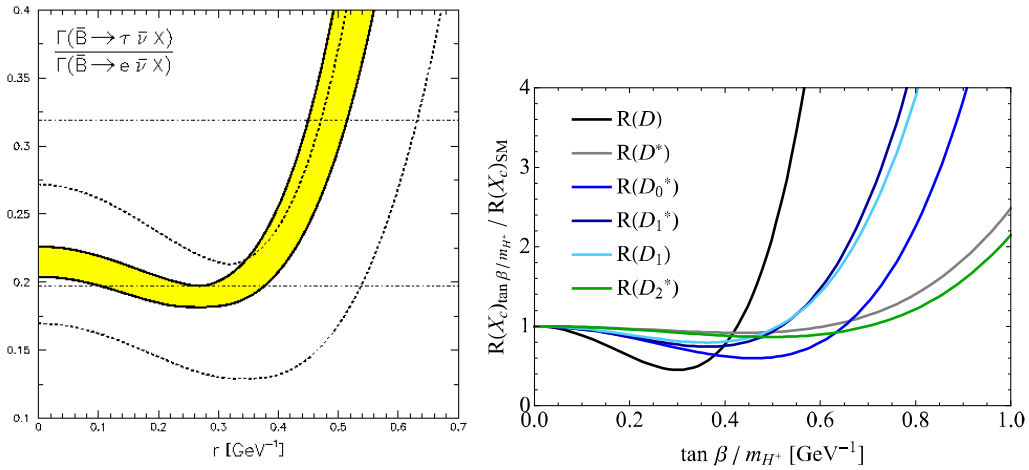


Figure 1.7: Predictions of the inclusive $R(X)$ (left) [51] and the exclusive branching fraction ratios (right) [48] for the 2HDM-II model. The inclusive prediction of $B \to X\tau\nu$ at $\mathcal{O}\left(m_b^{-2}\right)$ is given as the yellow band in the left plot and, for comparison, the 68% confidence interval of the predicted free-quark decay $b \to X\tau\nu$, is shown as dotted line. For certain values of r, destructive interference between the Higgs- and SM-currents appears, which is also visible in the right plot for the ratios of exclusive branching fractions.

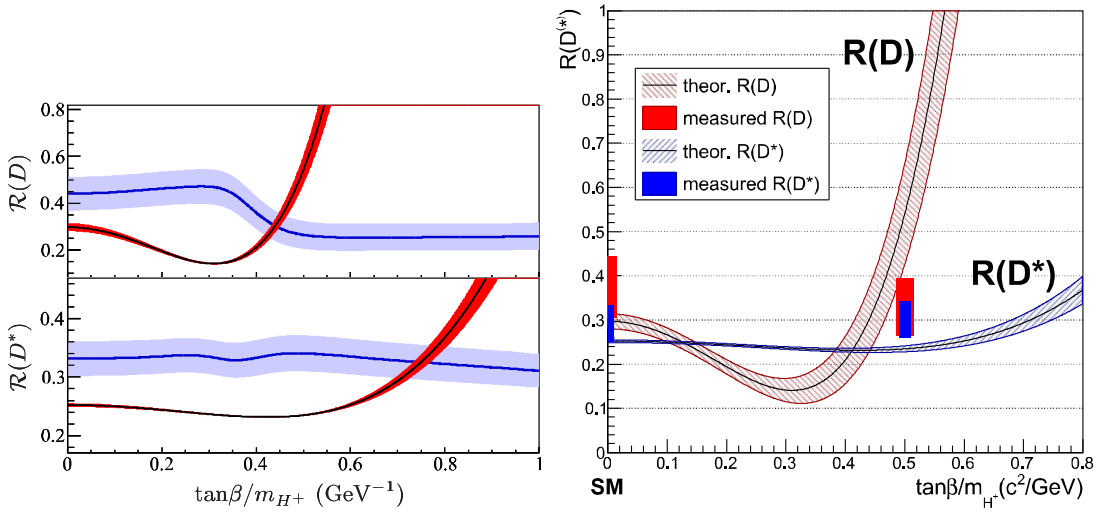


Figure 1.8: Comparison of 2HDM-II predictions with measurements of $R(D)$ and $R(D^*)$ by *BABAR* (left) [13] and Belle (right) [14]. *BABAR* estimated their branching fraction ratios over a large range of r (blue bands) which intersect with the 2HDM-II prediction (red band) at incompatible $R(D)$ and $R(D^*)$. Belle compares their results with the SM and a 2HDM-II with $r = 0.5$ which are both found to be compatible.

CHAPTER 2

The Belle experiment

The Belle experiment, conducted by an international collaboration of several hundred scientists from currently 20 countries, was one of the major B physics experiments. It was located at the KEK (short for Japanese "High Energy Accelerator Research Organisation") in Tsukuba, Japan and ran from 1999 till 2010. A competing experiment, the *BABAR* experiment, was conducted at SLAC in Stanford, USA, which ran at the same time. Nowadays, the LHCb experiment at CERN is the only running B physics experiment, but soon the Belle II experiment will start as the successor of Belle.

The main goal of the Belle and *BABAR* experiments was the observation and study of CP violation in the B sector, as predicted by the CKM mechanism. Thus, a large sample of B mesons was needed and decays of the $\Upsilon(4S)$ resonance in e^+e^- collisions are perfectly suited for this. The $\Upsilon(4S)$ is a $b\bar{b}$ resonance which decays with nearly 100% into a $B\overline{B}$ pair. The process of $B\overline{B}$ production via the $\Upsilon(4S)$ resonance is depicted in Fig. 2.1. As the 19 MeV mass difference between the $\Upsilon(4S)$ and a $B\overline{B}$ pair is very small, the B mesons are produced nearly at rest and there is no contamination from the heavier excited B^* or B_s mesons. Thus, a clean sample of $B\overline{B}$ pairs can be produced in e^+e^- collisions at a centre-of-mass energy of $m(\Upsilon(4S)) = 10.58\,\text{GeV}$. Using collisions of point-like elementary particles like e^+e^- has the advantage of having a well defined energy and momentum in the collision. In contrast, the proton-proton collisions at LHC which LHCb analyses, have no definite collision momentum because the colliding proton constituents have an intrinsic momentum distribution inside the proton.

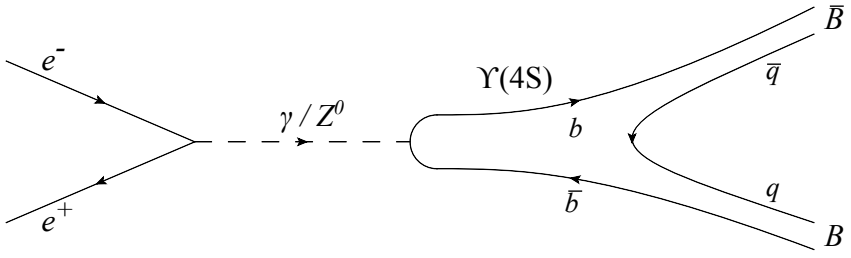


Figure 2.1: Faynman diagram of $B\overline{B}$ production in e^+e^- collisions. The electron and positron annihilate into a virtual photon or Z^0 which then produce a $b\bar{b}$ pair that forms the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ decays strongly into a $B\overline{B}$ pair. The strong interactions via gluons are not shown for simplicity.

The electrons and positrons for the Belle experiment are delivered by the KEKB accelerator, which is thus referred to as a B factory. The KEKB accelerator is shown in Fig. 2.2. It

consists of two storage rings, one for the electron and one for the positron beam, and a linear accelerator. The storage rings were rebuilt from the preceding TRISTAN accelerator and each has a circumference of 3 km. A pre-accelerator collides electrons with a Tantalum target to produce the positrons. The linear collider accelerates electrons and protons to their design energies of 8 GeV and 3.5 GeV, respectively, before they are injected into the storage rings.

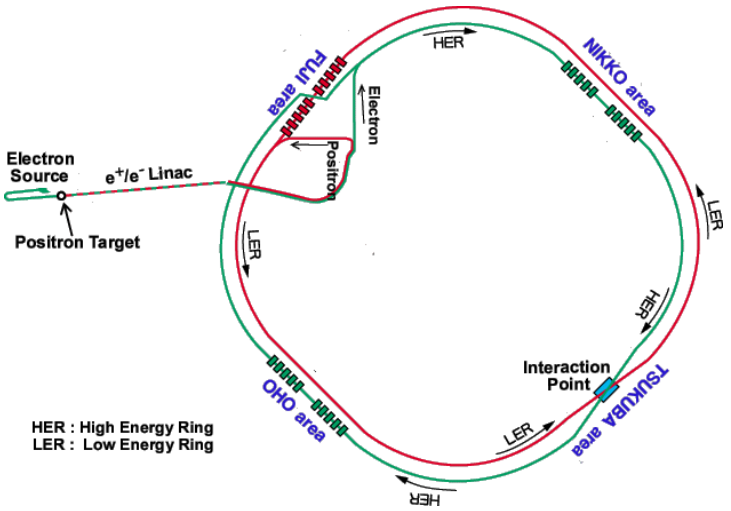


Figure 2.2: The KEKB accelerator at the KEK facility. Electrons are pre-accelerated and a part of them is brought to collision with a target to produce positrons. Both, e^+ and e^- are brought to their design energies by the linear accelerator (Linac) before being injected into the high (e^-) and low (e^+) energy rings. The collisions of the e^+ and e^- beams happen at the interaction point inside the Belle detector. Adapted from Belle's internal presentation supplements.

The asymmetric beam energies reflect the purpose of the Belle experiment to study time-dependent effects, in particular time-dependent CP violation. Such measurements benefit from the time-dilatation due to the boosted system and long flight distances between the production and the decay of the B mesons which allows to infer their life-time. The center-of-mass frame is boosted with respect to the lab or detector frame which is also reflected in the asymmetric shape of the detector.

The KEKB accelerator holds the world record of the highest luminosity of $2.11 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The high luminosity results in a large data set of $\sim$1000 fb^{-1}, of which 710 fb^{-1} were recorded at the $\Upsilon(4S)$ resonance, corresponding to roughly 770×10^6 $B\overline{B}$ pairs. Further data sets were taken right below the $\Upsilon(4S)$ to study light-quark continuum backgrounds and at the $\Upsilon(5S)$ for a sample of B^* and B_s mesons. The high luminosity comes with disadvantages like increased beam backgrounds from scattered electrons or positrons or synchrotron radiation which have to be dealt with in the analysis.

2.1 The Belle detector

The Belle detector is built up of several layers as shown in Fig. 2.3. Each layer was designed to provide complementary information of the particles produced in the e^+e^- collisions that allow identify them and to measure their energies and momenta. Below, a short introduction of the various detector components is given. A much more detailed description of the detector can be found in [53]. The detector is built asymmetrically to satisfy the requirements of asymmetric beam energies. Due to the barrel-like shape around the beam-pipe, polar coordinates (r, θ, ϕ) (or

(r, z, ϕ) with $z = r\cos\theta$) are used. The z-axis is placed parallel to the direction of the electron beam. The radius r is measured as the distance to the interaction point (IP) and the polar angle θ is the projection of the radial vector onto the z-axis. As the the detector is symmetric around the beam-line, particles are uniformly distributed along the the azimuthal angle ϕ. The acceptance region of the detector is $17° < \theta < 150°$ and thus covers 91% of the solid angle in the detector's rest frame.

In the following paragraphs, the detector components are described. The goal is to measure precisely the vertices, momenta and energies and to perform particle identification which is only possible by the interplay of several specialised devices.

The tracking and vertexing is done by the silicon vertex detector and the central drift chamber. Both are designed such that scattering in the detector material and energy loss are small and thus the flight path is not much disturbed. Charged particle trajectories are bent by a strong magnetic field which enables the determination of the charge and momentum. The separation of pions and kaons is done by the time-of-flight and Aerogel Cherenkov counters for low and high momenta, respectively, which are located further away from the IP. Photons and electrons cause electromagnetic showers in the electromagnetic calorimeter and are supposed to loose their energy there. The outer detector is the K_L muon detector which is very dense and thus only traversed by muons. All other detectable particle that passed through the inner layers are stopped and create showers.

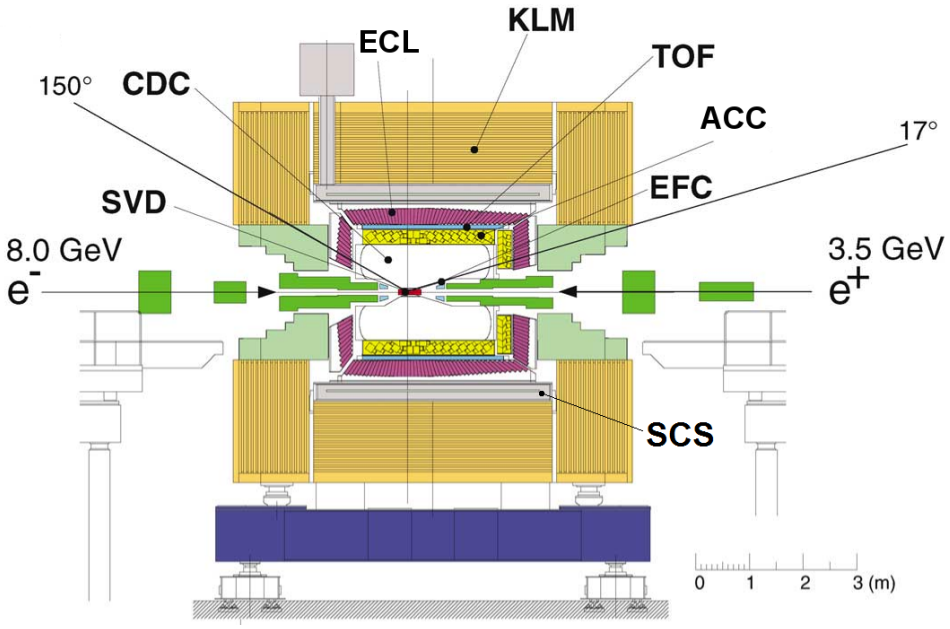


Figure 2.3: The Belle detector and its components. The asymmetric construction is due to the asymmetric beam energies. The components are from inside out: the silicon vertex detector (SVD), extreme forward calorimeter (EFC), central drift chamber (CDC), Aerogel Cherenkov counter (ACC), time-of-flight counter (TOF), electromagnetic calorimeter (ECL), superconducting solenoid (SCS) and the K_L muon detector (KLM). Adapted from [53].

2.1.1 Beam-pipe and silicon vertex detector

The first material particles have to traverse is that of the beam-pipe which separates the vacuum, in which the beams are circulating, from the detector. It is built as thin as possible with a

low-Z material to avoid scattering of the particles with the material as this would degrade the reconstruction of the true vertex and flight direction. Two beryllium pipes of 0.5 mm thickness each were used for this reason.

The first layer of the Belle detector is the silicon vertex detector (SVD) which allows for the precise vertexing of charged particles as they pass through. From the energy depositions in the layers, the particle's trajectory can be retraced and the vertex inferred. The Belle detector had a so-called SVD1 and a SVD2 phase, as the SVD1 was upgraded due to radiation damage after 140 fb^{-1} of recorded integrated luminosity. The SVD2 (SVD1) consists of four (three) layers of double-sided strip detectors (DSSD), placed in a radial range of $20 < r < 88$ mm ($30 < r < 60.5$ mm) and covering an polar range of $17° < \theta < 150°$ ($23° < \theta < 139°$). A sketch of the SVD2 layers and a DSSD module is given in Fig. 2.4. A DSSD is basically a pn-junction, a p-implant in an n-bulk, with strip-shaped readout nodes. If a charged particle crosses the n-bulk, it creates electron-hole pairs that move to the corresponding readout nodes due to the applied bias-voltage of 75V. The p-readout strips are oriented in z-direction and the n-readout nodes perpendicular to it to measure ϕ and θ, respectively. The achieved vertex resolution is inversely proportional to the track momentum and, for instance, 356 μm and 42 μm for a 0.1 GeV and a 1 GeV track, respectively.

The SVD is well suited to trace and to constrain the vertex of charged particles but, due to only four layers, it alone is barely able to identify them. For the identification, a larger volume is needed to measure e. g. the specific energy loss $\mathrm{d}E/\mathrm{d}x$ and thus, the following sub-detectors are needed. However, there are always low momentum or deflected tracks that do not reach further detector components and can even curl back to the IP due to the magnetic field. This causes poorly identified and even duplicated tracks in the tracking algorithm which is important and has to be dealt with, especially in inclusive analyses (see Sec. 5.1.1).

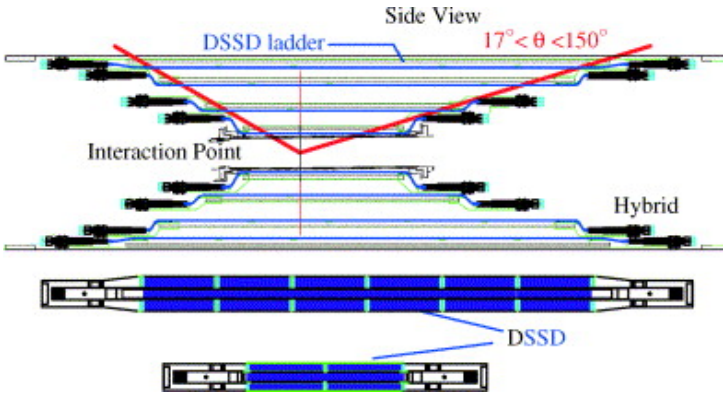


Figure 2.4: The Belle SVD2 (top) and the two DSSD ladders (bottom). The DSSD ladders are aligned in four radial layers with a small overlap to avoid lost tracks due to gaps between ladders. Each ladder consists of two to six DSSD modules, readout electronics and a mechanical support structure [54].

2.1.2 Central drift chamber

The next sub-detector is the central drift chamber (CDC), a gaseous ionisation detector, depicted in Fig. 2.5. It provides the main tracking and important particle identification information for charged particles. Charged particles ionise the gas when travelling through the chamber. Field wires inside the chamber, with a high voltage applied, separate the electrons and ionised atoms, which move to the next field or sensor wire, respectively, and cause a signal impulse. As most

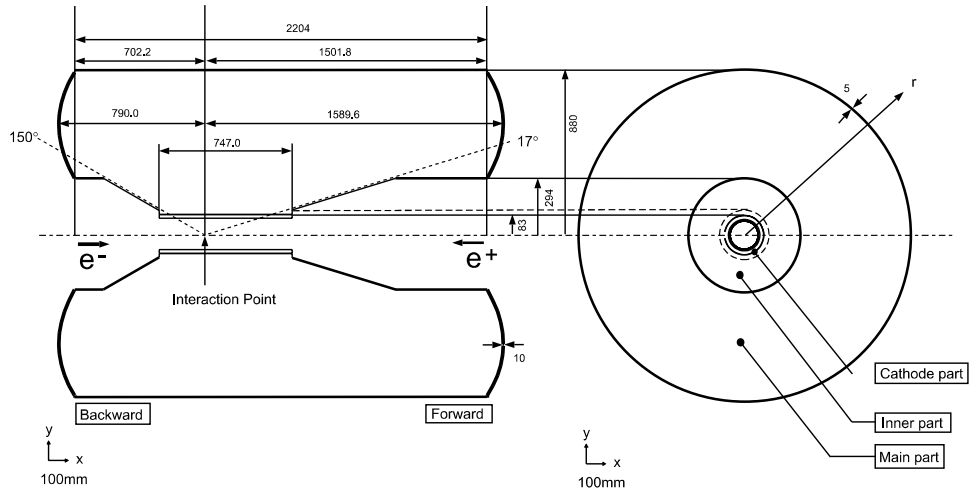


Figure 2.5: The CDC is the largest component of the inner detector. The innermost cathode part supports the track reconstruction in z-direction with strip detectors on the CDC wall and between the field wires. In the inner detector, the wires are placed with higher density. The lengths in the figure are given in mm [53].

of the tracks at Belle have momenta below 1 GeV and low momentum particles have a higher energy loss in matter and are more deflected, multiple scattering of the tracks with the gas atoms should be minimised. Thus, a low-Z, 50:50 gas-mixture of helium and ethane is used. The wires are aligned along the z-direction of the detector. Roughly half of the wires are slanted by 50 mrad to measure the polar angle, too. To support the track reconstruction in z-direction in the inner part of the detector, three cathode strip layers are placed within the innermost 10 mm of the CDC.

Furthermore, the CDC enables the determination of the energy loss $\frac{\mathrm{d}E}{\mathrm{d}x}$ which is specific for the different particle types and thus used for particle identification. The tracking information of the SVD and CDC together give a precise measurement of the trajectory and thus the charge and momentum of the particles. To improve the particle identification, especially towards the high track momenta, the following sub-detectors are needed.

2.1.3 Aerogel Cherenkov counter

The Aerogel Cherenkov counter (ACC) provides information for separating charged pions and kaons with momenta above 1.2 GeV. If charged particles travel trough a medium faster than the speed of light in this medium, Cherenkov-light is emitted. The necessary requirement on the diffraction index n of the medium and the velocity $\beta = v/c$ is

$$n \geq \frac{1}{\beta} = \frac{1}{\sqrt{1-(p/m)^2}} \tag{2.1}$$

and thus depends on the mass of the traversing particle. At Belle, a special silica aerogel with a small diffraction index is chosen. The diffraction index varies with the location of the ACC modules in the Belle detector (cf. Fig. 2.7). An ACC module (Fig. 2.6) consists of the aerogel medium and two readout photomultipliers and points toward the IP.

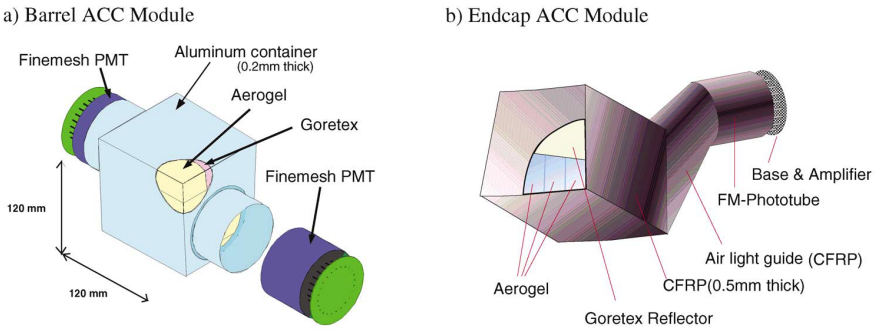


Figure 2.6: Schematic view of an ACC module for the barrel (left) and end-cap (right) region. The aluminium box contains the aerogel tiles to which the photomultiplier-tubes (FM-PMTs) are attached. The end-cap region module is built of a carbon-fibre reinforced plastic (CFRP) [53].

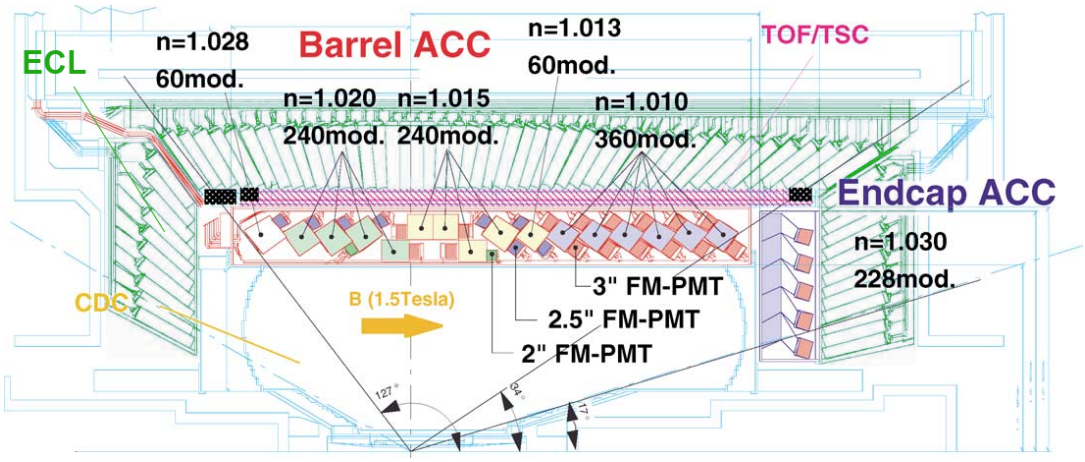


Figure 2.7: The central part of the Belle detector. The ACC Aerogel modules with attached photo-multipliers are shown with used diffraction index n and the number of modules of a certain kind. The time-of-tlight (TOF) and trigger scintilation (TSC) counters are depicted as a hatched purple bar. The electromagnetic calorimeter (ECL) crystals are shown in green and the CDC and the mechanical structure are shown in blue. Adapted from [53].

2.1.4 Time-of-flight counter

While the ACC is capable of separating high momentum pions and kaons, the time-of-flight counter (TOF) is well suited for low momentum tracks. The TOF is a fast scintillation detector which measures the time after the collision it is reached by charged particles. It is also depicted in Fig. 2.7. Thus, it is located further away from the IP than the ACC, at $r \approx 1.2\,\mathrm{m}$, because a longer travel distance of the particles improves the relative resolution. The TOF works well up to momenta of 1.2 GeV and is well suited for 90% of the particles produced in $\Upsilon(4S)$ decays. Particles have to have at least 0.3 GeV to reach the TOF. In addition, the TOF provides fast trigger signals for the hardware trigger (cf. Sec. 2.1.7) as well as the ACC and CDC. To avoid pile-up of trigger signals, the TOF is supported by thin trigger scintillation counters (TSC).

2.1.5 Electromagnetic calorimeter and extreme forward calorimeter

The electromagnetic calorimeter (ECL) is designed to measure the energy of photons and electrons and supports the electron identification. The ECL is located at a radius of 1.25 m (see Fig. 2.7) and particles need momenta above 0.3 GeV to reach it.

An array of 8763 CsI(Ti) crystals, pointing toward the IP, are chosen for this purpose. The fine segmentation gives a good spatial resolution that allows one to e. g. recover the final state radiation photon of an electron close to the electron track. The crystals are slightly tilted to avoid the escape of photons that hit the gap between two crystals. At the back of each crystal, two photo-diodes read out the scintillation light of the electromagnetic showers in the crystals. Photons are converted into e^+e^- pairs which then cause the electromagnetic shower. The set-up is designed for a wide range of photon energies due to the boosted centre-of-mass frame, from low energy photons from π^0 decays of a few 10 MeV up to 4 GeV from, e. g., $B \to X_s\gamma$ decays. In this large energy range, the ECL achieves a relative energy resolution of 4% to 1.6% at 100 MeV and 8 GeV, respectively [27]. Electron/hadron separation is achieved by measuring their energy E and considering the momentum p, measured by the inner detectors. Both are combined to the the ratio E/p which is approximately 1 for the light electron.

The extreme forward calorimeter (EFC) is placed close to the IP and beam-line in order to increase the coverage of the calorimeter down to $\theta > 6.4°$ and $\theta < 171.2°$. Its position is also depicted in Fig. 2.3. Due to the high-radiation environment it is made of radiation-hard BGO-crystals. Initially meant to support the ECL, the EFC is in practice mainly used by the KEKB control to monitor the beam and for luminosity measurements.

2.1.6 K_L Muon detector and superconducting solenoid

Outside the ECL a superconducting solenoid is located. It provides a nearly homogeneous magnetic field of $B = 1.5\,\mathrm{T}$ in z-direction and encloses the complete inner detector. The magnetic field bends the trajectories of tracks with charge q and velocity $\vec{v}$ according to the Lorentzian force $\vec{F} = q\vec{v} \times \vec{B}$.

Outside the solenoid, the K_L Muon detector (KLM) is located. It is the last active component of the detector and, at the same time, serves as the flux return for the magnetic field. The KLM consists of alternating layers of 4.7 cm thick iron plates and glass electrode resistive plate counters (RPC). An RPC module consists of a gas-filled volume between two glass electrodes with a bias voltage of 8 kV. A charged particle passing trough the volume ionises the gas and discharges the electrodes. Particles need to have at least 0.6 GeV transversal momentum to reach the KLM detector.

The thick iron plates and the ECL together have 4.7 interaction lengths. Particles interact and cause showers of ionising particles either already in the ECL or when entering the KLM. Such a shower, that matches no track in the inner detector, gives the direction of a K_L, however, it is not sufficient to measure its energy.

For muon identification, the fact is used that muons and hadrons interact differently with the detector material. Muons, as minimal ionising particles, pass the KLM further and are less deflected than hadrons which cause showers. A good muon track candidate needs to have KLM muon-hits assigned by the tracking algorithm to distinguish it from a pion. Thus, studies with muons can be less efficient compared to electrons because the minimal momentum threshold is higher as the additional information from the KLM is needed. Especially samples of slow muons with $p_\mathrm{T} < 1\,\mathrm{GeV}$ suffer from hadrons identified as muons ("fake muon candidates") that have to be dealt with and are important to this analysis.

2.1.7 Trigger-system and data-acquisition

Most hits in the Belle detector are not due to $B\overline{B}$ events or even due to collisions. They stem from beam-background due to KEKB's high luminosity. Roughly half of the 200Hz event rate is related to beam-background events. Furthermore, the large number of di-photon and Bhabha events has to be reduced, but not completely as these events are interesting for calibration and luminosity measurements. To reduce the rate of unwanted events and thus the size of data samples to be stored, multiple levels of triggers are used.

The level 1 and 2 triggers are hardware based and the former decides whether an event is interesting already within 2.2 μs. The sub-detector triggers report to the level 1 global decision logic (GDL) which takes the final decision whether an event is to be recorded or not. Not all sub-detectors contribute to the trigger logic, as visible in the scheme shown in Fig. 2.8. Without going into the details, the detector components deliver information on what kind of event might have occurred. For example, the total energy in the ECL has to be within a certain range or the CDC hits should match patterns of track segments. After a positive GDL decision, the sub-detector responses are read out and the event rate is further reduces by the level 2 trigger. Afterwards, the signals are digitalized and analysed by the level 3 and 4 real-time software triggers. At level 3, a fast track reconstruction is performed to select hadronic events. Before events are converted into the off-line data format, the level 4 off-line trigger further reduces the event rate by performing a fast event reconstruction. The trigger efficiency for $B\overline{B}$ events is above 99% after level 4. Afterwards, the data is available to analyses and is described in the next chapter.

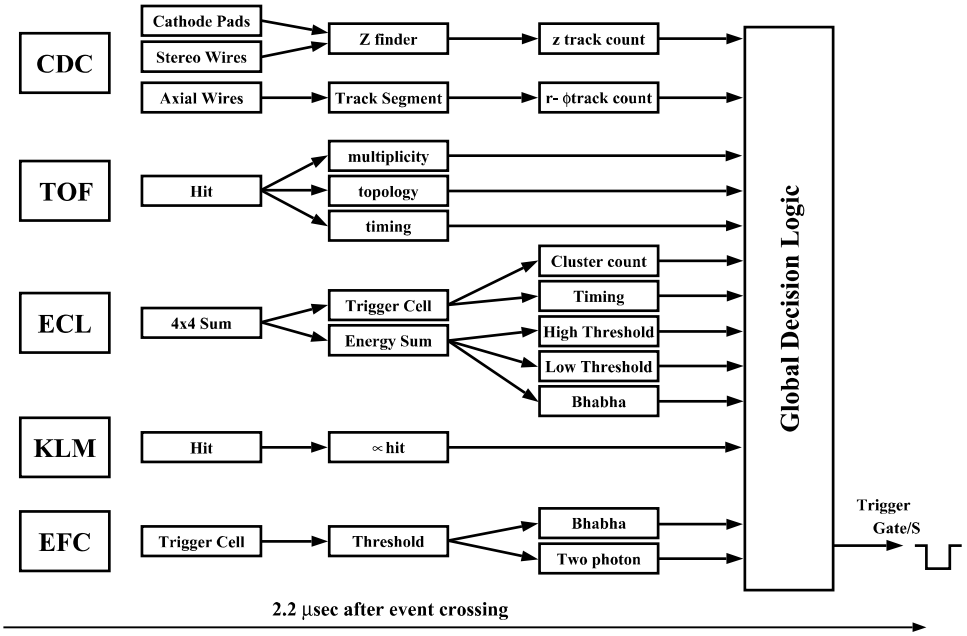


Figure 2.8: The level 1 hardware trigger and the global decision logic. The sub-detectors report various estimates of the events topology, visible energy and momentum. Examples are the track segments in the CDC or the total energy in the ECL. Based on these information, the global decision logic accepts or rejects an event within 2.2 μs [53].

CHAPTER 3

Data and Monte Carlo samples

Beside the data recorded by the Belle detector, analyses rely on Monte Carlo simulations (MC). The physics simulation of the signal and background processes as well as modelling the detector response is an important tool in high-energy physics. The MC simulation allows to study selection criteria and efficiencies of signal and background to find an optimal strategy for the signal extraction. In the signal extraction process itself, the MC simulations can be used as templates to the fitting procedure. As the MC samples the underlying probability distribution functions of observables, convolved with the detector response, the templates are directly comparable with the collision data. Furthermore, the MC allows to study properties the signal in detail, even in a "blinded" analysis. Thus it is crucial to have a comprehensive MC that models the data accurately.

At the beginning of this chapter, the Belle data set is introduced (Sec. 3.1). The following Sections 3.2 and 3.3 cover the MC and applied corrections, respectively.

3.1 The Belle data set

The Belle experiment recorded a huge data set of roughly $1\,\mathrm{ab}^{-1}$. The majority of it, a sample of $711\,\mathrm{fb}^{-1}$, was taken at the $\Upsilon(4S)$-resonance, dedicated to B meson studies. It corresponds to $(771.6 \pm 10.6) \times 10^6$ $B\overline{B}$ pairs produced in e^+e^- collisions. A smaller sample of $121\,\mathrm{fb}^{-1}$ was recorded at the $\Upsilon(5S)$-resonance to study decays of B^* and B_s mesons. In addition, a data sample to estimate light quark continuum production was recorded $\sim 60\,\mathrm{MeV}$ below the $\Upsilon(4S)$ resonance. The e^+e^- collisions do not only produce the $b\bar{b}$ part but also $q\overline{q}$ ($q = u, d, s, c$) "continuum" and $\ell^+\ell^-$ pairs. The cross sections of $q\overline{q}$ continuum and di-lepton production are together much higher than that of $b\bar{b}$. The trigger system of the Belle experiment already reduces the amount of unwanted events, however, continuum background keeps being relevant. Inclusive analyses as that of $B \to X\tau\nu$ typically suffer more from continuum background than exclusive ones. However, there are techniques as e. g. hadronic B tagging (cf. Sec. 3.4) to strongly suppress non-$b\bar{b}$ backgrounds. The data was recorded in units of so-called experiments which are data taking periods between detector and accelerator maintenance shut-downs [27].

The analysed $\Upsilon(4S)$ data set is preselected and referred to as the "HadronB-skim". The goal of the skim is to reduce the amount of data that needs to be processed by rejecting irrelevant events. The HadronB-skim selects hadronic events and suppresses (radiative) Bhabha, di-tau, di-photon and beam gas events. It is optimised to select $B\overline{B}$ events which thus have a high efficiency of $\epsilon = 0.991$. For continuum events, the skim is somewhat less efficient ($\epsilon = 0.795$) [55].

The further event reconstruction and selection is analysis specific and for this analysis described in Chapter 5.

3.2 The Belle Monte Carlo

As most analyses rely on simulated MC samples, the Belle Collaboration put big effort into the MC generation of common B decays. The MC simulation is split into dedicated samples, summarised in Tab. 3.1. The so-called "generic $B\overline{B}$" sample contains charmed semileptonic and most hadronic B decays and for many analyses it is signal and background sample at the same time due to the comprehensive B decay modelling. The B meson production and decay chains were mostly generated with the EvtGen package [56]. Inputs to EvtGen are two lists. One list describes all particles in terms of e. g. their mass, lifetime or spin and the second list describes their decays as a set of branching fractions, final-state particles and the theoretical models to be used for a certain decay. Thus, large parts of the B meson decays (and those of their daughter particles) are modelled as exclusive decays using measured or estimated branching fractions. To fill up the B decay width with hadronic decays, approximately a third of the B decay width is modelled at quark level without stating a specific final state using the Jetset/Pythia package [57].

The $q\overline{q}$ continuum MC sample models light quark production which is always present as a background. It is also generated with Jetset/Pythia. The modelling of quark fragmentation is very difficult and thus the modelling of the continuum background. Therefore, analyses with a high level of continuum background make use of the off-resonance data sample. However, in this analysis of $B \to X\tau\nu$ the continuum background is strongly suppressed and the continuum MC is used due to its six times larger sample size compared to the expectation in data.

Further dedicated MC samples are of interest depending on the type of analysis, such as samples dedicated to signal specific processes or large samples of certain background types.

After the MC generation, the detector geometry, material and readout properties are simulated using a modified version of GEANT3 [58] named *Gsim*. This step considers interactions with the sub-detectors as well as the material of the mechanical structure which causes scattering, showers or photon conversions. The simulated events are overlaid with real, randomly triggered background events after the detector simulation to improve the modelling of events recorded with the real detector. The MC production is in general experiment-number dependent, i. e. the slightly different detector and accelerator conditions in every experiment are taken into account. In addition, the generic $B\overline{B}$ MC models the tiny variations during an experiment (so-called "run-dependent" MC).

For this analysis of $B \to X\tau\nu$, the modelling of the inclusive signal and $B \to X_c\ell\nu$ background is crucial. The B meson decays $B \to D^{(*)}\ell\nu$ ($\ell = e, \mu, \tau$) are described sufficiently well by the generic $B\overline{B}$ MC (i. e. only efficiency corrections need to be applied, cf. Sec. 3.3) and are, therefore, taken from this sample. However, the decays of higher excitations of the charmed mesons (D^{**}), are revised in order to incorporate recent measurements of final states which are not modelled in the generic $B\overline{B}$ MC. The modelling of D^{**} mesons and its implications for $B \to D^{**}\ell\nu$ decays are discussed in Chapter 4. In addition, a dedicated sample of CKM-suppressed $B \to X_u\ell\nu$ decays is used in this analysis.

To overcome statistical limitations of the MC samples, multiples of the recorded data sample have been generated, referred to as "streams", i. e., one MC stream means that the MC sample corresponds to the size of the data sample recorded by Belle.

Name	#streams	rel. sample size
generic $B\overline{B}$	10	1
continuum	6	10^{-2}
$B \to X_u \ell\nu$	20	10^{-3}
$B \to D^{**}\ell\nu$	5	10^{-2}
$B \to D^{**}\tau\nu$	10	10^{-3}

Table 3.1: MC samples used in this analysis. The first three samples describe disjunct physical processes and add up to the expected data sample. The $B \to D^{**}\ell\nu$ and $B \to D^{**}\tau\nu$ samples are generated separately for this study (Chapter 4). Therefore, the corresponding decays have to be removed from the generic $B\overline{B}$ MC. The sample size is a rough estimate for this analysis of $B \to X_c\tau\nu$ relative to the size of the generic $B\overline{B}$ sample to illustrate the contribution to the total MC sample.

3.3 MC efficiency corrections

A large number of inputs enter the MC simulation, such as branching fractions, decay models and detector response models. Some of the physics inputs have become outdated since the MC was generated. In addition, the identification of a certain particle species' can be differently efficient in data and MC due to insufficient modelling of detector responses. Both effects are taken into account when performing analyses with MC samples. To correct for these effects, several studies have been performed by the Belle Collaboration which are, beside other corrections, discussed below.

Branching fractions Since the MC was generated, various decays have been measured with higher precision or even for the first time, leading to deviations from the assumed branching fractions in the MC. Numerous branching fractions are corrected as summarised in Tab. 3.2 and A.2 – A.5. The branching fractions given in the tables are selected by their importance to the analysis, based on their contribution after the event selection. However, often no measurement of the branching fraction is available, especially for hadronic B decays and charmed meson decays. In this case, it is not possible to correct the MC. An exception are the semileptonic D decays which are, due to measured final states and the known inclusive $D \to X\ell\nu$ branching fraction, possible to constrain.

The most important corrections regard the semileptonic B decays which are set to the current world average values [59]. When correcting these branching fractions, the current value of $\mathcal{B}(B \to X_c\ell\nu)$ from inclusive measurements is taken into account and the exclusive decays are scaled to match this constraint. More details on the semileptonic B meson composition are given in Chapter 4. Throughout this analysis, isospin symmetry is assumed a life-time ratio of $\tau_{B^+}/\tau_{B^0} = 1.076 \pm 0.004$ [20] is used to infer consistent branching fractions of B^+ and B^0.

Individual hadronic B decays are a minor contribution to the backgrounds in this analysis because of the variety of hadronic decays, each having a small branching fraction. However, the most common hadronic decays are corrected as well.

Decays, for which the branching fractions are not corrected as mentioned above, are scaled in a way that the overall B and D decay widths are preserved.

Form factors of semileptonic B decays In the generic Belle MC, semileptonic B decays are either generated with the `HQET2` (for $B \to D^{(*)}\ell\nu$, based on the CLN parametrisation in Sec. 1.2.1) or the `IGSW2` EvtGen module (for $B \to D^{**}\ell\nu$, based on the quark model of [60]). The latter is nowadays known to poorly model the data and superseded by the LLSW model

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$B^+ \to D\ell^+\nu$	2.31	2.30 ± 0.10
$B^+ \to D^*\ell^+\nu$	5.79	5.29 ± 0.12
$B^+ \to D_1\ell^+\nu$	0.81	0.77 ± 0.18
$B^+ \to D_2^*\ell^+\nu$	0.39	0.32 ± 0.10
$B^+ \to D_0^*\ell^+\nu$	0.91	0.58 ± 0.09
$B^+ \to D_1'\ell^+\nu$	0.81	0.32 ± 0.08
$B^+ \to D_{2S}\ell^+\nu$	0.02	0.10 ± 0.05
$B^+ \to D_{2S}^*\ell^+\nu$	0.05	0.20 ± 0.10
$B^+ \to D\pi\ell^+\nu$	0.16	0.06 ± 0.03
$B^+ \to D^*\pi\ell^+\nu$	0.10	0.00
$B^+ \to X_c\ell^+\nu$	11.35	10.8 ± 0.4
$B^+ \to X_u\ell^+\nu$	0.22	0.21 ± 0.03

Table 3.2: Updated branching fractions of charged semileptonic decays. The branching fraction for neutral B mesons can be obtained by dividing by the life-time ratio $\tau_{B^+}/\tau_{B^0} = 1.076 \pm 0.004$ [20]. The branching fractions (except for $B \to D^{(*)}\ell\nu$, see text for details) are scaled such that the sum over all given decay modes matches the inclusive branching fraction. See Chapter 4 for details on the derivation of the $B \to D^{**}\ell\nu$ branching fractions and their uncertainties.

(cf. Sec. 1.2.1). The CLN parametrisation in the `HQET2` module is still valid, however, the model parameters have changed since the MC generation. Thus, the `HQET2` parameters and the obsolete `ISGW2` model need to be changed as the semileptonic B decays are the dominant background to this analysis and the shape of, e. g., the lepton momentum spectrum is important. A mis-modelling can result in poor modelling of selection efficiencies and signal extraction variables.

The corrections are applied as per-event weights which are obtained from ratios of (analytic or MC) distributions of the new and old model. Thus, the width of every individual B meson decay is preserved and only the selection efficiency might change. Their impact is shown in Fig. 3.1.

The change for $B \to D\ell\nu$ is marginal since the MC and world average value of the form factor parameter ρ_D^2 are 1.16 and 1.19 ± 0.05 [59], respectively, but, nonetheless, it is corrected. In contrast, the MC form factor parameters of $B \to D^*\ell\nu$ ($\rho_{D^*}^2 = 1.3, R_1 = 1.18, R_2 = 0.71$) deviate more from the world average values ($\rho_{D^*}^2 = 1.21 \pm 0.03, R_1 = 1.40 \pm 0.03, R_2 = 0.85 \pm 0.02$ [59]) and result in a harder lepton momentum spectrum. The correction is performed in bins of lepton momentum $|\vec{p}_\ell^{\,*}|$ in the B rest frame and momentum transfer q^2.

