A Search for $B^- \rightarrow \tau^- \overline{\nu}_\tau$ using $\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ with the *BABAR* Detector

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
des Fachbereichs Physik
der Universität Dortmund

vorgelegt von

Denis Altenburg

Mai 2007

Abstract

This thesis presents a search for the purely leptonic decay of charged B mesons into a τ lepton and a neutrino (ν) using data collected with the BABAR detector located at the Stanford Linear Accelerator Center. The data sample used to perform the analysis corresponds to an integrated luminosity of $232.1\,\mathrm{fb}^{-1}$, where $210.6\,\mathrm{fb}^{-1}$ have been taken at a center-of-mass energy of about $10.58\,\mathrm{GeV}$ according to the peak mass of the $\Upsilon(4S)$ resonance. $21.5\,\mathrm{fb}^{-1}$ recorded about $40\,\mathrm{MeV}$ below the $\Upsilon(4S)$ resonance have been used for background studies. Due to the low expected number of signal events in the data sample facing a high expected background multivariate methods have been used to efficiently separate signal and background events. Three different selection strategies have been applied and compared with respect to the expected upper limit of the branching fraction of the decay $B^- \to \tau^- \overline{\nu}_\tau$ in the data sample. The best result has been achieved by a technique based on an exclusive reconstruction of semileptonic $B \to X_c \ell \nu$ ($X_c = D\,\mathrm{or}\,D^*$) decays.

13 events have been observed in the data sample, which is consistent with the background expectation of 11.16 events. An upper limit of the branching fraction of the decay $B^- \to \tau^- \overline{\nu}_\tau$ including systematic uncertainties has been calculated to

$$\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) < 9.4 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,.$$

Kurzfassung

Diese Arbeit beschreibt eine Suche nach dem rein leptonischen Zerfall geladener B-Mesonen in ein τ-Lepton und ein Neutrino (ν) mit Daten, die mit dem BABAR-Detektor am Stanford Linear Accelerator Center gesammelt wurden. Der benutzte Datensatz entspricht einer integrierten Luminosität von $232{,}1\,\mathrm{fb}^{-1}$, wobei $210{,}6\,\mathrm{fb}^{-1}$ bei einer Schwerpunktsenergie von $10{,}58\,\mathrm{GeV}$ genommen wurden. Diese Energie entspricht näherungsweise der Masse der $\Upsilon(4S)$-Resonanz. $21{,}5\,\mathrm{fb}^{-1}$ wurden bei einer Schwerpunktsenergie etwa $40\,\mathrm{MeV}$ unterhalb der $\Upsilon(4S)$-Resonanz aufgezeichnet und für Untergrundstudien verwendet. Aufgrund der geringen erwarteten Anzahl von Signalereignissen im benutzten Datensatz und des hohen erwarteten Untergrundes wurden multivariate Methoden benutzt, um Signal und Untergund effizient voneinander zu trennen. Dabei kamen drei verschiedene Selektionsstrategien zum Einsatz, die im Hinblick auf die erwartete obere Grenze für das Verzweigungsverhältnis des Zerfalls $B^- \to \tau^- \overline{\nu}_\tau$ im Datensatz verglichen wurden. Das beste Resultat wurde mit einer Methode erreicht, die auf der exklusiven Rekonstruktion semileptonischer $B \to X_c \ell \nu$ ($X_c = D\,\mathrm{or}\,D^*$) Zerfälle basiert.

13 Ereignisse wurden im Datensatz beobachtet, was innerhalb der Unsicherheiten mit der Untergrunderwartung von 11,16 Ereignissen verträglich ist. Eine obere Grenze für das Verzweigungsverhältnis des Zerfalls $B^- \to \tau^- \overline{\nu}_\tau$ inklusive systematischer Unsicherheiten wurde bestimmt zu

$$\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) < 9.4 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,.$$

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Nowadays, the Standard Model of particle physics is one of the most effective theories of modern physics. The validity of this model has been tested by a variety of experiments and in most cases it provides a sufficient description of the interactions of elementary particles. Nevertheless, the outstanding success in the prediction of nature on the level of elementary particles should not deceive the fact that this theory depends on a substantial number of free parameters, e.g. coupling constants and particle masses. In particular, to date the origin as well as the hierarchy of particle masses is still unexplained.

Furthermore, an enormous problem of the Standard Model lies in the disability to explain the observed matter-antimatter-asymmetry in our universe. Indeed, the Standard Model already contains a mechanism, which provides such an asymmetry. This CP violation necessarily results in different decay and production rates of particles compared to the corresponding antiparticles. Such an effect has been discovered in the system of neutral kaons in 1964 by *Christenson*, *Cronin*, *Fitch*, and *Turlay* [1]. However, supposing particle-antiparticle equilibrium in the early phase of the universe the Standard Model is not sufficient to explain the present asymmetry, even if one assumes maximal CP violation. Therefore, there are reasons to be sceptic and to probe the Standard Model.

Many experiments have been built at the end of the last century in order to perform precision measurements of Standard Model parameters and to search for new physics. Besides other important projects, the B meson factories, PEP-II in the United States and KEK-B in Japan with their detectors *BABAR* and *Belle*, play an important role for todays flavour physics. These projects have been explicitly designed to measure parameters of the weak interaction. In particular, the primary goal of these experiments is the measurement of CP violation in decays of B mesons with high precision and the comparison with Standard Model predictions.

After eight years of data taking no significant inconsistency of the collected data with Standard Model expectations of weak processes on the B sector has been found. Given the current precision, the results of measurements of the aforementioned parameters as well as CP asymmetries all fit together in the Standard Model picture. However, the high number of B mesons produced by the B factories allows searches for rare B decays. Such processes are very interesting since new physics could possibly contribute to the total decay rates resulting in deviations of the measured branching fractions from the corresponding Standard Model predictions.

Investigation of the purely leptonic B decay $B^- \to \ell^- \overline{\nu}_\ell$ is important for two reasons. At first, this decay provides direct access to parameters describing the inner structure of B mesons. In particular, the B meson decay constant f_B is currently poorly known from theoretical calculations. Since this parameter acts as input for other measurements, it introduces high theoretical uncertainties and therefore represents a limiting factor of these measurements. Moreover, additional contributions from charged *Higgs* bosons could possibly modify the decay rate of this process and a precise measurement of its branching fraction can be used to set limits on a charged *Higgs* mass predicted by theories beyond the Standard Model. Such effects are expected to be more prominent in the decay $B^- \to \tau^- \overline{\nu}_\tau$ compared to $B^- \to e^- \overline{\nu}_e$ or $B^- \to \mu^- \overline{\nu}_\mu$.

To date one observation of the decay $B^- \to \tau^- \overline{\nu}_\tau$ has been claimed by the *Belle* collaboration in 2006. All searches performed by many other groups including the *BABAR* collaboration did not find a significant signal and upper limits have been obtained from these analyses. This already indicates the challenge to search for this decay. Due to the enormous background all these measurements used reconstruction procedures, which provide high purities paired with low reconstruction efficiencies. In contrast, the analysis presented in this thesis aims at an increase of the significance of $B^- \to \tau^- \overline{\nu}_\tau$ using different reconstruction techniques.

Although former analyses already used the most prominent τ decay channels for the reconstruction of $B^- \to \tau^- \overline{\nu}_\tau$ decays, the potential of the decay mode $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ to perform such a search has not been investigated in detail. Therefore, all aforementioned analyses did not take advantage of the kinematical properties of this decay mode. On the other hand, this process shows a distinctive kinematics, which might be exploitable for an efficient signal selection. Moreover, the *BABAR* detector possesses a high performance tracking system and decay vertices can be reconstructed with high resolution. These facts motivate the choice of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay for the reconstruction of $B^- \to \tau^- \overline{\nu}_\tau$ events. This analysis investigates how vertex information and kinematical quantities can be used to select signal events with high reconstruction efficiency and a tolerable background level.

The theoretical considerations given in Chap. 2 start with a brief introduction of the Standard Model with emphasis on the weak interaction before the theoretical issues related to the decays $B^- \to \tau^- \overline{\nu}_\tau$ and $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ are discussed.[1] After a short overview of the *BABAR* detector (Chap. 3), Chap. 4 describes the reconstruction techniques used to perform the search for $B^- \to \tau^- \overline{\nu}_\tau$ including studies of systematic corrections and uncertainties. The results and their implications for B physics are summarized in Chap. 5 and a conclusion of the potential of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay for a search for $B^- \to \tau^- \overline{\nu}_\tau$ is given in the summary (Chap. 6).

---

[1] Throughout this thesis, the reconstruction of signal events is referred to as $B^- \to \tau^- \overline{\nu}_\tau$ with a subsequent $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay. This term includes the charge-conjugated decay chain.

Chapter 2

Theoretical Considerations

This chapter contains an overview of the most important theoretical issues related to the decay $B^- \to \tau^- \overline{\nu}_\tau$ with emphasis on the weak interaction within the framework of the ***Standard Model*** **(SM)** and the discovery potential of physics beyond the SM. The discussion results in a strong motivation to search for this decay. The chapter starts with some general considerations on the SM and the nature of weak interactions and explains the importance of the $B^- \to \tau^- \overline{\nu}_\tau$ decay for the understanding of B physics. The last part concerns with the properties of the $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ decay.

2.1 The Standard Model of Particle Physics

For a long period particle physicists have been dealing with the search for one fundamental theory explaining the properties of the elementary particles and all their interactions. Although this ultimate theory has not been discovered so far, the SM of particle physics represents a good description of the nature of interacting matter and is extremely successful in predicting the experimental data. It encloses three of the four fundamental interactions, i.e. the *Strong*, the *Electromagnetic* and the *Weak Interaction*, and is able to explain the composition of baryonic matter. The fourth one, namely the *Gravitational Interaction*, is of geometric nature and therefore not included. Within the SM there are two blocks of elementary particles, *Quarks* and *Leptons*, arranged in three *Generations* also called *Flavours*. These fundamental particles are summarized in Tab. 2.1. The existence of all these particles has already been confirmed by experiments and the measurements uncovered a mass pattern. The particle masses increase from one generation to the other. To date the explanation of this pattern is one of the remaining unknowns.