The correction factors for $B \to D\ell\nu$ have been obtained by C. Oswald and the factors for $B \to D^*\ell\nu$ are calculated with an implementation of HQET2 by S. Falke and F. Bernlochner.

$B \to D^{**}\ell\nu$ decays are modelled with LSSW instead of ISGW2 (cf. Fig. 3.1). A ratio of the analytic predictions of LLSW and MC distributions generated with `ISGW2` is used as a per-event weight in bins of the LLSW model parameters w and $\cos\theta_\ell$. Model parameters and parameter ranges are adapted from [34] and summarised in Tab. A.1. Details on the error estimation will be given in Sec. 5.4.2.

Particle identification efficiency (PID) The identification of tracks and clusters is based on the detector geometry and response and thus affected by the accuracy of the detector model in the MC simulation. To compensate for different efficiencies in data and MC in the process of the identification of particles, several studies have been performed by the Belle internal

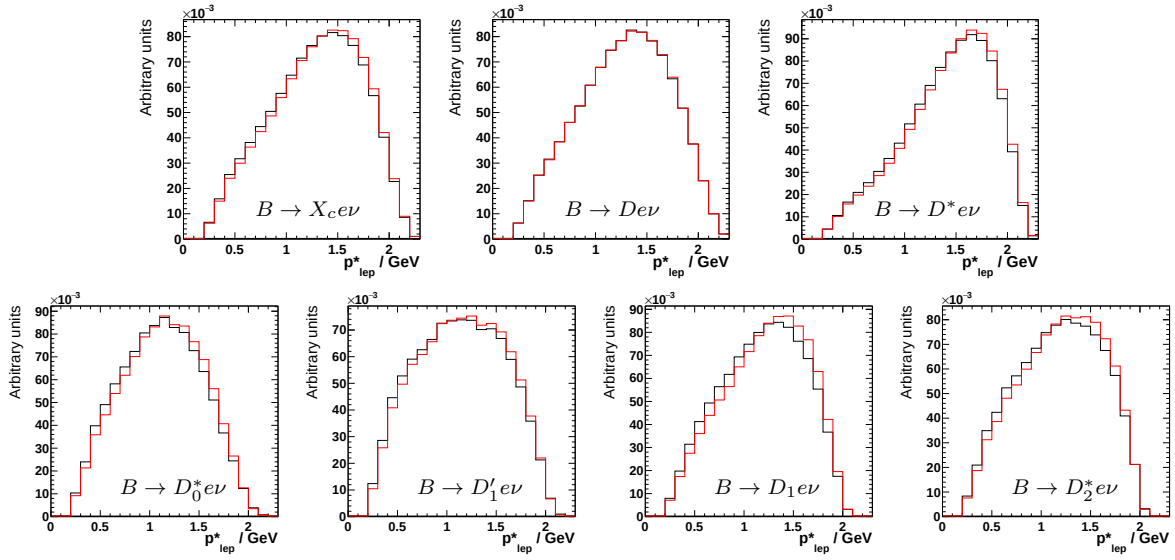


Figure 3.1: Reconstructed electron momentum distributions of the inclusive $B \to X_c \ell \nu$ spectrum and its exclusive components after updating the MC (red) and before (black). In general, the spectra become harder with the corrections.

PID group. There are dedicated studies of lepton and hadron selection efficiencies. The most important efficiency corrections for this analysis are those for the lepton identification. The normalised lepton momentum spectrum of this analysis (basically events with a single lepton; for details see Chapter 5), together with the uncertainties due to these corrections is shown in Fig. 3.2. The efficiency to identify a charged track as either an electron or a muon has been studied in low-multiplicity $e^+e^- \to e^+e^-\ell^+\ell^-$ events and average-multiplicity inclusive $e^+e^- \to J/\psi X \to \ell^+\ell^- X$ events [61]. To take the detector geometry and possible differences due to the momentum of the track into account, the lepton efficiency correction is available in ten bins of track momentum $|\vec{p}_\ell^{\,*}|$ and seven bins of azimuthal angle θ_ℓ, both measured in the detector's rest frame. There are separate corrections for electrons and muons which are both applied in this analysis. The uncertainties on the correction factors are statistical and systematic, the latter are derived from the differences between low- and average-multiplicity studies. These uncertainties are propagated into the final results.

Lepton mis-identification probability The probability of hadrons being identified as leptons is subject to this correction. Similar to the arguments in the previous paragraphs, the efficiencies are different in data and MC. The spectrum of fake lepton candidates is shown in Fig. 3.3, together with the efficiency correction uncertainty. The contamination of fake lepton candidates varies for every analysis, especially in dependence of the applied lower momentum threshold. For this analysis, the contamination is 4% and 12% for electrons ($p_\mathrm{T} > 0.3\,\mathrm{GeV}$) and muons ($p_\mathrm{T} > 0.6\,\mathrm{GeV}$), respectively (see Sec. 5.1 for details).

To estimate corrections to the MC, a dedicated $D^* \to D\pi$ sample is studied [62, 63]. The lepton mis-identification corrections are obtained for the four combinations of $\pi, K \to e, \mu$ mis-identification separately, where the arrow means that the hadron mimics a lepton. The correction depends on the track momentum $|\vec{p}_\ell^{\,*}|$ (eleven bins) and azimuthal detector angle θ_ℓ (six bins). The former two requirements are due to the fact that the mis-identification probability increases for low momenta and certain detector regions. Due to the four categories

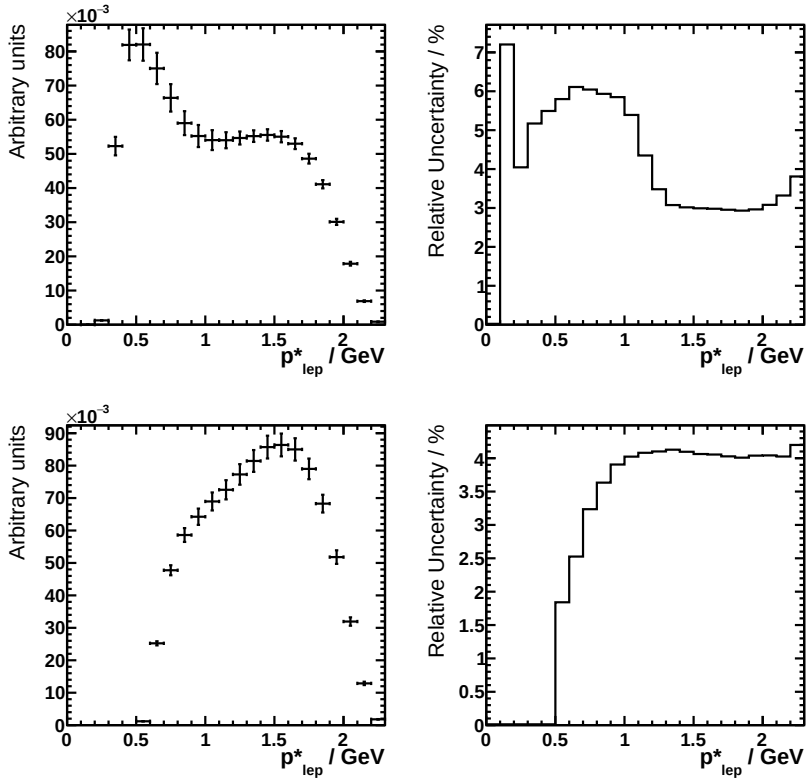


Figure 3.2: The lepton identification efficiency. Shown is the normalised, reconstructed lepton momentum spectrum of this analysis in the B meson rest frame for electrons (top left) and muons (bottom left). The absolute uncertainty on the efficiency is depicted as error bars. The relative uncertainty is given separately (right column). In the electron spectrum, the low momentum peak of secondary leptons is clearly visible, whereas it is not present for muons because of their small selection efficiency at low momenta.

of mis-identification ($\pi, K \to e, \mu$), the sample sizes of the fake rate study are low and dominate the uncertainty of the correction factors, whereas the systematic uncertainties are negligible. The correction is applied to all fake lepton candidates and the uncertainties are propagated into the results of this analysis.

Further corrections The generic MC on average misses 2.4% of generated luminosity compared to the collision data [64]. This happened in the generation process and is found to vary with the experiment-number. The generic MC is thus accordingly rescaled to match the intended number of generated events.

An equal production of charged and neutral $B\overline{B}$ at the $\Upsilon(4S)$ is assumed in the MC generation [64]. However, the experimental branching fractions are [20]

$$\mathcal{B}\left(\Upsilon(4S) \to B^+B^-\right) = (51.4 \pm 0.6)\% \quad \text{and} \quad \mathcal{B}\left(\Upsilon(4S) \to B^0\overline{B^0}\right) = (48.6 \pm 0.6)\%. \tag{3.1}$$

The $\Upsilon(4S) \to B\overline{B}$ branching fractions are corrected accordingly in the MC.

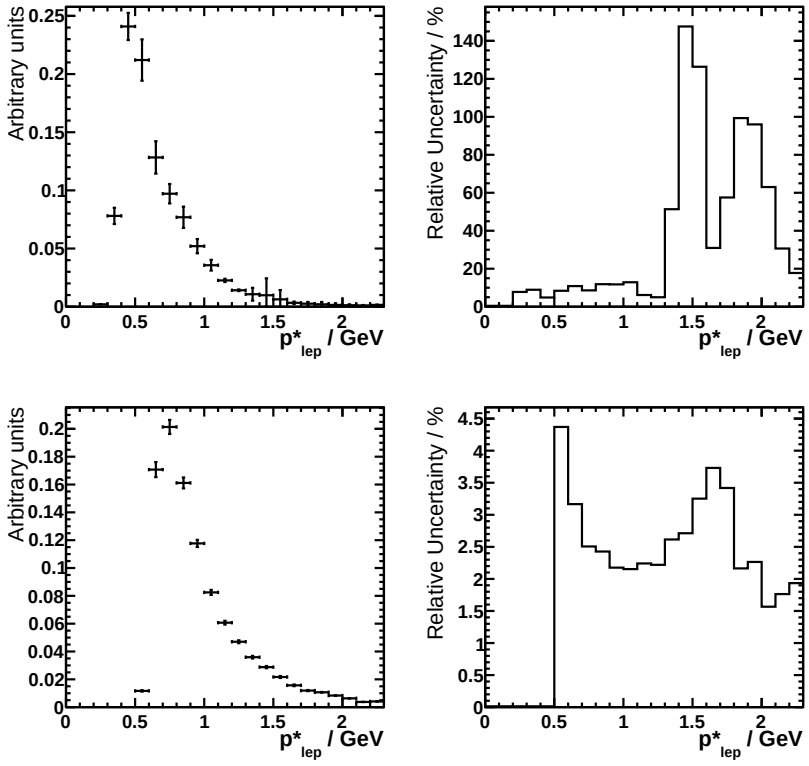


Figure 3.3: The lepton mis-identification probability in this analysis. Shown is the normalised, reconstructed lepton momentum spectrum of this analysis in the B meson rest frame for electrons (top left) and muons (bottom left). The absolute uncertainty on the efficiency is depicted as error bars. The relative uncertainty is given separately (right column). The uncertainty is purely statistical and, therefore, higher in bis with less events. Note that at high momenta especially the electron mis-identification is negligible.

3.4 Full reconstruction hadronic B tagging method

At the e^+e^- B-factories, there are several methods how to tag a $B\overline{B}$ event. Usually, only one of the two B mesons is analysed, the so-called signal B or B_{signal}; the other one can, however, provide additional information. It depends on the analysis which information of the tag B (B_{tag}) is used. The tighter the requirements on the tag B, the less efficient is the event selection of the analysis but purer samples are achieved. Commonly used in analyses are, e. g., untagged methods with high selection efficiencies or partially reconstructed B_{tag} which yield flavour information and reduce combinatorial backgrounds. The purest, but also the smallest, samples and most additional information is achieved by reconstructing the B_{tag} semileptonically or hadronically, with efficiencies of $\epsilon \sim 10^{-2}$ and $\epsilon \sim 10^{-3}$, respectively. The former yields less kinematic constraints due to the undetectable neutrino, whereas the latter suffers from even smaller efficiencies due to tiny hadronic B decay branching fractions. Nonetheless, the hadronic B_{tag} provides a wealth of information and constraints. As the B_{tag} is fully reconstructed from final state particles, the four-momentum, flavour and charge are known. Thus, in case of e^+e^- B-factories where the initial conditions of the collision are well defined, all these properties can be derived for the B_{signal}, too. Knowing the flavour and the full four-vector of the B_{signal} is a valuable (kinematic) constraint, independent of the information inferred from the B_{signal}

final-state particles. Furthermore, it allows to study observables in the B_{signal} rest frame, even if the B_{signal} is itself not fully reconstructable as it is the case in this analysis. Especially hadronic tagging strongly decreases the background level in an analysis. It suppresses continuum events as well as combinatorial background.

If the analysis relies on the correctly reconstructed flavour of the B_{signal}, there is the chance that either the B^0_{tag} or B^0_{signal} has oscillated into its anti-particle and thus the flavour prediction of the tagging is wrong. This gives rise to different backgrounds for B^+B^- and $B^0\overline{B^0}$ events. For the analysis of semileptonic B decays, such as $B \to X\tau\nu$, which relies on the detection of a single lepton, the backgrounds of B^+B^- and $B^0\overline{B^0}$ have to be treated differently. Decays of the type $B^0 \to X\ell^-$ have an irreducible background from $B^0 \to \bar{B}^0 \to X(\to X\ell^-)\pi$. For this analysis, the mixing is not problematic, but is considered in the background modelling of the signal extraction strategy (cf. Sec. 5.4).

As the branching fractions of individual hadronic B decays are small, hadronic B tagging yields a low tagging efficiency which is considered in the development of the tagging algorithm. At Belle, this aspect is approached in two ways: On the one hand, more than a thousand hadronic decay modes are considered. On the other hand, artificial neural networks (ANN) are introduced to assure that the considered hadronic decay modes have high purities. Feed-forward ANNs which are based on the "NeuroBayes" package [65] with three layers (one input, hidden and one output layer) are used. They are trained with samples of the generic Belle MC. The details of the algorithm are documented in [65], but shortly described below.

The algorithm reconstructs charged and neutral B mesons in up to four stages as shown in Fig. 3.4. The first stage is the identification of the final state particles, e, μ, γ, $\pi^\pm$ and $K^\pm$. Each candidate that passes the identification criteria is classified by a dedicated ANN whose inputs are observables like momentum, flight direction or specific energy loss. Multiple candidates are possible, i. e., one track signature can have two or more assigned particle hypotheses. B_{tag} candidates are formed from all possible combinations of particle hypothesis. However, every track signature can only contribute once to a B_{tag} candidate. For example, a track matches the hypothesis of a muon and a pion and thus can enter B_{tag} candidates as either a muon or a pion, but not as a muon and a pion in the same B_{tag}. In the following two stages, these particle candidates are combined to "intermediate" particles and resonances, e. g., $\pi^0, K_S, J/\psi$ and D. Vertex fits are performed for these candidates to improve the momentum and spatial resolution and, again, the candidates are classified by ANNs. The inputs to the ANNs are information on the daughter particles and the vertex fit results of the new composite particle. At the last stage, B_{tag} candidates are formed, classified and the best three charged and neutral B candidates (according to the ANN classification) are kept. To fine-tune the purity and efficiency of the tagged sample, the result of the final classification is available to the user as an observable in the interval $(0, 1)$ which can be interpreted as the Bayesian probability to select a well reconstructed B_{tag}. To obtain an uncorrelated observable to the ANN output for additional suppression of combinatorial background, the mass of the B candidate is not used in the ANN B_{tag} classification. In analyses, however, instead of the reconstructed B_{tag} mass, the "beam-energy constrained mass", m_{bc}, is considered for an improved resolution. This observable is uncorrelated to the ANN output and further reduces combinatorial background. It is defined as

$$m_{\text{bc}} = \sqrt{E'^2_{\text{Beam}}/4 - \left|\vec{p}'_B\right|^2}, \tag{3.2}$$

with the beam energy E'_{Beam} and the reconstructed B_{tag} candidate momentum $|\vec{p}'_B|$ in the centre-of-mass frame, denoted by the prime.

After the hadronic full reconstruction, a further suppression of $q\bar{q}$ continuum events is possible

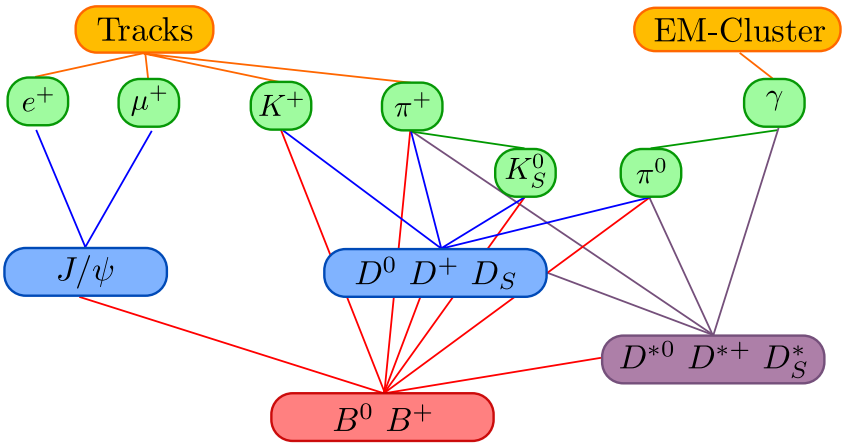


Figure 3.4: The four stages (tracks and clusters, final state particles, intermediate resonance, B_{tag} candidates) of the full reconstruction tagging algorithm. Final state particles are combined into intermediate particles and resonances and finally build up the B_{tag} candidate [65].

and provided by the Belle full reconstruction tool. Beside the classification of the reconstruction chain by the final ANN, there is the opportunity to take the event shape into account. Event shape variables parametrise the spatial distribution of the particles in an event and allow for a separation of $q\overline{q}$ from $B\overline{B}$ events. $B\overline{B}$ pairs produced at the $\Upsilon(4S)$ resonance nearly decay at rest in the centre-of-mass frame and distribute their daughter particles isotropically. In contrast, light $q\overline{q}$ production is more "jet-like". Several event shape variables are fed into a dedicated continuum-suppression ANN to further enhance the $B\overline{B}$ purity.

As already mentioned, the efficiency of a full reconstruction B tagging is rather low but the purity can be very high. Due to the ANN output available to the user, a continuous choice between the two is possible as illustrated for charged and neural B_{tag} in Fig. 3.5.

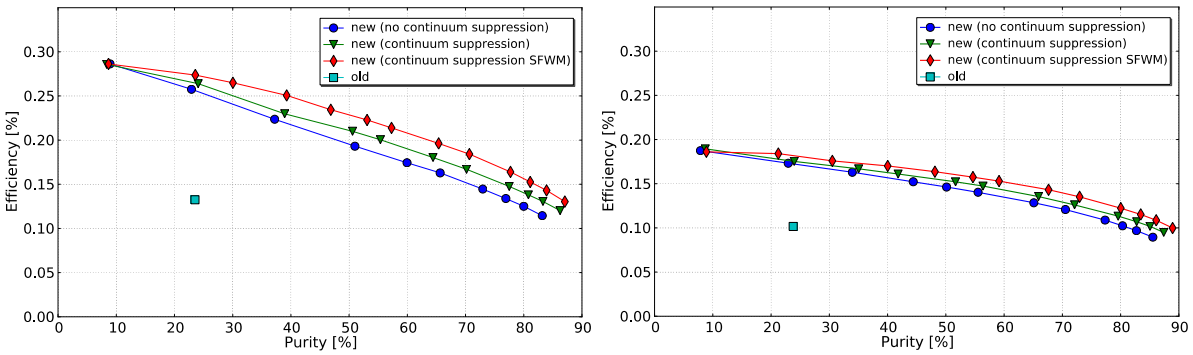


Figure 3.5: Efficiency and purity of the full reconstruction B tagging at Belle for charged (left) and neutral (right) B_{tag} [65]. The three lines correspond to different degrees of $q\overline{q}$ continuum suppression. The single data point marks to the old full reconstruction algorithm at Belle (before 2012). In this work, the most stringent and pure method with maximal $q\overline{q}$ suppression (red line) is used.

CHAPTER 4

A model for decays of excited charmed mesons

This analysis relies heavily on the modelling of semileptonic B decays. About one third of $B \to X_c \ell\nu$ decays go into higher excitations of the D meson ($B \to D^{**}\ell\nu$, cf. Tab. 3.2). Even though not dominant, semileptonic $B \to D^{**}\ell\nu$ decays are important to be correctly modelled because, for example, the heavy D^{**} soften the lepton momentum spectrum. Several final states of D^{**} meson decays were not considered in the generic $B\overline{B}$ MC, because the D^{**} mesons themselves have not been observed for a long time. An example are the two-pion final states $D^{**} \to D^{(*)}\pi\pi$ which have a different final state signature than $D^{**} \to D^{*}\pi$. The higher multiplicity leads to lower-momentum tracks which have a higher probability to be wrongly identified or to be missed in the reconstruction. The different event topology and mass distribution can have a significant impact in certain observables as e. g. the lepton momentum spectrum and thus on the signal extraction process.

The current experimental status implies, that there is a discrepancy between the $\mathcal{B}(B \to X_c \ell\nu)$ from inclusive measurements and the sum over exclusive measurements, $\sum_i \mathcal{B}(B \to D_i \ell\nu)$, where D_i stands for all considered charmed mesons. In an analysis like the one at hand, that models the inclusive decays as a sum of exclusive ones, this issue has to be addressed to correctly model the inclusive semileptonic $B \to X_c \ell\nu$ decay. In the generic $B\overline{B}$ MC, this gap is closed by scaling up the exclusive $B \to D_i \ell\nu$ branching fractions for all charmed mesons. However, this scaling leads to an overestimation of the branching fractions of $B \to D^{**}\ell\nu$ and especially of the $B \to D^{(*)}\pi\ell\nu$ final states.

A dedicated $B \to D^{**}\ell\nu$ MC with an extended set of D^{**} decay channels helps to model the data much more precisely than $B \to D^{**}\ell\nu$ decays implemented in the generic MC. The additional decay chains address the measured final states and reduce the inclusive-exclusive gap. Allowed but not yet measured modes are also included in this work. The details on how to model the new set of D^{**} decays and to generate a dedicated MC sample are discussed in this chapter. The work on the dedicated D^{**} MC started in collaboration with M. Huschle and P. Urquijo based on [66]. However, the first version of the MC is outdated and had several issues (e. g. some decay modes were missing) and thus the D^{**} modelling is completely revisited, improved and the MC generation is redone as described below.

4.1 The excited charmed meson states

The mass spectrum of charmed meson states is subject to dedicated studies and a topic of ongoing theoretical research (see [67, 68]). This chapter focuses only on the sub-set of charmed

mesons which are relevant for semileptonic B decays and are summarised in Tab. 1.2.

The mass spectrum of charmed mesons, i. e. heavy-light quark systems, has been introduced in Sec. 1.1.1. This chapter is about the excited 1P states, i. e. the four $l = 1$ angular excitations of the charmed meson. While usually only the 1P states (D_0^*, D_1', D_1 and D_2^*) are denoted as D^{**} mesons, in the following the term is extended to the radially excited 2S states for simplicity.

Numerous measurements have been carried out to obtain the branching fractions of $B \to D^{**}\ell\nu$ [20, 59, 69]. However, because of the limited knowledge on the D^{**} mesons themselves, only partial branching fractions of the 1P states are known and none for the 2S states. The total branching fraction of $B \to D^{**}\ell\nu$ hence depends on assumptions for the decay modes of the D^{**} mesons. The D^{**} meson decay chains can be constrained by measurements of semileptonic B decays with higher multiplicity final states (see Tab. 4.1). Additional information is available through ratios of hadronic B decay branching fractions, e. g. $\frac{\mathcal{B}\left(B\to D_2^*\pi\right)\times\mathcal{B}\left(D_2^*\to D\pi\right)}{\mathcal{B}\left(B\to D_2^*\pi\right)\times\mathcal{B}\left(D_2^*\to D^*\pi\right)}$, for which a summary of available measurements is given in Tab. 4.2.

Decay mode	$\mathcal{B}/\%$
$B^+ \to D\pi\ell^+\nu$	0.64 ± 0.08 [59]
$B^+ \to D^*\pi\ell^+\nu$	0.84 ± 0.09 [59]
$B^+ \to D\pi\pi\ell^+\nu$	0.30 ± 0.12 [70]
$B^+ \to D^*\pi\pi\ell^+\nu$	0.22 ± 0.10 [70]
$B^+ \to D_s^{(*)}K\ell^+\nu$	0.06 ± 0.01 [20]

Table 4.1: Measured final states of charmed semileptonic B decays. The values are obtained assuming isospin symmetry. The uncertainties on the $D^{(*)}\pi\pi$ final states contain an additional isospin uncertainty as the measurements are agnostic to whether the two pions are produced resonantly or non-resonantly. The $D\pi$ and $D\pi\pi$ final states do not stem from D^* and $D^*\pi$.

	D_0^*	D_1'	D_1	D_2^*	D_{2S}^*
$\mathcal{B}(B \to D^{**}\ell\nu) \times \mathcal{B}(D^{**} \to D\pi)/\%$	0.44 ± 0.08				
$\mathcal{B}(B \to D^{**}\ell\nu) \times \mathcal{B}(D^{**} \to D^*\pi)/\%$		0.20 ± 0.06	0.43 ± 0.03	0.12 ± 0.01	
$\mathcal{B}(D^{**} \to D\pi)/\mathcal{B}(D^{**} \to D^*\pi)$				1.54 ± 0.14	0.32 ± 0.09
$\mathcal{B}(D^{**} \to D\pi\pi)/\mathcal{B}(D^{**} \to D^*\pi)$			0.57 ± 0.23	< 0.7	

Table 4.2: Partial branching fractions and branching fraction ratios related to excited charmed mesons [20, 69]. The upper limit given for the D_2^* is an estimate at 90% confidence level, based on Ref. [71]. The values are obtained under the assumption of isospin conservation.

4.2 Modelling of D^{**} decays

In addition to "on-shell" decays of particles into light daughter particles, "off-shell" decays are possible because of the large width of the mother particle. Off-shell modes are decay chains whose final state particles together have a larger mass than the nominal mass of the mother particle. Although they are suppressed, their sum can significantly contribute to the total decay rate of D^{**} mesons and thus to the sum of exclusive semileptonic B decays. Off-shell decays are therefore included in the new model.

Building up the D^{**} model follows four major steps: Fist, the possible final states are determined by mass constraints. Afterwards, the transition matrix element is estimated by decomposing the angular configuration of the initial and final state and considering heavy-quark

spin symmetry (HQSS), which also reduces the possible final states. Third, the available phase space of the remaining decays is estimated. Finally, the results of steps two an three are combined to estimate the D^{**} branching fractions. The branching fractions are then compared with measurements to test the quality of the model and to check for consistency. All these steps are described in detail in the following sections.

4.2.1 Available final states

All combinations of D or D^* mesons with a light, unflavoured meson below the mass of the ω, $m_\omega = (782.7 \pm 0.1)\,\text{MeV}$ [20], are considered. In addition, $D_s^{(*)}K$ final states as well as $D^{(*)}\pi\pi$ are taken into account.

A final state is considered if $\sum m_i < (m_0 + \alpha\Gamma)$, i. e. the sum of the masses of daughter particles m_i is lower than the mass m_0 of the mother plus its width Γ multiplied by a cut-off parameter α. In the following $\alpha = 1.54$ is chosen, as for this value the cumulative distribution function of the Breit-Wigner distribution of the mother particle yields 90%: $\int_{-\infty}^{\alpha} P(m, m_0, \Gamma)\mathrm{d}m = 0.9$. Here,

$$P(m, m_0, \Gamma) = \frac{1}{2\pi} \frac{\Gamma}{(m - m_0)^2 + \Gamma^2/4} \tag{4.1}$$

is the non-relativistic Breit-Wigner distribution of a particle with mass m around the mean mass m_0 and width Γ. This procedure reduces the amount of final states, especially for the narrow 1P states.

4.2.2 Matrix element and heavy-quark spin symmetry

The matrix element of the strong interaction is assumed to factorise into a part describing the transition between the initial and final state spin configuration $|M_{\text{spin}}|$ and a general QCD matrix element including the hadronisation $|M_{\text{QCD}}|$. Furthermore, the assumption is made that $|M_{\text{QCD}}|$ cancels in the ratio $\mathcal{B}(D_i \to X) = \frac{\Gamma(D_i \to X)}{\Gamma(D_i)}$ as the hadronisation is difficult to calculate and beyond the scope of this model. Thus, the branching fractions are determined by angular momentum conservation at leading order.

The strong decay of the D^{**} conserves the spin orientation of the charm quark in the limit of a large charm quark mass, $m_c \to \infty$ (cf. Sec. 1.2.1). In this limit, some decay modes are prohibited and thus suppressed at finite masses. The general spin and angular momentum configuration is part of the matrix element and will be considered in the following. The procedure given below was introduced in Ref. [60] and is adapted to describe the light vector meson (ρ, ω) production as well.

As already stated, the initial D^{**} state is described by the heavy quark spin s_h and the total angular momentum of the light degrees of freedom $j_l = s_l + l$ which couple to the total D^{**} meson angular momentum J. The final state has the same quantum numbers and is denoted with a prime, e. g. J', from now on. In the transition, the light degrees of freedom de-excite by emitting a hadron h and loosing the angular momentum $J_h = K + L$, where K and L denote the angular momenta of the light meson and the system of charmed and light meson (charmed-light system in the following), respectively.

As the strong interaction conserves parity, only combinations of the initial and final state angular momenta fulfilling this requirement are considered. The parity of the initial state, a meson with angular momentum l, is $P = (-1)^{l+1}$ and that of the final state is $P' = (-1)^{l'+K+L}$. The transition $J \to J' + J_h$ can be decomposed and given in terms of Clebsh-Gordan coefficients,

whereas m_i denote the z-components of the angular momenta. Clebsh-Gordan-coefficients describe the decomposition of a coupled state into the uncoupled sub-states. They are thus the coefficients of the linear combination of the sub-states to the coupled state. The coefficients are non-vanishing for allowed $J \to J' + J_h$ transitions and the partial wave, i. e. the angular momentum of the charmed-light meson system can be inferred. Only the leading-order partial wave will be used later on and possible higher excitations of the charmed-light meson system are neglected. For clarity, the coefficients are grouped for the initial state (A), the final state charmed meson (A'), the transition (A_{trans}) and the light meson (A_{light}) contribution.

$$\begin{aligned} A &= \langle s_h, m_{s_h}; j_l, m_J - m_{s_h} | J, m_J \rangle \langle s_l, m_{s_l}; l, m_J - m_{s_h} - m_{s_l} | j_l, m_J - m_{s_h} \rangle \\ A' &= \langle s_h, m_{s_h}; j'_l, m_{J'} - m_{s_h} | J', m_{J'} \rangle \langle s'_l, m_{s'_l}; l', m'_J - m_{s_h} - m_{s'_l} | j'_l, m'_J - m_{s_h} \rangle \\ A_{\text{trans}} &= \langle J_h, m_J - m_{J'}; j'_l, m_{J'} - m_{s_h} | j_l, m_J - m_{s_h} \rangle \langle J_h, m_J - m_{J'}; J', m_{J'} | J, m_J \rangle \\ A_{\text{light}} &= \langle K, m_K; L, m_J - m_{J'} - m_K | J_h, m_J - m_{J'} \rangle. \end{aligned} \tag{4.2}$$

The first line, A, simply describes the initial state meson, whereas A' already exhibits the conservation of the initial state heavy-quark spin in the final charmed meson, i. e., the z-component of s_h is the same in A and A'. The momentum conservation in the decay is ensured by A_{trans} on the level of the light degrees of freedom as well as the total angular momentum. The angular momentum L between the final-state particles also needs to be in agreement with the state of the light meson which is ensured by A_{light}. In Ref. [60], only scalar mesons are considered for which $A_{\text{light}} = \delta_{J_h L}$. However, to incorporate vector mesons, A_{light} becomes a Clebsh-Gordan-coefficient as the vector meson's spin has to be considered in the momentum conservation. Due to angular momentum conservation, one of the three spin orientations can be expresses by the two others in the Clebsh-Gordan coefficients: $\langle j_1, m_1; j_2, m_2 | j, m_1 + m_2 \rangle$. Putting these together yields

$$M_{\text{spin}} \propto \sum_{m_J, m_{J'}, m_{s_l}, m_K} \delta_{PP'} A \cdot A' \cdot A_{\text{trans}} \cdot A_{\text{light}}. \tag{4.3}$$

The summation is over all possible combinations of spin orientations, except the heavy-quark spin whose orientation is a free choice as it only defines the spatial orientation and thus $M_{\text{spin}}(m_{s_h}) = M_{\text{spin}}(-m_{s_h})$. The Kronecker delta ensures parity conservation and further reduces the allowed final states. Together, momentum and parity conservation result in suppressed partial wave contributions because certain spin configurations are not possible.

Only the leading-order partial wave, i. e. the lowest L for which the matrix element is non-vanishing, is considered in the following. The partial wave of a three-body decay is not as trivial as the one for the two-body decay. The reason is the $\pi\pi$-pair in the final state which can be interpreted as a pseudo-particle $\tilde{\pi}$ with angular momentum $L_{\pi\pi}$. Because the $\pi\pi$ pair can have any relative angular momentum, $L_{\pi\pi}$ is not fixed and, therefore, the sum in Eq. 4.3 converges very slowly. An even more advanced model may consider the associated spherical harmonics and integrate out their angular dependence after the summation, which gives an additional suppression of higher order partial waves. However, this is beyond the scope of this model and the sensitivity of this analysis. Therefore, the partial-wave decomposition is omitted for the $D^{**} \to D^{(*)}\pi\pi$ transitions.

4.2.3 Phase space estimate

To model the relative branching fractions of the D^{**} mesons, the phase spaces of the decays are estimated. The two-body phase space Φ_2 depends only on the magnitude of the momentum p of the final state particles (in the rest frame of the mother particle), which is given by the masses in the initial (m_0) and final state (m_1, m_2)

$$\Phi_2 = \frac{p}{2M} \text{ with } p = \sqrt{\left(m_0^2 - (m_1+m_2)^2\right)\left(m_0^2 - (m_1-m_2)^2\right)}. \tag{4.4}$$

Following Ref. [72], the phase space is modified to take into account the combinatorics of the angular momentum L of the partial wave and an exponential suppression by a "barrier-penetrating" factor

$$\Phi_2^L = \frac{p^{2L+1}}{2M} e^{-p^2}. \tag{4.5}$$

However, so far the phase space does not allow for off-shell decays for which the width Γ of the mother particle needs to be taken into account. This is achieved by a convolution of Φ_2^L with the Breit-Wigner mass distribution (Eq. 4.1) of the mother particle

$$\tilde{\Phi}_2^L = \int_{m_1+m_2}^{\infty} P(m, m_0, \Gamma)\Phi_2^L(m, m_1, m_2)\mathrm{d}m. \tag{4.6}$$

In principle, this can be extended to convolutions with the widths of the daughter particles, too. However, the width of most of the final state particles is negligible. Only the ρ meson has a significant width which is neglected for simplicity.

So far, the procedure describes the case of two-body final states. In case of three-body final states, the calculation of the integrated phase space is not trivial and needs assumptions to obtain analytical solutions. Under the assumption that two of the three final-state particles are identical, the integrated phase space Φ_3 is given by [73]

$$\Phi_3 = \frac{1}{2^9\pi^4 m_0^2}\int_{4m_2^2}^{(m_0-m_1)^2} \sqrt{1-\frac{4m_2^2}{s}} \cdot \sqrt{\left(\frac{m_0^2-m_1^2+s}{2m_0}\right)^2 - s} \quad \mathrm{d}s. \tag{4.7}$$

In this case, it is not trivial to incorporate the angular momentum and the "barrier-penetrating" factor. The former is omitted anyhow, as the transition into a three-body final state is much less governed by a certain partial wave than the two-body decays (cf. Sec. 4.2.2). The latter is not considered because the momenta of the light mesons in $D\pi\pi$ and $D^*\pi\pi$ are similar and thus this factor should be of comparable size and cancel in the ratio. Due to the assumptions which are different from those in the two-body decays, the three-body phase space is only used to estimate the relative decay rate between $D^{**} \to D\pi\pi$ and $D^{**} \to D^*\pi\pi$. The width of the mother particle is considered in the same way as before for the two-body decays, i. e. by convolving with the mother's Breit-Wigner distribution.

4.3 Branching fraction calculation

With the above considerations, the branching fractions of D^{**} and $B \to D^{**}\ell\nu$ decays are estimated. First, the two-body decays are considered. Later, the three-body decays are incorporated which are treated separately because no three-body matrix element was calculated.

The results of the two-body decays are combined to the partial decay width according to Eq. 1.10

$$\Gamma(i \to f) \propto \tilde{\Phi}_2^L \cdot |\mathcal{M}_{\text{spin}}|^2. \tag{4.8}$$

The branching fraction of the charmed meson D_i^{**} to the final state X_j are calculated as

$$\mathcal{B}\left(D_i^{**} \to X_j\right) = \frac{\Gamma(D_i^{**} \to X_j)}{\sum_k \Gamma(D_i^{**} \to X_k)}, \tag{4.9}$$

thus, only decays of the same mother particle appear in the ratio and contributions from the QCD matrix element are assumed to cancel.

Using Eq. 4.8, the model predicts the ratios

$$\Gamma(D_2^* \to D\pi)/\Gamma(D_2^* \to D^*\pi) = 1.9\ , \tag{4.10}$$

$$\Gamma(D_{2\text{S}}^* \to D\pi)/\Gamma(D_{2\text{S}}^* \to D^*\pi) = 0.7\ , \tag{4.11}$$

whereas measurements find 1.54 ± 0.14 and 0.32 ± 0.09, respectively. Thus, the prediction is not entirely precise, but still remarkable for this simplistic model. To further improve the accuracy, the measured ratios are used instead of the predicted ones from now on and the $D\pi$ final state is updated accordingly. In principle, the $D^*\pi$ final state could be updated instead and would lead to enhanced $D^{**} \to D^{(*)}\pi$ transitions compared to the other D^{**} decays. The choice, however, was taken to enhance the non-$D^{(*)}\pi$ transitions because this helps to tackle the problem of the inclusive-exclusive gap.

Now that all two-body partial decay widths are set, the branching fractions of the D^{**} mesons and thus those of $B \to D^{**}\ell\nu$ are obtained. The three-body decays are included iteratively which is described in the next paragraph. Afterwards, the branching-fractions of $B \to D^{(*)}\pi\ell\nu$, $B \to D^{(*)}\pi\pi\ell\nu$ and $B \to D_s^{(*)}K\ell\nu$ are extracted from the model and compared with the measurements (Tab. 4.1 and 4.2). In all results, the indirect production of $D^{(*)}\pi\pi$ via intermediate resonances, like $D^{**} \to D\rho \to D\pi\pi$ or $D_{2\text{S}}^* \to D_2^*\pi \to D\pi\pi$, is taken into account.

Additional information stems from the total $B \to X_c\ell\nu$ branching fraction and the well measured $B \to D^{(*)}\ell\nu$ decays, as the $B \to D^{**}\ell\nu$ decays have to fill the difference between the two. As throughout the whole analysis, the exclusive semileptonic B branching fractions of Tab. 3.2 are used here.

The $B \to D^{(*)}\pi\pi\ell\nu$ contribution The contributions of decays into $D^{(*)}\pi\pi$ are estimated in the following way: First, the measured ratio $\mathcal{B}\left(D_1 \to D\pi\pi\right)/\mathcal{B}\left(D_1 \to D^*\pi\right)$ is used to set the decay width of $D_1 \to D\pi\pi$. Based on that, the decay width of $D_1 \to D^*\pi\pi$ is fixed with the ratio of phase spaces $\Phi_3(D_1 \to D^*\pi\pi)/\Phi_3(D_1 \to D\pi\pi)$.

Still, the contribution of the remaining D^{**} to the $D^{(*)}\pi\pi$ final state needs to be estimated. To do so, the $D^{**} \to D\pi\pi$ decay width is related to the $D^{**} \to D^*\pi$ decay width by a constant factor of 0.01. An exception is D_0^* which neither decays into $D^*\pi$ nor $D\pi\pi$ due to parity conservation. Here, $D^*\pi\pi$ and $D\pi$ are related by the same factor. The factor 0.01 is a rather arbitrary factor which is chosen such that the measurements of $\mathcal{B}\left(B^+ \to D^{(*)}\pi\pi\ell\nu\right)$ are not exceeded. It could be different for every D^{**} meson, but since there are no measurements available and the structure of $D^{**} \to D^{(*)}\pi\pi$ transitions is completely unknown, an equal factor for all D^{**} mesons is a reasonable choice. The $D^*\pi\pi$ final states of the D_1, D_2^*, $D_{2\text{S}}$ and $D_{2\text{S}}^*$ mesons are related to the $D\pi\pi$ final states by the ratio of the three-body phase space estimates as it was done above for the D_1. After all, this results in rather low branching fractions of the D^{**} mesons

into $D\pi\pi$ which is supported by the recent finding that $\mathcal{B}(D_2^* \to D\pi\pi)/\mathcal{B}(D_2^* \to D^*\pi) \lesssim 0.7$ (based on Ref. [71]).

Remarks on $B \to D_s^{(*)}K\ell\nu$ It is found that the model overestimates the amount of $B \to D_s^{(*)}K\ell\nu$ already with the contribution from the 1P states. This is understandable as these decays require the production of a heavier $s\bar{s}$-quark pair in contrast to the light $u\bar{u}$ and $d\bar{d}$ pairs in the other decays. The suppression from the production of the heavier quark pair is not part of the model. An additional suppression factor, which is tuned to match the measured $B \to D_s^{(*)}K\ell\nu$ branching fraction, might be one possible way to solve this issue. However, a different approach is used as it is found that the modelled decays $D_{2S}^{(*)} \to D_s^{(*)}K$ reproduce the measured $\mathcal{B}(B \to D_s^{(*)}K\ell\nu)$ quite well if the 1P states do not contribute at all. The 1P $\to D_s^{(*)}K$ contribution is thus set to zero which is arbitrary but not problematic for two reasons: First, the contribution of $D_s^{(*)}K$ is negligible (smaller than the uncertainty on the inclusive $\mathcal{B}(B \to X_c\ell\nu)$). Second, this model is primarily designed for an inclusive analysis and thus the production scheme is not relevant to such accuracy. Due to the small amount of $D_s^{(*)}K$ this even holds for most exclusive analyses of semileptonic B decays.

The branching fractions of $B \to D_{2S}^{(*)}\ell\nu$ There are no measurements of $B \to {D_{2S}}^{(*)}\ell\nu$ and thus the branching fraction needs to be assumed. The measured final states are used as a guideline to derive an estimate for the $\mathcal{B}\left(B \to D_{2S}^{(*)}\ell\nu\right)$. The main goal is to reduce the inclusive-exclusive gap and not to produce more of a specific final state than measured. The measured $B \to D^*\pi\ell\nu$ branching fraction is found to be the most limiting constraint. Therefore, $\mathcal{B}(B^+ \to D_{2S}\ell^+\nu) = 0.1\%$ and $\mathcal{B}(B^+ \to D_{2S}\ell^+\nu) = 0.2\%$ is chosen, whereas the $B^+ \to D_s^{(*)}K\ell^+\nu$ final state is the restricting factor. Possible deficits in one or more final states can be filled up with non-resonant semileptonic B decays as summarised in Tab. 4.3. However, as the non-resonant contributions are well within the uncertainties, the inclusive-exclusive gap is filled by slightly enlarged $B \to D^{**}\ell\nu$ branching fractions in this analysis.

Decay mode	Branching fraction / %		
	This model	Measured	Non-resonant
$B^+ \to D\pi\ell^+\nu$	0.58	0.64 ± 0.08 [59]	0.06
$B^+ \to D^*\pi\ell^+\nu$	0.87	0.84 ± 0.09 [59]	-
$B^+ \to D\pi\pi\ell^+\nu$	0.28	0.30 ± 0.12 [70]	0.02
$B^+ \to D^*\pi\pi\ell^+\nu$	0.16	0.22 ± 0.10 [70]	0.06
$B^+ \to D_s^{(*)}K\ell^+\nu$	0.04	0.06 ± 0.01 [20]	0.02
$B^+ \to X_c\ell^+\nu$	-	10.8 ± 0.4 [20]	-

Table 4.3: Comparison of measured and model predicted branching fractions as well as estimates of non-resonant branching fractions. The latte are estimated from the deficit of the model to the measured branching fraction.

MC simulation Based on the obtained model, a dedicated MC sample is generated with EvtGen and the full Gsim detector simulation. In the MC production itself, only one of the two B mesons in the generated MC events decays according to the new model, while the second B decays generically. There are two arguments why this is not relevant in the given circumstances. First, analyses that rely on a hadronic tagging method will highly suppress events where both B

mesons decay semileptonically (i. e. one with the new an one with the old model). Second, the hadronic reconstruction at Belle does not rely on D^{**} mesons in hadronic decays and one can expect that those are also suppressed. Hence, it is very unlikely to have semileptonic B decays with the old model of the generic MC in the dedicated D^{**} sample after hadronic tagging.

In the MC production, appropriate EvtGen modules are selected to describe the decays. As no `LLSW` module is available but corrections factors from `ISGW2` to `LLSW` are, `ISGW2` is used to model the semileptonic $B \to D^{**}\ell\nu$ decays. Without going into further details of EvtGen and its modules (an overview is given in [56, 74]), the subsequent decays of D^{**} mesons are modelled with the modules summarised in Tab. A.7. If possible, the obtained leading partial wave or at least the spin of the initial and final state particles is considered to model the angular distribution of the final state particles correctly. For decays where no appropriate model is available, a uniform phase-pace distribution (`PHSP` module) is assumed. This affects the three-body decays as well as some rare cases, e. g. $D_0^* \to D\pi$.

4.4 Conclusions

The results of the last sections are visualised in a term diagram (Fig. 4.1) and given in Tab. 4.4 and A.6. This new model for D^{**} meson decays has several advantages compared to the generic MC model. It considers new, measured final states as e. g. $D^{(*)}\pi\pi$ and $D_s^{(*)}K$ as well as possible but not yet measured modes as $D^{(*)}\eta$. The relative fraction of the decay mode is derived from estimates of the matrix element using HQSS as well as the available phase space. This results in a self-consistent model which matches the constraints from exclusive and inclusive B decays. It is thus well suited for this and future studies of inclusive and certain exclusive semileptonic B decays.

Decay mode	$\mathcal{B}$ / %	$\Delta\mathcal{B}_{\text{Meas.}}$ / %	$\Delta\mathcal{B}_{\text{Model}}$ / %
$B^+ \to D_0^*\ell^+\nu$	0.58	0.10	0.07
$B^+ \to D_1'\ell^+\nu$	0.32	0.16	0.06
$B^+ \to D_1\ell^+\nu$	0.77	0.05	0.17
$B^+ \to D_2^*\ell^+\nu$	0.32	0.03	0.10
$B^+ \to D_{2\text{S}}\ell^+\nu$	0.1		0.05
$B^+ \to D_{2\text{S}}^*\ell^+\nu$	0.2		0.10

Table 4.4: Obtained branching fractions of $B \to D^{**}\ell\nu$ decays and estimates of uncertainties. They are combinations of the model estimates in Tab. A.6 and the partial branching fractions in Tab. 4.2. $\Delta\mathcal{B}_{\text{Meas.}}$ is the propagated uncertainty of the measured partial branching fractions. A rough estimate of the model uncertainty is $\Delta\mathcal{B}_{\text{Model}}$ which is estimated as the half of the partial branching fraction into unmeasured modes.

The important aspect for an analysis using simulations based on this new D^{**} model is the change in the observables. Important quantities to most analyses are the momentum $|\vec{p}_\ell^{\,*}|$ of the prompt lepton in the rest frame of the decaying B meson and the squared missing mass, m^2_{miss}. The latter is the squared sum of the four-momenta of the undetected particles in an event and is introduced in detail in Sec. 5.2. Fig. 4.2 compares normalised distributions of reconstructed $|\vec{p}_\ell^{\,*}|$ and m^2_{miss} as they appear in this analysis for the generic Belle MC and the dedicated D^{**} MC. The impact of the additional decay modes is clearly visible. Especially the $B \to D_1'\ell\nu$ shows a softer lepton spectrum due to the numerous off-shell decays. In contrast, the lepton spectrum of the $B \to D_1\ell\nu$, which differs from the generic MC only due to the additional $D^{(*)}\pi\pi$ on-shell

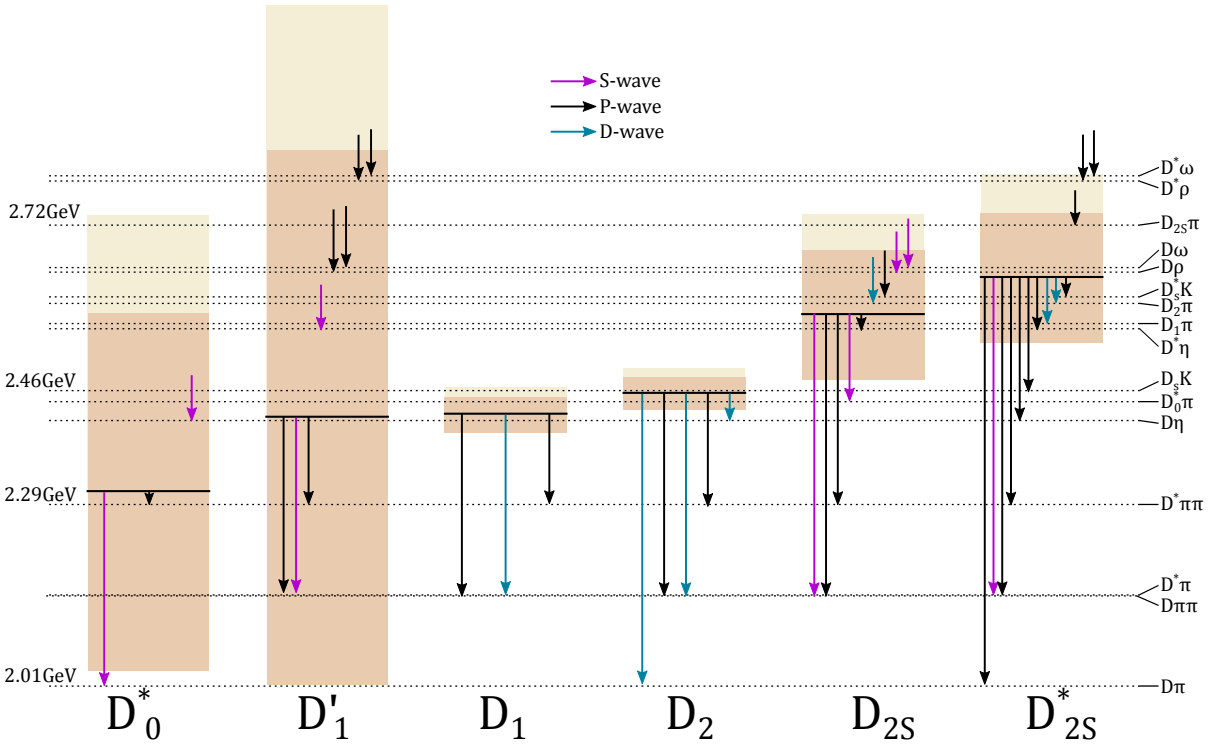


Figure 4.1: The modelled D^{**} decays depicted as arrows pointing to the final states (dotted lines). Black horizontal lines indicate the nominal D^{**} mesons masses and the red boxes their width. The maximum mass below which final states are considered is shown in grey. The starting point of the off-shell arrows is arbitrary.

final state, does not change. Similar are the plots of m^2_{miss} which show a broadening of the distribution for the new D^{**} MC. This is expected because the higher the final state multiplicity is, and thus the lower the momenta of the final state particles are, the higher is their probability to be missed in the track reconstruction.

A higher multiplicity and a softer $|\vec{p}^{\,*}_\ell|$ spectrum have an impact on semileptonic analyses, especially the measurements involving τ leptons, i. e. $R(D)$, $R(D^*)$ and $R(X)$. Typically, the τ is reconstructed in in the leptonic $\tau \to \ell\nu\nu$ ($\ell = e, \mu$) channels. Compared with prompt leptons produced in $B \to X_c\ell\nu$ decays, the signal lepton has a rather low momentum because it is a secondary particle from the three-body τ decay. In the lepton momentum spectrum, the signal thus resides in the low momentum region with backgrounds from other processes producing secondary leptons as well as the low momentum tail of the $B \to X\ell\nu$ lepton momentum distribution. Previous analyses used (amongst others) the lepton momentum spectrum to estimate the $B \to D^{**}\ell\nu$ background with on-shell D^{**} meson decays (the recent Belle analyses use the first version of the dedicated D^{**} MC, but with fewer off-shell decays than in the model used here). The model shows that a certain amount of off-shell decays causes a softer momentum spectrum than expected from simulations using pure on-shell D^{**} decays. Thus, the $B \to D^{**}\ell\nu$ lepton spectrum becomes softer and populates the signal region of inclusive and exclusive $B \to \tau$ analyses. A similar argumentation holds for the squared missing mass, which is, due the multiple neutrinos produced in $B \to \tau$ decays, expected to be large for signal events. The simulated missing mass distributions in Fig. 4.2 show a broadening of the $B \to D^{**}\ell\nu$ decays towards higher missing mass values, and thus again an expansion into the signal region.

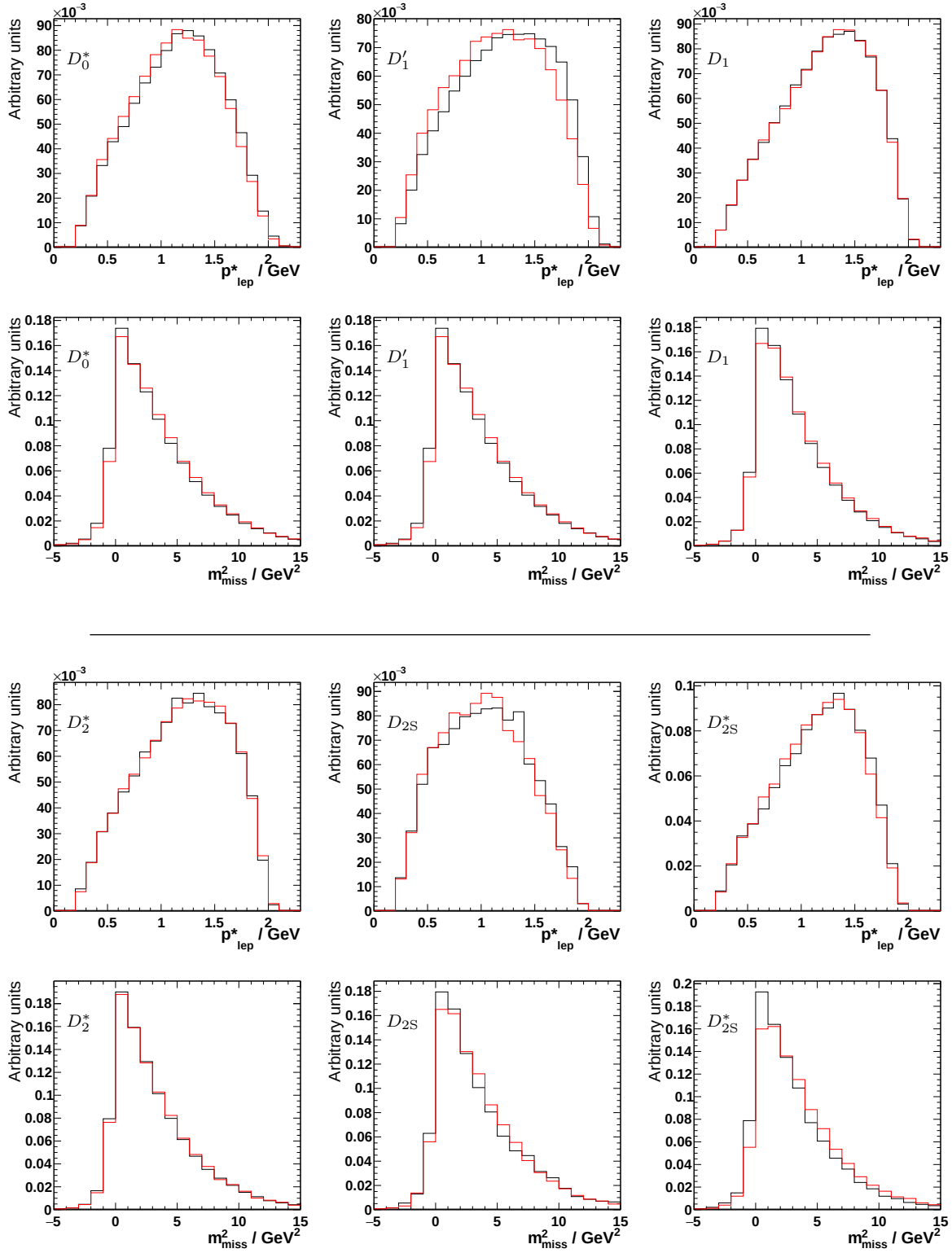


Figure 4.2: Lepton momentum spectra (in the reconstructed B_{signal} rest-frame, cf. Sec. 3.4) and missing mass distributions of $B \to D^{**}\ell\nu$ decays with new (red) and generic (black) modelling. Mesons with large off-shell contributions (D_0^*, D_1', D_{2S} and D_{2S}^*) show a softer lepton spectrum. All mesons show a reduced missing mass resolution because of final states with higher multiplicities.

Hence, it is a crucial aspect for B decays to τ final states and for the first time taken properly into account in this thesis.

So far, the current chapter sets the focus on $B \to X_c \ell \nu$ with a light lepton because measurements for these decays are available and the modelling is even more important than for the $B \to X_c \tau \nu$ signal. This first sounds contradictory, but is understandable because the momentum of the signal lepton is barely correlated with that of the τ lepton (cf. Sec. 5.2). The impact of the τ momentum on $|\vec{p}_\ell^{\,*}|$ or m^2_{miss} is thus reduced and so the importance of an exact τ momentum modelling. Nevertheless, a softer lepton spectrum due to off-shell decays also affects the momentum of the τ and that of its daughters. Therefore, the signal is modelled with the new D^{**} transitions, too.

There are some simplifications made in this model that might become important in future experiments which directly measure decays of D^{**} mesons. Of course, more measurements are needed to further constrain the decay modes of the individual D^{**} mesons. An extended set of branching fractions or ratios of them would be advantageous to constrain most of the decay modes. Additionally, a future extended model will benefit from measurements of branching fractions of the 2S states.

The estimates of the decay width in this work rely on phase space calculations and approximations of the matrix element in the heavy-quark scheme. A more advanced theory can, for example, take the actual angular distribution of the matrix elements into account or estimate more of the QCD matrix element, e. g. the $s\bar{s}$ production.

However, given the current knowledge, the presented model is sufficiently accurate as it reproduces the experimental constraints reasonably well.

CHAPTER 5

Analysis of $B \to X\tau\nu$ decays

Inclusive measurements of B decays are challenging and even more difficult with τ leptons in the final state. In general, inclusive analyses have a high signal efficiency but a rather low purity because they lack typical constraints from the reconstruction of the final state and the decay chain. In addition, the τ lepton decays too fast to interact with the detector and only secondary leptons or hadrons from the τ decay are in the final state. Semileptonic B decays into a τ lepton always include multiple neutrinos in the final state. Beside the neutrino produced in the semileptonic $B \to X\tau\nu$ decay, the semileptonic $\tau \to X\nu$ decay and the leptonic $\tau \to \ell\nu\nu$ decay produce one or two additional neutrinos, respectively. In this analysis, the $B \to X\tau\nu$ signal is selected in the leptonic τ decay mode because the final-state lepton has a distinct signature in the hadron dominated environment of $B\overline{B}$ decays. The three neutrinos produced in $B \to X\tau\nu$ decays leave the detector without energy depositions which makes a full reconstruction of the τ itself impossible. Due to the missing neutrinos in the final state, the total reconstructed energy and momentum differ from the initial state e^+e^- kinematics, which will be exploited in this analysis.

To keep this analysis of $B \to X\tau\nu$ inclusive, no intermediate resonances are reconstructed. For this reason, this work is an analysis of $R(X)$ rather than $R(X_c)$. Even though charmless $B \to X_u\ell\nu$ decays are not explicitly modelled in the MC simulation, these decays are not suppressed by the selection and thus part of the extracted signal yield.

The composition of the charmed mesons affects the shape of kinematic distributions, e. g. the lepton momentum spectrum, which is later considered in the systematic uncertainties (cf. Chapter 4 and Sec. 5.5). The inclusive approach of this analysis in combination with the τ lepton in the final state provides only a small set of useful observables to distinguish signal from background. In addition, due to the large analysed data set, even a slight mis-modelling of the data by the MC can become statistically significant. Together, these two aspects are the particular challenge of an inclusive $B \to X\tau\nu$ analysis.

This analysis does not directly measure the signal branching fraction but the ratio of branching fractions

$$R(X) = \frac{\mathcal{B}(B \to X\tau\nu)}{\mathcal{B}(B \to X\ell\nu)}, \tag{5.1}$$

and the $B \to X\ell\nu$ decay mode is thus also referred to as the *normalisation mode* in this analysis, whereas ℓ is either an electron or a muon. The extraction of $R(X)$ instead of the branching fraction $\mathcal{B}(B \to X\tau\nu)$ has several advantages. First, the overall MC normalisation is irrelevant. Any constant factor to the MC that applies to $B \to X\tau\nu$ and $B \to X\ell\nu$ cancels in the ratio. And

even if, e. g., the tagging efficiency is not entirely uniform in the signal extraction variables, it cancels partially. Additionally, systematic uncertainties which are positively correlated between the signal and the normalisation mode also cancel. Examples for this are the number of $B\overline{B}$ pairs, the tagging efficiency or the lepton identification and selection efficiencies. By measuring the ratio $R(X)$, this analysis is completely independent of the prior knowledge of $\mathcal{B}(B \to X\tau\nu)$ and $\mathcal{B}(B \to X\ell\nu)$.

The following sections describe the strategy of the inclusive measurement of $R(X)$ with the Belle data set. First, the event selection and reconstruction are presented (Sec. 5.1) and the key observables used in this analysis are defined (Sec. 5.2). Section 5.3 discusses the modelling of the signal extraction variables. Afterwards, the signal extraction strategy is introduced (Sec. 5.4) and the expected systematic uncertainty of this analysis is estimated (Sec. 5.5). Last, the fitting procedure is tested in a side-band region (Sec. 5.6).

The analysis is implemented using the ROOT data analysis framework (v5.34/30) [75]. The minimisation in the fitting procedure is performed with Minuit2 [76, 77] which is implemented as part of the ROOT framework.

5.1 Event reconstruction and selection

In order to perform the previously outlined analysis, the events recorded at Belle need to be selected according to their signal signature. The first step, recording and preselecting $B\overline{B}$ events, has been described in Chapter 2 and Sec. 3.4. The required event signature, as depicted in Fig. 5.1, is a hadronically decaying B_{tag} and a semileptonically decaying B_{signal} for both, signal and normalisation mode. The $B \to X\tau\nu$ signal is only selected by the leptonic τ decays, i. e. a single light lepton in the final state is required. With this choice, backgrounds from hadronic B decays are suppressed but the selection efficiency is reduced by $\mathcal{B}(\tau \to e/\mu\nu\nu) = (35.21 \pm 0.06)\,\%$. The analysis of e. g. $\tau \to \pi\nu$ or $\tau \to \rho\nu$ is difficult because of large hadronic backgrounds. Furthermore, the large number of pions in the final state of B decays makes it, even for signal events, difficult to select the τ daughter pion and not a pion from the charmed meson decay.

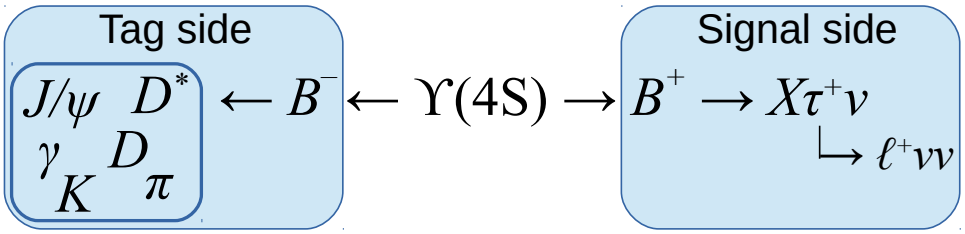


Figure 5.1: Illustration of the signal signature in an $B\overline{B}$ event. The $\Upsilon(4S)$ decays into a B^+B^- pair and one of the B mesons (here B^+ as an example) decays in the $B \to X\tau\nu$ signal mode (*signal side*). The τ subsequently decays, however, in this analysis only the leptonic final states of the τ are considered. The other B meson decays hadronically and is used in the B tagging of the event (*tag side*). Note the anti-correlation of the tag side flavour and the charge of final-state lepton. The same principle holds for the decay into a $B^0\overline{B^0}$ pair, however, $B^0\overline{B^0}$ mixing dilutes the anti-correlation between flavour and charge.

The event selection starts with the requirement of a B_{tag} by using the previously described full reconstruction technique (see Sec. 3.4). The best neutral or charged B_{tag} candidate is chosen in an event, according to the ANN classifier output. The hadronic tagging classifier, including an additional $q\overline{q}$ suppression $\text{ANN}_{q\overline{q}}$, is shown in Fig. 5.2. In the selection of hadronic

B events, the $\mathrm{ANN}_{q\bar{q}}$ is required to obey a loose selection of $\ln(\mathrm{ANN}_{q\bar{q}}) > -6$. However, after the reconstruction of final-state particles and the selection of signal-like events (see below), the selection is optimised by maximising the Figure of Merit $\mathrm{FoM} = S/\sqrt{S+B}$, whereas S and B are the event yields of well reconstructed and poorly reconstructed B_{tag}, respectively. As the hadronic tagging has different efficiencies for charged and neutral B_{tag}, the optimisation is performed separately (see Fig. A.1). However, both optimisations yield very similar results and thus a common selection of $\ln(\mathrm{ANN}_{q\bar{q}}) > -3$ is chosen.

Furthermore, combinatorial background is suppressed by demanding $m_{\mathrm{bc}} > 5.27\,\mathrm{GeV}$. The latter is the standard selection at Belle and keeps the peaking part of the m_{bc} distribution, while the combinatorial background populates the long tail below 5.27 GeV (see Fig. 5.2).

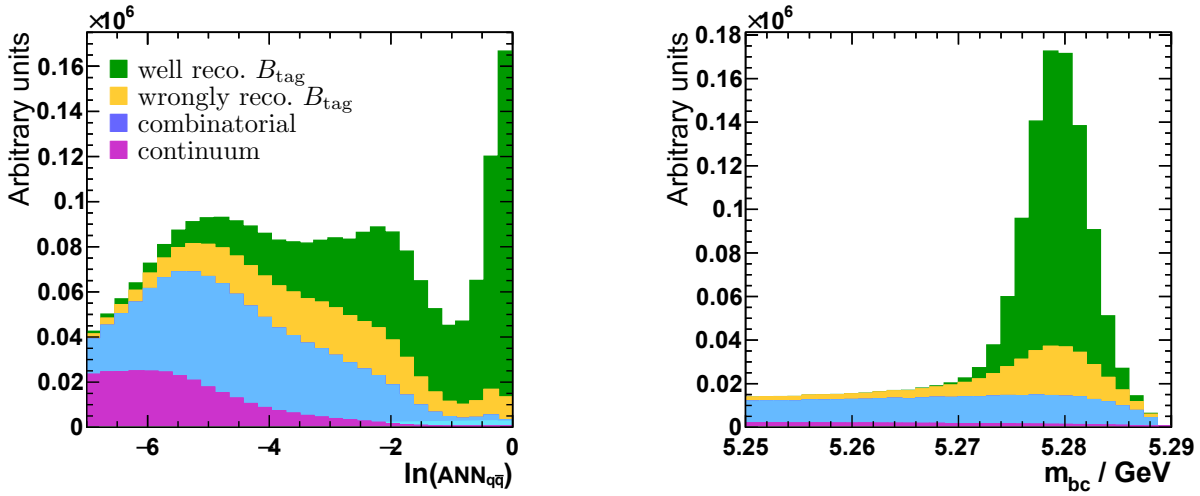


Figure 5.2: Distributions of $\ln(\mathrm{ANN}_{q\bar{q}})$ and m_{bc}. Well reconstructed B_{tag} (green) are located in the m_{bc} peak and have a peaking distribution in $\ln(\mathrm{ANN}_{q\bar{q}})$. A wrongly identified particle in the B_{tag} reconstruction (yellow) leads to a less distinct shape. In contrast, poorly reconstructed and combinatorial B_{tag} candidates (blue) and continuum (purple) are located in the tail region of m_{bc} and rise towards lower values of the neural network classifier output. For this analysis, events with $\ln(\mathrm{ANN}_{q\bar{q}}) > -3$ and $m_{\mathrm{bc}} > 5.27\,\mathrm{GeV}$ are selected. The distributions are shown after all other selection requirements.

5.1.1 Event clean-up and particle identification

The full reconstruction tagging is performed by an external module wich provides a sample of events containing B_{tag} candidates. Afterwards, the analysis-specific event clean-up and reconstruction of the final state is performed. Events are cleaned from duplicated tracks by the algorithm presented in [78]. In case of low momentum tracks, it is possible that the track curls back close to the IP and the tracking algorithm finds two tracks from the in hits in the detector. Those can be either two close tracks of the same charge or two oppositely charged tracks propagating back-to-back as depicted in Fig. 5.3. All low momentum tracks with transverse momentum $p_{\mathrm{T}} < 275\,\mathrm{MeV}$ are checked in pairs for being duplicated. Tracks are assumed to be a duplicate if the polar angle between the tracks is $\theta_{12} < 15°$ ($\theta_{12} > 170°$) for equally (oppositely) charged tracks, whose momenta differ less than $||\vec{p}_1| - |\vec{p}_2|| < 100\,\mathrm{MeV}$. The track which originates closer to the IP is kept, the other one is discarded as a duplicate. An exception are duplicates that have been used in the B_{tag} candidate. In this case, if only one of the two tracks in question is used on the tag side, the one on the signal side is discarded.

In the rare case that both tracks are used in the B_{tag} reconstruction, the event is discarded.

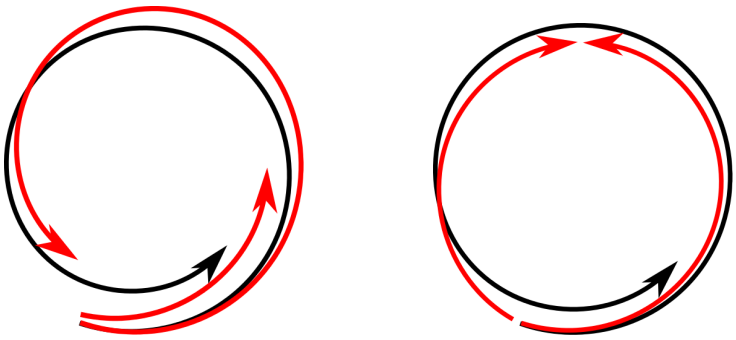


Figure 5.3: Sketch of track duplication of a single true track (black) leading to two reconstructed tracks (red). The view is in z-direction along the beam-line and the tracks start at the interaction point. The duplication either produces co-aligned (left) or back-to-back (right) tracks.

After the event is cleaned from duplicated tracks, the particle identification (PID) is performed. Charged tracks are tested on the hypothesis of being an 1. electron, 2. muon, 3. kaon or 4. pion. The first matching hypothesis is used. As this analysis heavily relies on a signal lepton from a τ decay, low momentum leptons play a crucial role and the PID criteria are chosen to be tight (see Tab. 5.1). The leptons are required to originate close to the IP and lie within the acceptance region of the tracking detectors. Electrons need to reach the ECL which requires a transverse momentum of $p_{\text{T}} > 0.3\,\text{GeV}$, whereas muons are required to reach the KLM, i. e., to have at least $p_{\text{T}} > 0.6\,\text{GeV}$.

	PID	min $\lvert\vec{p}\rvert$/ MeV	max $(\text{d}r, \text{d}z)$/cm	θ
Electron	> 0.90	300	$(0.5, 1.5)$	$17° < \theta < 150°$
Muon	> 0.97	600	$(0.5, 1.5)$	$25° < \theta < 145°$
Kaon	< 0.60	100	$(0.5, 1.5)$	-
Pion	> 0.60	100	$(0.5, 1.5)$	-
Photon	-	150	-	-

Table 5.1: The selection and identification criteria of all considered charged and neutral particles on the signal side. Charged tracks and neutral clusters that fail these criteria are discarded. PID refers to the particle identification probability (see text).

As a multivariate tool to separate leptons from hadrons, likelihood ratios are formed by the Belle PID group which are interpreted as the probability to select a true lepton (cf. Fig. 5.4). The highest probabilities which are still recommended by the Belle PID group are chosen: eID > 0.9 and μID > 0.97. For electrons, the ECL cluster shape and for muons the projection to a KLM hit is included in the likelihood.

Charged tracks that are not compatible with a lepton hypotheses are separated into kaons and pions at a likelihood ratio of 0.6, whereas lower values correspond to a Kaon. Also in the case of pions and kaons, the tracks need to originate from close to the IP. Tracks that fail all hypothesis tests are discarded.

Electrons have a high probability to emit final-state radiation (FSR) which reduces the reconstructed electron momentum and slightly changes its direction. The probability to emit FSR decreases rapidly with the number of emitted photons and their energy. Thus, only the possibility of a single photon with $E < 1\,\text{GeV}$ is considered. For every electron candidate, a search for a neutral cluster with $E < 1\,\text{GeV}$ in a 5° cone around the flight direction of the

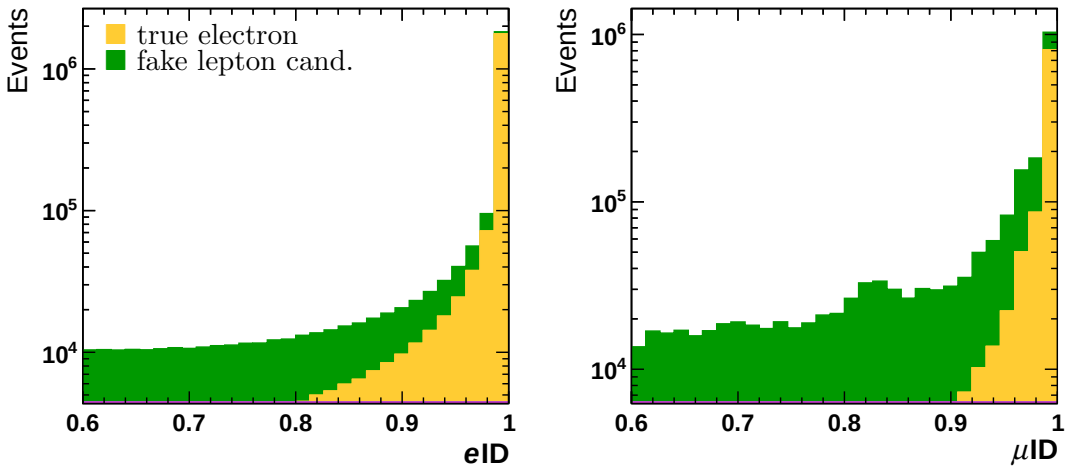


Figure 5.4: Distribution of the electron (left) and muon (right) identification probabilities. Electrons with eID > 0.9 and muons with μID > 0.97 are selected.

electron is performed. If successful, the electron four-momentum is modified according to $\vec{p}_e \to \vec{p}_e + \vec{p}_\gamma$, $E_e \to \sqrt{|\vec{p}_e + \vec{p}_\gamma|^2 + m_e^2}$. If multiple FSR photon candidates are found, the most energetic one is selected. The FSR of other particles is expected to be negligible and no FSR reconstruction is performed.

Neutral clusters are selected as photon candidates if they have an energy of $E_\gamma > 150\,\text{MeV}$ and are not used in the B_tag reconstruction or as an FSR photon. The energy threshold suppresses photons from beam background which is mostly located in the forward and backward direction of the detector. It is found that the high E_γ threshold of 150 MeV in the barrel region improves the data-MC agreement compared to the Belle standard of 50 MeV for exclusive analyses (cf. Sec. 5.3). Photons are assumed to originate from the IP.

5.1.2 Event selection

The signal lepton from the τ decay and the primary lepton of the normalisation mode can be either an electron or a muon and its charge is required to be compatible with the flavour of the decaying $B_\text{signal} = \overline{B}_\text{tag}$. Events with more than one lepton are rejected. In principle, a small but notabe sample is that of two oppositely charged leptons which accounts for ∼5% of the preselected sample. The second lepton stems from a semileptonic charm decay and allows to directly measure $B \to X_c\tau\nu$ rather than $B \to X\tau\nu$. However, this dilepton sample obeys different backgrounds and should be treated separately from the single-lepton sample. Therefore, it is not included in this analysis as it is negligible compared to the single lepton sample. Anyhow, it provides a future option to measure $R(X_c)$ statistically independent at Belle.

The correlation of the B_tag flavour and the signal lepton charge is a good choice for charged B_tag but in case of neutral B_tag, mixing and (to a small amount) CP eigenstates in the full reconstruction dilute this correlation. Therefore, the signal and background efficiencies ϵ of the requirement on the correlation are studied in MC for neutral B_tag. Fig. 5.5 shows the signal electron momentum with and without the requirement, as well as the sample of rejected events. It is clearly visible that semileptonic B decays are preserved, whereas secondary leptons from hadronic B decays are suppressed. The efficiencies are:

	$B \to X\tau\nu$	$B \to X\ell\nu$	other
ϵ	64.8%	78.5%	38.5%

Thus, despite the mixing in the neutral B sector, the flavour-charge correlation is a useful requirement, in particular to suppresses backgrounds which populate the low momentum signal region.

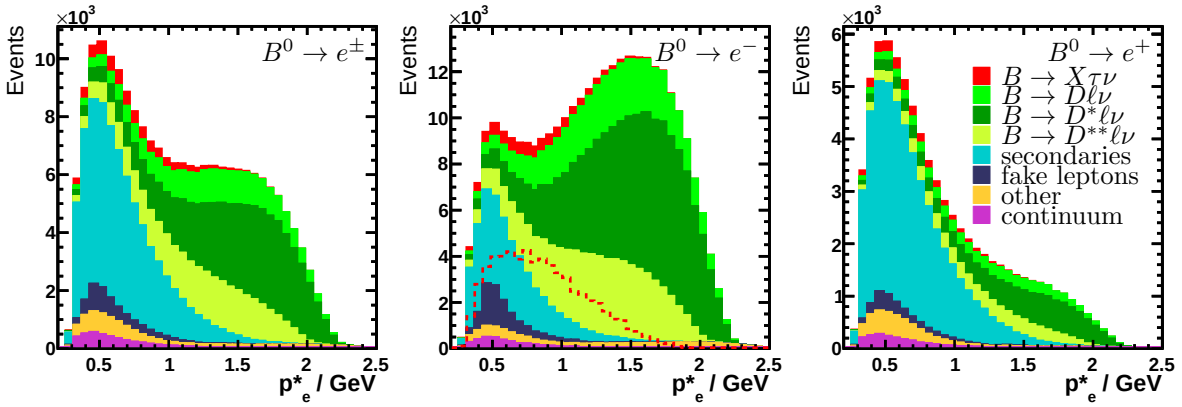


Figure 5.5: Impact of the requirement on the correlation of B^0_{tag} flavour and lepton charge. Shown is the lepton momentum without the requirement (left), with the requirement applied (centre) and inverted (right). Even though the $B^0\overline{B^0}$ events suffer from the dilution of this correlation due to mixing, a large amount of secondary leptons from hadronic B decays is rejected.

5.2 Observables

This section introduces the available observables to study the $B \to X\tau\nu$ decay. The hadronic full reconstruction tagging allows for boosting the kinematics into the rest frame of the B_{signal} and thus to study observables in this frame. Quantities in the B_{signal} rest frame are marked with a star, e. g. $|\vec{p}^*_\ell|$.

Lepton momentum An important observable for this analysis is the signal lepton momentum $|\vec{p}^*_\ell|$, shown in Fig. 5.6 (top). It allows one to separate highly energetic prompt leptons of the semileptonic B decays from the lower energetic secondary leptons, such as the signal lepton from $B \to X\tau\nu$ decays. The $B \to X\tau\nu$ signal lepton momentum spectrum does not carry as much information on the decay as that of $B \to X\ell\nu$ decays because $|\vec{p}^*_\ell|$ is not a good proxy for the τ momentum, as illustrated in Fig. 5.7. The correlation between the two is only ∼22% due to the two additional neutrinos in the τ decay. This leads to the fact that (in contrast to the prompt lepton in $B \to X\ell\nu$ decays) the τ daughter lepton momentum spectrum is not sensitive to the modelling of the τ momentum or the momentum transfer $q^2 = (p^\mu_\tau + p^\mu_\nu)^2$. The mean momentum of the light τ daughter signal lepton $\overline{|\vec{p}^*_\ell|} = 0.78\,\text{GeV}$ is lower compared to that of the τ $\overline{|\vec{p}^*_\tau|} = 1.12\,\text{GeV}$. This systematically lower momentum, on the one hand, can be exploited to extract the signal in $|\vec{p}^*_\ell|$ as it gives rise to a better separation of $B \to X\tau\nu$ from $B \to X\ell\nu$. However, on the other hand, the signal resides more in the regime of the less well modelled secondary and fake lepton candidates. Furthermore, particles with low momenta suffer from a poor selection efficiency which stems from the detector performance. For example, the requirements stated in Sec. 5.1 reduce the signal efficiency by 20% and 54%, only due to $p^e_{\text{T}} > 0.3\,\text{GeV}$ and $p^\mu_{\text{T}} > 0.6\,\text{GeV}$, respectively.

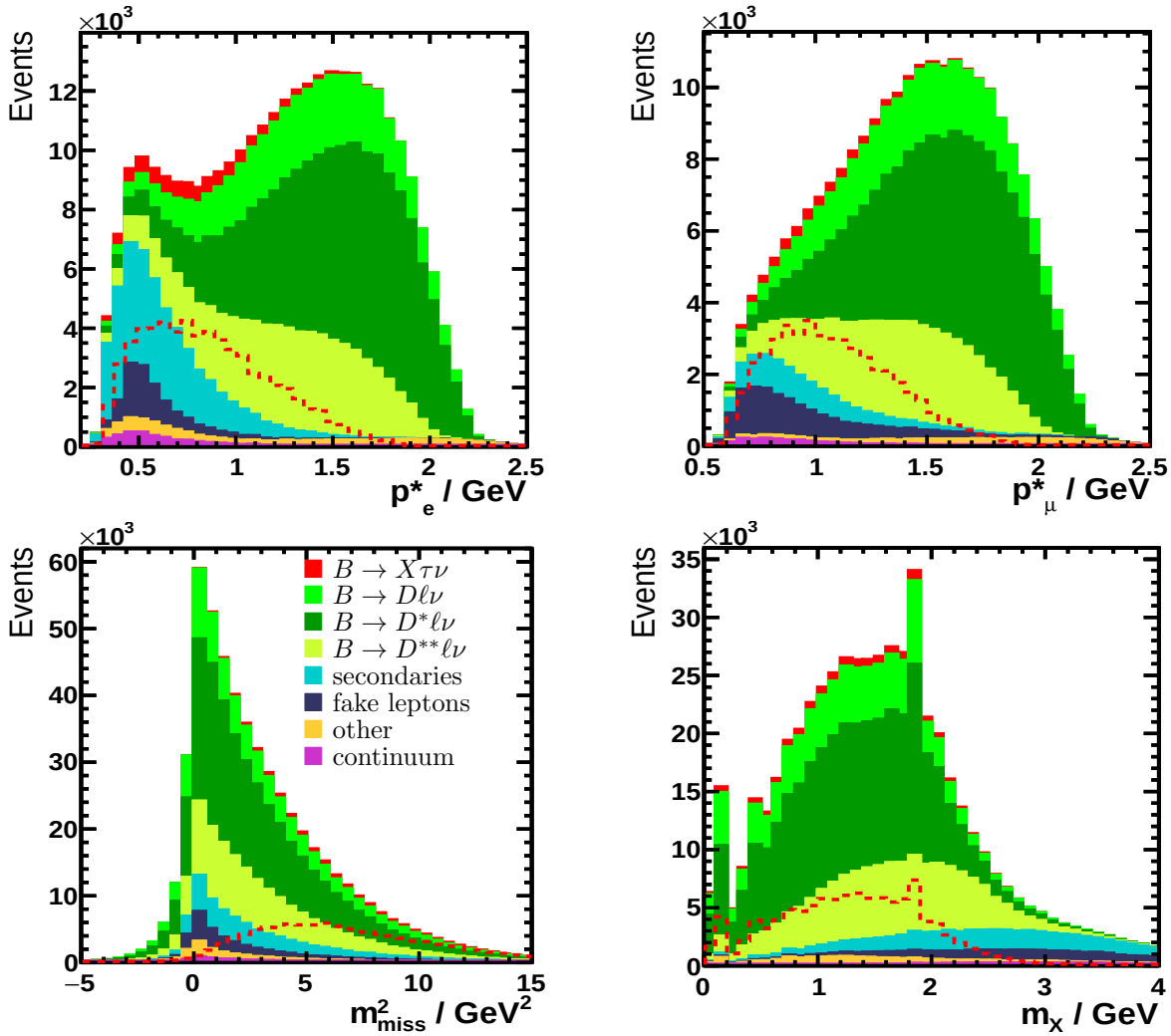


Figure 5.6: Distributions of the reconstructed lepton momentum spectrum, shown separately for electrons (top left) and muons (top right) as well as the squared missing mass, m^2_{miss}, distribution (lower left) and the hadronic mass, m_X, (lower right). The signal (red) is additionally shown as a dotted line of arbitrary normalisation to illustrate its distribution. It is the only component not peaking at $m^2_{\text{miss}} = 0\,\text{GeV}^2$. The prominent peak in the m_X spectrum belongs to hadronic $D^{(*)}$ decays. The momentum spectra of electrons and muons differ due to the low reconstruction efficiency of slow muons, which also leads to an enhanced amount of fake lepton candidates (dark blue).