The SM has been formulated as a local gauge theory and all interactions between quarks and leptons are mediated by gauge bosons with integer spin. The gauge symmetry group is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \,,$$

where $SU(3)_C$ is the symmetry group of ***Quantum Chromodynamics*** **(QCD)** representing the strong interaction. The colour charge C has been introduced as the generator of this symmetry group and can be *red*, *green* or *blue*. Every quark

		Interaction→	**Electroweak**				**Strong**
		Gauge Group→	$SU(2)_L \otimes U(1)_Y$				$SU(3)_C$
		Quantum Number→	T	T_3	Y	Q	C
Quarks:							
$Q_1 = \begin{pmatrix} u \\ d' \end{pmatrix}_L$	$Q_2 = \begin{pmatrix} c \\ s' \end{pmatrix}_L$	$Q_3 = \begin{pmatrix} t \\ b' \end{pmatrix}_L$	1/2	+1/2 −1/2	+1/6	+2/3 −1/3	r,g,b
u_R	c_R	t_R	0	0	+2/3	+2/3	
d_R	s_R	b_R	0	0	−1/3	−1/3	
Leptons:							
$L_1 = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$L_2 = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$L_3 = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	1/2	+1/2 −1/2	−1/2	0 −1	white
e_R^-	μ_R^-	τ_R^-	0	0	−1	−1	

Table 2.1: Quantum numbers of elementary particles within the SM.

carries one of these colour charges. Leptons are colour neutral and therefore don't take part on strong interaction processes. The second part of the SM gauge group ($SU(2)_L \otimes U(1)_Y$) reflects the *Electroweak Interaction*. Its theoretical description was formulated by *Glashow*, *Weinberg*, and *Salam* [2–4] as an unification of the electromagnetic and the weak interaction. The indices "L" and "R" in Tab. 2.1 denote the left-handed and right-handed components of the particle fields. The left-handed fields are weak isospin doublets Q_i and L_i, while all others are singlets with respect to the weak interaction. The electric charge Q is related to the electroweak hypercharge Y and the eigenvalue of the third component T_3 of the weak isospin T via

$$Q = T_3 + Y\,.$$

The abovementioned gauge bosons of the fundamental interactions follow from requiring invariance of the Lagrangian under local gauge transformations. In Tab. 2.2 the interactions with their corresponding gauge bosons are summarized. The strong interaction is mediated by eight massless *gluon* fields. Six of them are non-white combinations of colour and anti-colour and can therefore interact with other coloured objects, i.e. gluons can interact with quarks or themselves. This fact is one of the basic differences between the strong and the electromagnetic interaction. Since photons (γ) do not carry electric charge, they do not interact with other γ fields.

Interaction	**Gauge Boson**	**Mass**
electromagnetic	photon (γ)	0
weak	$W^\pm$	80.4 GeV/c^2
	Z^0	91.2 GeV/c^2
strong	8 gluons ($g_1 \ldots g_8$)	0

Table 2.2: The gauge bosons of the Standard Model

The physical gauge boson fields of the electroweak interaction $W^{\pm}$, Z^0 and γ are not equal to the fundamental generators W_1, W_2, W_3 of the $SU(2)_L$ and B of the $U(1)_Y$ gauge group but linear combinations of them

$$\begin{aligned} W^{\pm} &= \frac{1}{\sqrt{2}}(W_1 \mp iW_2)\,, \\ Z^0 &= \cos\theta_W W_3 - \sin\theta_W B\,, \\ \gamma &= \cos\theta_W W_3 + \sin\theta_W B \end{aligned} \tag{2.1}$$

with the weak mixing angle θ_W (*Weinberg* angle). Strong interactions or *Neutral Current* electroweak processes (Z^0, γ) cannot change quark flavours at tree level. Only $W^{\pm}$ boson exchange allows transitions between quarks of different generations. The properties and strengths of such processes are described by the ***Cabibbo-Kobayashi-Maskawa*** **(CKM)** matrix (Sec. 2.3).

The SM predicts neutrinos (ν) to be massless, but recent results from Super-Kamiokande [5] and SNO [6] are in contradiction to this expectation. However, the SM can be extended to account for neutrino masses, even though the current experimental data do not allow formulations of reliable extension models. It will be necessary to include right-handed neutrinos and a second mass mixing matrix on the lepton sector, similar to the CKM matrix on the quark sector. That would accommodate for transitions between the lepton generations and account for neutrino oscillations as they have been measured by the aforementioned experiments.

As already mentioned, the SM is grounded on invariance under local gauge transformations. On the other hand, this principle forbids mass terms for quarks and leptons in the Lagrangians and further requires $W^{\pm}$ and Z^0 bosons to be massless. Since this is in strong contrast to nature, a theoretical mechanism has been developed to generate quark, lepton, $W^{\pm}$, and Z^0 masses. The aforementioned CKM matrix is related to this *Higgs* mechanism (Sec. 2.2).

2.2 The Minimal *Higgs* Model

In order to solve the problem of massive quarks, leptons, and electroweak gauge bosons, various proposals have been made. The most popular procedure to introduce masses is the *Higgs* mechanism [7]. Within this theoretical model quarks and leptons interact with an omnipresent background field Φ. The mass generation then directly follows from the interaction with this "*Higgs*" field. In order to introduce such a new interaction, an additional term has been added to the SM Lagrangian

$$\mathcal{L}_{Higgs} = (\partial^{\alpha}\Phi)^{\dagger}(\partial_{\alpha}\Phi) - V(\Phi) \quad \text{with} \quad V(\Phi) = -\mu^2\Phi^{\dagger}\Phi + \lambda^2(\Phi^{\dagger}\Phi)^2\,, \tag{2.2}$$

where Φ represents a complex scalar isospin doublet $\Phi = (\phi_1, \phi_2)$. By construction, the *Higgs* potential $V(\Phi)$ has a non-zero vacuum expectation value $v = \mu/\lambda$ due to its "Mexican Hat" shape for $\mu^2 > 0$, $\lambda^2 > 0$ illustrated in Fig. 2.1. Since the *Higgs* field entails a non-vanishing energy-minimum at its ground state, it interacts with itself. The coupling of quarks of generation i to Φ is described by the *Yukawa* coupling

$$\mathcal{L}_{Yukawa} = \Lambda^u_{ij}\bar{Q}'_{Li}\widetilde{\Phi}u'_{Rj} + \Lambda^d_{ij}Q'_{Li}\Phi d'_{Rj} + h.c.\,, \tag{2.3}$$

where the Λ matrices represent the complex *Yukawa*-coupling constants and u'_{Rj}, d'_{Rj} denote the right-handed up-type (u,c,t) and down-type (d,s,b) quark fields, respectively. $\widetilde{\Phi}$ stands for the complex conjugated *Higgs* field. It is important to note that the quark states in Eq. 2.3 are the eigenstates with respect to the electroweak interaction, but different from the mass eigenstates. In order to explain the relationship between mass generation and CKM matrix, we need to have a closer look at the ***Charged Current* (CC)** Lagrangian, which implies the coupling to the $W^\pm$ bosons and represents the second part of the electroweak Lagrangian besides the already mentioned *Neutral Current*.

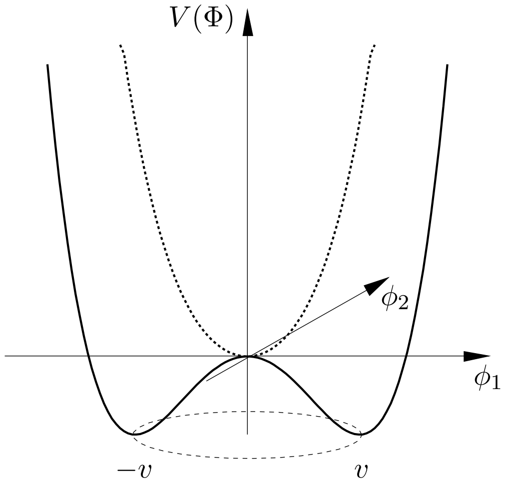


Figure 2.1: The *Higgs* potential: The dotted line illustrates $V(\Phi)$ for $\mu^2 < 0$, $\lambda^2 > 0$, while the solid line corresponds to $\mu^2 > 0$, $\lambda^2 > 0$. The dashed ellipse illustrates the circle of minima with radius v in the $\phi_1 - \phi_2$ plane.

In the SM transitions between down-type and up-type quarks via $W^\pm$ exchange in $SU(2)_L$ can be expressed by

$$\mathcal{L}_{\mathrm{CC}} = -\frac{1}{\sqrt{2}} g\bar{u}'_{Li}\gamma^\mu d'_{Li}W^+_\mu + h.c. \tag{2.4}$$

with the gauge coupling constant g. Up to this point symmetry is not broken in the sense that no particular realization of the *Higgs* field is preferred. On the other hand, if this model is really valid in nature, the choice of the representation of Φ is not fully free. It needs to be chosen that way, that masses are assigned to fermions, $W^\pm$, and Z^0 bosons, but photons explicitly have to remain massless. Within the ***Spontaneous Symmetry Breaking* (SSB)** the upper component of Φ is set to zero and ϕ_2 is expanded around the minimum v as displayed in Eq. 2.5

$$\Phi(x) = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \xrightarrow{\mathrm{SSB}} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}\,. \tag{2.5}$$

This particular choice of the *Higgs* field assigns mass terms to $W^\pm$ and Z^0 bosons resulting in masses of

$$m_W = \frac{gv}{2}, \qquad m_Z = \frac{gv}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W} \tag{2.6}$$

and simultanousely prevents the mechanism from giving mass to the γ [8]. The small effective *Higgs* field $h(x)$ is relevant for particle-*Higgs* interactions. In principle, such processes should be measurable and therefore provide potential for *Higgs* searches. Currently, the ***Large Hadron Collider*** **(LHC)** project is under construction, which has been explicitly designed to discover the abovementioned or other representations of a *Higgs* field by measuring various possible processes within and beyond the SM.

However, if Eq. 2.5 is inserted into the *Yukawa*-Lagrangian (Eq. 2.3), one obtains

$$\mathcal{L}_{Yukawa} = \mathcal{L}_{\text{mass}} + \mathcal{L}_{h(x)} \,, \tag{2.7}$$

where the mass terms of the up-type and down-type quarks are included in $\mathcal{L}_{\text{mass}}$

$$\mathcal{L}_{\text{mass}} = M^u_{ij} u'_{Li} u'_{Ri} + M^d_{ij} d'_{Li} d'_{Ri} + h.c. \tag{2.8}$$

In Eq. 2.8 $M^{u,d}$ denote the quark mass matrices given by

$$M^{u,d} = -\Lambda^{u,d} \frac{v}{\sqrt{2}} \,. \tag{2.9}$$

In general, $M^{u,d}$ have non-zero diagonal elements, but they can be diagonalized by multiplying them with a set of four unitary matrices

$$(V^u_L)^\dagger M^u V^u_R = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \,, \qquad (V^d_L)^\dagger M^d V^d_R = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \,. \tag{2.10}$$

This transformation rotates the eigenstates of the electroweak interaction into the mass eigenstates in the quark space and the electroweak eigenstates can be written as

$$u'_L = V^u_L u_L \,, \quad u'_R = V^u_R u_R \,, \quad d'_L = V^d_L d_L \,, \quad d'_R = V^d_R d_R \,. \tag{2.11}$$

If one introduces the mass eigenstates of the quarks in Eq. 2.4, the CC Lagrangian can be expressed as

$$\mathcal{L}_{\text{CC}} = -\frac{1}{\sqrt{2}} g \bar{u}_{Li} \gamma^\mu \bar{V}_{ij} d_{Li} W^+_\mu + h.c. \tag{2.12}$$

with the product of two unitary matrices $\bar{V} = V^u_L (V^d_L)^\dagger$. V denotes the already introduced CKM mass mixing matrix.

2.3 The CKM Mass Mixing Matrix

Sec. 2.2 showed that the W boson does not couple to the mass eigenstates of the quarks and the coupling strength depends on the CKM matrix, which translates the weak interaction eigenstates into the mass eigenstates:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.13}$$

Simply speaking, the total coupling constant for transitions like $q \rightarrow W^- Q$ is given by the product of the weak coupling constant g and the corresponding CKM matrix element V_{Qq} (Eq. 2.12), where Q stands for an up-type and q for a down-type quark. As can be seen in Eq. 2.13, this matrix is constructed as a 3×3 matrix and can be parametrized by three mixing angles θ_{ij} and one complex phase δ often called *Kobayashi-Maskawa* phase. It should be noted that this phase is the only source of ***CP* violation (~~*CP*~~)** in flavour changing interactions within the SM. In general, there are many possible parametrizations of the CKM matrix; the standard parametrization has been chosen as

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2.14}$$

with $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$. Measurements of CKM matrix elements $|V_{Qq}|$ yielded a clear hierarchy ($s_{13} \ll s_{23} \ll s_{12} \ll 1$), which has been exploited in a new parametrization; the *Wolfenstein* parametrization [9]

$$\begin{aligned} s_{12} &= \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} = \lambda\,, \quad s_{23} = \lambda\frac{|V_{cb}|}{|V_{us}|} = A\lambda^2\,, \\ s_{13}e^{i\delta} &= V_{ub}^* = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}\left[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})\right]} \end{aligned} \tag{2.15}$$

with a new set of parameters $A, \lambda, \bar{\rho}, \bar{\eta}$. Eqs. 2.15 ensure the reparametrized CKM matrix to be unitary in all orders of λ. Traditionally, the *Wolfenstein*-CKM matrix is expressed using ρ, η instead of $\bar{\rho}, \bar{\eta}$ leading to

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho + i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{2.16}$$

As mentioned above, the hierarchy of the CKM matrix elements is reflected in this parametrization. Since $\lambda = \sin\theta_C \approx 0.22$, where θ_C denotes the *Cabibbo* angle, the diagonal elements are of order 1, i.e. transitions within one quark generation are preferred. Such processes are often referred as *Cabibbo*-favoured decays. In contrast, the coupling of $W^{\pm}$ to quarks of different flavours is *Cabibbo*-suppressed by orders of the parameter λ. With the requirement of unitarity, i.e. $V^{\dagger}V = \mathbf{1}$, one obtains

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}\,, \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik} \quad \text{with} \quad \delta_{nm} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}\,. \tag{2.17}$$

Six of these relations are vanishing ($\delta_{nm} = 0$) and represent triangles in the complex plane. One of these ***Unitarity Triangles* (UT)**, namely

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.18}$$

or after dividing by $V_{cd}V_{cb}^*$

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \tag{2.19}$$

is commonly used to visualize and to judge the CKM formalism with respect to experimental results. The vertices of this triangle in the complex plane are equal to $(0,0)$, $(1,0)$ and $(\bar{\rho},\bar{\eta})$ (Fig. 2.2a), which directly follows from the *Wolfenstein* parametrization (Eq. 2.15).

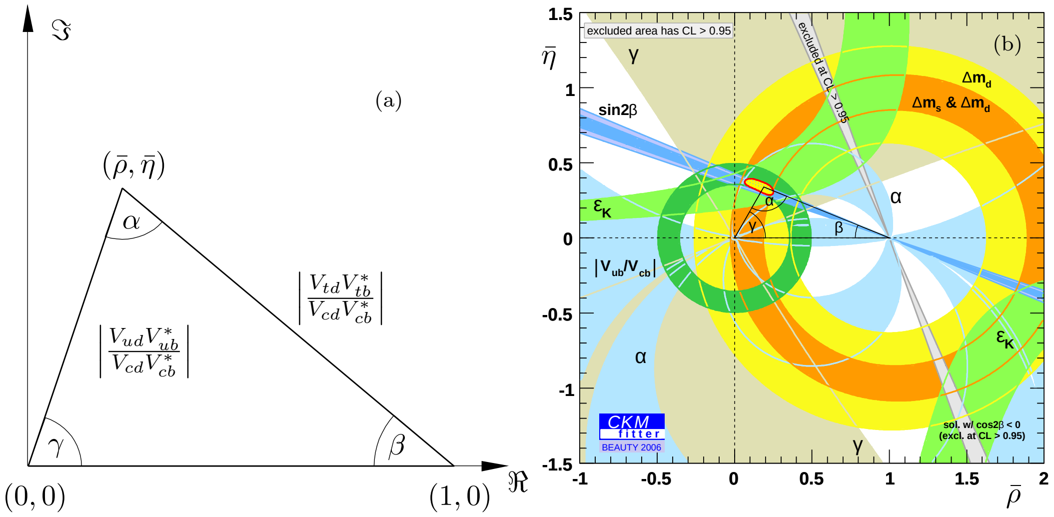


Figure 2.2: (a) Sketch of the Unitarity Triangle: The side lengths and angles do not correspond to the experimental values. (b) Global SM Fit to the $\bar{\rho}-\bar{\eta}$ plane: The red-bordered yellow area illustrates the 1σ contour of the apex directly extracted from the global fit likelihood function [10, 11]. The input and output values of the fit parameters are summarized in Ref. [12].

The precise measurement of all angles and side lengths of this UT is one important aim of todays flavour physics. The processes and experiments, the CKM parameters can be independently derived from, are not discussed here, but a detailed overview of the status of such measurements can be found in Ref. [13].

In principle, the experimental data could include non-SM physics, which could possibly enter, e.g. in loop processes. Such contributions result in inconsistencies with the CKM picture and the UT would not perfectly close at one point. The experimental results can therefore be used to constrain global fits in the $\bar{\rho}-\bar{\eta}$ plane to check for such deviations from the prediction. This tests the validity of the CKM model.

Fig. 2.2b illustrates the state of summer 2006. The experimental data are in almost perfect agreement with the CKM picture and there is not much space left for new physics on the given level of precision. However, some of the input measurements depend on parameters, which still include high theoretical uncertainties and therefore limit these measurements to some extent. As it is shown in Sec. 2.4, the decay

$B^- \to \tau^- \overline{\nu}_\tau$ provides important information to improve the knowledge of the physics on the B sector and supplies additional constraints to the global SM fit in the $\bar{\rho} - \bar{\eta}$ plane.

2.4 The Decay $B^- \to \tau^- \overline{\nu}_\tau$

This section discusses the leptonic decay $B^- \to \ell^- \overline{\nu}_\ell$ ($\ell = e, \mu, \tau$) and its relevance for SM parameters. It further shows, how this decay acts as a probe of new physics.

Within the SM the decay $B^- \to \ell^- \overline{\nu}_\ell$ proceeds through a weak annihilation diagram illustrated in Fig. 2.3. In the following the decay rate for this process is calculated.[1] The single steps are taken from a discussion on the charged pion decay $\pi^- \to \ell^- \overline{\nu}_\ell$ given in Ref. [8]. Since this process is in total analogy to leptonic B decays, the argumentation, apart of some small modifications, can be adopted.

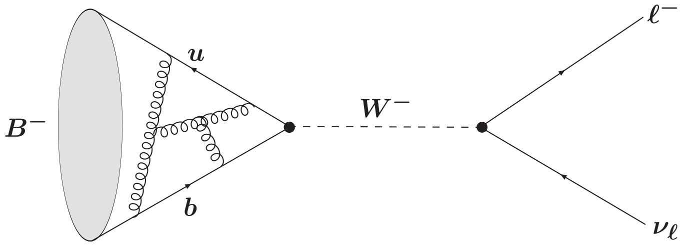


Figure 2.3: *Feynman* diagram of the decay $B^- \to \ell^- \overline{\nu}_\ell$. A horizontal time axis is assumed. The gluon lines illustrate the bound state of the quarks within the B meson.

The SM matrix element $\mathcal{M}_{\mathrm{SM}}$ consists of two independent parts; the leptonic current describing the $W^- \to \ell^- \overline{\nu}_\ell$ decay, and the hadronic part J_μ^{had}, which introduces the B meson current. $\mathcal{M}_{\mathrm{SM}}$ can be written as

$$\mathcal{M}_{\mathrm{SM}} = \frac{G_F}{\sqrt{2}} V_{ub} J_\mu^{had} \cdot \bar{u}_\ell \gamma^\mu (1 - \gamma^5) v_{\nu_\ell} \,, \tag{2.20}$$

where G_F denotes the well-known *Fermi* constant ($G_F = 1.166 \times 10^{-5}\,\mathrm{GeV}^{-2}$ [13]). As discussed in Sec. 2.3, V_{ub} needs to be inserted to account for the fact that the W boson couples to the weak interaction eigenstates of the u and b quark instead of their mass eigenstates. The weak coupling to the leptonic system with its *Dirac* spinors u_ℓ and v_{ν_ℓ} exhibits the typical $V - A$ structure, i.e. the W couples with a combination of a vector (V) and an axial vector (A) current with

$$V^\mu = \bar{u}_\ell \gamma^\mu v_{\nu_\ell} \,, \qquad A^\mu = \bar{u}_\ell \gamma^\mu \gamma^5 v_{\nu_\ell} \,. \tag{2.21}$$

One may expect the W coupling to the u and b quarks of the B^- to be of same nature as for the leptonic part, i.e. $\bar{u}_u \gamma^\mu (1 - \gamma^5) v_b$. But this is not correct, since the quarks are bound within the B meson and can therefore strongly interact. Furthermore, the hadronic part cannot be calculated within perturbation theory. On the other hand,

---

[1] Throughout this thesis, the unit system $\hbar = c = 1$ is used in kinematical calculations.