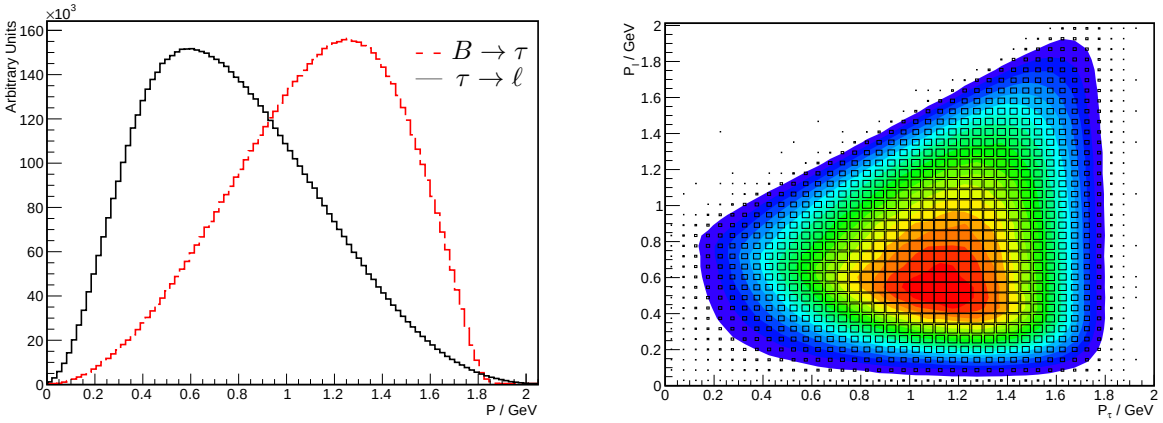


Figure 5.7: Momentum spectra in the B_{signal} rest frame of the τ and its daughter lepton (e and μ are shown together) for $B \to X\tau\nu$ decays at generator level. *Left:* Momentum distributions of the prompt τ (dashed red line) and its daughter lepton (solid black line). The momentum spectrum of the secondary daughter lepton is significantly softer than that of the prompt τ. *Right:* The 2D distribution of the mother and daughter particle momenta shows only a low correlation of ∼22%.

The lepton momentum spectrum of the non-semileptonic B decays have a rather diverse composition regarding the flavours of the B meson and the lepton. Fig. 5.8 shows the momentum spectra of leptons which are produced neither in $B \to X\tau\nu$ nor in $B \to X\ell\nu$ decays. In general, a large fraction of the leptons stems from hadronic B decays. For charged B decays, the decays into double-charmed final states are dominant because of the charge-flavour correlation requirement in the lepton selection. This requirement is diluted for neutral B events and thus the single-charmed final states occur as often as the double-charmed ones. Independent of the B flavour, the muon sample suffers from a high amount of fake muon candidates which usually stem from hadronic B decays. In contrast, electron fake candidates are a small component, however, a large amount of electrons is produced in $\gamma \to e^+e^-$ pair-production. The very diverse compositions for charged and neutral B mesons as well as electrons and muons will be considered in the fit to extract the signal (Sec. 5.4).

As visible in Fig. 5.6, the signal lepton has a slightly harder momentum spectrum than the non-semileptonic backgrounds. To reduce the dependence on these backgrounds without loosing significantly signal events, the signal-to-background ratio in the $|\vec{p}^*_\ell|$ spectrum is considered. This is done separately for electrons and muons and also individually for charged and neutral B_{tag}, because their non-semileptonic backgrounds differ. The optimal lower momenta (see also Fig. A.2) are $|\vec{p}^*_e| > 0.44\,\text{GeV}$ and $|\vec{p}^*_\mu| > 0.62\,\text{GeV}$ as well as $|\vec{p}^*_e| > 0.5\,\text{GeV}$ and $|\vec{p}^*_\mu| > 0.67\,\text{GeV}$ for charged and neutral B_{tag}, respectively. Only the very low momenta, which barely contain signal events, are removed by this procedure.

Squared missing mass Another important observable is the squared missing mass m^2_{miss} (see Fig. 5.6, lower left). It is derived from the detected, i. e. visible, four-momenta p^μ_{visible} as well as the known four-momenta of the colliding beam particles $p^\mu_{e^+e^-}$:

$$m^2_{\text{miss}} = p^2_{\text{miss}} = \left(p^\mu_{e^+e^-} - p^\mu_{\text{visible}}\right)^2 . \tag{5.2}$$

It corresponds to the squared four-momentum that is carried away by undetected particles. Thus, it vanishes for purely hadronic events and events with a single neutrino or lost photon. The m^2_{miss} distribution is broadened due to the detector resolution as well as not reconstructed

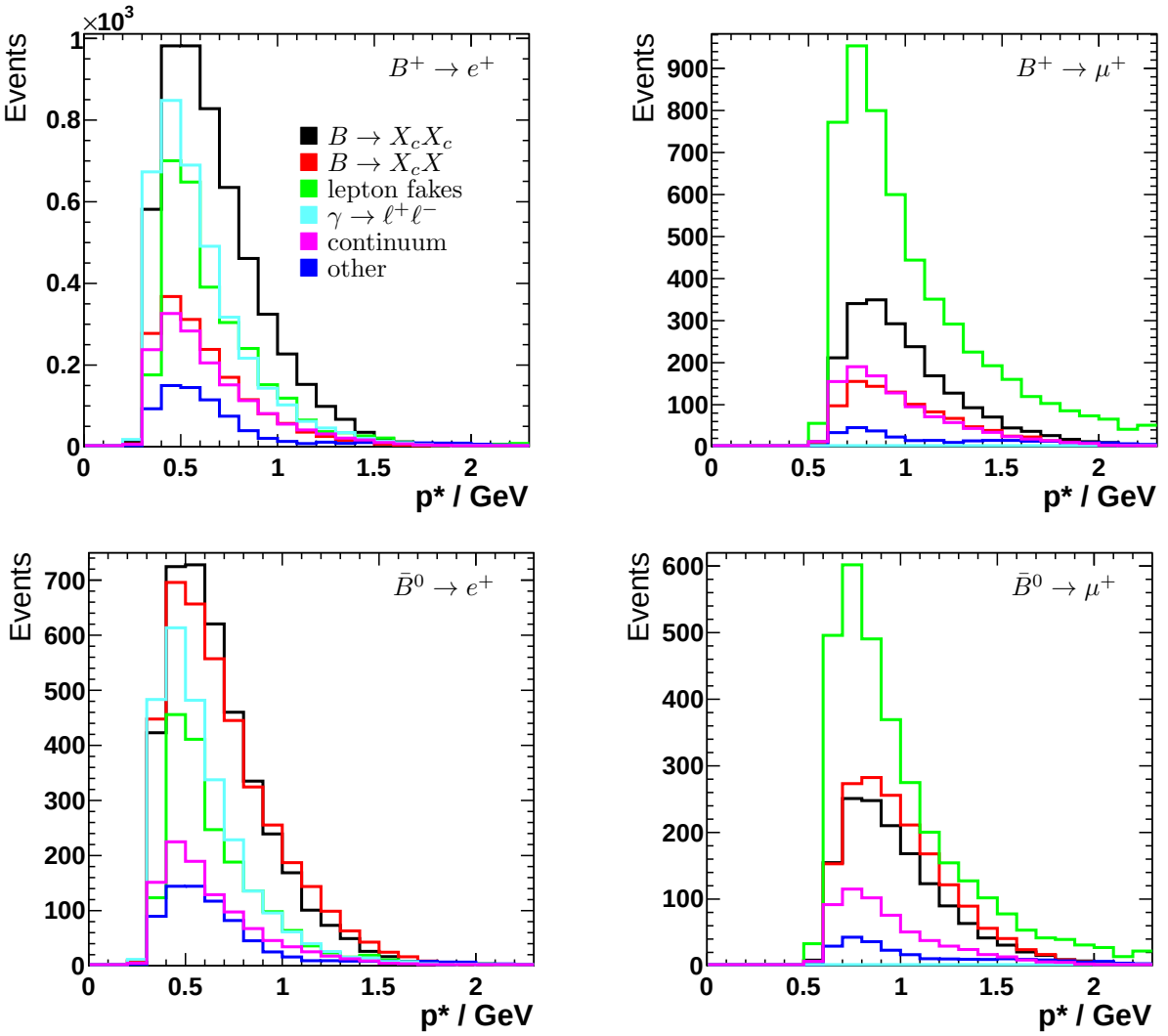


Figure 5.8: The composition of the lepton momentum spectrum of non-semileptonic B decays which varies strongly with the flavours of the B meson and the lepton. Charged B decays are dominated by double-charmed final states (black) which produce secondary leptons that pass the flavour-charge correlation requirement. Conversely, the strict charge correlation is diluted for neutral B mesons and thus the single- (red) and double-charmed final states are equally frequent. In general, muons suffer from a high rate of fake lepton candidates (green) whereas electrons have a high probability of being produced in pair-production (cyan). Other decays (blue) include leptons produced in e. g. charmonium or strange meson decays.

tracks and photons as illustrated in Fig. 5.9. While the former is a symmetric effect, the latter gives rise to a long tail towards high m^2_{miss} values.

In case of a single missing particle, m^2_{miss} equals the mass of the missing particle. Therefore, close to zero, missing mass is nearly independent of the kinematics in the event. The distribution is broadened by detector effects which smear the reconstructed energy and momentum in p^μ_{visible}. At high m^2_{miss} values, the situation changes. The undetected particles generate a squared mass of

$$m^2_{\text{miss}} = \left(\sum\nolimits_i E_i\right)^2 - \left|\sum\nolimits_i \vec{p}_i\right|^2 \tag{5.3}$$

$$= \sum\nolimits_i m_i^2 + 2\sum\nolimits_{i\neq j} \left(E_i E_j + p_i p_j \cos\theta_{ij}\right), \tag{5.4}$$

which depends directly on the masses m_i, energies E_i, absolute momenta p_i and the relative directions θ_{ij} of the particles and thus on the physical processes. This is important as it needs accurate knowledge of the physics to model the data accurately in the tail region. However, at the same time, the tail is smeared out and individual processes cannot be separated. This reduces the impact of particular decays and their kinematic modelling, however, as backgrounds have different resolutions in m^2_{miss}, their relative composition matters.

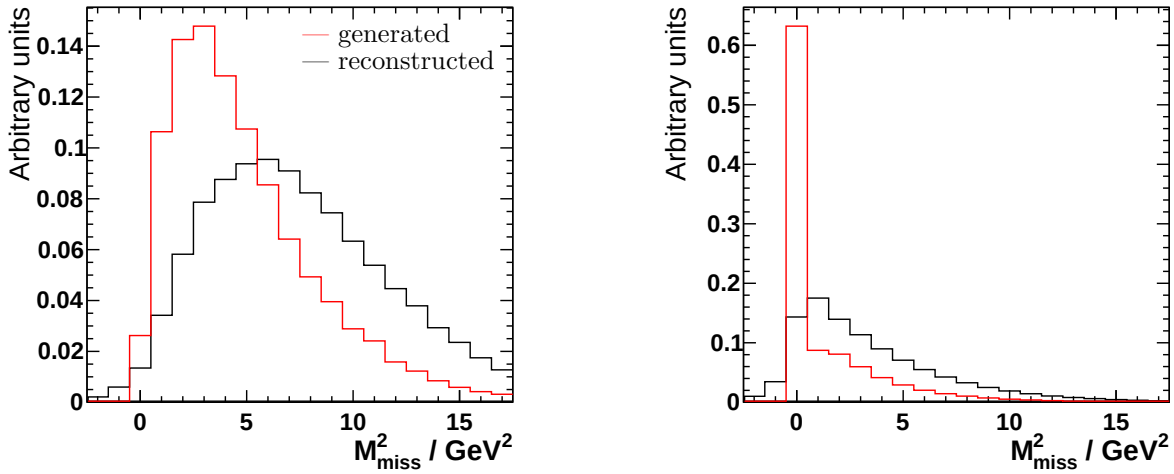


Figure 5.9: Distributions of generated (red) and reconstructed (black) m^2_{miss} for $B \to X\tau\nu$ (left) and $B \to X\ell\nu$ (right). The reconstructed distributions show a reduced resolution and increased mean value due to lost tracks and the detector resolution. For $B \to X\ell\nu$ decays, values of $m^2_{\text{miss}} > 0\,\text{GeV}^2$ at generator level stem from K_L or neutrinos from semileptonic decays of the charmed mesons which are assumed to be lost. All distributions have the same normalisation for a better comparison.

Hadronic mass Another interesting observable is the mass spectrum of the hadronic system X on the signal side m_X (Fig. 5.6, lower right). It shows the same distribution for signal decays and decays of the normalisation mode. Additionally, it exhibits peaks at the mass of the charmed meson and also of lighter mesons. The latter are mainly due to events where other final state particles are lost. Purely hadronic decays tend to masses closer to the B mass. These decays enter the analysis sample either due to a fake lepton candidate or due to a lepton produced later in the decay chain, usually a semileptonic charm decay. As for m^2_{miss}, the MC description of m_X depends on the modelling of the physics in the decay chains. It also shows

less discriminating power between the signal decay and backgrounds (especially between the signal and the normalisation mode) compared to m^2_{miss}. Thus the latter is used in the signal extraction.

Conclusions In an inclusive analysis as the one presented here, there are only a few observables which contain independent information. As the $B \to X\tau\nu$ decay produces three distinct classes of particles, the lepton, the hadronic system and the undetected neutrinos, the mentioned variables ($|\vec{p}^*_\ell|$, m_X and m^2_{miss}) contain almost all relevant information. As the signal lepton itself is a secondary τ daughter, its direction does not correspond to that of the prompt τ. Thus, angular distributions are not useful in this analysis as they do not offer any discriminating power. In this analysis, the lepton momentum and squared missing mass are chosen as signal extraction variables. The hadronic mass will be referred to later as an independent observable to confirm the data-MC agreement (Sec. 5.6).

5.3 Data-MC agreement studies

This analysis bases on a reliable modelling of the data by the MC simulation which is, therefore, investigated in this section. Discrepancies in the modelling can for instance arise from the physics models at generator level, which have impact on the simulated momentum and energy spectra or lead to different reconstruction efficiencies in data and MC, which would lead to selection biases. Two aspects should be emphasised again. First, this analysis is based on a large sample of inclusive semileptonic B decays and, in contrast to (exclusive) analyses with smaller sample sizes, even a slight mis-modelling in the MC can be statistically significant and clearly visible. The second point is the limited set of observables available to extract the signal due to the *inclusive* selection. As pointed out in the previous section, there are only three useful observables and thus their modelling is vital to this analysis.

The following sections discuss the data modelling with a focus on the signal extraction variables $|\vec{p}^*_\ell|$ and m^2_{miss}. Distributions of data and MC are shown in a signal-free side band. For a better comparison and to emphasise statistically significant deviations between data and MC, the residuals $(N_{\text{data}} - N_{\text{MC}})/\Delta$ are also shown in the figures, where (N_{data} and N_{MC} are the number of data and MC events, respectively. $\Delta = \sqrt{N_{\text{data}}}$ is the statistical uncertainty of the data.

5.3.1 Lepton momentum modelling

While the momentum spectrum of prompt leptons from semileptonic B decays (in the rest frame of the decaying B) is sensitive to the physics of the decay, this is much less the case for signal and other secondary leptons. Their spectra are diluted because of the significant momentum of the mother particle.

The momentum spectra of electrons and muons in data and MC after applying all MC corrections (cf. Sec. 3.3) are shown in Fig. 5.10. The disparate shapes of the electron and muon spectra are due to the different selection efficiencies and the larger probability of mis-identification for muons. Electrons are much more efficiently reconstructed at low momenta than muons and muons suffer from a high rate of fake lepton candidates. At large momenta, where $B \to X\ell\nu$ decays are expected, the spectra are similar.

The MC distributions show deviations from the data in all regions which is, to some extend, expected because of the following reasons. The secondary and fake lepton branching fractions

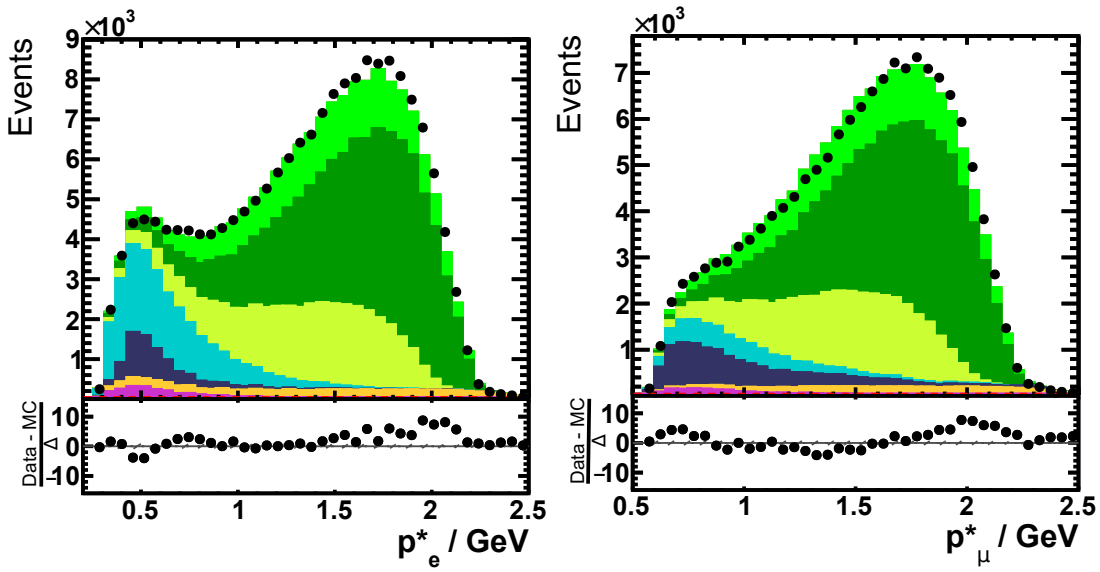


Figure 5.10: Comparison of data (black points) and MC of the lepton momentum spectrum of electrons (left) and muons (right). The spectrum shows events in the signal free $m^2_{\text{miss}} \leq 3\,\text{GeV}^2$ region. The colours have the same meaning as in Fig. 5.6. *Note:* At this point, only statistical uncertainties on data and MC are taken into account which are too small to be visible. In the residual plot, the MC statistics are shown as a grey hatched band (also barely visible). The MC is normalised to the number of events in data.

and efficiencies are not well known and likely have a wrong normalisation in the MC, even though the MC corrections have been applied. However, the normalisation of this component will be estimated in data as it is a free parameter in the signal extraction fit (cf. Sec. 5.4).

Of greater implication is the shape of the $B \to X\ell\nu$ decays. At large momenta, an excess of data is visible while there is a deficit of data at intermediate momenta. Considering the composition of the lepton momentum spectrum, large momenta are populated by $B \to D^{(*)}\ell\nu$ and $B \to X_u\ell\nu$ decays, however, the latter is a very small component. Thus, the deviation hints at a harder lepton spectrum which could in principle be addressed by varying the CLN form-factor parameters of the $B \to D^*\ell\nu$ decay. However, these parameters are highly constrained by measurements and the uncertainty on the parametrisation itself is stated to be $\lesssim$2% [32].

Other processes producing high momentum leptons are $q\bar{q}$ continuum events. They might be relevant at large momenta because high-momentum leptons could be produced and stem, in particular, from semileptonic charm decays in $c\bar{c}$ events. Although continuum is already highly suppressed in this analysis (the lowermost component in Fig. 5.10) it has a poorly known shape and could be underestimated. *Off-resonant* $q\bar{q}$ data was taken ∼60 MeV below the $\Upsilon(4S)$ mass and thus below the $B\overline{B}$ production threshold. This data can in principle be used instead of relying on the $q\bar{q}$ MC simulation. Unfortunately, the hadronic tagging algorithm is not capable of processing continuum data as it assumes $B\overline{B}$ production at the $\Upsilon(4S)$ energy which affects momenta and energies of tracks and, even worse, the particle identification. Therefore, the impact of continuum is estimated differently.

To check if continuum events are underestimated at large momenta, a tight selection on the continuum suppression classifier $\ln(\text{ANN}_{q\bar{q}}) > -1$ is applied. The residuals at high $|\vec{p}^{\,*}_\ell|$ are barely affected as plotted in Fig. 5.11. The recent inclusive $B \to X_s\gamma$ analysis by Belle uses a high momentum lepton as a flavour tag and estimates the $q\bar{q}$ background with the

continuum data sample [79–81]. The comparison (see Fig. 5.11) shows that the inclusive tag lepton momentum spectrum exhibits similar residuals. Hence, significant underestimation of high momentum leptons from $q\bar{q}$ continuum events is likely ruled out because the data excess is also visible in inclusive analyses using continuum data.

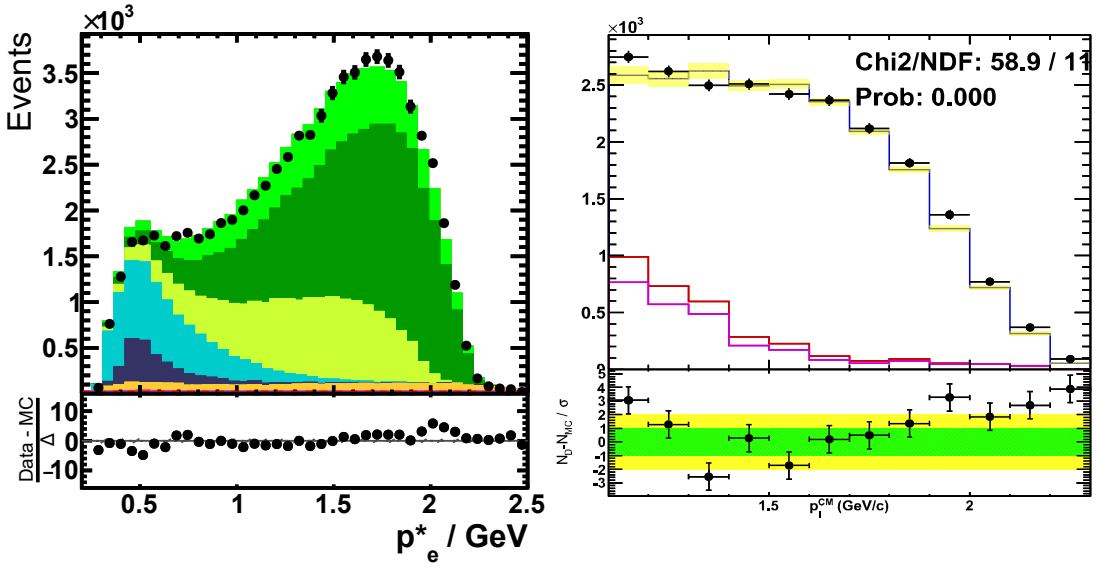


Figure 5.11: Impact of leptons from continuum background at large lepton momenta. *Left:* A tight selection on the continuum suppression classifier output reduces nearly all $q\bar{q}$ events, but the excess in data remains. *Right:* A similar excess is observed in an inclusive analysis by Belle [79–81] using a tag lepton and continuum data. Shown is the electron momentum spectrum in the centre-of-mass frame in data (black dots), $B \to X\ell\nu$ (blue line), secondaries and fake leptons (red line) and continuum data (purple line). The former two are taken from MC and floated to match the data. Taken from [81].

As pointed out, exclusive analyses have precisely measured the $B \to D^*\ell\nu$ branching fraction the CLN shape parameters. Inclusive analyses of $B \to X\ell\nu$ can obtain the lepton momentum spectrum by subtracting simulated backgrounds, i. e. these analyses are not sensitive to such modelling issues. However, analyses of $B \to X\tau\nu$ rely on the modelling of the semileptonic B decays due to the high background level. A recent analysis of inclusive charmed and charmless semileptonic B decays by *BABAR* models the $B \to X_c\ell\nu$ as a sum of exclusive modes and finds a harder lepton spectrum than expected [82]. The analysis yields a considerably larger $B \to D^*\ell\nu$ branching fraction which is $\sim 3.5\sigma$ above the world average value.

The question arises if the selection of a D^* affects the leptons spectrum or could underestimate the semileptonic B decay rate at large momenta compared to inclusive analyses. Possible contributions in this kinematic region could stem from non-resonant $B \to D\pi\ell\nu$ decays. These decays are modelled exclusively and separately from $B \to D^*\ell\nu$ in the MC simulation, but could in principle interfere with the resonant D^* production. Furthermore, in analyses measuring $B \to D^*\ell\nu$ decays, D^* mesons are reconstructed in $D^* \to D\pi, D\gamma$. The neutral D^* is $\sim 2\,\mathrm{MeV}$ too light to decay through $D^{*0} \to D^+\pi^-$ on-shell. Potential off-shell decays are expected to give only a small contribution to the decay width but are additionally suppressed by the narrow selection on the D^* mass in exclusive analyses. Hints for off-shell decays or unconsidered interferences can e. g. be found in the spectrum of the mass difference $m(D^{(*)}\pi) - m(D^{(*)})$ (see Fig. 5.12) of the $B \to D^{**}\ell\nu$ measurement by *BABAR* [83]. An excess of data at low $D\pi$ mass differences is observable which is larger for charged than neutral B decays and thus could match

the signature of unconsidered off-shell $D^{*0} \to D^+\pi^-$ decays.

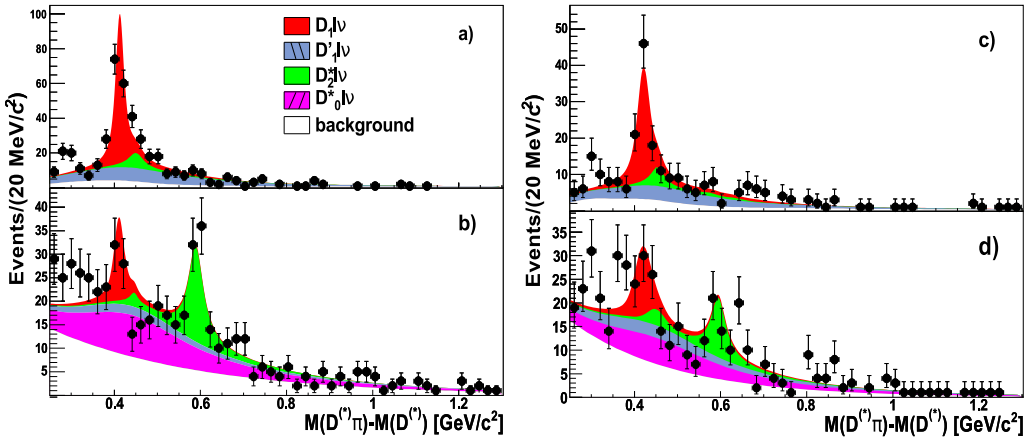


Figure 5.12: The $m(D^*\pi) - m(D^*)$ (a, c) and $m(D\pi) - m(D)$ (b, d) mass difference spectra for charged (a, b) and neutral (c, d) B meson decays of the $B \to D^{**}\ell\nu$ analysis by *BABAR* [83]. Clearly visible is an excess of data events at low mass differences, especially for $m(D\pi) - m(D)$. The excess is less significant for decays of neutral B mesons, however, also the data sample is smaller.

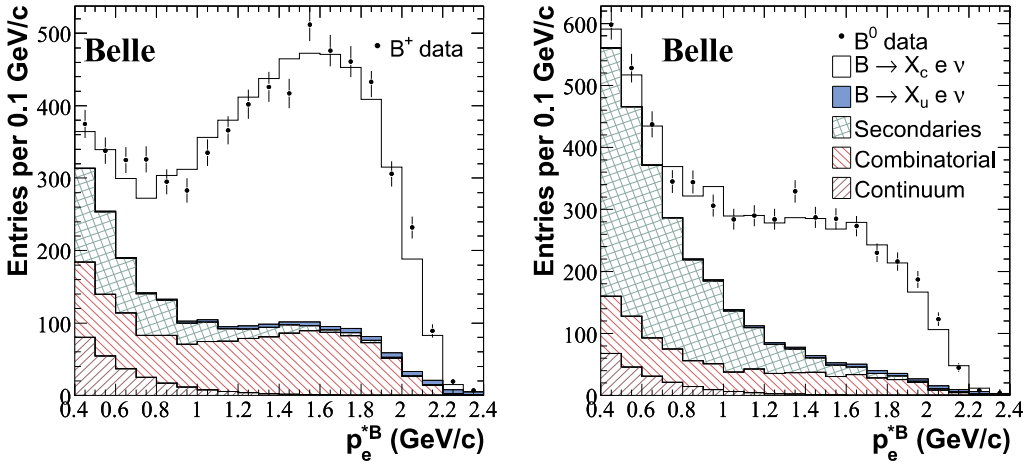


Figure 5.13: $|\vec{p}^*_\ell|$ spectrum of the inclusive $B \to X_c\ell\nu$ analysis at Belle [84], shown separately for charged (left) and neutral (right) B_{tag}. The analysis was performed with a smaller data set of 140 fb^{-1} (∼20% of the entire Belle data set) and uses a superseded B tagging method. Nonetheless, an underestimation of the data at high momenta is already visible.

Another example is the inclusive analysis of $B \to X_c\ell\nu$ by Belle which uses a much smaller data set of 140 fb^{-1} [84]. The relevant lepton momentum spectrum is shown in Fig. 5.13. Already in this statistically less significant analysis, an excess of the data at high momenta and a deficit at intermediate momenta is visible.

To conclude, an excess of data at large lepton momenta has been seen before in inclusive analyses. It is possible that interferences and/or off-shell decays contribute in this momentum region which are not considered in the sum of exclusive modes modelling in the MC. This could lead to an underestimation of the $B \to D^*\ell\nu$ branching fraction in exclusive measurements and thus a softer lepton momentum spectrum of $B \to X\ell\nu$ decays in the simulation. The lepton spectrum modelling is picked up again in Sec. 5.6, after the fitting procedure is introduced and

tested in a signal-free side-band.

5.3.2 The modelling of the square missing mass

Due to its three undetectable neutrinos, the signal lies in the tail region of the m^2_{miss} distribution. In contrast, all backgrounds have a more or less prominent peak at vanishing missing mass. However, the background distributions have long tails and thus populate the signal region, too. These long tails are, in any case, due to lost particles that cause an energy-momentum imbalance in the event. Lost particles can either stem from the production of undetectable particles, for example an additional neutrino from a secondary semileptonic decay (e. g. $D \to K\ell\nu$) or long-lived neutral kaons. Even though the Belle detector is equipped with the KLM to detect K_L, it cannot detect their full energy and only estimates their direction. Other particles are lost in the detector, either in the direction of the beam-pipe, by interactions with dead material in the detector or they are missed by the event reconstruction algorithms. Furthermore, wrongly assigned particle hypotheses or the modelling of the final state multiplicity have an impact on the squared missing mass.

The modelling of m^2_{miss} is subject to special attention in this analysis. Fig. 5.14 shows the data and MC expectation of the m^2_{miss} distribution for events fulfilling $|\vec{p}^{\,*}_\ell| \geq 1.2\,\text{GeV}$ or $m_X \geq 2.2\,\text{GeV}$ to exclude the signal region. Clearly visible is a jump in the residuals at $m^2_{\text{miss}} = 0\,\text{GeV}^2$. The MC deficit at $m^2_{\text{miss}} < 0\,\text{GeV}^2$ is subject to an extensive study which is outlined below.

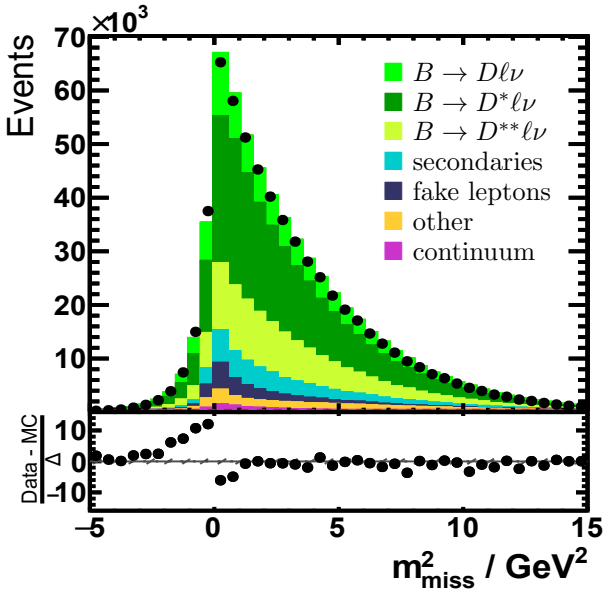


Figure 5.14: Distribution of the squared missing mass, m^2_{miss}, in the signal-depleted $|\vec{p}^{\,*}_\ell| \geq 1.2\,\text{GeV}$ and $m_X \geq 2.2\,\text{GeV}$ region. Clearly visible is the excess of data in the negative tail and the sharp step in the residual plot at $m^2_{\text{miss}} = 0\,\text{GeV}^2$. Only statistical uncertainties are taken into account in the residual plot. The statistical uncertainty of the MC is indicated as a grey hatched band. The MC is normalised to the number of events in data.

Events with $m^2_{\text{miss}} < 0\,\text{GeV}^2$ It is clear that the region of reconstructed $m^2_{\text{miss}} < 0\,\text{GeV}^2$ is not due to a physical process. To understand this, the definition of the squared missing mass is

considered

$$m^2_{\text{miss}} = \left| p^{\mu}_{e^+e^-} - p^{\mu}_{\text{visible}} \right|^2 \tag{5.5}$$

where all four-momenta of detected particles are summed up to p^{μ}_{visible}. To better see the origin of negative m^2_{miss} values, consider that $\vec{p}'_{\text{miss}} = -\vec{p}'_{\text{visible}}$ in the centre-of-mass frame (denoted with a prime) an thus

$$\begin{aligned} m^2_{\text{miss}} &= (E'_{\text{miss}} - |\vec{p}'_{\text{miss}}|)(E'_{\text{miss}} + |\vec{p}'_{\text{miss}}|) \\ &= (E'_{\text{miss}} - |\vec{p}'_{\text{visible}}|)(E'_{\text{miss}} + |\vec{p}'_{\text{visible}}|) \\ &= (m_{e^+e^-} - E'_{\text{visible}} - |\vec{p}'_{\text{visible}}|)(m_{e^+e^-} - E'_{\text{visible}} + |\vec{p}'_{\text{visible}}|) \\ &= m^2_{e^+e^-} + m^2_{\text{visible}} - 2E'_{\text{visible}} m_{e^+e^-}, \end{aligned} \tag{5.6}$$

where $m_{e^+e^-} = \sqrt{s}$ is the collision energy. Hence, to achieve $m^2_{\text{miss}} < 0\,\text{GeV}^2$, either the measured missing momentum is larger than the missing energy or the visible energy is overestimated compared to the visible mass.

As long as nothing or only a single particle p is lost, m^2_{miss} is independent of the specific decay in the event as either $m^2_{\text{miss}} = 0\,\text{GeV}^2$ or $m^2_{\text{miss}} = m^2_p$. Therefore, the $m^2_{\text{miss}} \approx 0\,\text{GeV}$ peak region is rather independent of the actual decay and kinematic modelling. However, the finite detector resolution results in a broadening of the peak and negative m^2_{miss} values occur.

To identify possible origins of $m^2_{\text{miss}} < 0\,\text{GeV}^2$, one can consider the following three general steps of the MC simulation. In the first step, the decay chains are simulated. In the second step, the particles are placed in the detector model and are reconstructed. And in the last step, particles which do not stem from the $\Upsilon(4S)$ decay are added. In the first step, m^2_{miss} equals zero because of four-momentum conservation. In the second step, where the particle candidates are reconstructed, the momentum, energy and mass hypothesis are affected by the detector resolution. Furthermore, particles can be missed or duplicated by the reconstruction algorithm. This broadens the m^2_{miss} peak at zero and also shift events into to the negative m^2_{miss} region. In the last step, background particles from e. g. beam-background are added. This reduces the missing energy and affects the missing momentum randomly and can thus enhance the tail towards negative m^2_{miss}. Such backgrounds are typically more likely in data than in the simulation.

This three-step picture provides the following insights: negative squared missing mass values either stem from the detector resolution or non-collision particles. Of course, events containing several neutrinos or low-momentum tracks, that are not found by the tracking algorithm, are less likely to contribute at negative m^2_{miss}. In contrast, low multiplicity processes like $B \to D\ell\nu$ have a higher probability to end up at negative m^2_{miss} because they are more likely to be reconstructed correctly and thus would have a narrow distribution around $m^2_{\text{miss}} = 0\,\text{GeV}$, which then is broadened by the previously discussed steps two and three.

At higher values of m^2_{miss}, the physics modelling becomes more important because $m^2_{\text{miss}} = E'^2_{\text{miss}} - |\vec{p}'^2_{\text{miss}}| \approx E^2_{\text{miss}} \approx (m_b - E'_{\text{visible}})^2$. The distribution of the visible energy is relevant to estimate the kinematics of the lost particles in region of large m^2_{miss} values. Thus, beside the detector simulation, the physics modelling is important. However, as can be seen in Fig. 5.14, the tail region is modelled sufficiently well.

Track multiplicity There are several observables beside m^2_{miss} which might help to understand the origin of the poor modelling of the negative m^2_{miss} tail. Some of them describe the entire event (e. g. multiplicities) and others only the tag or signal side, e. g. the B_{tag} momentum and energy. An interesting observable is the total number of reconstructed tracks in the event. A

wrong modelling of the number of tracks can indeed result in a poor m^2_{miss} modelling. To further discuss this, m^2_{miss} is plotted in bins of the track multiplicity as shown in Fig. 5.15. Events with less than six reconstructed tracks are mostly located at $m^2_{\text{miss}} > 0\,\text{GeV}^2$ as there are usually some tracks lost. However, these events are only a small fraction of the entire data sample. The peak at zero becomes more narrow the more tracks are in an event, i. e. the m^2_{miss} resolution improves as the relative impact of a single lost track reduces. However, the excess in data at $m^2_{\text{miss}} < 0\,\text{GeV}^2$ is still present, quite independent of the multiplicity, and can thus not explain the mis-modelling of the track multiplicity cannot cause the residuals.

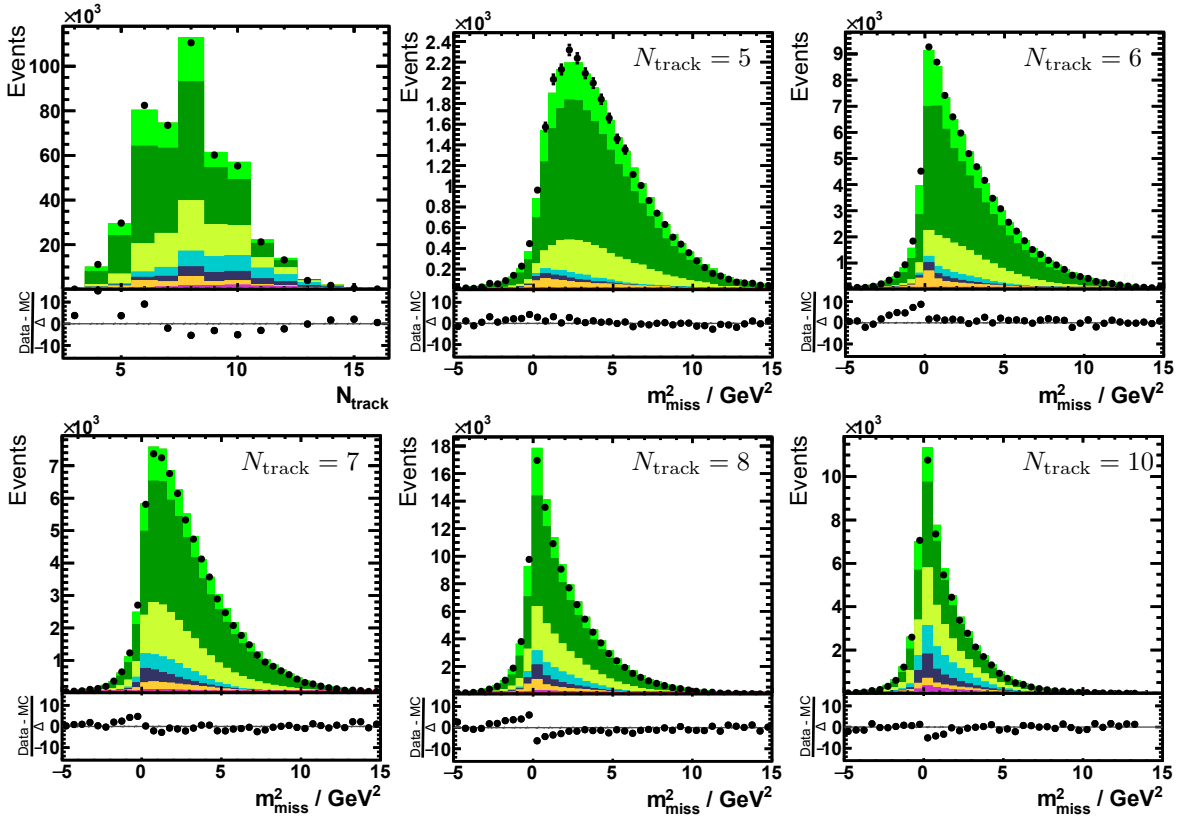


Figure 5.15: Number of reconstructed tracks in the event, N_{track} (top left), and m^2_{miss} in bins of N_{track}. Signal events are rejected by selecting only events with $|\vec{p}^{\,*}_\ell| \geq 1.2\,\text{GeV}$ and $m_X \geq 2.2\,\text{GeV}$. The shape changes with the number of tracks, but the excess in data and the step in the residual plot is always present. The latter is more pronounced in multiplicity bins of high statistical significance. The MC is normalised to the total number of events in data. The colours have the same meaning as in Fig. 5.14.

A more comprehensive picture is given by the number of lost tracks, i. e. considering the generated and reconstructed tracks. Defining the number of true tracks, N_{true}, is not trivial because the tracks and photons generated by EvtGen are not the ones at the end of the decay chain. By interactions with the detector material, pair-production and bremsstrahlung occur, which are simulated in the Gsim detector simulation (cf. Sec. 3.2). Also, neutral kaon decays are simulated by Gsim, of which K_L likely decay outside, but the K_S inside the detector. Thus, N_{true} only counts charged $\Upsilon(4S)$ final states before the Gsim detector simulation. K_L are assumed to be lost, but the two charged pions from $K_S \to \pi^+\pi^-$ decays are taken into account.

Fig. 5.16 shows the distribution of lost tracks $N_{\text{lost}} = N_{\text{true}} - N_{\text{reco}}$. Negative N_{lost} values are possible because of e. g. $\gamma \to e^+e^-$ photon conversion or duplicated tracks, but only account

for ∼2% of the sample. To assess whether a wrong modelling of the number of lost tracks is related to the m^2_{miss} residuals, a fit is performed to estimate N_{lost} in data. As the information of lost tracks is not available in data, m^2_{miss} is used to extract N_{lost}. In the fit, which is a binned likelihood fit, normalisations of three different samples are floating. The three samples contain events with $N_{\text{lost}} < 1$, $1 \leq N_{\text{lost}} < 2$ and $2 \leq N_{\text{lost}}$, respectively. The two former samples also have to fulfil the requirement that the true $m^2_{\text{miss}} = 0\,\text{GeV}^2$, i. e. only events without K_L or multiple neutrinos are considered to improve the resolution. In Fig. 5.16, the post-fit distribution is shown. The peak region of m^2_{miss} is still not well modelled, but from the shape of the residuals it seems the component without lost tracks is too narrow to describe that data accurately. To take this into account, the component without lost tracks is smeared by a Gaussian distribution. This indeed improves the modelling after the fit, however, the residuals at negative m^2_{miss} remain. Even though m^2_{miss} is correlated with N_{lost}, this test shows that events without lost tracks cannot provide a sufficient explanation of the mis-modelling.

Impact of the collision energy Starting wit Eq. 5.6, there are three basic event quantities that can an have impact on the m^2_{miss} modelling: the centre-of-mass energy, E_{visible} and m_{visible}. The centre-of-mass energy or invariant mass of the colliding e^+e^- is given by the energies of the electron E_{e^-} and positron E_{e^+}, as well as the fixed crossing angle $\theta = 0.022\,\text{rad}$ under which they collide. Their four-momenta are

$$p^\mu_{e^-} = (E_{e^-}, E_{e^-}\sin\theta, 0, E_{e^-}\cos\theta)^T \text{ and } p^\mu_{e^+} = (E_{e^+}, 0, 0, -E_{e^+})^T \tag{5.7}$$

and $\sqrt{s} = m_{e^+e^-} = \sqrt{(p^\mu_{e^-} + p^\mu_{e^+})^2}$. Distributions of the energies of the electron and positron beam and the centre-of-mass energy are shown in Fig. 5.17. The agreement of data and MC is rather poor. However, the width of these distributions has to be considered. The root-mean-square (RMS) is <1 MeV which is far too small to affect m^2_{miss} in a significant way and can be ruled out at this point regarding the mis-modelling at $m^2_{\text{miss}} < 0\,\text{GeV}^2$.

Improving the particle selection The visible energy, momentum, mass and the squared missing mass are shown in Fig. 5.18 in the centre-of-mass frame for a loose and an optimised selection. Before using an optimised track and photon selection, a slope in the residuals of all the variables is visible. As these variables depend on reconstructed tracks and photons, their selection is revised. Successively, photon and track selection criteria are tested and the sum of squared residuals of m^2_{miss} is minimised. In general, there is no standard track selection at Belle because it is always analysis specific. However, a loose selection of tracks is no momentum requirements and $(\text{d}r, \text{d}z) < (0.5, 2.0)\,\text{cm}$ and for photons $E_\gamma > 150, 50, 100\,\text{MeV}$ in the forward ($\theta < 33°$), barrel ($33° < \theta < 127°$) and backward ($127° < \theta$) region, respectively. The optimised selection is given in Tab. 5.1 and is used throughout the analysis. It includes a tighter photon selection which removes low energetic photons with $E_\gamma < 150\,\text{MeV}$. Tracks need to have a minimum transversal momentum of 100 MeV and are required to originate closer from the IP ($(\text{d}r, \text{d}z) < (0.5, 1.5)\,\text{cm}$). This removes wrongly reconstructed tracks which populate the low momentum end. The tight selection is only available for signal side photons and tracks, because the hadronic tagging algorithm comes with its own particle identification and already built up the B_{tag} out of its track candidate list. The new selection reduces the average number of photons on the signal side in data (MC) from $\bar{N}_\gamma = 4.4 \to 2.2$ ($4.3 \to 2.2$) and average number of tracks $\bar{N}_{\text{track}} = 3.3 \to 3.2$ ($3.4 \to 3.3$). The selection efficiencies of signal and background decays are only negligibly reduced because of the following reasons: The signal lepton momentum requirement is already higher

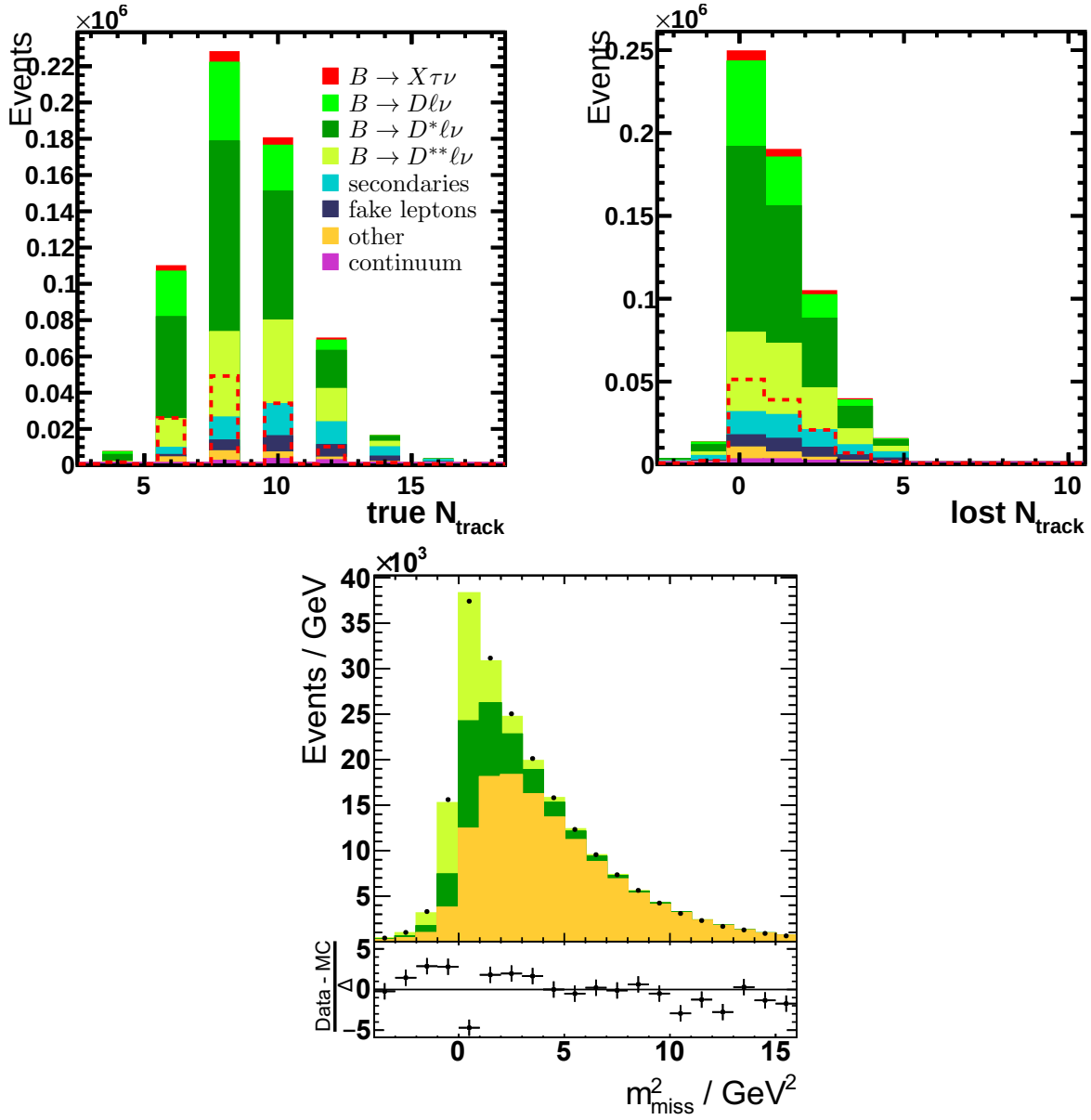


Figure 5.16: Distributions of the number of true (top left) and lost (top right) tracks. The lower plot shows m^2_{miss} after fitting the normalisations of the samples with zero (light green), one (green) and several (orange) lost tracks. Still, a poor modelling of the peak region is visible and a further smearing with a Gaussian is not able to remove the residuals in the negative m^2_{miss} range.

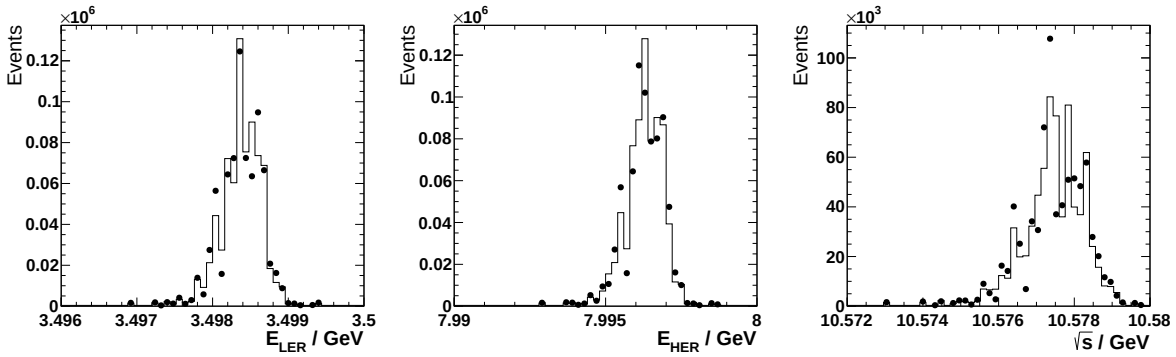


Figure 5.17: Comparing data (black dots) and MC (solid line) for beam observables. E_{LER} and E_{HER} are the energies of the positrons in the low- and the electrons in the high-energy-ring of the KEKB accelerator, respectively. The statistical uncertainties are too small to be visible. The MC is normalised to the number of events in data.

than the value in the new selection and thus does not affect the selection efficiency. Additionally, well reconstructed leptons almost always originate from the IP and the new requirement on $(\mathrm{d}r, \mathrm{d}z)$ is only slightly tighter than the standard selection.

The overall modelling of the data improves due to the new selection, visible in the m^2_{miss}, E'_{visible} and m_{visible} distributions (cf. Fig. 5.18). In contrast, the visible momentum is still poorly modelled and shows strong residuals which might be a hint towards the m^2_{miss} mis-modelling. In the centre-of-mass frame $|\vec{p}'_{\mathrm{miss}}| = |\vec{p}'_{\mathrm{visible}}|$ and, as stated in Eq. 5.6, too large missing momentum can cause negative missing mass. However, investigating the visible momentum by splitting the MC in two components, representing $m^2_{\mathrm{miss}} < 0\,\mathrm{GeV}^2$ and $m^2_{\mathrm{miss}} \geq 0\,\mathrm{GeV}^2$, respectively, reveals that the $|\vec{p}'_{\mathrm{visible}}|$ residuals are, at least not completely, those of m^2_{miss} as shown in Fig. 5.19.

On the modelling of track momenta The visible energy and mass seem to be sufficiently well modelled but the visible momentum requires further studies. In MC, the visible momentum distribution is slightly shifted to smaller momenta. To investigate the total momentum, the impact of the momentum of single tracks is studied. Therefore, the total momentum of all tracks which fall into a certain momentum bin $p + \mathrm{d}p$ (i. e. the summed momenta p in a momentum range $p+\mathrm{d}p$) is considered and shown in Fig. 5.20 (left). The plot illustrates that low momentum tracks in the MC carry more momentum than in data and vice versa for large momenta. A possible reason for this is a mis-modelling of the average track momenta already at generator level, which, however, cannot explain the m^2_{miss} residuals alone as it should have been visible in different kinematic observables. A naive solution is to shift the momenta in MC by a tiny momentum-dependent amount to match the distribution of summed track momenta. However, although the m^2_{miss} residuals are nearly gone after applying the shift (cf. Fig. 5.20, right), this procedure leads to a poor modelling of the mass spectra of e. g. $K^+\pi^-$ or $\pi^+\pi^-$ pairs, i. e. such a shift would have been discovered in earlier analyses. Nevertheless, the test indicates that the problem is likely related to the tracks.

To summarise, track momenta are on average larger in data than in MC and contribute more to the total momentum and could contribute to the mis-modelling of m^2_{miss}. However, the underlying effect is complicated and could be due to a wrong track or momentum reconstruction efficiency in the MC, but the track finding is very efficient and has a small uncertainty of 0.35% per track. Also, the momentum reconstruction cannot be significantly mis-modelled as it is easily verified in the mass spectra. This implies that the momentum distribution is modelled

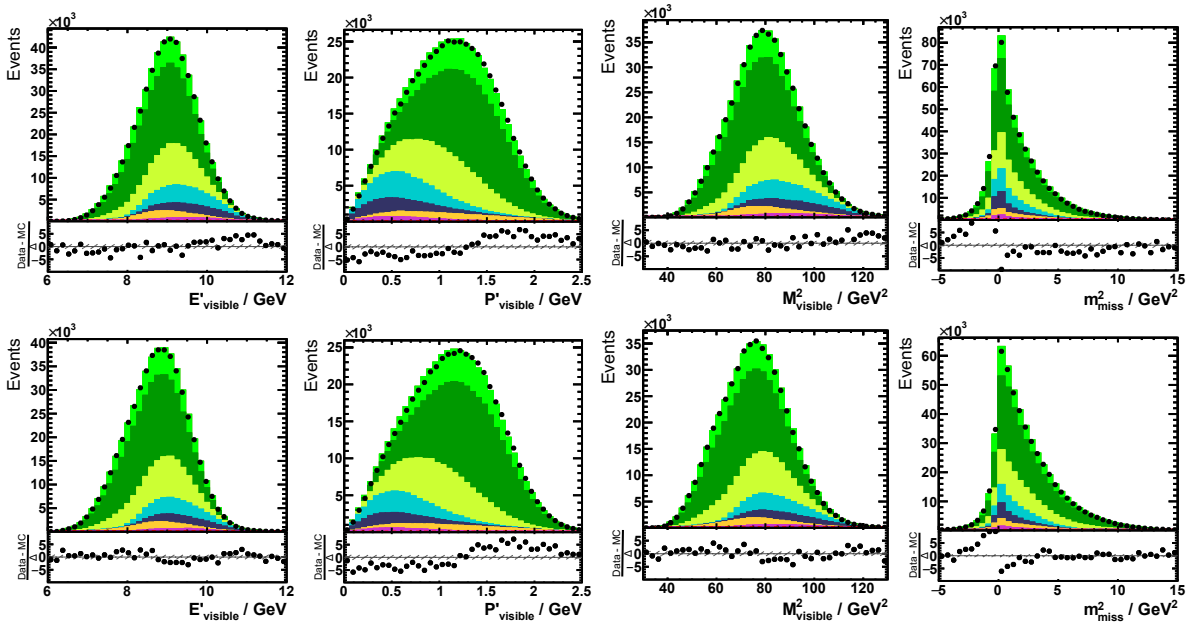


Figure 5.18: The loose standard (top row) and the improved, tighter particle selection (bottom row) for the total visible energy, momentum, squared mass and squared missing mass. Energy and momentum are shown in the centre-of-mass frame. The data-MC agreement of the total energy and mass is improved by the enhanced selection, however, the total momentum and m^2_{miss} still show a mis-modelling. The MC is normalised to the number of events in data. The colours have the same meaning as in Fig. 5.14

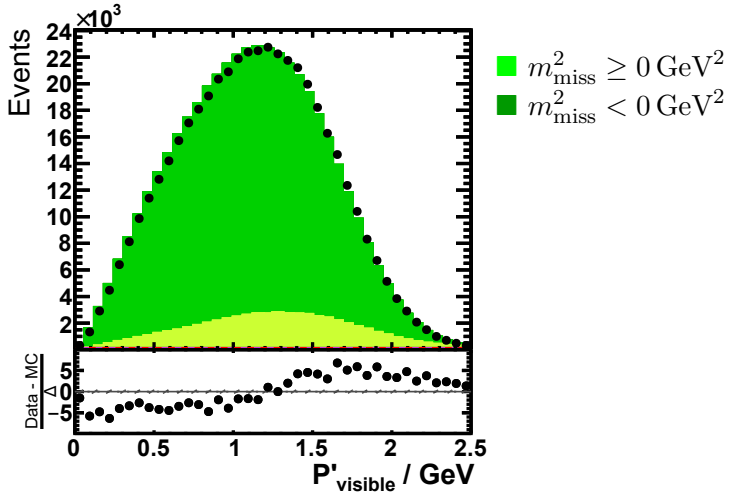


Figure 5.19: Shape comparison for the visible momentum of events with negative (light green) and positive (dark green) squared missing mass. The distribution of $m^2_{\text{miss}} < 0\,\text{GeV}^2$ events does not follow the residuals of the visible momentum. The MC is normalised to the number of events in data.

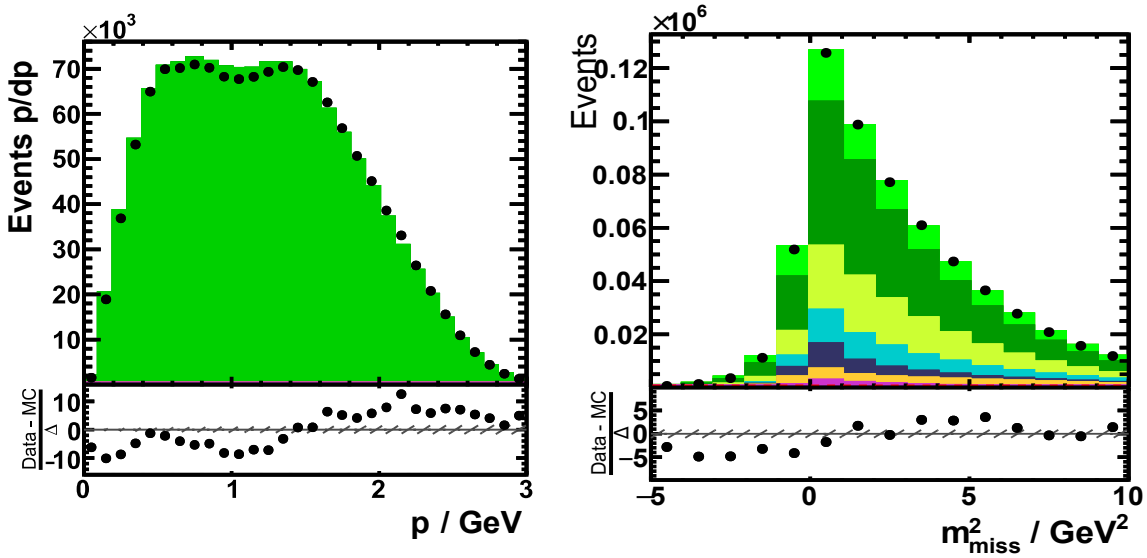


Figure 5.20: *Left:* The total momentum of tracks carrying a certain momentum, where an excess of data for large momenta is visible. Correcting this observation with a momentum shift improves the m^2_{miss} modelling (right) at the expense of distorted mass spectra. The MC is normalised to the number of events in data.

imperfectly earlier in the MC production chain, namely at generator level. Even if the physics modelling cannot cause $m^2_{\text{miss}} < 0\,\text{GeV}^2$ events, it could produce more slow tracks on average in MC than in data which more often are lost.

Photon energies In addition to the tracks, individual photons from the beam or mis-identified FSR could cause a wrong modelling of the missing mass if they pass the energy threshold of $E_\gamma > 150\,\text{MeV}$. An overall shift in the photon energies E_γ would affect the mass peak of e. g. π^0 similar to the discussion of shifted track momenta. Thus, such a shift could only affect certain continuous contributions, e. g. beam-backgrounds. First, the distributions of the total E'_γ and $\vec{p}'_\gamma$ in the centre-of-mass-frame are examined which show a reasonable modelling (cf. Fig. 5.21). To further study the impact of photons, the distribution of the sum of photons energies from all events in a certain energy bin $E_\gamma + \mathrm{d}E_\gamma$ is obtained, which is also shown in Fig. 5.21. In contrast to the analogous distribution for the track momenta, no large residuals are observed, which is likely due to the rather high photon threshold in this analysis. As a conclusion, photons are thus not further considered as problematic at this point.

Semi-inclusive study An interesting fact is that exclusive analyses quite commonly use m^2_{miss} without noticing residuals in the negative tail region. However, an imperfect modelling is indeed observed, e. g. a too narrow m^2_{miss} MC peak in the recent analysis of $B \to D\ell\nu$ decays by Belle [85]. Thus, for comparison, a semi-inclusive sample of $B \to DX\ell\nu$ is selected. For this purpose, $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ decays are reconstructed from charged kaon and pion combinations which have to fall into the mass windows of $1.855\,\text{GeV} < m_{D^0} < 1.895\,\text{GeV}$ and $1.85\,\text{GeV} < m_{D^+} < 1.89\,\text{GeV}$, respectively. An event is required to have either a D^0 or a D^+ candidate. The selection efficiency is small due to the rather low branching fractions of $\mathcal{B}\left(D^0 \to K^-\pi^+\right) = (3.93 \pm 0.04)\%$ and $\mathcal{B}\left(D^+ \to K^-\pi^+\pi^+\right) = (9.5 \pm 0.3)\%$ and the selection is not optimised for this test. The $B \to DX\ell\nu$ reconstruction selects 2.8% of the $B \to X\ell\nu$

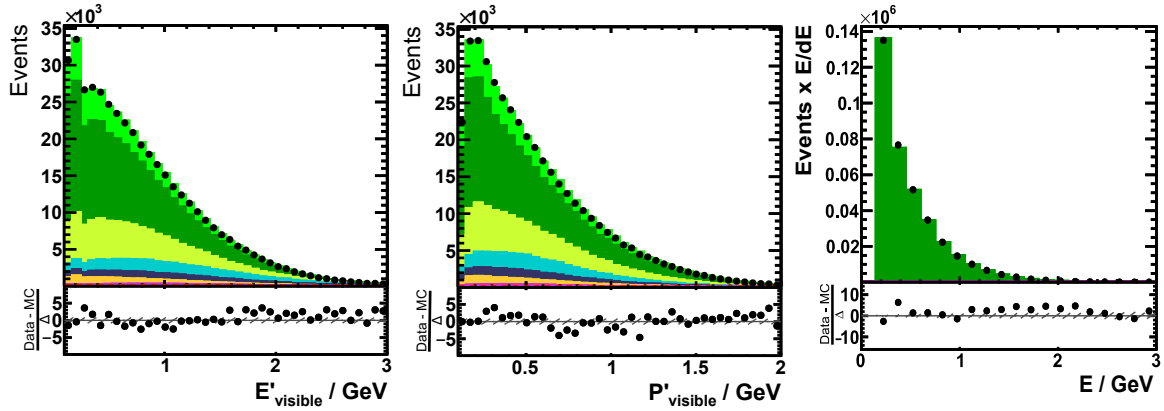


Figure 5.21: Total visible energy (left) and momentum (centre) of photons as well as the sum of photon energies E_γ from all events in a certain energy bin $E_\gamma + \mathrm{d}E_\gamma$ (right). The modelling is sufficiently well and therefore photons are not further considered regarding the m^2_{miss} mis-modelling. The MC is normalised to the number of events in data.

sample. The m^2_{miss} distribution for the $B \to DX\ell\nu$ sample is shown in Fig. 5.22. The residuals are much smaller which is expected because the effect is much less statistically significant in such a small sample compared to the inclusive sample. Nonetheless, there is still a small jump in the residuals, implying that the modelling of $m^2_{\mathrm{miss}} < 0\,\mathrm{GeV}^2$ is still insufficient.

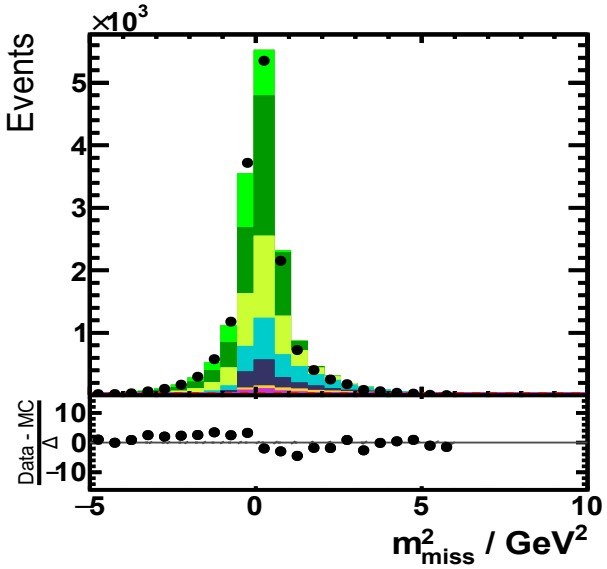


Figure 5.22: Distribution of squared missing mass for the semi-inclusive sample of $B \to DX\ell\nu$ decays. The MC is normalised to the number of events in data.

Further studies In addition to these studies, further tests have been performed which are outlined below. Many of these tests are related to the tag side because in principle a selection bias of the B_{tag} could cause the poor modelling. If there is a selection bias due to the hadronic tagging, it is not expected to be the same in all B_{tag} reconstruction modes. Thus, the m^2_{miss} distribution is studied in bins of the B_{tag} reconstruction mode (see Fig. 5.23). Most reconstruction modes only contribute with a small number of events to the full tagged sample and thus do not show

significant residuals, but the few modes which contribute significantly show imperfect modelling which hints at a reconstruction mode independent issue.

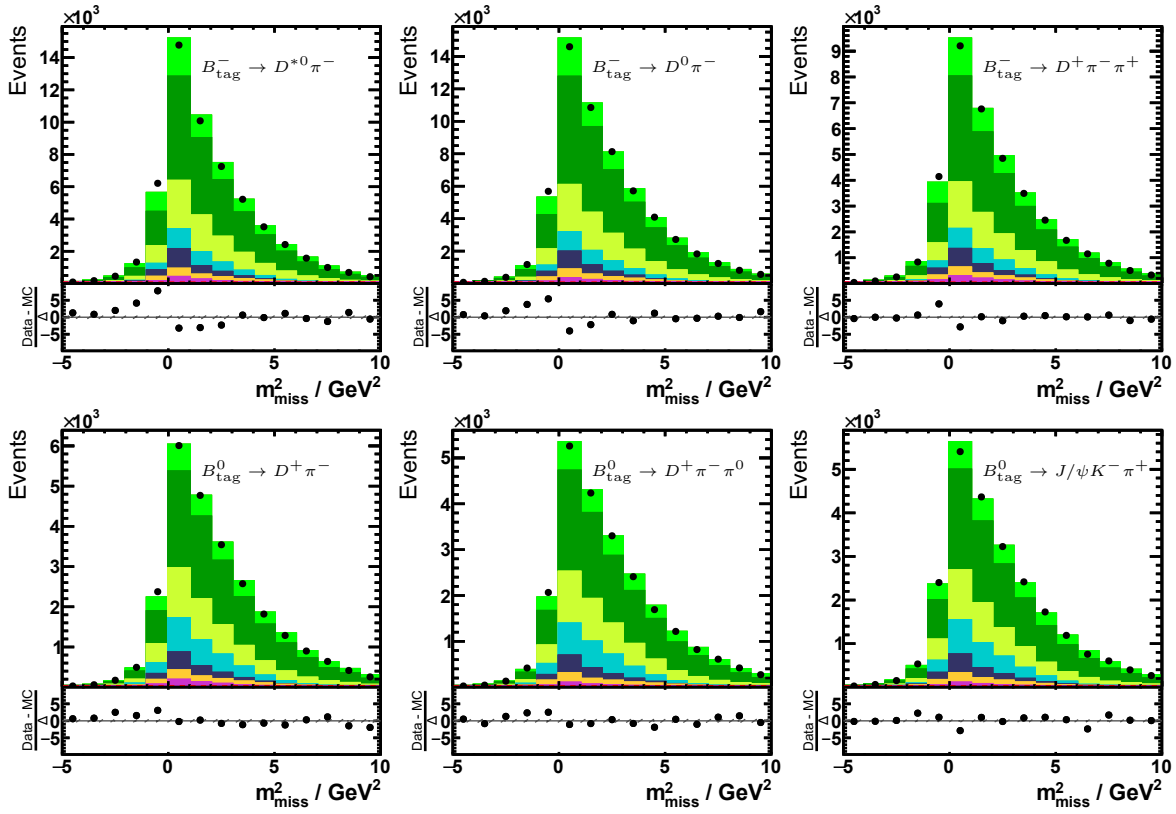


Figure 5.23: Dependence of m^2_{miss} modelling on the B_{tag} reconstruction mode. Shown are the three most common reconstruction modes of charged (top) and neutral (bottom) B_{tag}. The jump in the residuals is much more pronounced for charged than for neutral B_{tag}, but the much lower sample size of the latter has to be considered. The MC is normalised to the number of events in data.

The impact of photons on the tag side and thus on the m^2_{miss} modelling is studied. Photons are reconstructed from neutral clusters in the ECL that are known to be imperfectly modelled. For example, it can happen that a cluster is not assigned to its track and is treated as a photon. Furthermore, the reconstruction of neutral pions, which is always performed in the $\pi^0 \to \gamma\gamma$ channel, has a poor efficiency and purity. In contrast to the signal side with its tight photon energy requirement, the tagging algorithm can consider photons with energies down to 30 MeV. To estimate if photons and neutral pions, used in the B_{tag} reconstruction, have any impact on the description of m^2_{miss}, two samples are studied, one with and one without neutral particles in the reconstruction modes (see Fig. 5.24). The effect of neutral particles in the B_{tag} is found to be negligible as the shapes of the m^2_{miss} distributions are the same.

The impact of the particle identification is investigated, too. Several requirements on the PID classifier to distinguish kaons and pions are tested, but without a noticeable impact on the modelling of m^2_{miss}. Also, the definition of poor tracks is revisited and tracks which do not have a high probability to be either a kaon or a pion are discarded. Such a selection increases the mean of the missing mass distribution because many low momentum tracks are rejected. However, even though the shape of m^2_{miss} changes, the jump in the residuals at $m^2_{\text{miss}} = 0\,\text{GeV}^2$ remains. The best test to ensure that particle identification of hadrons is not responsible for this, is to not identify particles at all, except the signal lepton. All tracks are treated as pions

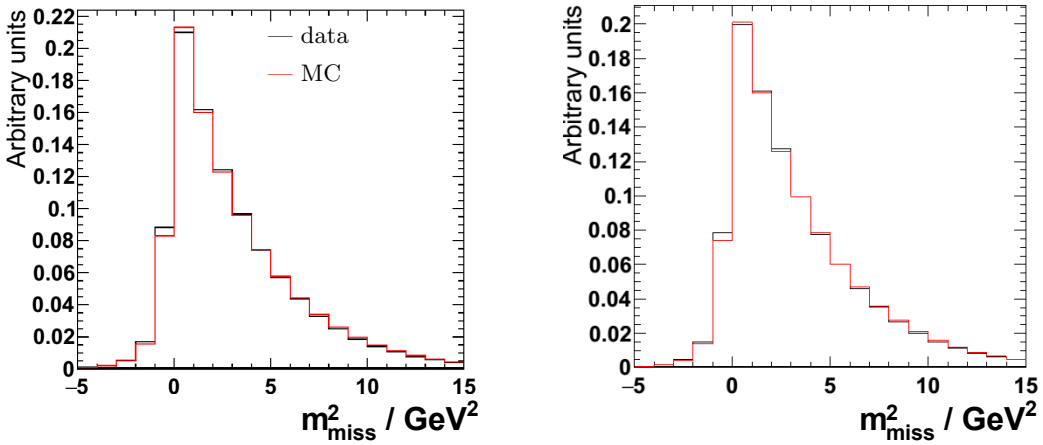


Figure 5.24: Impact of B_{tag} reconstruction modes without (left) and with (right) photons in the reconstruction chain. No difference, especially considering the excess of data at $m^2_{miss} < 0\,\text{GeV}^2$, is visible.

and a difference between data and MC due to wrongly assigned particle hypotheses should vanish. However, the poor m^2_{miss} modelling is not affected.

In addition, studies are performed to find efficiency corrections that reduce the m^2_{miss} residuals. Examples are corrections to the photon, track or kaon multiplicity and the total visible momentum. The corrections are obtained from the ratio of data and MC histograms in the considered observable. These correction factors are then applied as event weights and the m^2_{miss} distribution is examined. However, none of the tested efficiency corrections are able to remove the observed residuals. Possible reasons for that have been given in the previous paragraphs. Searches for efficiency corrections are extended to multiple dimensions to take correlations between the observables into account. However, the tests showed that per-event corrections on the basis of observables unlikely account for the m^2_{miss} mis-modelling.

The influence of branching fractions has been intensively studied. As already mentioned, the most common B and D meson branching fractions in the MC sample after the signal selection are updated to recent measurements. They indeed improve the overall data-MC agreement, however, not the mis-modelling at $m^2_{miss} < 0\,\text{GeV}^2$.

Conclusions The discrepancy between data and MC at $m^2_{miss} < 0\,\text{GeV}$ has a complex origin. Numerous tests have been performed to identify the underlying processes and obtain correction factors. The presented steps of a tightened particle selection improved the description of the data (and is used is this analysis). However, efficiency corrections of e. g. track and photon multiplicities or tag observables are not to improve the m^2_{miss} mis-modelling. This suggests that the origin already lies on the track level,

For these reasons, and because the signal populates the large m^2_{miss} region, the analysis is performed only with events which have positive missing mass.

The requirement of $m^2_{miss} \geq 0\,\text{GeV}^2$ could potentially lead to a selection bias which is estimated here. In contrast to signal events, which populate the high m^2_{miss} region and are thus not affected by the selection, events of the $B \to X\ell\nu$ normalisation mode are rejected with different efficiencies in data and MC. As less $B \to X\ell\nu$ events are rejected in MC than in data, the normalisation mode could be slightly overestimated in the remaining $\geq 0\,\text{GeV}^2$ sample. From the number of observed events in data and the MC prediction, the bias on $R(X)$

is estimated to be $\lesssim$1% which is negligible compared to the total uncertainty on $R(X)$ of $\sim$7.5% (see Chapter 6).

5.4 Signal extraction strategy

The procedure of extracting the signal from the data is presented in the following sections.

The branching fraction ratio $R(X)$ is measured, i. e., the signal branching fraction normalised to the semileptonic $B \to X\ell\nu$ decay, which is referred to as the normalisation mode. Events which contain neither signal nor normalisation decays are denoted as *other decays*. Signal and normalisation events are required to have a true lepton from a $B \to X\tau\nu$ and $B \to X\ell\nu$ decay and its mother must be a true τ or B, respectively. All other events are part of the "other decays" component which are mostly secondary leptons from hadronic B decays, fake lepton candidates and continuum events.

The signal extraction is performed in a binned 2D χ^2 fit in $|\vec{p}^{\,*}_\ell|$ and m^2_{miss}. The normalisation mode is well separated from the signal in the lepton momentum spectrum and in the squared missing mass distribution. The latter is needed to constrain the other decays because they have a similar lepton momentum spectrum as the signal.

5.4.1 Fit templates and model parameters

The fit templates are taken from the 2D MC distributions of $|\vec{p}^{\,*}_\ell|$ and m^2_{miss} which are shown in Fig. 5.25. For every fit component c, the 2D MC histogram, denoted as a vector of bins $\vec{f}^{\,c}$, is obtained by using MC truth information. The templates are normalised to the predicted MC yields, i. e. $\sum_i f^{\text{c}}_i = N^{\text{c}}$.

Free parameters of the fit are the template scale factors

$$r_X = \frac{R(X)}{R(X)_{\text{MC}}} \qquad A^{B\to X\ell\nu} = \frac{N^{B\to X\ell\nu}}{N^{B\to X\ell\nu}_{\text{MC}}} \qquad A^{\text{other}} = \frac{N^{\text{other}}}{N^{\text{other}}_{\text{MC}}} \tag{5.8}$$

for the signal, the $B \to X_c\ell\nu$ and the *others* component, respectively. The MC is not expected to have the right overall normalisation because of, e. g., the systematic uncertainty on the efficiency of the hadronic tagging and the number of produced $B\overline{B}$ pairs. Both uncertainties are expected to cancel in the ratio $R(X)$, but the former can be different for charged and neutral B mesons. For this reasons, separate normalisations A_j ($j = B^+, B^0$) are considered. The component of "other decays" is not only different between charged and neural B mesons, but also between electrons and muons because of their very different mis-identification probabilities (cf. Fig. 5.8). For this reason, the "other decays" component is additionally separated into electron and muon templates with individual normalisations A^{other}_{jk} ($k = e, \mu$) which are, however, not of particular interest for the analysis. The complete binned template $\vec{f}(\vec{\theta}) = (f_0(\vec{\theta}), f_1(\vec{\theta}), \dots)$ as a function of parameters $\vec{\theta}$ reads

$$\vec{f}(r_X, A^{B\to X\ell\nu}_{B^+}, \dots, A^{\text{other}}_{B^0\mu}) = \sum_{j=B^+,B^0} \left\{ A^{B\to X\ell\nu}_j \left(r_X \vec{f}^{\,\text{signal}}_j + \vec{f}^{\,\text{norm}}_j \right) + \sum_{k=e,\mu} A^{\text{other}}_{jk} \vec{f}^{\,\text{other}}_{jk} \right\}. \tag{5.9}$$

This formulation is independent of the $B \to X\tau\nu$ and $B \to X\ell\nu$ branching fractions assumed in the MC. The formula leads to the desired cancellation effects on $R(X)$ because uncertainties on

the MC normalisation are uncertainties on the $A_j^{B\to X\ell\nu}$ parameters and thus r_X parametrises only the ratio of $B \to X\tau\nu$ and $B \to X\ell\nu$.

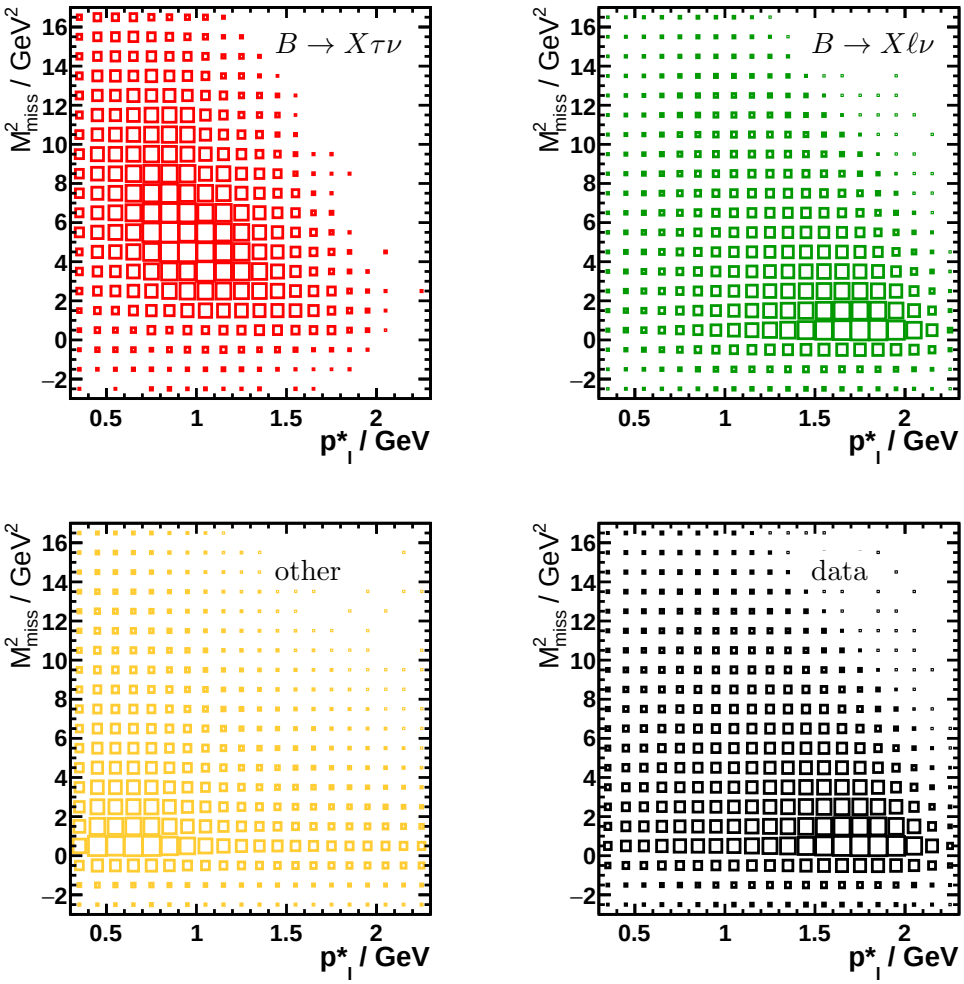


Figure 5.25: Distributions of signal (red), normalisation (green) and other decays (yellow) as well as data (black) in the two signal extraction variables m^2_{miss} and $|\vec{p}^{\,*}_\ell|$. The signal is clearly separated from the backgrounds which are, however, still present in the signal region. The data illustrates that the $B \to X\ell\nu$ component is by far the dominant decay mode in the analysed sample.

5.4.2 Fit set-up

The fitting procedure is an optimisation problem to find the set of best model parameters $\vec{\theta}_0$ which yield the best agreement with the observed data. The number of data events n_i in bin i, and the total number $n = \sum_i n_i$ are distributed according to a Poissonian distribution. In the limit of large n_i, which is a good assumption in this analysis, the Poissonian is well approximated by a Gaussian distribution. To satisfy this assumption and to avoid biases, bins with $n_i < 30$ in data or in the sum of templates are ignored in the fitting procedure. The optimal $\vec{\theta}_0$ which maximises the probability $p(\vec{n}\,|\,\vec{f}(\vec{\theta}_0))$ of observing the data $\vec{n}$ with a given model $\vec{f}(\vec{\theta})$ is given

as the maximum of likelihood-function $\mathcal{L}$, i. e. the products of the probabilities in every bin i,

$$\mathcal{L}(\vec{\theta}) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\left(n_i - f_i(\vec{\theta})\right)^2}{2\sigma_i^2} \right), \tag{5.10}$$

with standard deviation $\sigma_i = \Delta n_i = \sqrt{n_i}$.

To get rid of the computationally expensive products and exponentials, the log-likelihood is used,

$$-2\ln\mathcal{L}(\vec{\theta}) = \sum_i \left(\frac{n_i - f_i(\vec{\theta})}{\sigma_i} \right)^2 =: \chi^2(\vec{\theta}) \tag{5.11}$$

which is the well-known least-squares or χ^2 error-function and constant terms in the sum are neglected. Hence, for the optimal parameter set $\vec{\theta}_0$, $\chi_0^2 \equiv \chi^2(\vec{\theta}_0)$ is minimal.

So far, statistically independent bins have been assumed. To consider e. g. systematic uncertainties which are correlated across various bins, the χ^2 can be generalised from a multivariate Gaussian in Eq. 5.10 with covariance matrix C to

$$\chi^2(\vec{\theta}) = \left(\vec{n} - \vec{f}(\vec{\theta})\right)^T C^{-1} \left(\vec{n} - \vec{f}(\vec{\theta})\right). \tag{5.12}$$

Additional uncertainties are considered in the covariance matrix $C = C_{\text{stat}} + C_{\text{sys},1} + \cdots + C_{\text{sys},n}$ and assumed to be independent of $\vec{\theta}$. However, the latter is not necessarily true when considering uncertainties on the templates $\vec{f^c}(\vec{\theta})$ and thus the matrix needs to be inverted each time $\vec{\theta}$ changes during the minimisation process. This is computationally expensive and can become numerically unstable. Therefore, only bin-wise independent uncertainties, e. g. the statistical MC uncertainty, are included in the covariance matrix because this keeps the matrix diagonal and the inversion is simply $C_{ii}^{-1} = 1/C_{ii}$. By doing so, the covariance matrix includes not only the variance of the data, but also of the MC which is given by the sum of squared event weights, $\text{var}_{\text{MC}} = \sum_i w_i^2(\vec{\theta})$, as well as the statistical uncertainties on the lepton efficiency and the lepton mis-identification correction factors.

Systematic uncertainties are included in the fit by introducing a set of *nuisance parameters* (NP) $\vec{\lambda}$ which are restricted by a set of prior probabilities of the same size. Consequently, they do not change the number of degrees of freedom of the fit. The prior probabilities are usually external measurements and are expected to be normally distributed (which is not a requirement, however) and hence enter the χ^2 quadratically.

The sum of priors χ^2_{NP} changes the χ^2 to

$$\chi^2(\vec{\theta}) \to \chi^2(\vec{\theta}, \vec{\lambda}) = \chi^2(\vec{\theta}) + \chi^2_{\text{NP}}(\vec{\lambda}), \quad \text{with} \quad \chi^2_{\text{NP}}(\vec{\lambda}) = \sum_j \lambda_j^2, \tag{5.13}$$

where $\vec{\lambda} = 0$ corresponds to the unaffected MC, i. e. the expectation values of the priors. As an example, the nuisance parameter for a certain branching fraction affects the weight w of an event as

$$w \to w \cdot (1 + \delta\mathcal{B} \cdot \lambda) \tag{5.14}$$

with the relative uncertainty on the branching fraction $\delta\mathcal{B} = \Delta\mathcal{B}/\mathcal{B}$. The NPs are subject to the minimisation, however, they are constrained by the extra sum in the χ^2. Hence, the fit can change the shape of the MC templates by changing e. g. branching fractions to improve

the overall agreement of MC and data, but is restricted by the prior knowledge, i. e. external measurements.

The uncertainty of a nuisance parameter reported by the fitting procedure states how constrained the NP is by the data. Without any constraints beside its prior probability, the nuisance parameter has, by construction, an uncertainty of one. In contrast, a tight constraint on a NP from the data results in a smaller uncertainty. Furthermore, its uncertainty reflects the dependence of the fit on that particular NP and its contribution to the total uncertainty of other parameters.