the matrix element $\mathcal{M}_{\mathrm{SM}}$ needs to be invariant under *Lorentz*-Transformation, i.e. J_μ^{had} has to be a vector or axial vector. Since the B meson is a spinless particle, J_μ^{had} can only be constructed from the four-momentum vector P_μ of the B leading to the following approach

$$J_\mu^{had} = f_{B^-} P_\mu^{B^-} \,. \tag{2.22}$$

The factor f_{B^-} is a constant usually called the B *Decay Constant* and $f_{B^-}^2$ can be interpreted as the probability that the quarks inside the B meson come close enough for W^- exchange. Using the ansatz of Eq. 2.22 $\mathcal{M}_{\mathrm{SM}}$ can be expressed in the B rest frame ($P_\mu^{B^-} = (m_{B^-}, \vec{0})$) with the charged B meson mass m_{B^-} as

$$\mathcal{M}_{\mathrm{SM}} = \frac{G_F}{\sqrt{2}} V_{ub} f_{B^-} m_{B^-} \cdot \bar{u}_\ell \gamma^0 (1 - \gamma^5) v_{\nu_\ell} \,. \tag{2.23}$$

In order to calculate the decay rate of $B^- \to \ell^- \overline{\nu}_\ell$, one has to take into account the phase space factor of this two-body decay. In the B rest frame the differential rate can be written as

$$\frac{\mathrm{d}\Gamma_{\mathrm{SM}}(B^- \to \ell^- \overline{\nu}_\ell)}{\mathrm{d}\Omega} = \frac{|\vec{p}_\ell|}{32\pi^2 m_{B^-}^2} |\mathcal{M}_{\mathrm{SM}}|^2 \tag{2.24}$$

with the spatial angle Ω and the magnitude of the lepton three-momentum

$$|\vec{p}_\ell| = \frac{1}{2m_{B^-}} (m_{B^-}^2 - m_\ell^2) \,. \tag{2.25}$$

The quantitative evaluation of the leptonic current in Eq. 2.23 is straight-forward as given in Ref. [8]. At the end $|\mathcal{M}_{\mathrm{SM}}|^2$ is of very simple form

$$|\mathcal{M}_{\mathrm{SM}}|^2 = 2G_F^2 |V_{ub}|^2 f_{B^-}^2 m_{B^-}^2 m_\ell^2 \left[1 - \frac{m_\ell^2}{m_{B^-}^2}\right] \,. \tag{2.26}$$

The integration of Eq. 2.24 over Ω results in an additional factor of 4π and after insertion of Eq. 2.26 the rate of the decay $B^- \to \ell^- \overline{\nu}_\ell$ can be expressed as

$$\Gamma_{\mathrm{SM}}(B^- \to \ell^- \overline{\nu}_\ell) = \frac{G_F^2}{8\pi} |V_{ub}|^2 f_{B^-}^2 m_{B^-} m_\ell^2 \left[1 - \frac{m_\ell^2}{m_{B^-}^2}\right]^2 \,. \tag{2.27}$$

It is not surprising that we find the well-known helicity suppression of light lepton production, which directly follows from the $V - A$ structure of the weak interaction. The decay rates obtained for different lepton flavours compare as given in Eqs. 2.28

$$\begin{aligned} \frac{\Gamma_{\mathrm{SM}}(B^- \to \tau^- \overline{\nu}_\tau)}{\Gamma_{\mathrm{SM}}(B^- \to \mu^- \overline{\nu}_\mu)} &= \frac{m_\tau^2}{m_\mu^2} \left[\frac{m_{B^-}^2 - m_\tau^2}{m_{B^-}^2 - m_\mu^2}\right]^2 \approx 230 \,, \\ \frac{\Gamma_{\mathrm{SM}}(B^- \to \tau^- \overline{\nu}_\tau)}{\Gamma_{\mathrm{SM}}(B^- \to e^- \overline{\nu}_e)} &= \frac{m_\tau^2}{m_e^2} \left[\frac{m_{B^-}^2 - m_\tau^2}{m_{B^-}^2 - m_e^2}\right]^2 \approx 10^7 \,, \end{aligned} \tag{2.28}$$

where m_{B^-} and the lepton masses m_ℓ have been taken from Ref. [13]. Hence, the τ channel is of higher experimental interest than the other lepton channels, even though the analysis itself is much more challenging (Chap. 4).

In general, the decay constants of charged (f_{B^-}) and neutral B mesons (f_{B^0}) are different induced by the difference of the u and the d quark masses. On the other hand, this mass difference is extremely small compared to the mass of the b quark within the B meson and it is convenient to assume *Isospin symmetry.* In the following we assume

$$f_{B^-} = f_{B^0} \equiv f_B \,. \tag{2.29}$$

The branching fraction of $B^- \to \tau^-\overline{\nu}_\tau$ arises from Eq. 2.27 with the lifetime of charged B mesons τ_{B^-} to

$$\begin{aligned} \mathcal{B}_{\text{SM}}(B^- \to \tau^-\overline{\nu}_\tau) &= \Gamma_{\text{SM}}(B^- \to \tau^-\overline{\nu}_\tau)\tau_{B^-} \\ &= \frac{G_F^2}{8\pi}|V_{ub}|^2 f_B^2 m_{B^-} m_\tau^2 \left[1 - \frac{m_\tau^2}{m_{B^-}^2}\right]^2 \tau_{B^-} \,. \end{aligned} \tag{2.30}$$

2.4.1 Branching Fraction, f_B, and $|V_{ub}|$

The branching fraction of $B^- \to \tau^-\overline{\nu}_\tau$ (Eq. 2.30) depends on the product $|V_{ub}|f_B$, while all other parameters are already measured very precisely. Thus, within the SM a precise measurement of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ directly translates into $|V_{ub}|f_B$. These two parameters are of substantial importance for the understanding of B physics, but to date neither $|V_{ub}|$ nor f_B are precisely known.

Currently, most analyses to extract the CKM matrix element $|V_{ub}|$ are based on branching fraction measurements of semileptonic B decays $B \to X_u\ell\nu_\ell$. Generally, such measurements can be classified into *inclusive* and *exclusive.* The term inclusive refers to methods without explicit reconstruction of the hadronic system X_u, while within the exclusive measurements X_u is reconstructed in certain meson states, e.g. $B \to \pi\ell\nu_\ell$ or $B \to \rho\ell\nu_\ell$. All of them have to deal with very high $B \to X_c\ell\nu_\ell$ background, which leads to substantial statistical and systematic uncertainties of the measured branching fractions. Additionally, there are high theoretical uncertainties in the translation of the measured $B \to X_u\ell\nu_\ell$ branching fractions into $|V_{ub}|$.

The ***Heavy Flavour Averaging Group*** **(HFAG)** calculated and compared $|V_{ub}|$ from recent results from inclusive and exclusive measurements using different theoretical frameworks [14]. They quoted an average value for $|V_{ub}|$ from the inclusive measurements of

$$|V_{ub}^{\text{incl}}| = (4.52 \pm 0.19_{\text{ex}} \pm 0.27_{\text{th}}) \times 10^{-3} \,, \tag{2.31}$$

where a model by *Lange, Neubert*, and *Paz* [15] with experimental input from $b \to c\ell\nu$ and $b \to s\gamma$ moment measurements [16] has been used for the calculation. The total uncertainty is given to be of 7.3 %. The theoretical uncertainties were simply added in quadrature, but a different and more realistic treatment of these uncertainties can easily increase the total error to a level of about 10 %. On the other hand, HFAG also quoted $|V_{ub}|$ from exclusive $B \to \pi\ell\nu_\ell$ measurements, but if one wants to calculate $|V_{ub}|$ from the partial branching fraction of this process, it depends on form factor normalizations. These quantities cannot be exactly extracted but only determined by theoretical calculations, e.g. ***Lattice QCD*** **(LQCD)** [17–19] and ***Light-Cone Sum Rules*** **(LCSR)**[20]. Using these different calculations $|V_{ub}^{\text{excl}}|$ ranges between 3.38×10^{-3} and 3.93×10^{-3} strongly differing from the inclusive average value (Eq. 2.31). An averaging of the inclusive and exclusive results is very difficult, since it is not clear how to deal with the high uncertainties related to the theories used to extract $|V_{ub}|$ from the measurements. However, the global SM fit to the $\bar{\rho} - \bar{\eta}$ plane illustrated in Fig. 2.2b (Sec. 2.3) used an averaged input value of [12]

$$|V_{ub}| = (4.10 \pm 0.09_{\text{ex}} \pm 0.39_{\text{th}}) \times 10^{-3}\,. \tag{2.32}$$

Thus, this value has been used for further calculations within this thesis. In summary, to date $|V_{ub}|$ is known with large uncertainties of the order 10 % and its precision is limited by theory.

The situation for f_B is even worse compared to $|V_{ub}|$. It is poorly known and the most precise values have been determined within numerical calculations in the frameworks of LQCD and LCSR. Tab. 2.3 summarizes most recent results of such calculations. LQCD and LCSR calculations are in agreement and of compareable precision. f_B is known from these numerical calculations on the 10 % level. However, these values strongly depend on the theoretical description of QCD and it is indispensable to determine f_B from real data. Once again it is mentioned, that with an improved knowledge of $|V_{ub}|$ and a precise measurement of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ f_B is directly measured and can be compared to the theoretical calculations.

f_B/MeV	Method
$216 \pm 9 \pm 19 \pm 4 \pm 6$	LQCD (unquenched, improved staggered quarks) [21]
$191 \pm 10^{+12}_{-22}$	LQCD (unquenched) [22]
$204 \pm 8 \pm 29^{+44}_{-0}$	LQCD (unquenched) [23]
206 ± 20	LCSR [24]

Table 2.3: Theoretical Calculations of f_B: The term "unquenched" implicates that the calculations use the full QCD. In contrast, lattice calculations in "quenched" approximation omit sea quarks effects. The concept of "improved staggered quarks" is an alternative discretisation of the quark fields and results in more efficient numerical calculations. A detailed overview of the concepts and mathematics of LQCD is given in Refs. [25, 26].

If one now calculates the SM prediction $\mathcal{B}_{\text{SM}}(B^- \to \tau^- \overline{\nu}_\tau)$ from Eq. 2.30 using the $|V_{ub}|$ average from Eq. 2.32 and f_B from an unquenched LQCD calculation [22], one obtains

$$\mathcal{B}_{\text{SM}}(B^- \to \tau^- \overline{\nu}_\tau) = (1.08 \pm 0.05\,^{+0.27}_{-0.34}) \times 10^{-4} \left[\frac{|V_{ub}|}{0.00410}\right]^2 \left[\frac{f_B}{191\,\text{MeV}}\right]^2 , \tag{2.33}$$

where the first error is induced by the experimental error of $|V_{ub}|$ and the second reflects the combined theoretical uncertainties of f_B and $|V_{ub}|$. This leads to the question: "What is the experimental status of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$?"

Collaboration	**Year**	$\mathcal{L}/\,\mathrm{fb}^{-1}$	$\sqrt{s}/\,\mathrm{GeV}$	$\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)/10^{-4}$
L3	1997	0.005	91.2	< 5.7 [27]
ALEPH	2000	–	91.2	< 8.3 [28]
CLEO	2000	9	10.58	< 8.4 [29]
BABAR	2003	82	10.58	< 4.1 [30]
BABAR	2004	82	10.58	< 4.2 [31]
BABAR	2006	211	10.58	< 2.6 [32]
BABAR	2006	288	10.58	< 1.8 [33]
BABAR	2007	346	10.58	< 1.7 [34]
Belle	2004	140	10.58	< 2.9 [35]
Belle	2005	253	10.58	< 1.8 [36]
Belle	2006	414	10.58	$1.06\,^{+0.34}_{-0.28}\,^{+0.18}_{-0.16}$ [37]
Belle	2006	414	10.58	$1.79\,^{+0.56}_{-0.49}\,^{+0.46}_{-0.51}$ [38]

Table 2.4: Former results on branching fraction of $B^- \to \tau^-\overline{\nu}_\tau$: The upper limits are given at the 90 % confidence level (C.L.). The results of the LEP experiments ALEPH and L3 were extracted from data taken at the Z^0 pole at a center-of-mass (c.m.) energy of $\sqrt{s} \approx m_Z$. The CLEO, *BABAR* and *Belle* data were taken on the $\Upsilon(4S)$ resonance. $\mathcal{L}$ denotes the integrated luminosity used to perform the analyses.

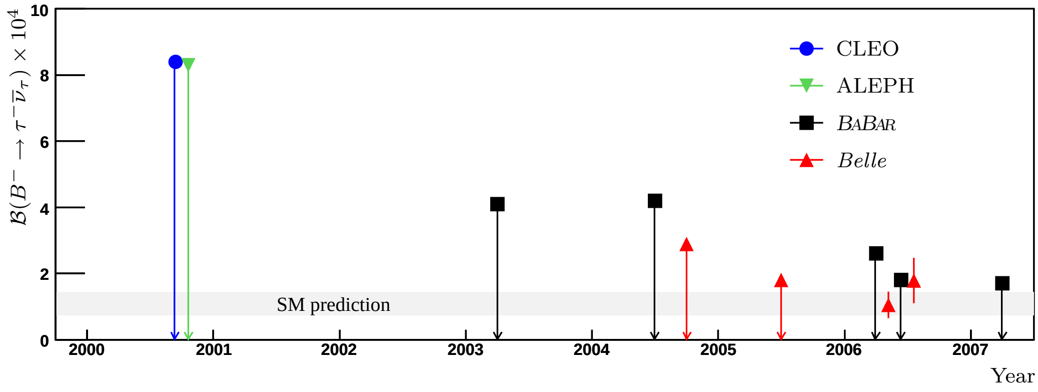


Figure 2.4: The $B^- \to \tau^-\overline{\nu}_\tau$ History: The plot shows former results of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ listed in Tab. 2.4.

Many searches have been performed to observe this important process (Tab. 2.4, Fig. 2.4). Since 2003 the $B^- \to \tau^-\overline{\nu}_\tau$ search is dominated by the B factories and in fact, the *Belle* collaboration claimed an observation at the conference of ***Flavour Physics and CP violation 2006*** **(FPCP06)** [37]. The central value of $(1.06\,^{+0.34}_{-0.28}\,^{+0.18}_{-0.16}) \times 10^{-4}$ has been corrected at the ***International Conference of High Energy Physics 2006*** **(ICHEP06)** to be $(1.79\,^{+0.56}_{-0.49}\,^{+0.39}_{-0.46}) \times 10^{-4}$ [39][2]. At the same conference *BABAR* presented a new measurement with a central value

---

[2] The systematic errors given here differ from that given in Ref. [38].

of $(0.88\,^{+0.68}_{-0.67} \pm 0.11) \times 10^{-4}$ [33]. HFAG combined the full likelihoods of these two independent measurements and determined an average value of [40]

$$\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) = (1.34\,^{+0.48}_{-0.47}) \times 10^{-4}\,. \tag{2.34}$$

In spring 2007 an update of the *BABAR* analysis given in Ref. [33] has been published [34]. Furthermore, a second *BABAR* analysis has been developed, which is independent to the aforementioned *BABAR* measurement since different selection techniques have been used. Therefore, the likelihoods of these two independent measurements have been combined and a new *BABAR* result of

$$\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) = (1.20\,^{+0.40}_{-0.38}\,^{+0.29}_{-0.30} \pm 0.22) \times 10^{-4} \tag{2.35}$$

has been presented at FPCP07 [41], where the first error is statistical, the second is due to the systematic background prediction uncertainty, and the third is due to other systematic sources.

However, all former measurements of the *BABAR* and *Belle* collaborations are limited by statistics, which directly follows from the techniques used to filter out signal events from the enormous background. The essential procedures of these measurements described in Chap. 4 suffer from low reconstruction efficiencies and therefore can be possibly improved by different reconstruction methods. But before such new methods are discussed in detail, some additional implications of the decay $B^- \to \tau^-\overline{\nu}_\tau$ are given.

2.4.2 Constraints on the $\bar{\rho} - \bar{\eta}$ plane

An additional aspect reveals by considering flavour oscillations on the B sector, which can proceed via box diagrams as illustrated in Fig. 2.5a. B^0 mesons can turn into $\overline{B}^0$ with opposite flavour composition and vice versa. Such oscillations are described

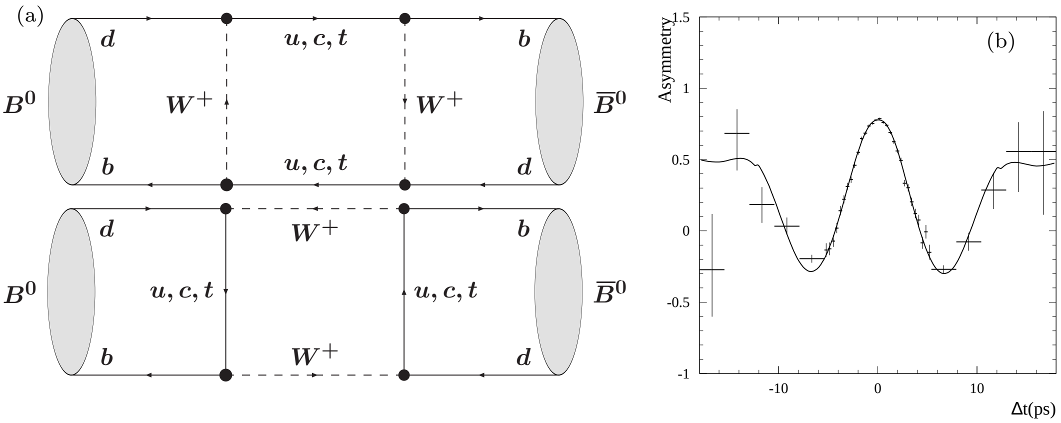


Figure 2.5: $B^0 - \overline{B}^0$ mixing: (a) possible *Feynman* diagrams, (b) oscillation from a time-dependent asymmetry measurement by *BABAR* [42]

by the oscillation frequency

$$\Delta m_d = \frac{G_F^2}{6\pi^2} m_{B^0} m_W^2 f_{B^0}^2 B_{B^0} \eta_{B^0}(x_t) S_0(x_t) |V_{td}|^2 , \tag{2.36}$$

where $\eta_{B^0}(x_t)$ denotes a QCD correction factor depending on Λ_{QCD} and the top quark mass m_t. $S_0(x_t)$ is called *Inami-Lim function* [43] and is given as [44]

$$S_0(x_t) = \frac{4x_t - 11x_t^2 + x_t^3}{4(1-x_t)^2} - \frac{3x_t^3 \ln x_t}{2(1-x_t)^3} \quad \text{with} \quad x_t = \frac{m_t^2}{m_W^2} . \tag{2.37}$$

The subscript "0" indicates that Eq. 2.37 does not include QCD corrections to the box diagrams. In general, all three up-type quarks can contribute to the $B^0 - \overline{B}^0$ mixing, but u and c contributions are dominated by the t contribution. This is no consequence of the CKM mixing matrix, since all relevant combinations $V_{ib}^* V_{id}$ with $i = u, c, t$ are of compareable size in magnitude (Sec. 2.3). But rather it arises from the fact, that the top quark is much heavier than the others [45]. It is shown in Ref. [46] that the product $\eta_{B^0}(x_t) S_0(x_t)$ does not sensitively depend on the definition of the top quark mass anymore and therefore $\eta_{B^0}(x_t) = \eta_{B^0} = (0.55 \pm 0.01)$ can be assumed in the calculation of Δm_d (Eq. 2.36).

The calculation of the hadronic matrix element $\mathcal{M}_{B^0 \to \overline{B}^0}$ for $B^0 - \overline{B}^0$ transitions suffers from the standard difficulties of strong interactions since a correct determination of such a matrix element requires a complete theory of structure and interactions of hadrons [47]. In order to solve this problem, a sum over a complete set of possible states n_i is inserted between the $V - A$ currents (Eqs. 2.38). This is still exact since this complete set of states gives 1.

$$\begin{aligned} \mathcal{M}_{B^0 \to \overline{B}^0} &= \left\langle B^0 \left| \left[\bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right]^2 \right| \overline{B}^0 \right\rangle \\ &= \sum_i \left\langle B^0 \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| n_i \right\rangle \left\langle n_i \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| \overline{B}^0 \right\rangle \\ &\overset{\text{VIA}}{\approx} \left\langle B^0 \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| 0 \right\rangle \left\langle 0 \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| \overline{B}^0 \right\rangle \end{aligned} \tag{2.38}$$

The sum is then truncated with the contribution of the vacuum state corresponding to an annihilation of a B^0 and a subsequent generation of a $\overline{B}^0$ (Eq. 2.38). Since the vacuum state is naturally not equal to a sum over all possible states n_i, this technique is an approximation (***Vacuum Insertion Approximation*** **(VIA)**). The ratio of $\mathcal{M}_{B^0 \to \overline{B}^0}$ and the approximated matrix element $\mathcal{M}^{\text{VIA}}_{B^0 \to \overline{B}^0}$ is parametrized by the *Bag Parameter* B_{B^0} as shown in Eq. 2.39.

$$\begin{aligned} \mathcal{M}_{B^0 \to \overline{B}^0} &= B_{B^0} \cdot \mathcal{M}^{\text{VIA}}_{B^0 \to \overline{B}^0} \\ &= B_{B^0} \cdot \left\langle B^0 \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| 0 \right\rangle \left\langle 0 \left| \bar{u}_b \gamma^\mu (1-\gamma^5) v_d \right| \overline{B}^0 \right\rangle \\ &= \frac{8}{3} B_{B^0} f_{B^0}^2 m_{B^0}^2 \end{aligned} \tag{2.39}$$

B_{B^0} cannot be calculated within perturbation theory. Hence, again theoretical estimates are needed. Using LQCD a value of $B_{B^0} = 0.836 \pm 0.027\,^{+0.056}_{-0.062}$ [22] has been calculated corresponding to a 5 % uncertainty.