The considered systematic uncertainties and the corresponding nuisance parameters are discussed in the following. The sources of systematic uncertainties and their effect on the measurement are discussed in Sec. 5.5 while here the more technical part is covered. A complete list of the considered parameters is given in Tab. 5.2.

Branching fraction nuisance parameters Of outstanding importance is the inclusive $B \to X\ell\nu$ normalisation mode and its composition. Especially the $B \to D^{**}\ell\nu$ decays have large experimental uncertainties and thus also the sum of exclusive modes (cf. Tab. 3.2). Therefore, the branching fractions of the exclusive X decays are considered as nuisance parameters to account for their uncertainties and hence their effect on the composition of the inclusive sample. A further constraint stems from the inclusive $B \to X\ell\nu$ branching fraction and, therefore, the sum of exclusive semileptonic B decays is forced to match the total $B \to X\ell\nu$ branching fraction. In other words, the sum of the N exclusive branching fractions must not change and consequently

$$\sum_{j=1}^{N} \Delta\mathcal{B}_j \cdot \lambda_j = 0 \Leftrightarrow \lambda_N = -\frac{1}{\Delta\mathcal{B}_N} \sum_{j=1}^{N-1} \Delta\mathcal{B}_j \cdot \lambda_j, \tag{5.15}$$

which implies that one NP is not a free parameter but determined by the $N-1$ other NPs. Even though it is not a free parameter, the associated prior probability is taken into account. This method is also applied to the exclusive signal components. This allows one to work with a fixed inclusive branching fraction, which is necessary to extract the branching fraction ratio $R(X)$. No inclusive branching fraction is assumed and also its uncertainty is not considered in this constraint. Only when calculating $\mathcal{B}(B \to X\tau\nu)$ from $R(X)$, the uncertainty of the assumed inclusive $B \to X\ell\nu$ branching fraction enters.

In principle, negative branching fractions can occur. Especially, if the relative uncertainty is large (as it is the case for e. g. $B \to D_{2S}\ell\nu$), the event weight $(1 + \delta\mathcal{B}_j\lambda_j)$ can easily become negative. Therefore, an exponential cut-off in the prior probability is introduced that is implemented in the χ^2. The additive penalty to the χ^2 is chosen to be $10 \cdot |1 + \delta\mathcal{B}_j\lambda_j|$ and is applied if $(1 + \delta\mathcal{B}_j\lambda_j) < 0$. Through this procedure, no negative branching fractions are observed.

Counting the considered charmed meson states in the MC, the signal and normalisation modes contribute with seven and nine independent free NPs, respectively, and each with an additional constraint due to Eq. 5.15. Furthermore, a NP is considered for the inclusive $B \to X_u\ell\nu$ branching fraction.

A set of nuisance parameters is implemented to account for the new model of D^{**} decays. For each decay in the model, a 5% relative uncertainty is assumed. However, the total decay width of the individual D^{**} mesons is preserved by Eq. 5.15.

A considerable amount of low momentum electrons stems from photon conversions. To take potential mis-modelling of the detector material into account, a 3% [86] relative uncertainty on

their production is considered.

To consider the uncertainty on semileptonic charmed meson decays, which produce secondary leptons in the $|\vec{p}_\ell^{\,*}|$ signal region, a NP is assigned separately for charged and neutral D mesons. The inclusive semileptonic charm branching fractions $\mathcal{B}\left(\bar{D}^0 \to X\ell^+\nu\right)$ and $\mathcal{B}\left(D^+ \to X\ell^+\nu\right)$ are considered and implemented such that the total charm decay width is constant.

Charged and neutral B decays have the same nuisance parameters and relative uncertainties because isospin symmetry is still assumed and the isospin factor cancels in relative uncertainty $\delta\mathcal{B} = \Delta\mathcal{B}/\mathcal{B}$. Nevertheless, the exclusive $B \to X\ell\nu$ and $D \to X\ell\nu$ branching fraction NPs are implemented separately for charged and neutral B, which will be introduced and motivated in Sec. 5.6.

Form-factor parameters The lepton momentum spectrum is sensitive to the decay model used in the MC simulation, especially that of the $B \to D^{(*)}\ell\nu$ decays. The CLN / HQET2 model parameters (cf. Sec. 1.2.1) are allowed to vary in the fit by assigning a single nuisance parameter to each of them. In case of $B \to D\ell\nu$, the single CLN parameter ρ_D^2 is directly assigned a NP.

The experimental correlations between the $B \to D^*\ell\nu$ form factor parameters $\rho_{D^*}^2$, R_1 and R_2 are taken into account by relating the independent nuisance parameters with the principal components of the covariance matrix, i. e. scaling the axes of the error ellipsoid. The vector of CLN parameters $\vec{r} = (\rho^2, R_1, R_2)^T$ is rotated into the eigenbasis of the covariance matrix with the inverse matrix of eigenvectors Γ. In this basis, the nuisance parameters scale the independent principal components which are, afterwards, rotated back into the CLN representation. Written in vector components, the whole transformation reads

$$r^j \to \Gamma^{ji}\left(r^k(\Gamma^{-1})_{ik} + \lambda_i\sqrt{\xi}_i\right), \tag{5.16}$$

where $\vec{\xi}$ is the vector of eigenvalues which represents the variance along the principal axis i.

The semileptonic decays into 1P states are described by two independent parameter sets of the LLSW model; one for the broad $j_l = 1/2$ and one for the narrow $j_l = 3/2$ states, each with three independent nuisance parameters. The nuisance parameters represent the slope of the leading Isgur-Wise function $\mathrm{d}\tau/\mathrm{d}w$ and the kinetic energy operator η_{ke}. The third parameter enables the fit to vary between approximations B1 and B2 of the LLSW model, i. e. $\Gamma = c\Gamma_{\mathrm{B1}} + (1-c)\Gamma_{\mathrm{B2}}$ (cf. Sec. 1.2.2).

Semileptonic decays into 2S states follow the parameterless `ISGW2` model and thus have no nuisance parameters assigned.

In total, there are ten nuisance parameters related to the semileptonic form factors. In principle, the same could be applied to $B \to X\tau\nu$ signal decays. However, due to the fact that the signal lepton is secondary, such details of the semileptonic model are negligible.

Nuisance parameters for lepton efficiency The uncertainties on the lepton efficiency and the lepton mis-identification corrections (cf. Sec. 3.3) are considered. For the lepton efficiency correction factors, a systematic and a statistical uncertainty is reported [61]. The systematic uncertainty is assumed to be completely correlated across all bins and, therefore, is covered by a single nuisance parameter. In contrast, the statistical uncertainty is bin-wise independent and hence is included in the covariance matrix. The dependence on the model parameters $\vec{\theta}$ of this contribution to the covariance matrix is considered, but influence of any NP is neglected. The weights of events with true leptons in bin i are thus related to the relative efficiency

uncertainties $\delta\epsilon_i$ by

$$w_i \to w_i \cdot (1 + \delta\epsilon_i^{\text{sys}} \lambda^{\text{sys}}). \tag{5.17}$$

The lepton mis-identification probability correction factors have negligible systematic uncertainties and therefore only statistical uncertainties are taken into account in the covariance matrix. The described set-up is separately implemented for electrons and muons.

5.4.3 Fit validation

The described fit set-up is tested for being stable, unbiased and reliable regarding the estimates of the total uncertainties. This is done by pseudo-experiments with pseudo-data drawn from the MC templates. A sample of pseudo-data is generated in two steps. In the first step, the MC is fluctuated bin by bin according to its Poissonian distribution to simulate statistical fluctuations. Since the statistical fluctuations shall represent the data statistics, the MC is scaled to the expected number of events in data before applying the Poissonian variations. The second step generates the systematic variations covered by the nuisance parameters by obtaining a set of new external measurements, i. e. expectation values are drawn from the prior probabilities. The new expectation values $\vec{\lambda}^0$ are considered in the χ as

$$\sum_j (\lambda_j - \lambda_j^0)^2 \tag{5.18}$$

The fit is validated by performing the analysis with a large sample of pseudo-data sets and obtaining the pull of every parameter θ_i in each fit:

$$\text{pull} = \frac{\theta_i^{\text{fit}} - \theta_i^{\text{true}}}{\Delta\theta_i^{\text{fit}}}. \tag{5.19}$$

The distribution of pull values for a certain parameter θ_i should follow a normal distribution with zero mean and a width of one. The mean of zero represents an unbiased fit that on average reproduces the input $\vec{\theta}^{\text{true}}$ while a width equal to one confirms a correctly estimated uncertainty.

For this analysis 1000 pseudo experiments are drawn and the signal extraction fit is carried out. The nuisance parameters of the D^{**} meson decays are skipped for this test, as they are found to have negligible impact on the fit in terms of negligible correlations with other parameters and are computationally expensive. In Fig. 5.26, the pull distributions of $R(X)$ and the normalisation are shown. Clearly visible is the unbiased fit result and the correctly estimated total uncertainty. Pull distributions of further parameters can be found in Fig. A.4.

A standard procedure used at Belle to validate the fitting procedure as well as the MC, is a so-called "stream test". One of the MC streams is used as pseudo-data and all remaining streams are used as fit templates. As the generic $B\overline{B}$ MC consists of ten streams, just as many tests are performed. For MC components that have fewer than ten streams available (continuum, $B \to X_u \ell\nu$, and the $B \to D^{**}\ell\nu$ MC), the following procedure is applied. The $B \to X_u \ell\nu$ MC is the same in the pseudo-data and the MC templates, as it is only a small component. For continuum MC, the six available streams are used like the ten generic $B\overline{B}$ MC streams but start again with the first stream after the sixth test. The $B \to D^{**}\ell\nu$ MC is split into two halves that are used alternately. Fig. 5.27 shows a summary of the stream test. As no systematic difference between the MC streams are expected, only the statistical uncertainty is considered. The obtained scale factors r_X and $A_j^{B\to X\ell\nu}$ fluctuate around the nominal MC value, as expected. Neither a bias nor a wrongly estimated uncertainty is observed for r_X. Only the statistical

1	r_X	49†	$D_1 \to D^*\pi\pi$
2	$A_{B^+}^{B\to X\ell\nu}$	50	$D_2^* \to D\pi$
3	$A_{B^+e}^{other}$	51	$D_2^* \to D\pi\pi$
4	$A_{B^+\mu}^{other}$	52	$D_2^* \to D^*\pi$
5	$A_{B^0}^{B\to X\ell\nu}$	53	$D_2^* \to D^*\pi\pi$
6	$A_{B^0e}^{other}$	54†	$D_2^* \to D\eta$
7	$A_{B^0\mu}^{other}$	55	$D_{2S} \to D^*\pi$
8	$\gamma \to e^+e^-$	56	$D_{2S} \to D\pi\pi$
9	$B \to D\tau\nu$	57	$D_{2S} \to D^*\pi\pi$
10	$B \to D^*\tau\nu$	58	$D_{2S} \to D_0^*\pi$
11	$B \to D_1\tau\nu$	59	$D_{2S} \to D^*\eta$
12	$B \to D_2^*\tau\nu$	60	$D_{2S} \to D_2^*\pi$
13	$B \to D_0^*\tau\nu$	61	$D_{2S} \to D_s^*K$
14	$B \to D_1'\tau\nu$	62	$D_{2S} \to D\rho$
15	$B \to D_{2S}\tau\nu$	63†	$D_{2S} \to D\omega$
16†	$B \to D_{2S}^*\tau\nu$	64	$D_{2S}^* \to D\pi$
17	$B^+ \to D\ell\nu$	65	$D_{2S}^* \to D\pi\pi$
18	$B^+ \to D^*\ell\nu$	66	$D_{2S}^* \to D^*\pi$
19	$B^+ \to D_1\ell\nu$	67	$D_{2S}^* \to D^*\pi\pi$
20	$B^+ \to D_2^*\ell\nu$	68	$D_{2S}^* \to D\eta$
21	$B^+ \to D_0^*\ell\nu$	69	$D_{2S}^* \to D_sK$
22	$B^+ \to D_1'\ell\nu$	70	$D_{2S}^* \to D^*\eta$
23	$B^+ \to D\pi\ell\nu$	71	$D_{2S}^* \to D_1\pi$
24	$B^+ \to D_{2S}\ell\nu$	72	$D_{2S}^* \to D_2^*\pi$
25†	$B^+ \to D_{2S}^*\ell\nu$	73	$D_{2S}^* \to D_s^*K$
26	$B^0 \to D\ell\nu$	74	$D_{2S}^* \to D_{2S}\pi$
27	$B^0 \to D^*\ell\nu$	75	$D_{2S}^* \to D^*\rho$
28	$B^0 \to D_1\ell\nu$	76	$D_{2S}^* \to D^*\omega$
29	$B^0 \to D_2^*\ell\nu$	77†	$D_{2S}^* \to D_{2S}\gamma$
30	$B^0 \to D_0^*\ell\nu$	78	$B^+ : D^0 \to X\ell\nu$
31	$B^0 \to D_1'\ell\nu$	79	$B^+ : D^+ \to X\ell\nu$
32	$B^0 \to D\pi\ell\nu$	80	$B^0 : D^0 \to X\ell\nu$
33	$B^0 \to D_{2S}\ell\nu$	81	$B^0 : D^+ \to X\ell\nu$
34†	$B^0 \to D_{2S}^*\ell\nu$	82	Electron efficiency
35	$B \to X_u\ell\nu$	83	Muon efficiency
36	$D_0^* \to D\pi$	84	ρ_D^2
37	$D_0^* \to D^*\pi\pi$	85	$B \to D^*\ell\nu$: ff r_1
38†	$D_0^* \to D\eta$	86	$B \to D^*\ell\nu$: ff r_2
39	$D_1' \to D^*\pi$	87	$B \to D^*\ell\nu$: ff r_3
40	$D_1' \to D\pi\pi$	88	$1P_{3/2} : d\tau/dw$
41	$D_1' \to D^*\pi\pi$	89	$1P_{3/2} : \eta_{ke}$
42	$D_1' \to D^*\eta$	90	$1P_{3/2} : c_{B1\text{-}B2}$
43	$D_1' \to D\rho$	91	$1P_{1/2} : d\tau/dw$
44	$D_1' \to D\omega$	92	$1P_{1/2} : \eta_{ke}$
45	$D_1' \to D^*\rho$	93	$1P_{1/2} : c_{B1\text{-}B2}$
46	$D_1' \to D^*\omega$		
47	$D_1 \to D^*\pi$		
48	$D_1 \to D\pi\pi$		

Table 5.2: List of floating parameters $\vec{\theta}$ (1 – 7) and nuisance parameters ($\vec{\lambda}$) (8 – 93) of the simultaneous fit. Nuisance parameters 85–87 are associated to the decorrelated CLN parameters (cf. Tab. 1.1) and 88–93 are associated to the parameters of the LLSW model.
† : These parameters are not free in the fit but determined by the inclusive constraint (cf. Eq. 5.15).

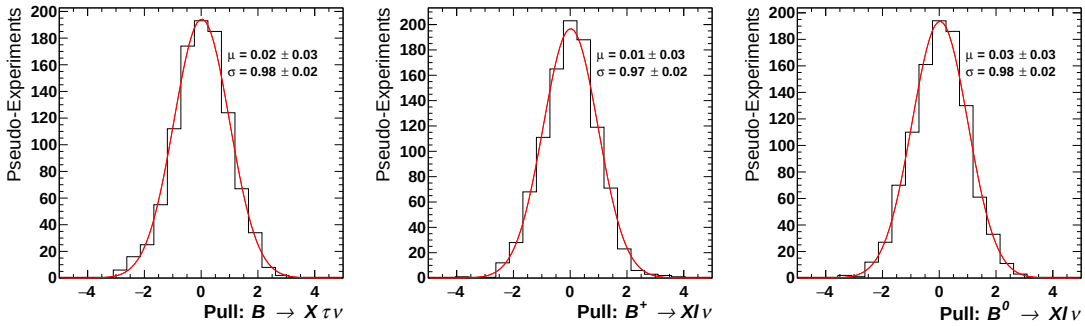


Figure 5.26: Pull distributions of $R(X)$ and the normalisation mode $B \to X\ell\nu$. A Gaussian distribution is fitted to confirm the unbiased and well reproduced uncertainties.

uncertainty of the normalisation mode seems slightly underestimated which is, however, not of concern for the signal extraction. To conclude, there is no significant difference between individual streams and in their analysis processing.

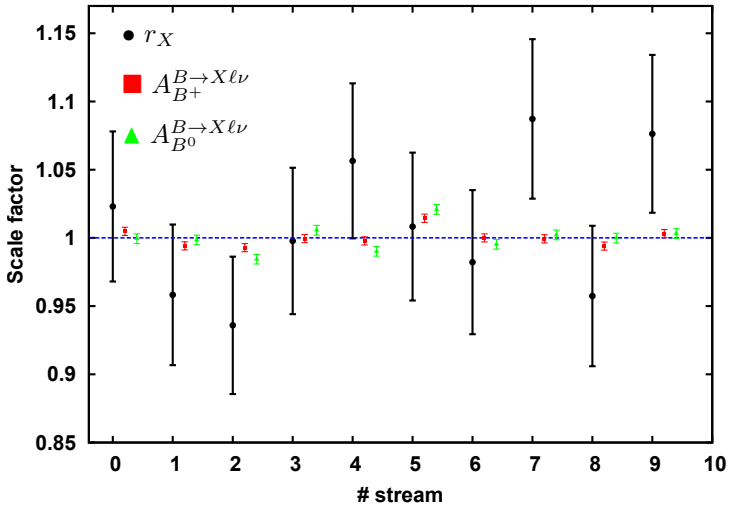


Figure 5.27: The signal and normalisation mode scaling factors (cf. Eq. 5.9) in the stream test fits to pseudo-data from disjunct MC sub-samples. The error bars represent the statistical uncertainty of the fit parameters.

5.5 Systematic uncertainties

The systematic uncertainties of this analysis are estimated from pseudo-data before the fitting procedure is confronted with data.

Due to the implementation of nuisance parameters and considering uncertainties in the covariance matrix, the fitting procedure estimates the total variance $\mathrm{var}_{\mathrm{tot}} = \mathrm{var}_{\mathrm{stat}} + \mathrm{var}_{\mathrm{sys}}$. A breakdown of the statistical as well as the systematic uncertainties is obtained to understand the impact of the individual sources. Therefore, the nominal fit is repeated while the nuisance parameter of interest, i, is fixed to its best-fit value to obtain the variance without i, $\mathrm{var}_{\mathrm{w/o}\ i}$. Afterwards, the variance due to NP i is calculated by $\mathrm{var}_i = \mathrm{var}_{\mathrm{tot}} - \mathrm{var}_{\mathrm{w/o}\ i}$. The estimated statistical uncertainties and the list of individual systematic uncertainties are shown in Tab. 5.3.

A discussion of the individual contributions is given below.

Rel. uncertainty $\delta R(X)/\%$	
Statistical	± 5.2
PID	± 1.1
$\mathcal{B}(B \to X\tau\nu)$ composition	± 0.6
$\mathcal{B}(B \to D\ell\nu)$	± 0.2
$\mathcal{B}(B \to D^*\ell\nu)$	$^{+5.5}_{-5.0}$
$\mathcal{B}(B \to D^{**}\ell\nu)$ composition	± 3.7
$\mathcal{B}(D \to X\ell\nu)$	± 4.7
D^{**} decay model	± 0.2
$\mathrm{FF}_{\mathrm{CLN}}(B \to D^{(*)}\ell\nu)$	± 0.7
$\mathrm{FF}_{\mathrm{LLSW}}(B \to D^{**}\ell\nu)$	$^{+5.5}_{-5.1}$
MC statistics	± 2.6
Total systematic	$^{+8.2}_{-7.9}$
Total	$^{+9.7}_{-9.4}$

Table 5.3: Relative expected uncertainty on $R(X)$, estimated from pseudo-experiments. *Composition* means the summary of the contributions from the individual exclusive decays. Due to correlations, the individual uncertainties do not necessarily sum up in quadrature to the total uncertainty.

Branching fractions The influence of the branching fraction uncertainties is diverse. The individual exclusive branching fractions of the signal ($\mathcal{B}(B \to X\tau\nu)$ composition) itself are rather unimportant. This is due to fact that the signal particle is a secondary lepton which causes a rather undistinguishable mixture of exclusive components in $|\vec{p}_\ell^{\,*}|$ and m^2_{miss}. As the sum of exclusive decays is constrained to the inclusive prediction, i. e. only the composition can vary, they barely affect the signal efficiency.

In contrast, the $B \to X\ell\nu$ normalisation mode components are clearly distinguishable in the $|\vec{p}_\ell^{\,*}|$ spectrum (and to a lesser extend also in m^2_{miss}). The $B \to X\ell\nu$ components extend into the signal region and thus can change the shape of the background in this region. This results in a relatively large impact of the $B \to X\ell\nu$ composition on the $R(X)$ measurement.

Furthermore, the branching fraction of inclusive semileptonic charm decays ($\mathcal{B}(D \to X\ell\nu)$) is a noticeable source of systematic uncertainty. Indirectly, it also accounts for the uncertainty on single- and double-charmed hadronic B decays because their subsequent charm decay is the dominant process to produce secondary leptons in the background. As these leptons mainly populate the signal region in $|\vec{p}_\ell^{\,*}|$, their uncertainty is a significant contribution to the systematic uncertainty on $R(X)$.

The D^{**} decay model branching fractions cause only a minor systematic uncertainty. The single D^{**} decay modes are almost indistinguishable in the signal extraction variables and thus their relative composition is not important. The fitting procedure rather generates the D^{**} related uncertainties via the branching fractions ($\mathcal{B}(B \to D^{**}\ell\nu)$ composition) and the form-factor parameters of $B \to D^{**}\ell\nu$ decays ($\mathrm{FF}_{\mathrm{LLSW}}(B \to D^{**}\ell\nu)$).

Form-factor parameters The form-factor parameters of the CLN model of $B \to D^{(*)}\ell\nu$ decays ($\mathrm{FF}_{\mathrm{CLN}}(B \to D^{(*)}\ell\nu)$) are only minor sources of systematic uncertainties. These parameters have small experimental uncertainties and therefore possible variations in the shapes are small.

However, the much less constrained parameters of the LLSW model for $B \to D^{**}\ell\nu$ decays ($\mathrm{FF}_{\mathrm{LLSW}}(B \to D^{**}\ell\nu)$) affect the signal significantly and are one of the dominant sources of uncertainty.

Lepton efficiency and lepton mis-identification The uncertainties on the lepton identification and mis-identification corrections are minor systematic uncertainties in this analysis. Separated into lepton efficiency and fake lepton candidate corrections, both contribute approximately equally to the total PID uncertainty.

The lepton momentum distributions and the relative uncertainties of true and fake lepton candidates have been shown in Figs. 3.2 and 3.3. Their relative uncertainty is not completely uniform over the lepton momentum spectrum and thus contributes differently to the signal and the normalisation mode. Therefore, no better cancellation of the lepton efficiency correction uncertainties is achieved.

Conclusions The systematic uncertainties of this analysis are, as expected, dominated by the modelling of the normalisation mode. Compared with previous analyses of the $B \to X\tau\nu$ decay, this work is expected to provide the most precise estimate of the inclusive branching fraction ratio $R(X)$. The uncertainty of the branching fraction $\mathcal{B}(B \to X\tau\nu)$ will additionally depend on the uncertainty of $\mathcal{B}(B \to X\ell\nu)$, which is currently at about 3% [20].

5.6 Control studies

The event selection of this analysis is designed to be as inclusive as possible in the sense that only a few selection requirements are used. On the one hand, this has the advantage of a rather model-independent selection and a large data sample to analyse. However, on the other hand, the downside is the lack of control modes and control regions because nearly the whole phase space of $B \to X\tau\nu$ and $B \to X\ell\nu$ decays is considered. As $B \to X\ell\nu$ is already considered as the normalisation mode, it is not available as a control mode. A signal-free control region can be obtained by either selecting a pure $B \to X\ell\nu$ sample (requiring large $|\vec{p}_\ell^{\,*}|$) or a sample with all contributions except signal decays (small m^2_{miss}). The latter is the most interesting case because the composition of backgrounds is expected to be similar to the nominal sample.

Another control selection can make use of the hadronic system X, namely selecting large m_X above the mass of the D^*_{2S} meson. In this regime, semileptonic B decays can be completely neglected, leaving a rather pure hadronic B decay sample. However, single- or double-charmed hadronic B decays, as well as their subsequent semileptonic D decays, are poorly known and do not serve well as a control. A similar selection is possible by inverting the requirement on the flavour-charge correlation between the B_{tag} and the signal lepton. Such a sample contains mainly secondary leptons from charmed hadronic B decays, but also signal and normalisation mode decays due to wrongly reconstructed B_{tag} and $B^0\overline{B^0}$ mixing. However, the best choice is the side-band of small m^2_{miss} values because it has a relatively large sample size and is expected to preserve the shape and composition of the lepton spectrum.

To test the fitting procedure, events in the region $0 < m^2_{\mathrm{miss}} < 3\,\mathrm{GeV}^2$ are selected. The fit is then mainly sensitive to the lepton momentum spectrum and can help to understand the impact of nuisance parameters before considering the signal region. The fit model presented in

Eq. 5.9 is used, but without the signal component and remaining signal events in the control sample are merged into in the "others" component. As the MC normalisation is not precisely known, the resulting value of the normalisation mode should not be related to the inclusive $B \to X\ell\nu$ branching fraction.

The fit is performed both, simultaneously and separately for the charged and neutral B samples. It yields a large pull on the $\mathcal{B}(B \to D^*\ell\nu)$ nuisance parameter ($\lambda_{B\to D^*\ell\nu} = 4.9$), which is stronger for the charged B fit ($\lambda_{B\to D^*\ell\nu} = 4.7$) than for the fit to the neutral B sample ($\lambda_{B\to D^*\ell\nu} = 3.7$). This result can be explained by the pre-fit excess of data at large momenta which was already discussed in Sec. 5.3.1. A larger $B \to D^*\ell\nu$ contribution in the X composition matches the data close to the kinematic endpoint. Arguments why $B \to D^*\ell\nu$ decays could be underestimated in inclusive studies have been given previously in Sec. 5.3.1. Additionally, the nuisance parameters of the exclusive branching fractions may also account for non-uniform tagging and selection efficiencies. For this reasons, it has been decided to drop the tight constraint on the $B \to D^*\ell\nu$ branching fraction from the exclusive measurements. In practice this means, that the NP associated with $\mathcal{B}(B \to D^*\ell\nu)$ becomes a free parameter, i. e. assuming a flat prior. Nevertheless, the change in $\mathcal{B}(B \to D^*\ell\nu)$ is still considered in the inclusive $\mathcal{B}(B \to X_c\ell\nu)$ constraint.

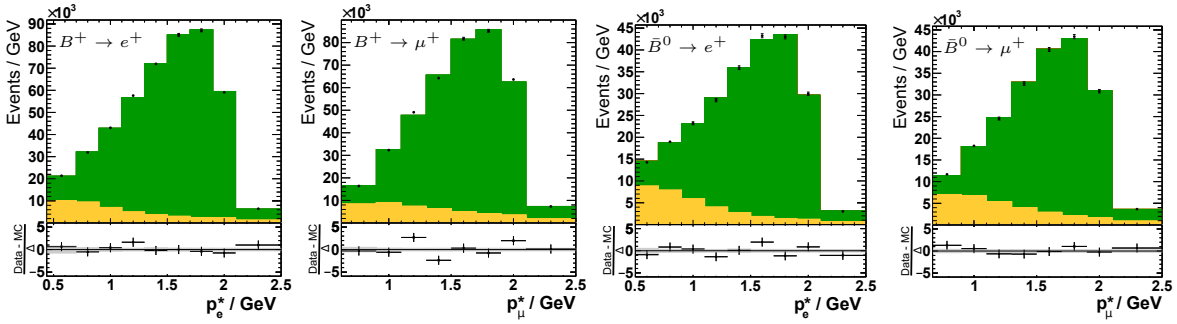


Figure 5.28: Signal electron and muon momentum spectra after the simultaneous fit to the charged (two left) and neural B (two right) samples.

The separate fits to the charged and neutral B samples exhibit different pulls on the exclusive X compositions and the semileptonic $D \to X\ell\nu$ decays (cf. Tab. 5.4).

	λ_j	pull
B^+	$D^+ \to X\ell\nu$	0.18
	$D^0 \to X\ell\nu$	1.20
B^0	$D^+ \to X\ell\nu$	0.75
	$D^0 \to X\ell\nu$	0.35

Table 5.4: Pulls on the nuisance parameters of the semileptonic charm decay in the m^2_{miss} side-band fit.

As long as the rather unknown MC normalisation and efficiencies are uniformly distributed in $|\vec{p}^*_\ell|$ and m^2_{miss}, they are covered by the scaling $A_j^{B\to X\ell\nu}$ in Eq. 5.9 and the nuisance parameters of the exclusive semileptonic branching fractions are the same for charged and neutral B. However, the results show a different set of best-fit NPs for the X composition. This hints at a non-constant efficiency due to e. g. the tag reconstruction which is not the same for charged and neutral B_{tag}. Also, the event topology in terms of charged track and neutral multiplicities is different in B^+B^- and $B^0\overline{B^0}$ events. Thus, the NPs associated with exclusive semileptonic B

decays are separated in the simultaneous fitting procedure to avoid a bias from the different exclusive X compositions. The other nuisance parameters, however, are kept the same for B^+_{tag} and B^0_{tag} because they are found (and expected) not to be significantly different (e. g. lepton efficiency systematic uncertainty or the form factor parameters).

Using the separate $B \to X\ell\nu$ and $D \to X\ell\nu$ NPs and the two $B \to D^*\ell\nu$ parameters, the fits are repeated. The post-fit distributions are shown in Fig. 5.28. As the $B \to D^*\ell\nu$ branching fraction parameter is no longer constrained, the $B \to D^*\ell\nu$ yield further increases. To compare the yields to the constrained $B \to D^*\ell\nu$ nuisance parameters, the unconstrained parameters can be converted to multiples of the experimental uncertainty on $\mathcal{B}(B \to D^*\ell\nu)$ which yields $\lambda_{B\to D^*\ell\nu} = 7.6$ and $\lambda_{B\to D^*\ell\nu} = 6.7$ for charged and neutral B parameters, respectively. Notably, the absolute normalisation of $B \to X\ell\nu$ is very stable and deviates less than 1% between all fits in the small m^2_{miss} side-band. This gives confidence in the fitting procedure as the parametrisation of the normalisation mode barely affects its yield.

To further validate the results of the fitting procedure, especially the impact of the nuisance parameters, the pre- and post-fit distributions of m_X are examined. Fig. 5.29 shows the hadronic mass distribution before and after the fit in $|\vec{p}^*_\ell|$ and m^2_{miss}. After the fit, the data are much better modelled by the MC. The improvement is remarkable because m_X is independent of $|\vec{p}^*_\ell|$ and even though m_X is correlated with m^2_{miss}, the latter is fitted only in three bins. The well modelled m_X distribution shows that the fitted scale factors and NPs, in particular the increased $B \to D^*\ell\nu$ yield, truly improve the data modelling in general, beyond the fitted distributions.

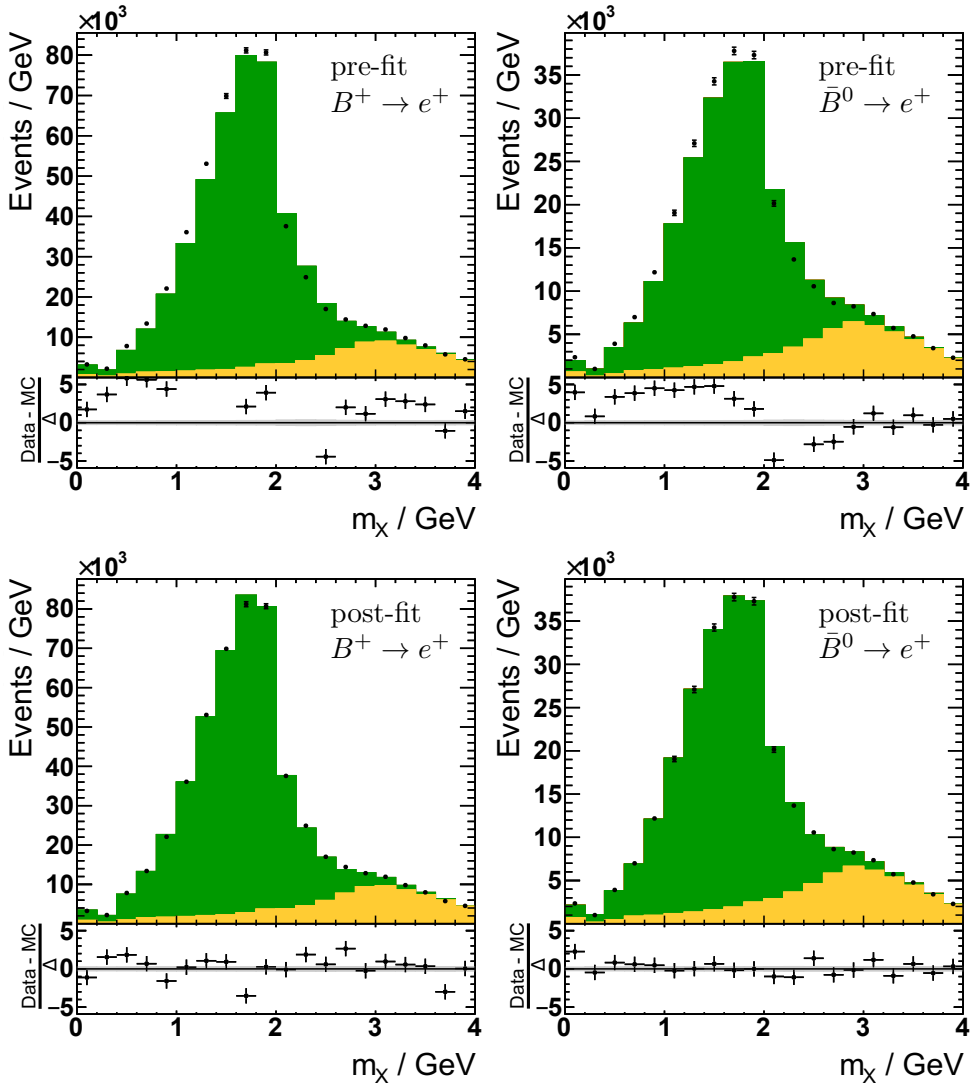


Figure 5.29: Pre-fit (top) and post-fit (bottom) distributions of the hadronic mass m_X in the low m^2_{miss} side-band. Shown are the distributions of the charged (left) and neutral (right) B sample for electrons. Similar plots for the muon sample can be found in Fig. A.5.

CHAPTER 6

The branching fraction of $B \to X\tau\nu$ decays

The final signal extraction is carried out as described in the previous chapter. The fit yields the final result of

$$R(X) = 0.298 \pm 0.012_{\text{stat}} \pm 0.018_{\text{sys}}. \tag{6.1}$$

Using the isospin-average branching fraction $\mathcal{B}(B \to X\ell\nu) = (10.86 \pm 0.16)\%$ [20], the $R(X)$ result is translated to a $B \to X\tau\nu$ branching fraction of

$$\mathcal{B}(B \to X\tau\nu) = (3.23 \pm 0.13_{\text{stat}} \pm 0.20_{\text{sys}} \pm 0.05_{B\to X\ell\nu})\,\%. \tag{6.2}$$

This is the most precise single measurement of $\mathcal{B}(B \to X\tau\nu)$ to date. The obtained value of $R(X)$ is independent of the assumed signal and semileptonic B decay branching fractions as well as the MC normalisation. Post-fit projections for $|\vec{p}^{\,*}_\ell|$ and m^2_{miss} are shown in Fig. 6.1 and slices of the 2D distribution in Figs. 6.2 and A.6. The fit has 244 degrees-of-freedom and converged at $\chi^2_0 = 279.5$ which corresponds to a p-value of 5.9%.

The total uncertainty of 0.022 on $R(X)$ is well consistent with the estimate from pseudo-experiments in Sec. 5.5 and corresponds to a relative uncertainty of 7.3%. The breakdown of the systematic uncertainties for the final signal extraction is shown in Tab. 6.1.

From the obtained results and the tree-level quark decay rate $\Gamma_{b\to q} = |V_{qb}|^2 G_F^2 m_b^5/(192\pi^3)$ as well as $|V_{ub}|/|V_{cb}| = 0.096 \pm 0.007$ [20], the inclusive charmed ratio

$$R(X_c) = 0.308 \pm 0.084 \tag{6.3}$$

is estimated. The SM predicts $R(X_c) = 0.223 \pm 0.005$ [36] and is thus exceeded by the $R(X_c)$ estimated in this work by 3.2σ. Thus, this analysis supports the previously observed enhanced signal yield of $B \to D^{(*)}\tau\nu$ decays at the B-factories. A deeper discussion of the result is given in the next chapter.

The separate signal extraction in the charged and neural B_{tag} samples yields:

$$R(X^0) = 0.31 \pm 0.02 \quad \Rightarrow \quad \mathcal{B}\left(B^+ \to X^0\tau^+\nu\right) = (3.51 \pm 0.29)\,\% \tag{6.4}$$

$$R(X^-) = 0.26 \pm 0.04 \quad \Rightarrow \quad \mathcal{B}\left(\bar{B}^0 \to X^-\tau^+\nu\right) = (2.67 \pm 0.38)\,\%. \tag{6.5}$$

The isospin average $R(X)$ is closer to $R(X^0)$, which reflects the fact that the charged B_{tag} sample is dominant in this analysis as it is nearly twice as large as the neutral sample.

A complete list of the results for the fit parameters is given in Tab. A.8.

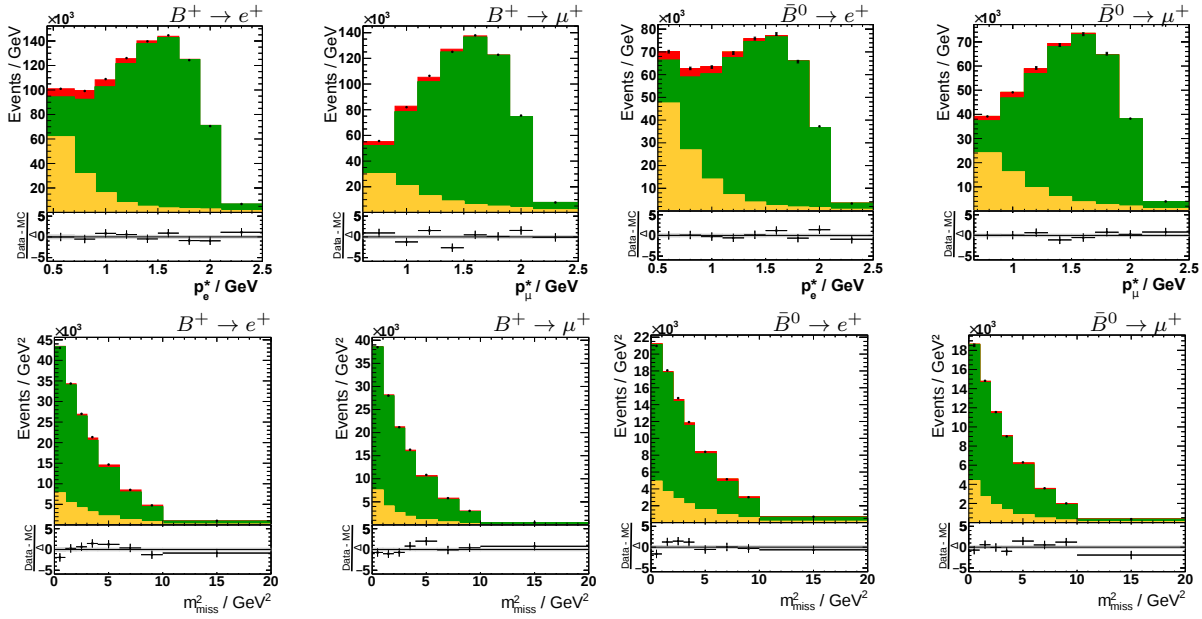


Figure 6.1: Post-fit distributions of the lepton momentum spectrum (top row) and m^2_{miss} (bottom row). Data is shown as black points with error bars, $B \to X\tau\nu$ in red, $B \to X\ell\nu$ in green and other decays in orange. The grey band in the residual plot shows the statistical uncertainties of the MC, lepton efficiency and lepton mis-identification corrections. The shown distributions are 2D projections onto the signal extraction variables.

	Rel. uncertainty $\delta R(X)/\%$
Statistics	± 3.9
PID	± 1.1
$\mathcal{B}(B \to X_c\tau\nu)$ composition	± 0.6
$\mathcal{B}(B \to D\ell\nu)$	± 0.6
$\mathcal{B}(B \to D^*\ell\nu)$	$^{+4.9}_{-4.3}$
$\mathcal{B}(B \to D^{**}\ell\nu)$ composition	± 3.0
$\mathcal{B}(D \to X\ell\nu)$	± 3.3
D^{**} decay model	± 0.5
$\text{FF}_{\text{CLN}}(B \to D^{(*)}\ell\nu)$	± 0.6
$\text{FF}_{\text{LLSW}}(B \to D^{**}\ell\nu)$	$^{+4.6}_{-4.2}$
MC statistics	± 1.9
Total systematics	$^{+6.6}_{-6.3}$
Total	$^{+7.7}_{-7.4}$

Table 6.1: Relative uncertainties on $R(X)$. *Composition* means the summary of the contributions from the individual exclusive decays. Due to correlations, the individual uncertainties do not necessarily sum up in quadrature to the total uncertainty.

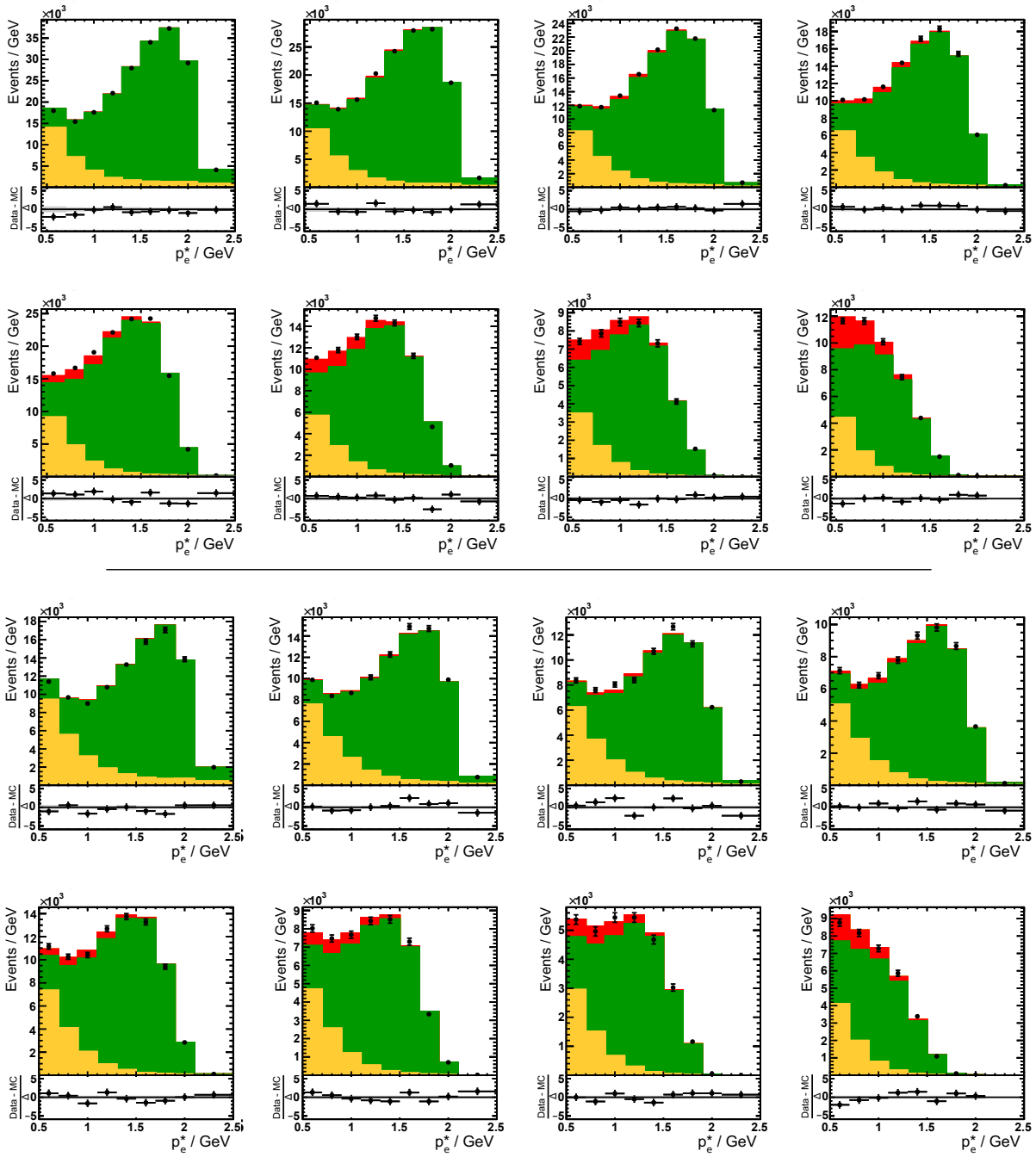


Figure 6.2: Post-fit electron momentum spectra in bins of m^2_{miss} of the nominal, simultaneous fit. The upper two rows show the charged B and the lower two rows the neutral B sample. The corresponding distributions for the muon final state are given in Fig. A.6. The colours have the same meaning as in Fig. 6.1.