The expressions for the branching fraction of $B^- \to \tau^- \overline{\nu}_\tau$ (Eq. 2.30) and the $B^0 - \overline{B}^0$ mixing frequency (Eq. 2.36) can be compared

$$\begin{aligned} \frac{\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^- \overline{\nu}_\tau)}{\Delta m_d} &= \frac{3\pi}{4} \frac{m_{B^-} m_\tau^2 \left[1 - \frac{m_\tau^2}{m_{B^-}^2}\right]^2 \tau_{B^-}}{\eta_{B^0} m_{B^0} m_W^2 B_{B^0} S_0(x_t)} \cdot \frac{f_{B^-}^2}{f_{B^0}^2} \cdot \frac{|V_{ub}|^2}{|V_{td}|^2} \\ &= F \cdot \frac{f_{B^-}^2}{f_{B^0}^2} \cdot \frac{|V_{ub}|^2}{|V_{td}|^2} \,. \end{aligned} \qquad (2.40)$$

Consequently, exploiting Eq. 2.29 the uncertain parameter f_B drops out and the ratio of CKM elements $|V_{ub}|/|V_{td}|$ becomes accessible with precise measurements of the branching fraction and the oscillation frequency. Δm_d has already been measured by various experiments with high accuracy and the HFAG average is given as [48]

$$\Delta m_d = (0.507 \pm 0.004)\,\mathrm{ps}^{-1} \,. \qquad (2.41)$$

The uncertainty of the prefactor F in Eq. 2.40 is of the order 6 % mainly driven by B_{B^0}. Thus, this is an additional motivation to improve $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ since it essentially relates the lengths of the two sides inferring at the apex of the UT (Sec. 2.3). Geometrically, the ratio of the radii of the circles labeled with Δm_d and $|V_{ub}/V_{cb}|$ around $(0,0)$ and $(1,0)$ in the complex *Wolfenstein* plane are not independent but constrained by $\sqrt{\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)/(F \cdot \Delta m_d)}$ (Fig. 2.2).

2.4.3 New Physics in $B^- \to \tau^- \overline{\nu}_\tau$

As shown in Sec. 2.2, within the SM masses are generated by a coupling to one scalar *Higgs* field, but to date the existence of a *Higgs* boson has not been proved experimentally and it is prudent to explore different possibilities of *Higgs* models, e.g. models with more than one isospin doublet. However, there are two major constraints on the *Higgs* sector; $m_W^2/m_Z^2 \cos^2\theta_W \approx 1$ and the strong limits on the existence of tree-level *Flavour Changing Neutral Currents* [49]. A very popular extension of the minimal *Higgs* model is represented by ***Two-Higgs Doublet Models*** **(2HDM)**, which are realized in low-energy supersymmetric models, such as the ***Minimal Supersymmetric Standard Model*** **(MSSM)** [49].

2.4.3.1 The Two-*Higgs* Doublet Model

Within the general 2HDM an additional isospin doublet is introduced that way, that the aforementioned constraints are satisfied. The *Higgs* potential now depending on two scalar fields Φ_1, Φ_2 becomes much more complicated compared to the minimal model and leads to five physical *Higgs* states; one pseudoscalar A^0, two neutral scalars H^0 and h^0, and two charged scalars $H^\pm$ (Tab. 2.5) [49]. In contrast to the minimal model, this model has six free parameters: four *Higgs* masses, the ratio of the vacuum expectation values $\tan\beta = v_2/v_1$ and the mixing angle α between the two neutral scalar states. The important parameter $\tan\beta$ is predicted by different models to be of the order m_t/m_b.

Higgs Boson	**Mass Eigen State**
$H^{\pm}$	$-\Phi_1^{\pm}\sin\beta + \Phi_2^{\pm}\cos\beta$
A^0	$\sqrt{2}(-\Im(\Phi_1^0)\sin\beta + \Im(\Phi_2^0)\cos\beta$
H^0	$\sqrt{2}\left[(\Re(\Phi_1^0) - v_1)\cos\alpha + (\Re(\Phi_2^0) - v_2)\sin\alpha\right]$
h^0	$\sqrt{2}\left[-(\Re(\Phi_1^0) - v_1)\sin\alpha + (\Re(\Phi_2^0) - v_2)\cos\alpha\right]$

Table 2.5: Physical *Higgs* states of 2HDM model [49]: β is given by $\tan\beta = v_2/v_1$ with the vacuum expectation values $v_{1,2}$ of the $\Phi_{1,2}$ fields. α denotes the mixing angle between the H^0 and h^0 states.

In 2HDM the *Higgs*-fermion coupling is model dependent and one has to choose how to couple the quarks and leptons to the *Higgs* doublets. In the so-called **Model-II** it is assumed that Φ_1 only couples to up-type quarks and leptons, while Φ_2 only couples to down-type quarks and neutrinos [49]

$$\Phi_1 = \begin{pmatrix} \Phi_1^{0*} \\ -\Phi_1^- \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} \Phi_2^+ \\ \Phi_2^0 \end{pmatrix}. \tag{2.42}$$

The vacuum expectation values are given as

$$\langle\Phi_1\rangle = \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \quad \langle\Phi_2\rangle = \begin{pmatrix} 0 \\ v_2 \end{pmatrix}. \tag{2.43}$$

Given this assumption, the Lagrangian of charged *Higgs*-fermion interaction takes the form [50]

$$\mathcal{L}_{H^{\pm}ff'} = \frac{g}{2\sqrt{2}m_W}\left[H^+ V_{ff'}\bar{u}_f(A + B\gamma^5)v_{f'} + H^- V_{ff'}^*\bar{u}_{f'}(A - B\gamma^5)v_f\right]. \tag{2.44}$$

Here, f denotes an up-type quark or neutrino, while f' stands for a down-type quark or lepton. It is remarkable, that similarly to the W coupling to quarks the *Higgs*-quark coupling includes the corresponding CKM matrix element $V_{ff'}$. The parameters A, B are related to the ratio of the vacuum expectation values $\tan\beta$

$$A = m_{f'}\tan\beta + m_f\cot\beta, \quad B = m_{f'}\tan\beta - m_f\cot\beta. \tag{2.45}$$

Given such an extended *Higgs* model, new phenomena and effects can arise even on the low energy scale. In fact, the existence of charged *Higgs* bosons would be relevant for the decay $B^- \to \tau^-\overline{\nu}_\tau$.

2.4.3.2 Charged *Higgs* Boson Effects in $B^- \to \tau^-\overline{\nu}_\tau$

Within the SM the weak annihilation is the only diagram contributing to the matrix element for $B^- \to \tau^-\overline{\nu}_\tau$ (Fig. 2.3). This is only true if the minimal *Higgs* model is realized. More complicated models, e.g. 2HDM, can easily extend the *Higgs* sector resulting in new physical *Higgs* states. An existence of charged *Higgs* bosons opens the possibility of annihilation of the quarks inside the B meson into $H^{\pm}$ instead of $W^\pm$ bosons (Fig. 2.3), which can decay via $H^- \to \tau^-\overline{\nu}_\tau$. Such a contribution modifies the decay rate of $B^- \to \tau^-\overline{\nu}_\tau$ and results in a measureable deviation of the branching fraction from the SM prediction $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)$.

Within the 2HDM the transition amplitude of $B^- \to \tau^-\overline{\nu}_\tau$ including a W^- and H^- part can be written as [51, 52]

$$\begin{aligned}\mathcal{M}_{\mathrm{2HDM}} = \frac{G_F V_{ub}}{\sqrt{2}} \big[\big\langle 0 \left|\bar{u}_u \gamma_\mu (1-\gamma_5) v_b\right| B^- \big\rangle\, \bar{u}_\tau \gamma^\mu (1-\gamma_5) v_{\nu_\tau} \\ - R_\tau \big[\big\langle 0 \left|\bar{u}_u (1+\gamma_5) v_b\right| B^- \big\rangle\, \bar{u}_\tau (1-\gamma_5) v_{\nu_\tau}\big]\ .\end{aligned} \tag{2.46}$$

The hadronic current $\langle 0 \left|\bar{u}_u \gamma_\mu (1-\gamma_5) v_b\right| B^- \rangle$ for weak quark-quark annihilation inside the B meson is equal to the previously defined J^{had}_μ (Eq. 2.22). R_τ depends on $\tan\beta$ as well as the charged *Higgs* boson mass m_{H^-}

$$R_\tau = \tan^2\beta \cdot \frac{m_b m_\tau}{m^2_{H^-}}\ . \tag{2.47}$$

In contrast to the axial vector current of the W exchange, the pseudoscalar current of the *Higgs* coupling to the constituents of the B meson can be expressed by [50]

$$\big\langle 0 \left|\bar{u}_u (1+\gamma_5) v_b\right| B^- \big\rangle = g_{B^-} \frac{m^2_{B^-}}{m_b}\ . \tag{2.48}$$

In analogy to f_{B^-}, the parameter g_{B^-} describes the probability that the quarks inside the B can annihilate via H^- exchange. Due to its property as a scalar meson it is expected that $f_B = (m_{B^-}/m_b) g_{B^-}$ if the quarks are in rest. Hence, in literature g_{B^-} is mostly presumed to be equal to f_B. Applying these substitutions and after calculation of the leptonic axial and pseudoscalar currents one gets a very simple expression for the branching fraction within the type-II 2HDM model

$$\mathcal{B}_{\mathrm{2HDM}}(B^- \to \tau^-\overline{\nu}_\tau) = \mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau) \cdot \left[1 - \tan^2\beta \frac{m^2_{B^-}}{m^2_{H^-}}\right]^2\ . \tag{2.49}$$

Eq. 2.49 implies that the presence of a H^- boson exchange in $B^- \to \tau^-\overline{\nu}_\tau$ simply modifies the SM prediction by a factor

$$r_H = \left[1 - \tan^2\beta \frac{m^2_{B^-}}{m^2_{H^-}}\right]^2\ . \tag{2.50}$$

r_H does not depend on the lepton mass, which retains lepton universality since the branching fractions of all $B^- \to \ell^-\overline{\nu}_\ell$ modes are modified by the same factor. Furthermore, r_H strongly depends on the meson mass. Thus, there is a higher potential to measure charged *Higgs* effects in leptonic B decays than e.g. in D or K decays.