CHAPTER 7

Conclusions and Outlook

The inclusive branching faction of $B \to X\tau\nu$ obtained in this analysis exceeds the SM prediction by 3.2σ as well as the measurements performed at LEP by 2.3σ, as shown in Fig. 7.1. This measurement is compatible with the findings of the B-factories for exclusive measurements of $B \to D^{(*)}\tau\nu$, respectively $R(D^{(*)})$. Hence, this analysis continues the path of the B-factories to measure enhanced branching fractions in the τ sector of semileptonic B decays. The sum of exclusive decays yields a lower bound on the inclusive branching fraction ratio of $R(D + D^*) = R(X)_{\text{min}} = 0.238 \pm 0.046$. In contrast to $R(X)_{\text{LEP}} = 0.233 \pm 0.72$ derived from the LEP measurements, the lower bound is well exceeded by this analysis as also visualised in Fig. 7.1.

The sum of exclusive branching fractions $B \to D^{(*)}\tau\nu$ slightly exceeds the inclusive LEP measurements such, that $B \to D^{**}\tau\nu$ decays are suppressed as $\mathcal{B}(B \to D^{**}\tau\nu) \leq 0.3\%$ at 90% confidence level. However, from the difference of the SM predictions of the inclusive $B \to X\tau\nu$ and the exclusive $B \to D^{(*)}\tau\nu$ branching fractions, a prediction for the $B \to D^{**}\tau\nu$ branching fraction is inferred: $\mathcal{B}(B \to D^{**}\tau\nu) = (0.37 \pm 0.11)\%$. The $B \to X\tau\nu$ branching fraction obtained in this work results in a branching fraction of $\mathcal{B}(B \to D^{**}\tau\nu) = (0.62 \pm 0.29)\%$. Although this estimate is not very significant, the value is compatible with the SM prediction in contrast to the estimate obtained from the LEP measurements.

Background modelling The post-fit values of the nuisance parameters, including the parameters associated to the $B \to D^*\ell\nu$ decay of charged and neutral B mesons, are shown in Fig. A.3. The latter is an important background parameter in this analysis. As already observed in the side-band, the fit yields a high value to match the excess of data at large signal lepton momenta. Also the nuisance parameters assigned to the $B \to D^{(*)}\ell\nu$ form-factor parameters are found to prefer a harder lepton spectrum, which is also observed in the side-band. The value of the unconstrained $B \to D^*\ell\nu$ parameter for the charged B sample of the simultaneous fit is similar to the value obtained in the side-band fit (cf. Sec. 5.6). It is six times larger than the experimental uncertainty on $\mathcal{B}(B \to D^*\ell\nu)$, i. e. $\lambda_{B\to D^*\ell\nu} = 6$. The values found in the neutral B sample are slightly lower. The uncertainties on the $\mathcal{B}(B \to D^*\ell\nu)$ parameters of the charged and neutral B sample are small ($\sim$1.5%) as they solely account for the large lepton momenta and are part of the inclusive $\mathcal{B}(B \to X\ell\nu)$ constraint (cf. Eq. 5.15).

As for the side-band fit, the post-fit distribution of the hadronic mass m_X is used to check if the fit is able to improve the data-MC agreement in this rather uncorrelated observable. The resulting distributions are given in Figs. 7.2 and A.8 for the electron and muon sample,

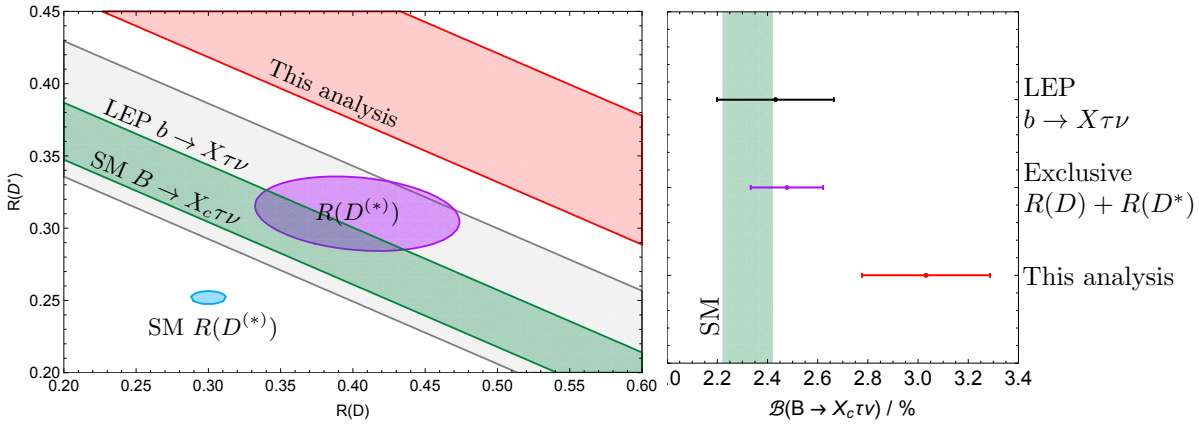


Figure 7.1: $R(X)$ in comparison to $R(D^{(*)})$ and $B \to X\tau\nu$ measurements. The semileptonic branching fractions of Tab. 3.2 are used to convert the measurements appropriately. *Left:* Measurements of exclusive $B \to D^{(*)}\tau\nu$ and inclusive $B \to X\tau\nu$ decays in the $R(D)$–$R(D^*)$ plane. The ellipses represent the 68% contours of the $R(D^{(*)})$ SM prediction (light blue) and measurements (purple). The ellipse of measured $R(D^{(*)})$ are the averaged results by *BABAR*, Belle an LHCb (cf. Fig. 1.6). The inclusive LEP measurements (grey band) match well the SM prediction (green band) which are both exceeded by this analysis (red band). *Right:* The result of this study and previous related measurements in terms of the isospin-averaged $B \to X_c\tau\nu$ branching fraction, based on Refs. [20, 59].

respectively. They show a well modelling of the data. Even the region of the D^* mass peak around $m_{D^*} \approx 2.0\,\text{GeV}$ shows a reasonable modelling and not, as one could expect from the large pull on the $B \to D^*\ell\nu$ parameters, an excess of MC. This shows, that the large value of $B \to D^*\ell\nu$ is not simply preferred to fill up the data excess in the lepton spectrum, but it is compatible with the hadronic mass system and clearly improves the data modelling.

The impact of the D^{**} model branching fractions is fortunately very small. In the analysis, a relative uncertainty of 5% on the individual D^{**} model branching fractions is assumed. As a test, the fit is repeated with 10% uncertainties to check the impact of this assumption. However, only a negligible increase of the uncertainty on $R(X)$ is observed. Other nuisance parameters are of minor importance and give only small contributions to the total uncertainty.

The full correlation matrix of the nominal fit is depicted in Fig. A.7. For the signal parameter $R(X)$, the correlations stay well below 45% and are largest for the $B \to D^*\ell\nu$ parameters. A few nuisance parameters are highly correlated (e. g. the D^{**} decays among themselves) which means that their impact on the modelling is very similar to the fit.

Implication on New Physics The measured $R(X)$ can be further investigated in view of New Physics. Assuming that New Physics do not significantly change the signal efficiency, the 2HDM-II prediction is confronted with the measured $R(X)$. The former is not a strong assumption because the lepton momentum and missing mass are barely sensitive to the specific decay kinematics. Using the 2HDM-II prediction of Ref. [51] (cf. Sec. 1.4 and Fig. 1.7), the 2HDM-II parameter $r = \tan\beta/m_H$ is obtained to be $r = (0.47 \pm 0.05)\,\text{GeV}^{-1}$. This value is compatible with the findings for $R(D)$ and $R(D^*)$ of Belle's hadronic tag analysis which states to be compatible with $r \sim 0.45\,\text{GeV}^{-1}$ [14]. The value of r derived in this work is also compatible with the value obtained from the $R(D)$ measurement by *BABAR* [13] of $r = (0.44 \pm 0.02)\,\text{GeV}^{-1}$. However, it is not compatible with $r = 0.75 \pm 0.05\,\text{GeV}^{-1}$ which *BABAR* reports for their $R(D^*)$ measurement. However, the latter is generally incompatible with the $R(D)$ measurements in the 2HDM-II framework.

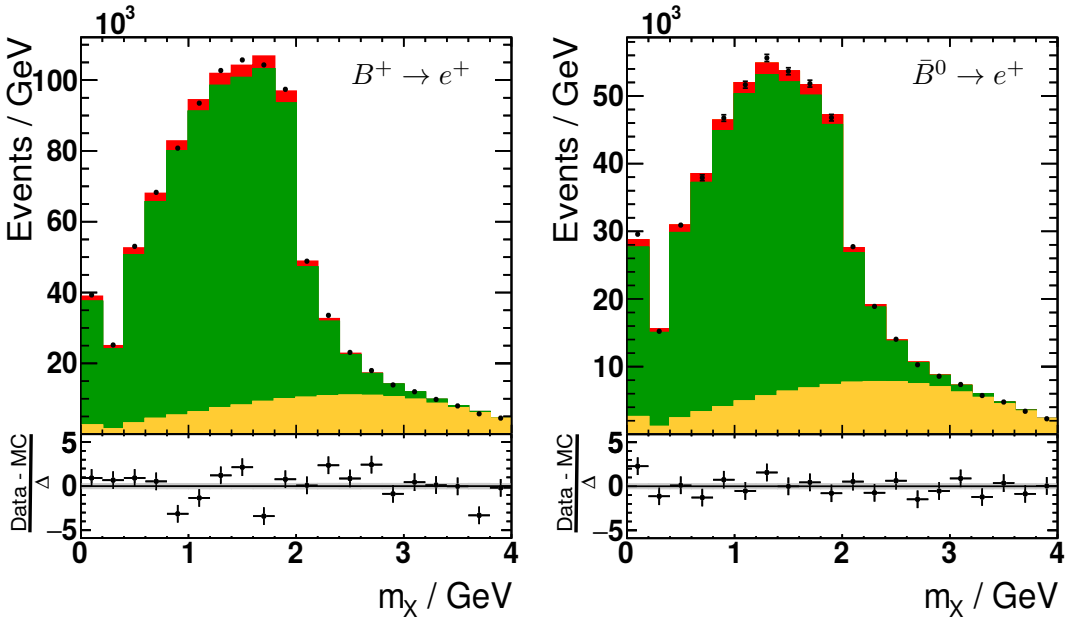


Figure 7.2: The post-fit hadronic mass spectrum m_X for charged (left) and neutral (right) B mesons and the electron signal lepton. The corresponding plot for the muon final state is shown in Fig. A.8. The data is well modelled and even the D^* peak region at $m_{D^*} \approx 2\,\text{GeV}$ shows a good data-MC agreement. Note, no fit is performed in m_X but only in $|\vec{p}^{\,*}_\ell|$ and m^2_{miss}.

Outlook A considerable part of this analysis falls upon the investigation and understanding of the signal extraction variables. Because of the *inclusive* selection, there are only a few uncorrelated observables that also provide sufficient power to discriminate signal and background. Hence, the choice of signal extraction variables is very limited which, on the one hand, makes it easier to find an optimal selection. But, on the other hand, the MC modelling is more difficult as a small mis-modelling can become statistically significant due to the large sample size. The modelling of the lepton momentum and especially the squared missing mass are challenging.

Using the set-up and results of this analysis, further aspects can be studied in the future. From the theoretical point of view, the unfolded differential decay rate is interesting to constrain the parameters of the form-factor expansion. A measurement of $\mathrm{d}\Gamma/\mathrm{d}q^2$ is a recent ongoing side project in the Belle working group in Bonn and is based on this analysis. The momentum transfer q^2 is estimated from the reconstructed four-momentum of the hadronic system as $q^\mu = p^\mu_{B_{\text{signal}}} - p^\mu_X$ which is challenging due to the low resolution of p^μ_X.

At the beginning of 2018, Belle II, the successor of the Belle experiment, will start its first data-taking period (*Phase II*), however, yet without the new vertex detector which will be installed a year later, before the main data taking (*Phase III*) starts in 2019. A huge data set of $50\,\text{ab}^{-1}$ will be recorded until the end of 2024 [87]. Not only the statistical precision will improve, but also the detector underwent a major upgrade. In particular, the improved tracking and PID system will help to measure the branching fraction of $B \to X\tau\nu$ more precisely. The new pixel detector, installed in addition to the SVD, is located even closer to the beam-pipe than the SVD in Belle. It thus allows to precisely track and vertex charged particles which will also be beneficially to inclusive analyses.

Even though the overall improvement can only be roughly assessed without a full detector simulation, the impact of the larger data set can be estimated. The latter does not only reduce

the statistical uncertainty on the data, but also allows one to study corrections to the MC simulation, e. g. efficiency corrections, with higher precision. Fig. 7.3 shows a projection of the relative uncertainty on $R(X)$ as a function of the analysed integrated luminosity. For this projection, statistical uncertainties have been scaled according to the increased data set. This includes also the statistical uncertainties on the efficiency corrections. The projection shows that the increasing significance of the measurement and the uncertainty on $R(X)$ falls below 5% already for $4\,\mathrm{ab}^{-1}$ of analysed data. For the complete Belle II data set, the systematic uncertainties are the limiting factor and the relative uncertainty reaches ∼4.2%. Theoretical and experimental progress can decrease the systematic uncertainties even further, which is not considered at this point.

The detector and physics modelling always plays a crucial role. This work shows that high-statistics analyses can reveal issues in the data modelling that are rather unnoticed by exclusive analyses with their relatively low sample sizes. In case of the huge Belle II data set, this will become even more relevant. Therefore, from the beginning, an effort should be put into the understanding and modelling of fundamental observables such as the lepton momentum and the squared missing mass. This work can give useful guidance and serve as a starting point for Belle II measurements of $R(X)$, more detailed studies of the data modelling as well as the modelling of charmed meson decays for which the model developed in this work can be adopted for future precision measurements.

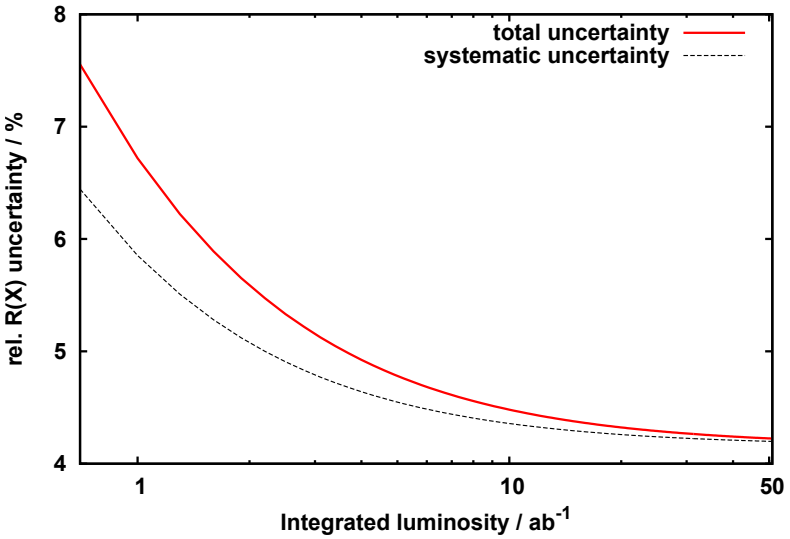


Figure 7.3: Projection of the relative uncertainty on $R(X)$, as a function of integrated luminosity for the Belle II experiment. The statistical and some of the systematic uncertainties scale with the analysed luminosity. With the full data set of $50\,\mathrm{ab}^{-1}$, Belle II will reach an uncertainty of ∼4.2%.

CHAPTER 8

Summary

In this thesis, the first measurement of the inclusive branching fraction of $B \to X\tau\nu$ decays at a B-factory is presented.

While the inclusive branching fraction of the b-hadron admixture $b \to X\tau\nu$ has been measured at the LEP experiments to be consistent with the SM expectation, exclusive analyses of $B \to D^{(*)}\tau\nu$ decays, performed at the B-factory experiments *BABAR*, Belle and LHCb, together find a 4σ tension with the SM, which might be a hint for New Physics. Therefore, this inclusive analysis is of particular interest and a valuable cross check of the exclusive B-factory measurements, and also tests the inclusive branching fraction measurements performed at LEP.

The analysis is performed with the large Belle data set of 770×10^6 $B\overline{B}$ pairs, recorded at the $\Upsilon(4S)$ resonance. The large size of the data-set allows to obtain a clean sample of hadronically tagged B meson events, where one of the two B mesons is fully reconstructed in a hadronic mode. Due to the clean experimental environment of e^+e^- collisions, the other B meson in the event is separately studied with respect to the signal decay.

Signal events are selected in the leptonic decay channels of the τ by requiring a single electron or muon in the final state which leads to three undetectable neutrinos. Furthermore, the decay chain of the hadronic system X is not reconstructed to perform an *inclusive* analysis. Together, these two aspects make the inclusive analysis of the $B \to X\tau\nu$ decay an exceptional challenge as this particular final state provides only few kinematic constraints and a limited set of independent observables.

As the inclusive decays of $B \to X\tau\nu$ and $B \to X\ell\nu$ are modelled as a sum of exclusive modes, the modelling of semileptonic B decays into D^{**} and the decays of the D^{**} mesons themselves are revised for this analysis. Especially the $B \to D^{**}\ell\nu$ decays can significantly populate the signal region and it is thus crucial that they are well modelled. A model of hadronic 1P and 2S charmed meson decays is developed by using constraints from heavy-quark spin symmetry and phase-space calculations. Besides the few measured decay modes of the D^{**} mesons, new but yet unmeasured modes are added. These modes also help to reduce the observed gap between the inclusive and the sum of exclusive semileptonic B branching fraction measurements. The new model is implemented in a dedicated MC sample which is used in this analysis.

The lepton momentum $|\vec{p}^*_\ell|$ and the squared missing mass m^2_{miss} are used to extract the signal and constrain backgrounds. The modelling of both observables is intensively studied in this thesis. The lepton momentum spectrum shows an excess of data at high values, which was also observed in previous inclusive analyses, at *BABAR* and Belle. For m^2_{miss}, the tail towards negative values is underestimated by the MC which was subject of an extensive study. As this

region is far away from the signal region and the processes behind this issue are complex, the analysis rejects events in the negative m^2_{miss} region.

The signal contribution is determined in a 2D χ^2 fit in $|\vec{p}^{\,*}_\ell|$ and m^2_{miss}. A large set of nuisance parameters is introduced to incorporate systematic effects that are constrained by external measurements. The fit is tested with pseudo-data and using the side-band of small m^2_{miss}. The side-band fit yields a significant excess of $B \to D^*\ell\nu$ decays with respect to the expectation, which can be explained by the previously mentioned excess in the lepton momentum spectrum. Therefore, to avoid a potential bias, $B \to D^*\ell\nu$ is a free parameter in the final fit.

This analysis measures the signal branching fraction, normalised to $\mathcal{B}(B \to X\ell\nu)$ with $\ell = e, \mu$, to be

$$R(X) = \frac{\mathcal{B}(B \to X\tau\nu)}{\mathcal{B}(B \to X\ell\nu)} = 0.298 \pm 0.012_{\text{stat}} \pm 0.018_{\text{sys}} \tag{8.1}$$

which translates to a branching fraction of

$$\mathcal{B}(B \to X\tau\nu) = (3.23 \pm 0.24)\,\%, \tag{8.2}$$

assuming the isospin-average branching fraction $\mathcal{B}(B \to X\ell\nu) = (10.86 \pm 0.16)\%$. This measurement is, so far, the most precise single measurement of $B \to X\tau\nu$ and the first measurement of this inclusive branching fraction at a B-factory.

In the future, the Belle II experiment will record a fifty times larger data set with an improved detector. A projection to this large data set is obtained and shows that Belle II can achieve a relative uncertainty of $\sim$4.2% on the branching fraction ratio $R(X)$.

APPENDIX A

Appendix

A.1 LLSW model

Decay rates and form-factor expansions of the LLSW model [34] (cf. Sec. 1.2.2).

Differential decay rates

$$\frac{1}{\Gamma_b}\frac{\mathrm{d}^2\Gamma(B\to D_2^*\ell\nu)}{\mathrm{d}w\mathrm{d}\cos\theta} = \frac{3}{2}r^3(w^2-1)^{3/2}\left\{\frac{4}{3}\sin^2\theta\left[(w-r)k_{A_1}+(w^2-1)(k_{A_3}+rk_{A_2})\right]^2 \right.$$
$$\left. +(1-2rw+r^2)\left[(1+\cos^2\theta)[k_{A_1}^2+(w^2-1)k_V^2]-4\cos\theta\sqrt{w^2-1}k_{A_1}k_V\right]\right\} \tag{A.1}$$

$$\frac{1}{\Gamma_b}\frac{\mathrm{d}^2\Gamma(B\to D_0^*\ell\nu)}{\mathrm{d}w\mathrm{d}\cos\theta} = 3r^3(w^2-1)^{3/2}\sin^2\theta\left[(1+r)g_+-(1-r)g_-\right]^2 \tag{A.2}$$

$$\frac{1}{\Gamma_b}\frac{\mathrm{d}^2\Gamma(B\to D_1'\ell\nu)}{\mathrm{d}w\mathrm{d}\cos\theta} = 3r^3\sqrt{w^2-1}\left\{\sin^2\theta\left[(w-r)g_{V_1}+(w^2-1)(g_{V_3}+rg_{V_2})\right]^2 \right.$$
$$\left. +(1-2rw+r^2)\left[(1+\cos^2\theta)[g_{V_1}^2+(w^2-1)g_A^2]-4\cos\theta\sqrt{w^2-1}g_{V_1}g_A\right]\right\} \tag{A.3}$$

Form-factors of $B\to D_2^*\ell\nu$

$$\begin{aligned}
k_V =& -\tau-\epsilon_b[(\bar\Lambda'+\bar\Lambda)\tau-(2w+1)\tau_1-\tau_2+\eta_b]-\epsilon_c(\tau_1-\tau_2+\eta_{\mathrm{ke}}-2\eta_1+\eta_3),\\
k_{A_1} =& -(1+w)\tau-\epsilon_b\{(w-1)(\bar\Lambda'+\bar\Lambda)\tau-(2w+1)\tau_1-\tau_2]+(1+w)\eta_b\}\\
&-\epsilon_c[(w-1)(\tau_1-\tau_2)+(w+1)(\eta_{\mathrm{ke}}-2\eta_1+\eta_3)],\\
k_{A_2} =& -2\epsilon_c(\tau_1+\eta_2),\\
k_{A_3} =& \tau+\epsilon_b[(\bar\Lambda'+\bar\Lambda)\tau-(2w+1)\tau_1-\tau_2+\eta_b]-\epsilon_c(\tau_1+\tau_2-\eta_{\mathrm{ke}}+2\eta_1-2\eta_2-\eta_3).
\end{aligned} \tag{A.4}$$

Form-factors of $B \to D_0^* \ell\nu$

$$
\begin{aligned}
g_+ =& \epsilon_c \left[2(w-1)\zeta_1 - 3\zeta \frac{w\bar{\Lambda}^* - \bar{\Lambda}}{w+1} \right] - \epsilon_b \left[\frac{(2w+1)\bar{\Lambda}^* - (w+2)\bar{\Lambda}}{w+1} \zeta - 2(w-1)\zeta_1 \right], \\
g_- =& \zeta + \epsilon_c \left[\chi_{\mathrm{ke}} + 6\chi_1 - 2(w+1)\chi_2 \right] + \eta_b \chi_b .
\end{aligned} \tag{A.5}
$$

Form-factors of $B \to D_1' \ell\nu$

$$
\begin{aligned}
g_A =& \zeta + \epsilon_c \left[\frac{w\bar{\Lambda}^* - \bar{\Lambda}}{w+1} \zeta + \chi_{\mathrm{ke}} - 2\chi_1 \right] - \epsilon_b \left[\frac{(2w+1)\bar{\Lambda}^* - (w+2)\bar{\Lambda}}{w+1} \zeta - 2(w-1)\zeta_1 - \chi_b \right] \\
g_{V_1} =& (w-1)\zeta + \epsilon_c \left[(w\bar{\Lambda}^* - \bar{\Lambda})\zeta + (w-1)(\chi_{\mathrm{ke}} - 2\chi_1) \right] \\
& - \epsilon_b \left\{ [(2w+1)\bar{\Lambda}^* - (w+2)\bar{\Lambda}]\zeta - 2(w^2-1)\zeta_1 - (w-1)\chi_b \right\}, \\
g_{V_2} =& 2\epsilon_c(\zeta_1 - \chi_2), \\
g_{V_3} =& -\zeta - \epsilon_c \left[\frac{w\bar{\Lambda}^* - \bar{\Lambda}}{w+1} \zeta + 2\zeta_2 + \chi_{\mathrm{ke}} - 2\chi_1 + 2\chi_2 \right] \\
& - \epsilon_b \left[\frac{(2w+1)\bar{\Lambda}^* - (w+2)\bar{\Lambda}}{w+1} \zeta - 2(w-1)\zeta_1 - \chi_b \right]
\end{aligned} \tag{A.6}
$$

$$
\begin{gathered}
\tau(1) \sim 1 \\
-2 < \tau'(1)/\tau(1) < -1 \\
\bar{\Lambda} = 0.4\,\mathrm{GeV} \\
\bar{\Lambda}' = 0.79\,\mathrm{GeV} \\
\bar{\Lambda}^* = 0.75\,\mathrm{GeV} \\
m_b = 4.8\,\mathrm{GeV},\ m_c = 1.4\,\mathrm{GeV} \\
\eta_1 = \eta_2 = \eta_3 = \chi_1 = \chi_2 = 0 \\
-0.5 < \eta_{\mathrm{ke}} = \zeta_{\mathrm{ke}} < 0.5 \\
\eta_b = \eta_{\mathrm{ke}},\ \chi_b = \zeta_{\mathrm{ke}}
\end{gathered}
$$

Table A.1: Parameters and parameter ranges of the LLSW model used in this analysis. The values and ranges are taken from Ref. [34].

A.2 Updated branching fractions

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$B^+ \to D\tau^+\nu$	0.77	0.782 ± 0.078
$B^+ \to D^*\tau^+\nu$	1.54	1.520 ± 0.201
$B^+ \to D_1\tau^+\nu$	0.0013	0.077
$B^+ \to D_2^*\tau^+\nu$	0.0008	0.025
$B^+ \to D_0^*\tau^+\nu$	0.0016	0.039
$B^+ \to D_1'\tau^+\nu$	0.0012	0.079
$B^+ \to D_{2S}\tau^+\nu$	-	0.002
$B^+ \to D_{2S}^*\tau^+\nu$	-	0.005
$B^+ \to X_c\tau^+\nu$	2.3149	2.53 ± 0.4

Table A.2: Assumed branching fractions for the exclusive signal decays. The $B \to D^{(*)}\tau\nu$ branching fractions are based on the SM predictions of $R(D)$ and $R(D^*)$ and all other branching fractions are based on Chapter 4. For the latter, a 50% relative uncertainty is assumed.

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$D^+ \to K^0 e\nu$	6.8	8.83
$D^+ \to K^{*0} e\nu$	3.1	5.53
$D^+ \to K_2^{*0} e\nu$	0.9	0.15
$D^+ \to K\pi e\nu$	1.0	0.18
$D^+ \to K^0\pi^0 e\nu$	1.0	0.18
$D^+ \to \pi^0 e\nu$	0.31	0.405
$D^+ \to \eta e\nu$	0.3	0.114
$D^+ \to \eta' e\nu$	0.05	0.022
$D^+ \to \rho^0 e\nu$	0.31	0.218
$D^+ \to \omega e\nu$	0.31	0.182
$D^+ \to X e\nu$	14.08	16.07 ± 0.3

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$D^+ \to K^0 \mu\nu$	6.8	9.4
$D^+ \to K^{*0} \mu\nu$	3.1	5.53
$D^+ \to K_2^{*0} \mu\nu$	0.9	0.15
$D^+ \to K\pi \mu\nu$	1.0	0.18
$D^+ \to K^0\pi^0 \mu\nu$	1.0	0.18
$D^+ \to \pi^0 \mu\nu$	0.31	0.405
$D^+ \to \eta \mu\nu$	0.3	0.114
$D^+ \to \eta' \mu\nu$	0.05	0.022
$D^+ \to \rho^0 \mu\nu$	0.31	0.218
$D^+ \to \omega \mu\nu$	0.31	0.182
$D^+ \to \mu\nu$	0.08	0.0374
$D^+ \to X \mu\nu$	14.16	17.6 ± 3.2

Table A.3: Updated branching fraction of semileptonic decays of charged D mesons. The branching fraction is scaled such, that it matches in inclusive branching fraction.

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$D^0 \to Ke\nu$	3.41	3.57
$D^0 \to K^* e\nu$	2.07	2.16
$D^0 \to K_2^* e\nu$	0.18	0.18
$D^0 \to K\pi^0 e\nu$	0.22	0.12
$D^0 \to K^0\pi e\nu$	0.22	0.17
$D^0 \to \pi e\nu$	0.26	0.289
$D^0 \to \rho e\nu$	0.19	0.177
$D^0 \to b_1 e\nu$	0.05	0.05
$D^0 \to Xe\nu$	6.6	6.49 ± 0.11

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$D^0 \to K\mu\nu$	3.41	3.33
$D^0 \to K^*\mu\nu$	2.07	1.92
$D^0 \to K_2^*\mu\nu$	0.18	0.18
$D^0 \to K\pi^0\mu\nu$	0.22	0.12
$D^0 \to K^0\pi\mu\nu$	0.22	0.17
$D^0 \to \pi\mu\nu$	0.26	0.289
$D^0 \to \rho\mu\nu$	0.19	0.177
$D^0 \to b_1\mu\nu$	0.05	0.05
$D^0 \to X\mu\nu$	6.6	6.7 ± 0.6

Table A.4: Updated branching fraction of semileptonic decays of neutral D-mesons. The branching fraction is scaled such, that it matches in inclusive branching fraction.

	$\mathcal{B}_{\text{old}}/\%$	$\mathcal{B}_{\text{new}}/\%$
$B^0 \to D^+ D_s^-$	0.65	0.72
$B^0 \to D^{*+} D^{*0} K^-$	1.18	1.06
$B^0 \to D^+ a_1^-$	0.75	0.6
$B^+ \to D^{*0}\rho^+$	1.55	1.0
$B^+ \to D^0 3\pi^\pm$	0.46	0.5
$B^+ \to D^0 D_s^+$	1.0	0.9
$B^+ \to D^{*0} D^{*+} K^0$	0.78	0.92

Table A.5: Updated branching fractions of hadronic B meson decays.

A.3 Excited charmed meson decays

	$\mathcal{B}(D_0^*)$	$\mathcal{B}(D_1')$	$\mathcal{B}(D_1)$	$\mathcal{B}(D_2^*)$	$\mathcal{B}(D_{2S})$	$\mathcal{B}(D_{2S}^*)$	$\mathcal{B}\left(B^+ \to X_i \ell\nu\right)$
$D\pi$	75.6			56.6		11.6	0.638
$D^*\pi$		60.6	55.8	36.8	56.1	36.2	0.868
$D\eta$	16.8			6.1		4.7	0.126
$D^*\eta$		17.2			10.3	25.1	0.116
D_sK						9.0	0.018
D_s^*K					6.4	5.5	0.017
$D\rho$		4.5			8.6		0.023
$D^*\rho$		4.9				1.1	0.018
$D\omega$		4.4			7.9		0.022
$D^*\omega$		7.2				1.1	0.025
$D\pi\pi$		0.6	31.8	0.4	0.6	0.4	0.248
$D^*\pi\pi$	7.6	0.5	12.4	0.1	0.4	0.2	0.141
$D_1\pi$						1.7	0.003
$D_2^*\pi$					3.3	1.6	0.006
$D_0^*\pi$					6.5		0.006
$D_{2S}\pi$						0.7	0.001
$D_{2S}^*\gamma$						1.2	0.002

Table A.6: Estimated branching fractions of excited charmed mesons. The inclusive branching fraction in the last column gives the contribution of the particular final state to the semileptonic B decays. This value is derived in combination with the $B \to D^{**}\ell\nu$ branching fractions from Tab. 4.4. All branching fractions are given in percent.

EvtGen module	Example decay mode
`ISGW2`	$B \to D^{**}\ell\nu$
`VVS_PWAVE`	$D_1 \to D^*\pi$
`VSP_PWAVE`	$D^*_{2S} \to D_{2S}\gamma$
`TVS_PWAVE`	$D^*_2 \to D^*\pi$
`HELAMP`	replaces `VVS_PWAVE` for $L = 1$
`SVS`	$D_{2S} \to D^*\pi$
`STS`	$D_{2S} \to D^*_2\pi$
`VSS`	$D^*_{2S} \to D\pi$
`TSS`	$D^*_2 \to D\pi$
`PHSP`	$D^{**} \to D^{(*)}\pi\pi$

Table A.7: EvtGen modules used in the D^{**} MC generation. The modules are chosen to match the initial and final state particles in terms of their total momentum $J = 0, 1, 2$ as `S` (scalar), `V` (vector) and `T` (tensor), respectively. The given decays are examples for those that match the same structure. The modules named "`PWAVE`" allow to specify the partial wave of the decay. `VVS_PWAVE` is not implemented for $L = 1$ in the Belle MC and replaced by `HELAMP` in the style of the generic MC. Details on these modules given in [74].

A.4 Event selection

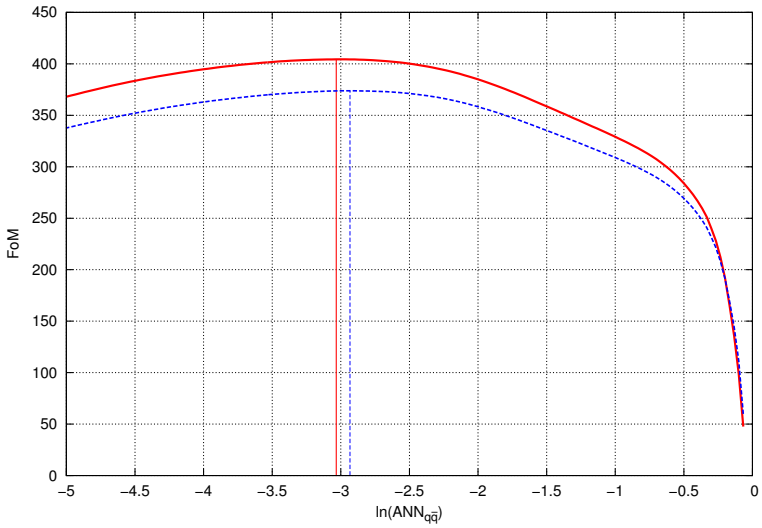


Figure A.1: Figure of Merit for the optimal selection on the hadronic tagging classifier output $\ln(\text{ANN}_{q\bar{q}})$. Maximised is $S/\sqrt{S+B}$, with S and B the event yield of well and poorly reconstructed B_{tag}, respectively. The optimisation is done separately for charged (solid red line) and neutral (dashed blue line) B_{tag}. However, a consistent value of -3 seams to be the optimal selection.

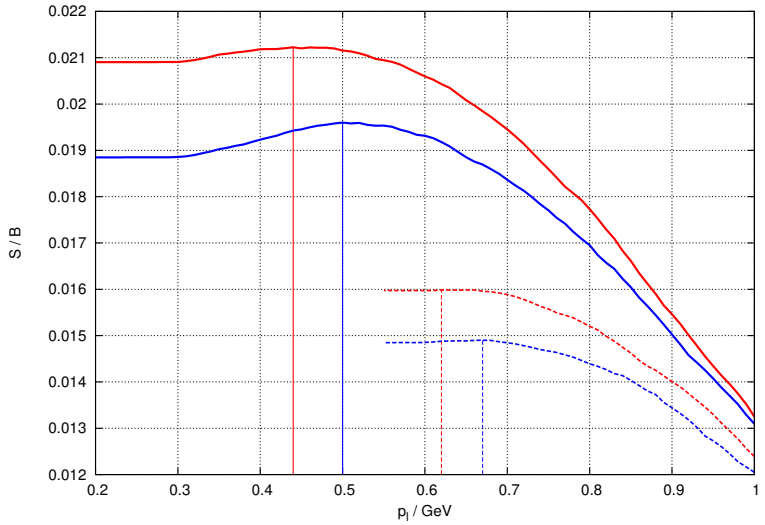


Figure A.2: Optimisation of the signal to background ratio in $|\vec{p}^{\,*}_{\ell}|$ to reduce the non-semileptonic component which has a even softer lepton momentum spectrum. This is done separately for electrons (solid lines) and muons (dashed lines) as well as charged (red lines) and neutral (blue lines) B_{tag}.

A.5 Nuisance parameters

1	$\gamma \to e^+e^-$
2	$B \to D\tau\nu$
3	$B \to D^*\tau\nu$
4	$B \to D_1\tau\nu$
5	$B \to D_2^*\tau\nu$
6	$B \to D_0^*\tau\nu$
7	$B \to D_1'\tau\nu$
8	$B \to D_{2S}\tau\nu$
9	$B \to D_{2S}^*\tau\nu$
10	$B^+ \to D\ell\nu$
11	$B^+ \to D^*\ell\nu$
12	$B^+ \to D_1\ell\nu$
13	$B^+ \to D_2^*\ell\nu$
14	$B^+ \to D_0^*\ell\nu$
15	$B^+ \to D_1'\ell\nu$
16	$B^+ \to D\pi\ell\nu$
17	$B^+ \to D_{2S}\ell\nu$
18	$B^+ \to D_{2S}^*\ell\nu$
19	$B^0 \to D\ell\nu$
20	$B^0 \to D^*\ell\nu$
21	$B^0 \to D_1\ell\nu$
22	$B^0 \to D_2^*\ell\nu$
23	$B^0 \to D_0^*\ell\nu$
24	$B^0 \to D_1'\ell\nu$
25	$B^0 \to D\pi\ell\nu$
26	$B^0 \to D_{2S}\ell\nu$
27	$B^0 \to D_{2S}^*\ell\nu$
28	$B \to X_u\ell\nu$
29	$D_0^* \to D\pi$
30	$D_0^* \to D^*\pi\pi$
31	$D_0^* \to D\eta$
32	$D_1' \to D^*\pi$
33	$D_1' \to D\pi\pi$
34	$D_1' \to D^*\pi\pi$
35	$D_1' \to D^*\eta$
36	$D_1' \to D\rho$
37	$D_1' \to D\omega$
38	$D_1' \to D^*\rho$
39	$D_1' \to D^*\omega$
40	$D_1 \to D^*\pi$
41	$D_1 \to D\pi\pi$
42	$D_1 \to D^*\pi\pi$
43	$D_2^* \to D\pi$
44	$D_2^* \to D\pi\pi$
45	$D_2^* \to D^*\pi$
46	$D_2^* \to D^*\pi\pi$
47	$D_2^* \to D\eta$
48	$D_{2S} \to D^*\pi$
49	$D_{2S} \to D\pi\pi$
50	$D_{2S} \to D^*\pi\pi$
51	$D_{2S} \to D_0^*\pi$
52	$D_{2S} \to D^*\eta$
53	$D_{2S} \to D_2^*\pi$
54	$D_{2S} \to D_s^*K$
55	$D_{2S} \to D\rho$
56	$D_{2S} \to D\omega$
57	$D_{2S}^* \to D\pi$
58	$D_{2S}^* \to D\pi\pi$
59	$D_{2S}^* \to D^*\pi$
60	$D_{2S}^* \to D^*\pi\pi$
61	$D_{2S}^* \to D\eta$
62	$D_{2S}^* \to D_sK$
63	$D_{2S}^* \to D^*\eta$
64	$D_{2S}^* \to D_1\pi$
65	$D_{2S}^* \to D_2^*\pi$
66	$D_{2S}^* \to D_s^*K$
67	$D_{2S}^* \to D_{2S}\pi$
68	$D_{2S}^* \to D^*\rho$
69	$D_{2S}^* \to D^*\omega$
70	$D_{2S}^* \to D_{2S}\gamma$
71	$B^+ : D^0 \to X\ell\nu$
72	$B^+ : D^+ \to X\ell\nu$
73	$B^0 : D^0 \to X\ell\nu$
74	$B^0 : D^+ \to X\ell\nu$
75	Electron efficiency
76	Muon efficiency
77	ρ_D^2
78	$B \to D^*\ell\nu$: ff r_1
79	$B \to D^*\ell\nu$: ff r_2
80	$B \to D^*\ell\nu$: ff r_3
81	$1P_{3/2}$: $d\tau/dw$
82	$1P_{3/2}$: η_{ke}
83	$1P_{3/2}$: $c_{B1\text{-}B2}$
84	$1P_{1/2}$: $d\tau/dw$
85	$1P_{1/2}$: η_{ke}
86	$1P_{1/2}$: $c_{B1\text{-}B2}$

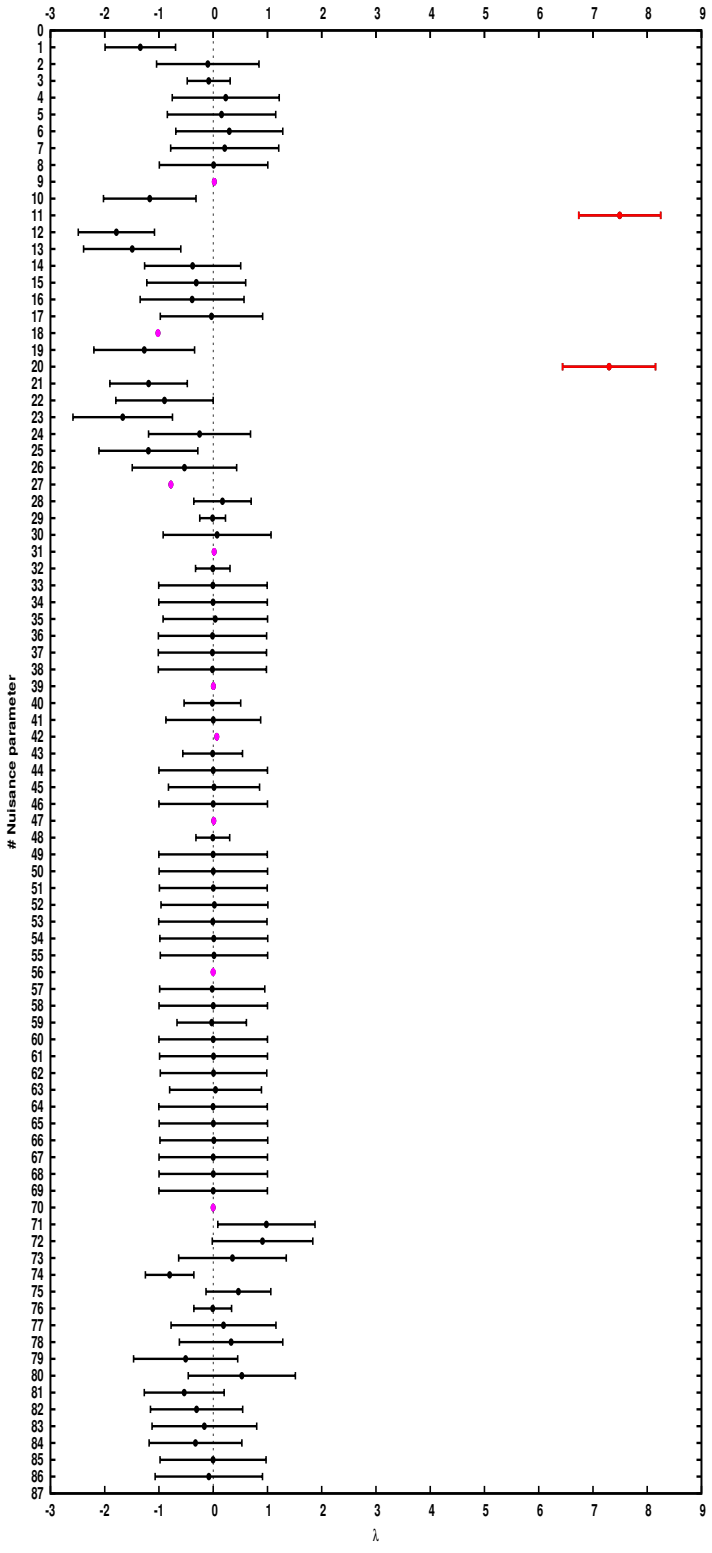


Figure A.3: Best-fit values of the nuisance parameters in the signal extraction. Purple points without errorbars are the parameters which are determined due to inclusive constraints (see Eq. 5.15). The $B \to D^*\ell\nu$ parameters (red) are free parameters but are included as they participate in the inclusive $B \to X\ell\nu$ constraint. Their position and errorbars are scaled such that they represent deviations from the external measurement.

A.6 Control studies

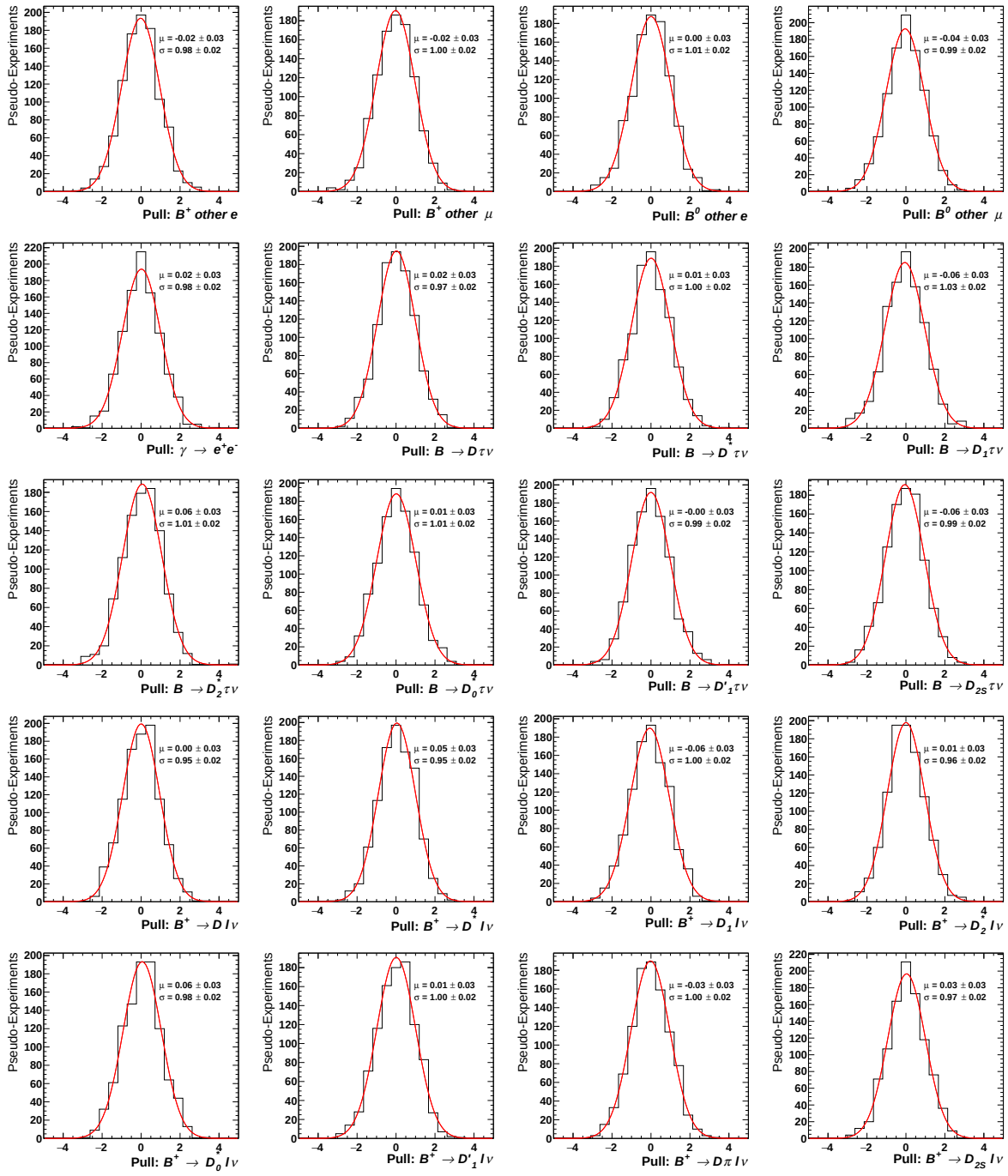
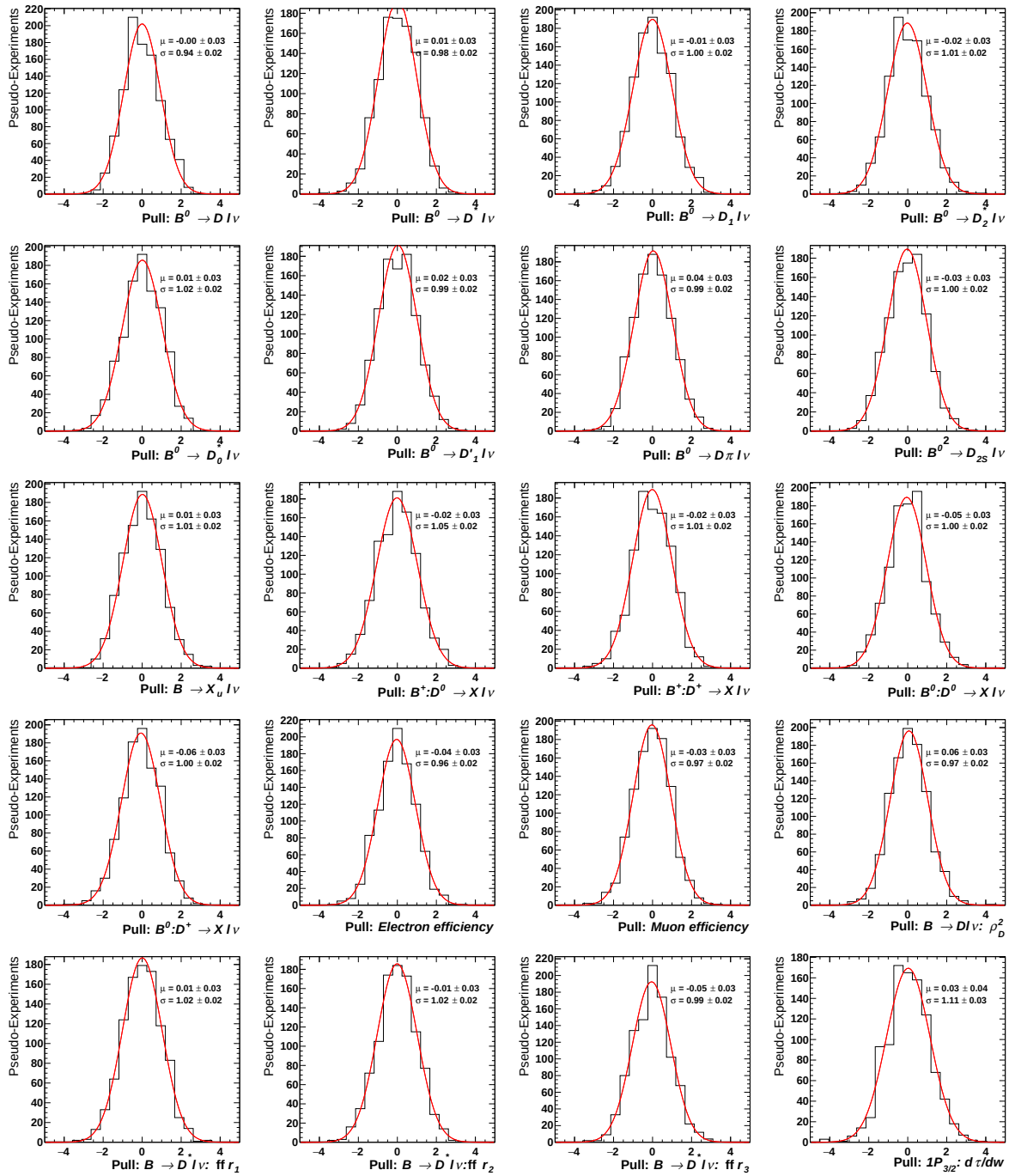
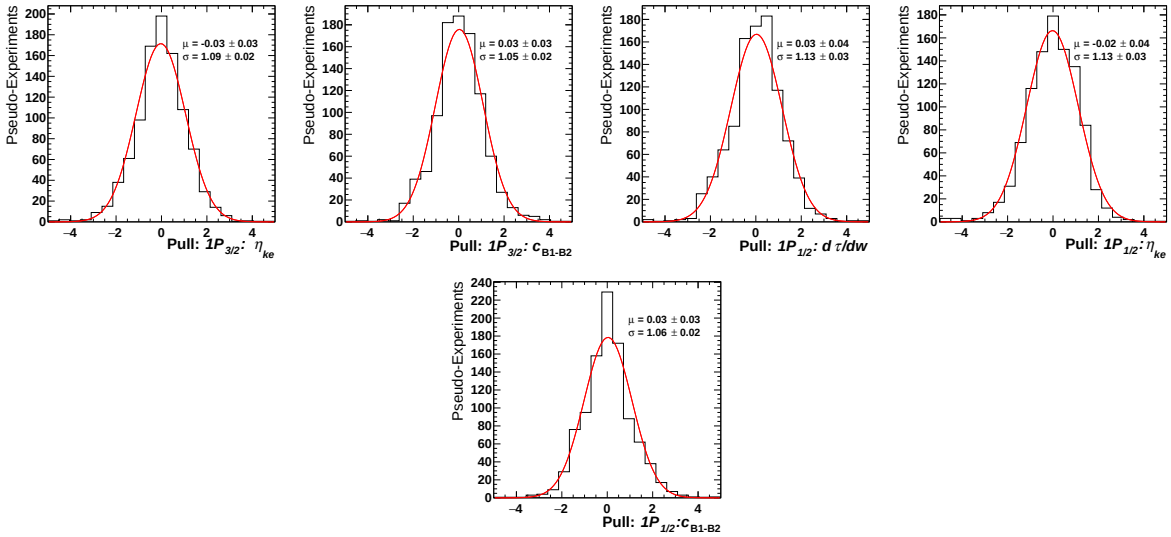


Figure A.4: Pull distributions of floating fit parameters. A Gaussian distribution is fitted to obtain the mean μ and width σ.



Continuation of Fig. A.4



Continuation of Fig. A.4

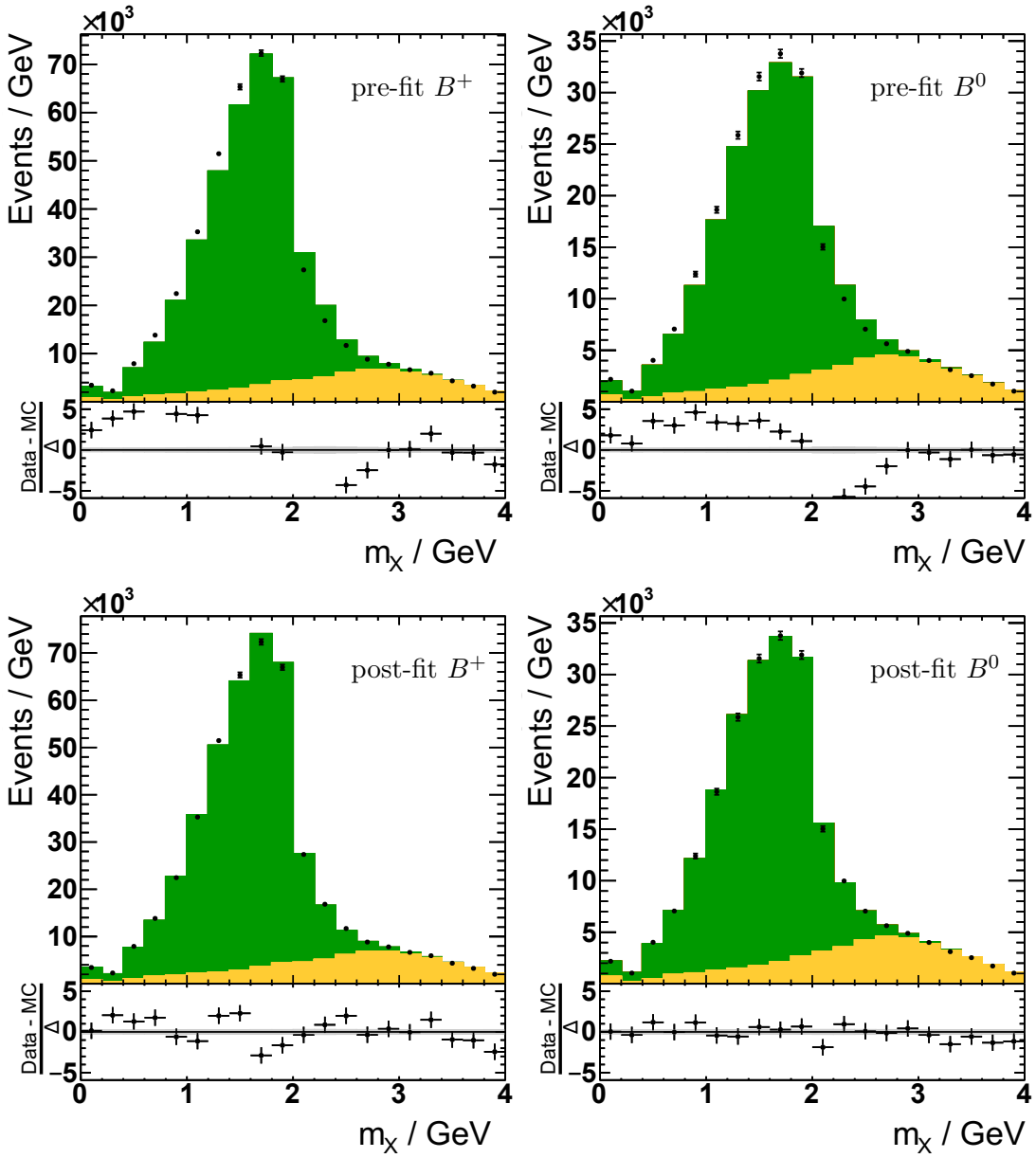


Figure A.5: Pre- (top) and post-fit (bottom) distribution of the hadronic mass m_X in the low m^2_{miss} side-band. Shown are the distributions of the charged (left) and neutral (right) B_{tag} muon sample. Note, the fit is not performed in m_X, but $|\vec{p}^{\,*}_\ell|$ and m^2_{miss}.

A.7 Post-fit results

$R(X)$	0.30 ± 0.02
A_{B^+}	0.98 ± 0.01
A_{B^0}	0.06 ± 0.02
A_{B^+e}	1.02 ± 0.02
$A_{B^0\mu}$	0.99 ± 0.02
A_{B^+e}	1.04 ± 0.03
$A_{B^0\mu}$	1.05 ± 0.03

Table A.8: Best-fit parameters of the $R(X)$ extraction. Normalisation A_j belongs to the $B \to X\ell\nu$ normalisation mode template and A_{jk} to the others template. Their absolute value has no physical meaning as the MC normalisation is rather unknown.

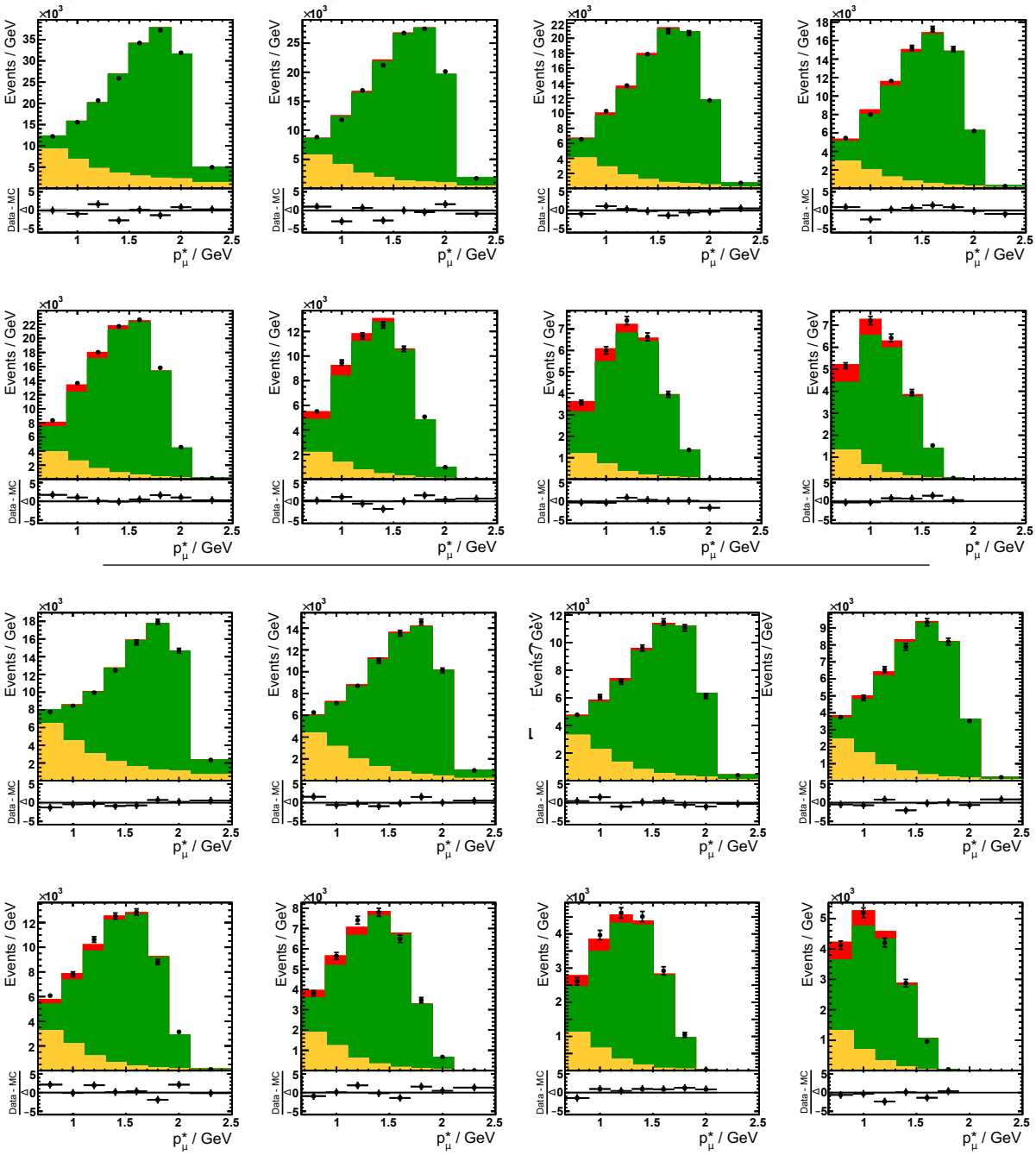


Figure A.6: Post-fit muon momentum spectra in bins of m^2_{miss}. The upper two rows show the charged B and the lower rows the neutral B samples. The corresponding plots for the muon final state are given in Fig. 6.2. The colours have the same meaning as in Fig. 6.2.

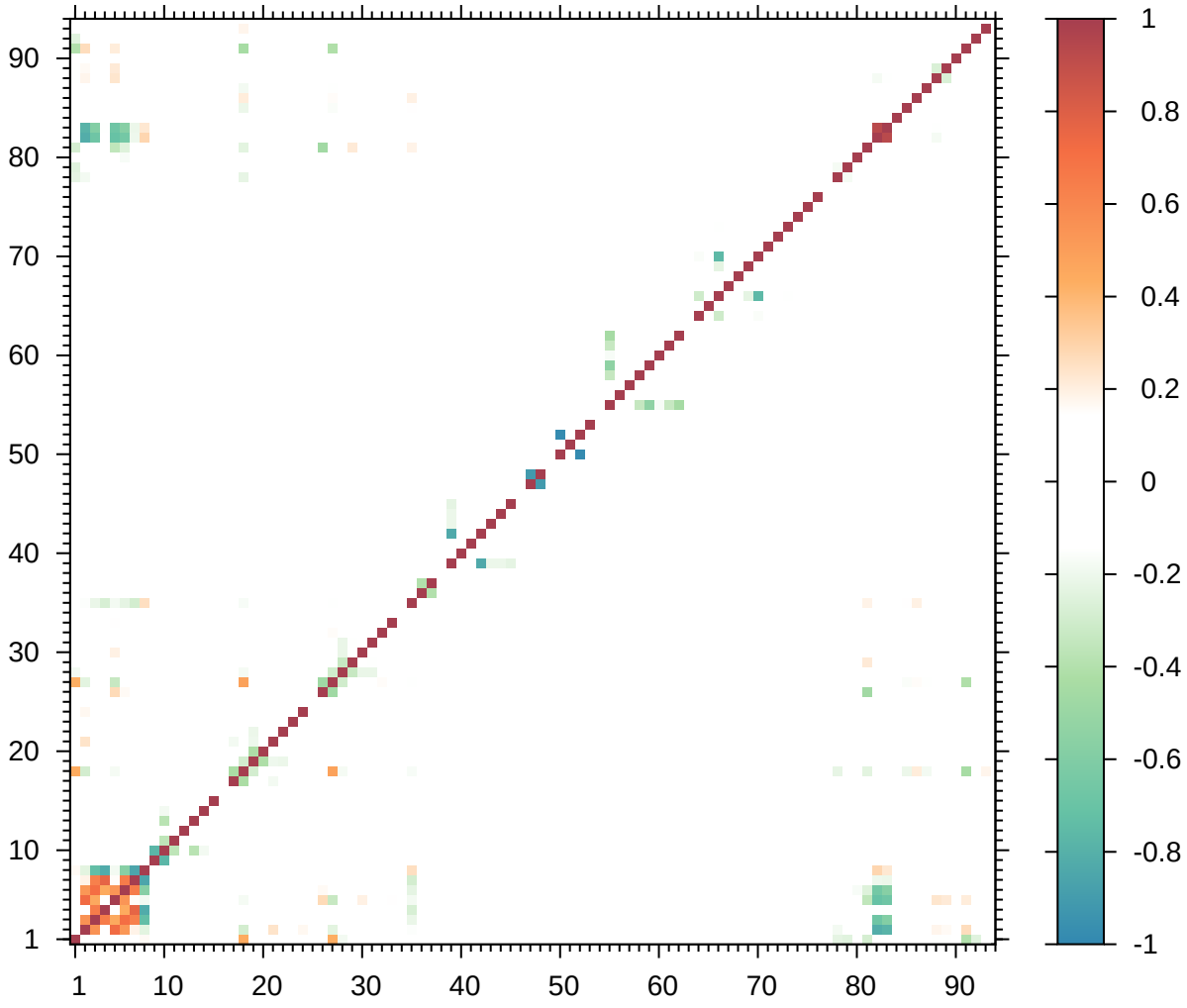


Figure A.7: Full correlation matrix of the nominal $R(X)$ fit. The signal (row = column = 0) is less than 50% correlated to any parameter. The parameter numbering is the same as in Tab. 5.2.

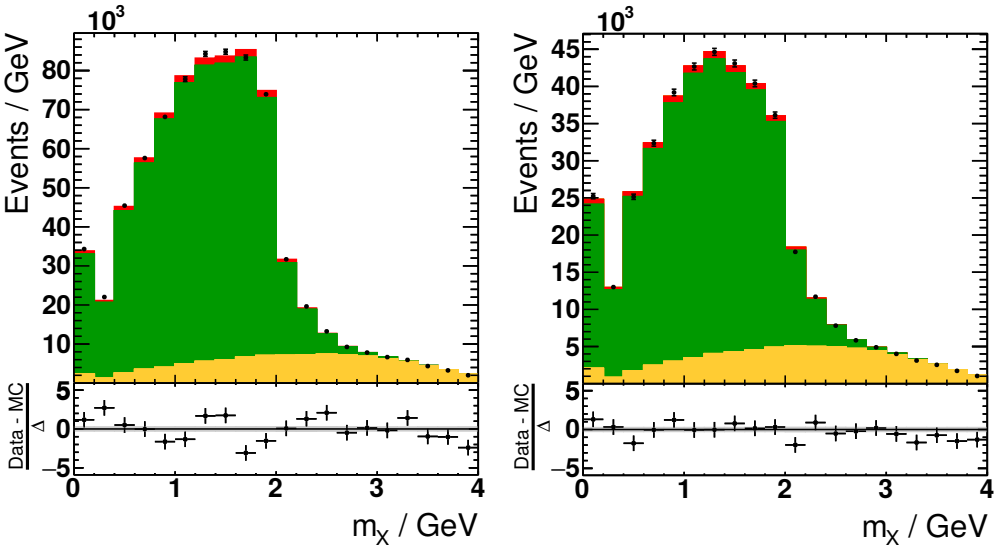


Figure A.8: The post-fit hadronic mass spectrum m_X for charged and neutral B mesons and the muon signal lepton. The corresponding plot for the electron final state is shown in Fig. 7.2. The data modelling is clearly enhanced and even the D^* peak region is reasonably modelled.

List of Figures

List of Tables

Bibliography

[1] S. L. Glashow, J. Iliopoulos and L. Maiani.
"Weak Interactions with Lepton-Hadron Symmetry".
In: *Phys. Rev. D* 2 (7 1970), pp. 1285–1292.

[2] N. Cabibbo. "Unitary Symmetry and Leptonic Decays".
In: *Physical Review Letters* 10 (June 1963), pp. 531–533.

[3] M. Kobayashi and T. Maskawa.
"CP-Violation in the Renormalizable Theory of Weak Interaction".
In: *Progress of Theoretical Physics* 49.2 (1973), p. 652. eprint:
`/oup/backfile/Content\_public/Journal/ptp/49/2/10.1143/PTP.49.652/2/49-2-652.pdf`.

[4] S. W. Herb et al.
"Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions".
In: *Phys. Rev. Lett.* 39 (5 1977), pp. 252–255.

[5] F. Abe et al. "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2626–2631.

[6] S. Abachi et al.
"Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} =$ 1.8 TeV".
In: *Phys. Rev. Lett.* 74 (13 1995), pp. 2422–2426.

[7] K. Abe et al. "Observation of Large *CP* Violation in the Neutral B Meson System".
In: *Phys. Rev. Lett.* 87 (9 2001), p. 091802.

[8] B. Aubert et al. "Observation of *CP* Violation in the B^0 Meson System".
In: *Phys. Rev. Lett.* 87 (9 2001), p. 091801.

[9] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons".
In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509.

[10] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons".
In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323.

[11] "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: (). arXiv: `1207.7214 [hep-ex]`.

[12] S. Chatrchyan et al.
"Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC".
In: *Physics Letters B* 716 (Sept. 2012), pp. 30–61. arXiv: `1207.7235 [hep-ex]`.

[13] J. P. Lees et al. "Measurement of an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays and implications for charged Higgs bosons". In: *Phys. Rev. D* 88 (7 2013), p. 072012.

[14] M. Huschle et al. "Measurement of the branching ratio of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ relative to $\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell$ decays with hadronic tagging at Belle".
In: *Phys. Rev. D* 92 (7 2015), p. 072014.

[15] Y. Sato et al. "Measurement of the branching ratio of $\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau$ relative to $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays with a semileptonic tagging method".
In: *Phys. Rev. D* 94 (7 2016), p. 072007.

[16] S. Hirose et al.
"Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$".
In: (2016). arXiv: `1612.00529 [hep-ex]`.

[17] R. Aaij et al. "Measurement of the Ratio of Branching Fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$". In: *Phys. Rev. Lett.* 115 (11 2015), p. 111803.

[18] HFAG Semileptonics group. *Summer 2016 results.*
URL: `http://www.slac.stanford.edu/xorg/hfag/semi/index.html` (visited on 03/01/2017).

[19] Wikicommons. *Standard Model of Elementary Particles.* (Visited on 22/04/2017).

[20] C. P. Particle Data Group and et al. "Review of Particle Physics".
In: *Chinese Physics C* 40.10, 100001 (Oct. 2016), p. 100001.

[21] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the {ATLAS} detector at the {LHC}".
In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693.

[22] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay".
In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415.

[23] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson".
In: *Phys. Rev. Lett.* 13 (4 1964), pp. 138–140.

[24] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix".
In: *Physical Review Letters* 51 (Nov. 1983), pp. 1945–1947.

[25] J. Charles et al.
"Current status of the standard model CKM fit and constraints on $\Delta F = 2$ new physics".
In: *Phys. Rev. D* 91 (7 2015), p. 073007.

[26] UTfit Collaboration. *UTfit Homepage.* URL: `utfit.org` (visited on 03/01/2017).

[27] A. J. Bevan et al. "The Physics of the B Factories".
In: *The European Physical Journal C* 74.11 (2014), p. 3026. ISSN: 1434-6052.

[28] M. Neubert. "Model-independent extraction of Vcb from semi-leptonic decays".
In: *Physics Letters B* 264.3 (1991), pp. 455–461. ISSN: 0370-2693.

[29] A. Sirlin. "Radiative corrections in the standard model".
In: *AIP Conf. Proc.* 300 (1994). 0093(1984), pp. 331–341.

[30] M. Neubert. "B Decays and the Heavy-Quark Expansion". In: *Heavy Flavours II.* Ed. by A. J. Buras and et al. World Scientific Publishing Co, 1998, pp. 239–293.

[31] A. G. Grozin. "Introduction to the Heavy Quark Effective Theory".
In: *ArXiv High Energy Physics - Phenomenology e-prints* (Aug. 1999).
eprint: `hep-ph/9908366`.

[32] I. Caprini, L. Lellouch and M. Neubert.
"Dispersive bounds on the shape of $B \to D^{(*)}\ell\nu$ form factors".
In: *Nuclear Physics B* 530.1 (1998), pp. 153–181. ISSN: 0550-3213.

[33] F. Bernlochner. Private Communication. 2016.

[34] A. K. Leibovich et al. "Semileptonic B decays to excited charmed mesons".
In: Phys. Rev. D 57 (Jan. 1998), pp. 308–330. eprint: `hep-ph/9705467`.

[35] A. F. Falk et al. "Heavy quark expansion for the inclusive decay $B \to \tau\nu X$".
In: *Physics Letters B* 326.1 (1994), pp. 145–153. ISSN: 0370-2693.

[36] Z. Ligeti and F. J. Tackmann. "Precise predictions for $B \to X_c\tau\overline{\nu}$ decay distributions".
In: *Phys. Rev. D* 90 (3 2014), p. 034021.

[37] R. Barate et al. "Measurements of $\mathcal{B}(b \to \tau^-\bar{\nu}_\tau X)$ and $\mathcal{B}(b \to \tau^-\bar{\nu}_\tau D^{*\pm}X)$ and upper limits on $\mathcal{B}(B^- \to \tau^-\bar{\nu}_\tau)$ and $\mathcal{B}(b \to s\nu\bar{\nu})$".
In: *The European Physical Journal C - Particles and Fields* 19.2 (2001), pp. 213–227.
ISSN: 1434-6052.

[38] P. Abreu et al.
"Upper limit for the decay $B \to \tau\nu_\tau$ and measurement of the $b \to \tau\nu_\tau X$ branching ratio".
In: *Physics Letters B* 496.1–2 (2000), pp. 43–58. ISSN: 0370-2693.

[39] M. Acciarri et al. "Measurement of the branching ratios $b \to e\nu X$, $\mu\nu X$, $\tau\nu X$ and νX".
In: *Z. Phys.* C71 (1996), pp. 379–390.

[40] M. Acciarri et al. "Measurement of the inclusive $b \to \tau\nu X$ branching ratio".
In: *Physics Letters B* 332.1–2 (1994), pp. 201–208. ISSN: 0370-2693.

[41] G Abbiendi et al. "Measurement of the branching ratio for the process $b \to \tau\nu_\tau X$".
In: *Physics Letters B* 520.1–2 (2001), pp. 1–10. ISSN: 0370-2693.

[42] OPAL Collaboration. *OPAL Event Displays*. URL:
`http://opal.web.cern.ch/Opal/events/opalpics.html` (visited on 15/03/2017).

[43] D. Bigi and P. Gambino. "Revisiting $B \to D\ell\nu$". In: *Phys. Rev. D* 94 (9 2016), p. 094008.

[44] S. Fajfer, J. F. Kamenik and I. Nišandžić. "$B \to D^*\tau\overline{\nu}_\tau$".
In: *Phys. Rev. D* 85 (9 2012), p. 094025.

[45] HFAG Semileptonics group. *Average of $R(D)$ and $R(D^*)$ for Moriod EW 2017*.
URL: `http://www.slac.stanford.edu/xorg/hfag/semi/moriond17/RDRDs.html`
(visited on 10/04/2017).

[46] L. Dhargyal. "$R(D^{(*)})$". In: *Phys. Rev. D* 93 (11 2016), p. 115009.

[47] P. Ko, Y. Omura and C. Yu.
"$B \to D^{(*)}\tau\nu$ and $B \to \tau\nu$ in chiral U(1)$'$ models with flavored multi Higgs doublets".
In: *Journal of High Energy Physics* 2013.3 (2013), p. 151. ISSN: 1029-8479.

[48] F. U. Bernlochner et al. "Combined analysis of semileptonic B decays to D and D^*:
$R(D^{(*)})$, $|V_{cb}|$, and new physics". In: (2017). arXiv: `1703.05330 [hep-ph]`.

[49] M. Freytsis, Z. Ligeti and J. T. Ruderman. "Flavor models for $\overline{B} \to D^{(*)}\tau\overline{\nu}$".
In: *Phys. Rev. D* 92 (5 2015), p. 054018.

[50] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models".
In: Phys. Rep. 516 (July 2012), pp. 1–102. arXiv: `1106.0034 [hep-ph]`.

[51] Y. Grossman and Z. Ligeti.
"The inclusive $B \to \tau\nu X$ decay in two Higgs doublet models".
In: *Physics Letters B* 332.3 (1994), pp. 373–380. ISSN: 0370-2693.

[52] A. Crivellin, C. Greub and A. Kokulu.
"Explaining $B \to D\tau\nu$, $B \to D^*\tau\nu$ and $B \to \tau\nu$ in a two Higgs doublet model of type III".
In: *Phys. Rev. D* 86 (5 2012), p. 054014.

[53] A. Abashian et al. "The Belle detector".
In: *Nuclear Instruments and Methods in Physics Research A* 479 (Feb. 2002), pp. 117–232.

[54] Z. Natkaniec et al. "Status of the Belle silicon vertex detector".
In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 560.1 (2006). Proceedings of the 13th International Workshop on Vertex DetectorsVERTEX 200413th International Workshop on Vertex Detectors, pp. 1–4. ISSN: 0168-9002.

[55] B. Chasey. "Belle Note 390: HadronB". Nov. 2001.

[56] D. J. Lange. "The EvtGen particle decay simulation package".
In: *Nucl. Instrum. and Meth.* A462 (2001), p. 152.

[57] T. Sjostrand. "PYTHIA 5.7 and JETSET 7.4 Physics and Manual".
In: *ArXiv High Energy Physics - Phenomenology e-prints* (Aug. 1995).
eprint: `hep-ph/9508391`.

[58] J. Allison et al. "An application of the GEANT3 geometry package to the description of the opal detector". In: *Computer Physics Communications* 47 (Oct. 1987), pp. 55–74.

[59] HFAG Semileptonics group. *Summer/PDG 2014 results.*
URL: `http://www.slac.stanford.edu/xorg/hfag/semi/index.html` (visited on 16/10/2016).

[60] N. Isgur et al. "Semileptonic B and D decays in the quark model".
In: *Phys. Rev. D* 39 (3 1989), pp. 799–818.

[61] L. Hinz. "Belle Note 954: Lepton ID efficiency correction and systematic error".
Sept. 2006.

[62] Belle PID group. *Joint PID group homepage.*
URL: `http://belle.kek.jp/group/pid_joint` (visited on 03/01/2017).

[63] P. Urquijo. Private Communication. 2013.

[64] Belle MC Group. *Remarks on generic MC (internal).*
URL: `belle.kek.jp/group/software/MC/genericmc.html` (visited on 28/04/2017).

[65] M. Feindt et al. "A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories".
In: *Nuclear Instruments and Methods in Physics Research A* 654 (Oct. 2011), pp. 432–440.
arXiv: `1102.3876 [hep-ex]`.

[66] F. U. Bernlochner, Z. Ligeti and S. Turczyk.
"A proposal to solve some puzzles in semileptonic B decays".
In: Phys. Rev. D 85.9, 094033 (May 2012), p. 094033. arXiv: `1202.1834 [hep-ph]`.

[67] R. Aaij et al. "Amplitude analysis of $B^- \to D^+\pi^-\pi^-$ decays".
In: *Phys. Rev. D* 94 (7 2016), p. 072001.

[68] R. Klein et al.
"$1/m$ corrections for orbitally excited heavy mesons and the $1/3 - 2/3$ puzzle".
In: *Phys. Rev. D* 91 (9 2015), p. 094034.

[69] HFAG B to Charm group. *Summer 2016 results.*
URL: `http://www.slac.stanford.edu/xorg/hfag/btocharm/2016/index.html`
(visited on 03/01/2017).

[70] J. P. Lees et al.
"Observation of $\overline{B} \to D^{(*)}\, \pi^+\pi^-\ell^-\overline{\nu}$ Decays in e^+e^- Collisions at the $\Upsilon(4S)$ Resonance".
In: *Phys. Rev. Lett.* 116 (4 2016), p. 041801.

[71] R. Aaij. "Erratum: Measurements of the branching fractions for $B_{(s)} \to D_{(s)}\pi\pi\pi$ and $\Lambda_b^0 \to \Lambda_c^+\pi\pi\pi$ [Phys. Rev. D 84, 092001 (2011)]". In: *Phys. Rev. D* 85 (3 2012), p. 039904.

[72] N. Isgur and M. B. Wise. "Spectroscopy with heavy-quark symmetry".
In: *Phys. Rev. Lett.* 66 (9 Mar. 1991), pp. 1130–1133.

[73] M. J. Savage. "Lecture 3, Introduction to Nuclear and Particle Physics". 1995.

[74] EvtGen Project. *EvtGen Module documentation.*
URL: `http://evtgen.warwick.ac.uk/docs/models/` (visited on 08/03/2017).

[75] R. Brun and F. Rademakers. "ROOT - An Object Oriented Data Analysis Framework".
In: *AIHENP'96 Workshop, Lausane.* Vol. 389. 1996, pp. 81–86.

[76] F. James.
"MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1".
In: (1994).

[77] M. Hatlo et al.
"Developments of mathematical software libraries for the LHC experiments".
In: *IEEE Transactions on Nuclear Science* 52.6 (2005), pp. 2818–2822. ISSN: 0018-9499.

[78] H. Kakuno. "Belle Note 615: Measurement of $\mathcal{B}(B \to X_u\ell\nu)$ and $|V_{ub}|$ through M_x reconstruction with Simulated Annealing technique". Nov. 2003.

[79] L. Pesántez et al.
"Measurement of the Direct CP Asymmetry in $B \to X_{s+d}\gamma$ Decays with a Lepton Tag".
In: *Physical Review Letters* 114.15, 151601 (Apr. 2015), p. 151601.
arXiv: `1501.01702 [hep-ex]`.

[80] L. Pesantez. "Study of radiative $B \to X_{s+d}\gamma$ decays at Belle". PhD thesis. 2017.

[81] L. Pesantez. "Belle Note 1294: Measurement of A_{CP} in the fully inclusive $B \to X_{s+d}\gamma$ channel with a lepton tag". Sept. 2015.

[82] J. P. Lees et al. "Measurement of the inclusive electron spectrum from B meson decays and determination of $|V_{ub}|$". In: *Phys. Rev. D* 95 (7 2017), p. 072001.

[83] B. Aubert et al. "Measurement of the Branching Fractions of $\overline{B} \to D^{**}\ell^-\overline{\nu}_\ell$ Decays in Events Tagged by a Fully Reconstructed B Meson".
In: *Phys. Rev. Lett.* 101 (26 2008), p. 261802.

[84] P. Urquijo et al. "Moments of the electron energy spectrum and partial branching fraction of $B \to X_c e\nu$ decays at the Belle detector".
In: Phys. Rev. D 75.3, 032001 (Feb. 2007), p. 032001. eprint: `hep-ex/0610012`.

[85] R. Glattauer et al. "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$". In: *Phys. Rev. D* 93 (3 2016), p. 032006.

[86] H. Hayashii and Y. Hirano. "Belle Note 547: Check on the material using γ-conversion in the τ-pair sample". July 2002.

[87] Jaroslaw Wiechczynski. *Presentation at Moriond QCD 2017: Future Belle II experiment at the KEK laboratory.* URL: `docs.belle2.org/record/553/files/BELLE2-TALK-CONF-2017-041.pdf` (visited on 28/04/2017).

Acknowledgements

First of all, I would like to thank Jochen Dingfelder for his professional assistance and for offering me the opportunity to work in his research group on the fascinating and current topic of B decays.
I also would like to thank Prof. Klaus Desch for being the second referee of this thesis.
Furthermore, I want to express my gratitude to Phillip Urquijo and Florian Bernlocher for their continuous support, countless helpful discussions and constructive ideas. I owe special thanks to Phill for paving a smooth and easy way into the Belle and HEP community, already during my master's thesis.
Furthermore I would like to thank Christoph Schwanda, as the chair of the Belle CKM working group, and the Belle internal referees of my analysis Andrezj Bozek, Tom Browder and Toru Iijima for their constructive and critical review of my work.

Many thanks to my working group for the past years and in particular my Belle office mates Christian Oswald, Luis Pesantez, Cesar Beleno and Stephan Duell for all the helpful discussions and the nice atmosphere in the Belle office.

Special thanks to my mother for supporting my studies in many ways. My deepest gratitude to my girlfriend Sophia for all the long nights and weekends of working together.