Eq. 2.50 further implies charged *Higgs* contributions to the $B^- \to \tau^-\overline{\nu}_\tau$ amplitude to result in either enhancement ($r_H > 1$) or suppression ($r_H < 1$). Destructive interference ($r_H < 1$) leads to suppression with two solutions for $\tan\beta$

$$\tan\beta = (1 \pm \sqrt{r_H})^{1/2} \frac{m_{H^-}}{m_{B^-}}\ . \tag{2.51}$$

On the other hand, in the case of enhancement $\tan\beta$ is given by the "+" solution of Eq. 2.51.

Consequently, a precise measurement of the branching fraction could possibly give hints on physics beyond the SM on the *Higgs* sector. Moreover, assuming 2HDM a comparison with the SM prediction can be used to set limits on the charged *Higgs* boson mass depending on $\tan\beta$ (Eq. 2.50). One should note that although 2HDM is required in MSSM, the existence of charged *Higgs* bosons does not necessarily imply supersymmetry but only reflects a non-minimal *Higgs* model.

2.5 Properties of the τ Lepton

The last part of the theory chapter deals with the properties of the τ lepton and particularly discusses the structure of the τ decay into three pions and one neutrino.

2.5.1 Mass, Lifetime, and Main Decays

In Sec. 2.1 three lepton isospin doublets L_i have been introduced and it further has been mentioned that the lepton masses increase with increasing generation index i. The τ lepton belongs to the third family and therefore it is the heaviest lepton within the three-generation SM. Since its discovery in 1975 [53], the τ mass has been measured by a variety of experiments and nowadays the world average is [13]

$$m_\tau = (1776.99^{+0.29}_{-0.26})\,\mathrm{MeV}/c^2\,. \tag{2.52}$$

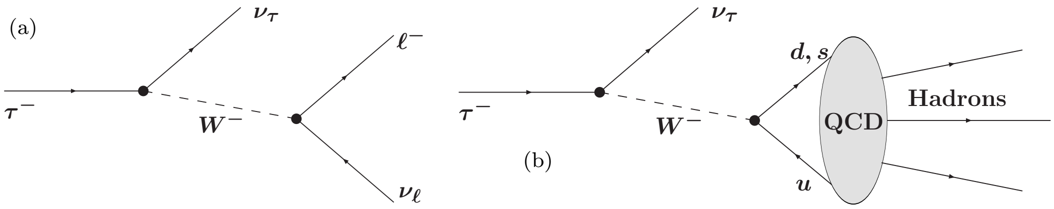


Figure 2.6: Topology of τ decays: *Feynman* graphs for (a) purely leptonic decays $\tau^- \to \ell^-\overline{\nu}_\ell\nu_\tau$ ($\ell = e, \mu$) and (b) hadronic decays $\tau^- \to (nh)^-\nu_\tau$. The ellipse illustrates the non-calculable QCD effects in the final state.

Due to its large mass τ's not only decay leptonically (Fig. 2.6a) but there is the possibility of production of light hadrons (h), e.g. π, η, or even strange mesons (K, $\phi(1020)$) as illustrated in Fig. 2.6b. The rich spectrum of τ decay modes (Tab. 2.6) results in a mean lifetime τ_τ of [13]

$$\tau_\tau = (290.6 \pm 1.0) \times 10^{-15}\,\mathrm{s}\,, \tag{2.53}$$

which is many orders of magnitude below the μ lifetime $\tau_\mu = 2.197 \times 10^{-6}\,\mathrm{s}$ [13]. However, the mean flight length of τ leptons acts as one experimental motivation to

Decay Mode	**Branching Fraction**
$\tau^- \rightarrow e^-\overline{\nu}_e\nu_\tau$	$(17.36 \pm 0.05)\,\%$
$\tau^- \rightarrow \mu^-\overline{\nu}_\mu\nu_\tau$	$(17.84 \pm 0.05)\,\%$
$\tau^- \rightarrow \pi^-\nu_\tau$	$(10.90 \pm 0.07)\,\%$
$\tau^- \rightarrow \pi^-\pi^0\nu_\tau$	$(25.50 \pm 0.10)\,\%$
$\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$	$(9.33 \pm 0.08)\,\%$
$\tau^- \rightarrow \pi^-\pi^0\pi^0\nu_\tau$	$(9.25 \pm 0.12)\,\%$
$\tau^- \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$	$(4.59 \pm 0.07)\,\%$
other	$\approx 5\,\%$

Table 2.6: τ decay channels [13]

use the decay $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ to search for $B^- \rightarrow \tau^-\overline{\nu}_\tau$ (Chap. 4). Hence, in the following the mean flight length of τ's in the B rest frame is estimated.

The calculation of the magnitude of the three-momentum vector $|\vec{p}_\tau|$ and the energy E_τ of τ leptons originated in the two-body decay $B^- \rightarrow \tau^-\overline{\nu}_\tau$ is straightforward. Assuming massless neutrinos ($m_{\nu_\tau} = 0$) and $|\vec{p}_{B^-}| = 0$ for a B^- in rest one obtains

$$E_\tau = \frac{m_{B^-}^2 + m_\tau^2}{2m_{B^-}}\,, \quad |\vec{p}_\tau| = \sqrt{E_\tau^2 - m_\tau^2} = \frac{m_{B^-}^2 - m_\tau^2}{2m_{B^-}}\,. \tag{2.54}$$

The mean flight length of τ leptons in the B rest frame can now be calculated as

$$d_\tau^{B^-} = \beta\gamma c\tau_\tau = \frac{|\vec{p}_\tau|}{m_\tau}c\tau_\tau \tag{2.55}$$

with

$$\beta = \frac{|\vec{p}_\tau|}{E_\tau}\,, \quad \gamma = \sqrt{\frac{1}{1-\beta^2}} = \frac{E_\tau}{m_\tau}\,. \tag{2.56}$$

Given the value of $|\vec{p}_\tau|$ from Eq. 2.54 and $c\tau_\tau = 87.11\,\mu\text{m}$ [13], the mean flight length in the B rest frame arises to

$$d_\tau^{B^-} = 114.71\,\mu\text{m}\,. \tag{2.57}$$

Thus, in the B rest frame the B^- and the subsequent τ^- decay are spatially separated. This can be exploited experimentally by reconstructing the decay vertex of the τ, but reliable decay vertex reconstruction cannot be applied with only one charged particle produced in the final state. Hence, decays of so-called *1-prong*[3] topology do not provide decay vertex information and it is indispensable to use other τ channels, e.g. $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$.

Theoretical descriptions of hadronic τ decays are difficult since they rely on hadronic dynamics in the intermediate energy range. Indeed, $\tau^- \rightarrow (nh)^-\nu_\tau$ decays show a rich structure of resonances, which need to be described kinematically,

---

[3] The term "1-prong" has been originally introduced by τ physics ($e^+e^- \rightarrow \tau^+\tau^-$) and denotes τ decays into one charged particle, i.e. $\tau^- \rightarrow \ell^-\overline{\nu}_\ell\nu_\tau$ and $\tau^- \rightarrow h^-(nh^0)\nu_\tau$ ($n = 0, 1, 2...$). In contrast, "3-prong" decays denote processes with three charged particles involved.

but the hadronic physics illustrated by the gray ellipse in Fig. 2.6b is not calculable by fundamental theories and its description is strongly model-dependent. Sec. 2.5.2 discusses in detail the situation for $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$.

2.5.2 The Decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$

The decay of τ leptons to an odd number of pions occurs almost exclusively through the axial-vector current. This directly follows from transformation properties of the weak current under G-parity [54]. Furthermore, assuming conservation of parity the 3π system is forced to be either a $J^P = 0^-$ or a $J^P = 1^+$ state. Indeed, the hadronic structure of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ has been studied by the OPAL collaboration [55, 56] in a model-independent analysis. No evidence for vector or scalar currents has been found.

The invariant amplitude for the hadronic decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ can be written in form of a factorized current-current interaction [57]

$$\mathcal{M}_{\tau^-\to\pi^-\pi^+\pi^-\nu_\tau} = \frac{G_F}{\sqrt{2}} |V_{ud}| \bar{u}_{\nu_\tau} \gamma_\mu (1-\gamma_5) v_\tau J^\mu_{had} \tag{2.58}$$

with the 3π hadronic transition current J^μ_{had} and the known leptonic $V-A$ current. The hadronic physics described by J^μ_{had} is the piece of interest since it probes the matrix element of the left-handed charged current between the vacuum and the 3π final state. It is of the general form [54][4]

$$J^\mu_{had} = -\left[g^{\mu\nu} + \frac{p^\mu_{3\pi} p^\nu_{3\pi}}{p^2_{3\pi}} \right] \times (q^1_\nu F^1 + q^2_\nu F^2 + q^3_\nu F^3) + p^\mu_{3\pi} F^4 \tag{2.59}$$

with the four-momentum of the 3π system $p^\mu_{3\pi}$ and the four-momentum transfers $q^\mu_i = p^\mu_{\pi_j} - p^\mu_{\pi_k}$ $(i \neq j \neq k)$. The form factors F_i model the unknown hadronic physics. The term proportional to $p^\mu_{3\pi}$ corresponds to scalar contributions predicted to be strongly suppressed as has been proven by the aforementioned OPAL measurement. Since the momentum transfers in τ decays are small, the form factors F_i are expected to be dominated by resonances. Different models to describe the form factors F_i are on the market [58–61]. Two popular models are discussed in the following.

2.5.2.1 The KS Model

In the model of ***Kühn*** and ***Santamaria*** **(KS)** [58] the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ is believed to be dominated by the $J^P = 1^+$ state a_1^-. This resonance is known to decay essentially through $a_1^- \to \rho^0\pi^-$ or $a_1^- \to \rho^-\pi^0$. Referring to Eq. 2.59 this assumption implies $F_3 = F_4 = 0$ and the hadronic current can be written as

$$\left\langle \pi^-(p^\mu_{\pi_1})\pi^-(p^\mu_{\pi_2})\pi^+(p^\mu_{\pi_3}) \left| A^\mu(0) \right| 0 \right\rangle \equiv J^\mu_{had} = F(s_1, s_2, s) V^\mu_1 + F(s_2, s_1, s) V^\mu_2 \tag{2.60}$$

with

$$V^\mu_j = p^\mu_{\pi_j} - p^\mu_{\pi_3} - p^\mu_{3\pi} \frac{p^\nu_{3\pi}(p^{\pi_j}_\nu - p^{\pi_3}_\nu)}{s} \qquad \text{with} \qquad j = 1, 2\,. \tag{2.61}$$

---

[4] Sum over repeated indices following the *Einstein* summation convention.

s is the square of the c.m. energy of the 3π system ($s = (p^{\mu}_{\pi 1} + p^{\mu}_{\pi 2} + p^{\mu}_{\pi 3})^2$) and s_1, s_2 denote the squares of the c.m. energies of the two neutral $\pi^-\pi^+$ combinations

$$s_j = (p^{\mu}_{\pi_j} + p^{\mu}_{\pi_3})^2 = m^2_{\pi_j \pi_3} \, . \tag{2.62}$$

The form factors $F(s_1, s_2, s)$ and $F(s_2, s_1, s)$ are now assumed to be entirely described by the $a_1 \to \rho^0\pi$ s-wave as well as a small admixture of $a_1 \to \rho'\pi$ s-wave

$$J^{\mu}_{had} \propto BW_{a_1}(s)(B_{\rho^0\rho'}(s_2)V^{\mu}_1 + B_{\rho^0\rho'}(s_1)V^{\mu}_2) \, , \tag{2.63}$$

where $B_{\rho^0\rho'}$ is given by a combination of two relativistic *Breit-Wigner* (BW) functions

$$B_{\rho^0\rho'}(s) = \frac{1}{1+\beta_{\rho'}} \left[BW_{\rho^0}(s) + \beta_{\rho'} BW_{\rho'}(s) \right] \tag{2.64}$$

with the parameter $\beta_{\rho'}$ related to the ρ' contribution. The *Breit-Wigner* functions describing the ρ^0 and ρ' resonances depend on the invariant mass squared s_j of the corresponding $(\pi^-\pi^+)_j$ combination

$$BW_{\rho}(s_j) = \frac{m^2_{\rho}}{m^2_{\rho} - s_j - i\sqrt{s_j}\,\Gamma_{\rho}(s_j)} \quad \text{with} \quad \Gamma_{\rho}(s_j) = \Gamma_{\rho}\frac{m^2_{\rho}}{s_j}\left[\frac{p(s_j)}{p(m^2_{\rho})}\right]^3 \tag{2.65}$$

and the model parameters $m_{\rho^0}, m_{\rho'}$ and $\Gamma_{\rho}, \Gamma_{\rho'}$ as well as

$$2p(s_j) = \sqrt{s_j - 4m^2_{\pi}} \, , \qquad 2p(m^2_{\rho}) = \sqrt{m^2_{\rho} - 4m^2_{\pi}} \, , \tag{2.66}$$

where $p(s_j)$ and $p(m^2_{\rho})$ can be derived from the assumption of a constant $\rho\pi\pi$ coupling corresponding to a relativistic p-wave phase space for the $\rho \to \pi\pi$ decay.

The last missing part to calculate J^{μ}_{had} (Eq. 2.63) is the *Breit-Wigner* function describing the a_1 resonance $BW_{a_1}(s)$. It is given as

$$BW_{a_1}(s) = \frac{m^2_{a_1}}{m^2_{a_1} - s - i\sqrt{s}\,\Gamma_{a_1}(s)} \quad \text{with} \quad \sqrt{s}\,\Gamma_{a_1}(s) = m_{a_1}\Gamma_{a_1}\frac{g(s)}{g(m^2_{a_1})} \, . \tag{2.67}$$

Here, m_{a_1} and Γ_{a_1} are parameters of the KS model. The energy dependence of the imaginary part of the a_1 propagator is fixed through the three-pion phase space factor $g(s)$. A convenient parametrization of this g function is given in Ref. [58] as

$$g(x) = \begin{cases} 4.1(x - 9m^2_{\pi})^3 \left[1 - 3.3(x - 9m^2_{\pi}) + 5.8(x - 9m^2_{\pi})^2\right] & x < (m_{\rho^0} + m_{\pi})^2 \\ x(1.623 + 10.38/x - 9.32/x^2 + 0.65/x^3) & x \geq (m_{\rho^0} + m_{\pi})^2 \end{cases} \tag{2.68}$$

with a $\pi^{\pm}$ mass of $m_{\pi} = 0.140\,\text{GeV}/c^2$. This parametrization is only valid for the given model including the ρ^0 and ρ' resonances in the region up to $s = 3\,\text{GeV}^2/c^4$.

The KS parametrization of the $\tau^- \to \pi^-\pi^+\pi^-\nu_{\tau}$ decay relies on a set of seven parameters; the mass and the total width of the a_1 resonance m_{a_1} and Γ_{a_1}, the mass and the total width of the ρ^0 resonance m_{ρ^0} and Γ_{ρ^0}; the mass and the total width of the ρ' resonance $m_{\rho'}$ and $\Gamma_{\rho'}$, and the previously mentioned $\beta_{\rho'}$. The formalism of

Parameter	m_{a_1}	Γ_{a_1}	$m_{\rho'}$	$\Gamma_{\rho'}$	m_{ρ^0}	Γ_{ρ^0}	$\beta_{\rho'}$
Value	1.251	0.599	1.370	0.510	0.773	0.145	−0.145

Table 2.7: Parameters of KS model used in `TAUOLA`: Masses and widths are given in GeV/c^2.

the KS model, as it has been briefly sketched, is part of the MC simulation package `TAUOLA` [62], which has been written to simulate $e^+e^- \rightarrow \tau^+\tau^-$ events including the subsequent τ decays. The parameter values used in the `TAUOLA` package are summarized in Tab. 2.7. However, besides this model there are other descriptions of the decay $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$. One of them is discussed in the next section.

2.5.2.2 The IMR Model

Contrary to the KS model, ***Isgur***, ***Morningstar***, and ***Reader*** **(IMR)** [59] tried to perform a description of the given decay within a time-ordered perturbation theory. The authors show that the decay chain $\tau^- \rightarrow a_1^-\nu_\tau \rightarrow \rho^0\pi^-\nu_\tau$ can be described by two time-ordered graphs as illustrated in Fig. 2.7 taken from Ref. [59]. The contribution of the first graph (Fig. 2.7a) is calculable since only a single quark-antiquark pair is generated. The second one (Fig. 2.7b) involves the creation of three $q\overline{q}$ pairs from the vacuum, which complicates the calculation. On the other hand, it is known from direct measurements that the second graph is strongly suppressed and the effects induced by the this part together with uncertainties arising from non-pointlike hadrons are assumed to be describable by a non-resonant smooth background [59].

Within the IMR model the differential decay width is proportional to an axial-vector spectral function

$$\frac{d\Gamma(s)}{ds_1 ds_2 ds} \propto f_{a_1}^2 P_{a_1}(s) F_{a_1}(s_1, s_2, s)\,. \tag{2.69}$$

Here, f_{a_1} describes the weak decay $W^- \rightarrow a_1^-\nu_\tau$ and is assumed to be energy-independent ($f_{a_1}(s)$ = const). The a_1 propagator $P_{a_1}(s)$ introducing the *Breit-Wigner* function includes the s dependence of m_{a_1} and Γ_{a_1} as well as the access of the $K^*\overline{K} + \overline{K}^*K$ channel. $F_{a_1}(s_1, s_2, s)$ describes the entire decay chain $a_1^- \rightarrow \rho^0\pi^- \rightarrow \pi^-\pi^+\pi^-$ and therefore includes the *Dalitz*-plot density distribution in the $s_1 - s_2$ space. In addition, this model explicitly allows orbital angular momenta of zero (S-wave) and two (D-wave) between ρ and π. The IMR parametrization allows determination of the ratio of the orbital angular amplitudes D/S. However, KS and IMR model need to be judged by their quantitative description of the $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ kinematics.

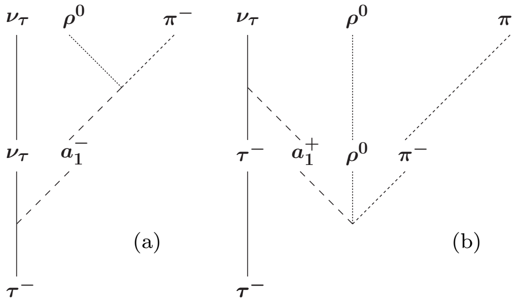


Figure 2.7: Two time-ordered graphs, which would combine to form a covariant *Feynman* graph for pointlike particles: (a) strong decay following the weak creation of a_1^- and (b) an $a_1\rho\pi$ vacuum fluctuation followed by annihilation of a_1^+.

2.5.3 a_1 Lineshape and $\rho - \rho$ Interference

Nowadays, the hadronic structure of $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ has been investigated by many different experiments [55, 56, 63–67] and the results have been compared to model predictions, e.g. *Dalitz*-plot projections and 3π mass spectrum.

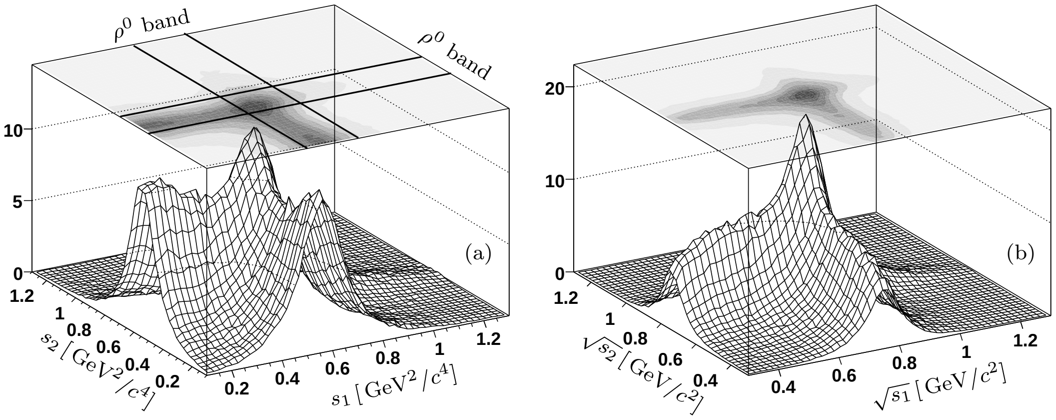


Figure 2.8: *Dalitz*-plot of $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$: (a) shows s_1 vs. s_2 from MC events generated with `TAUOLA`, while (b) illustrates the corresponding $\sqrt{s_1}$ vs. $\sqrt{s_2}$ distribution. The vertical axes have arbitrary units. The lines on top of (a) illustrate the ρ^0 bands as used by OPAL (Figs. 2.10b-d).

Since *Dalitz*-plots contain information about intermediate resonances participating on hadronic decays as well as non-resonant contributions, they are very useful to judge the validity of the aforementioned models. Fig. 2.8 illustrates the *Dalitz*-plot s_1 vs. s_2 (Fig. 2.8a) and the corresponding $\sqrt{s_1}$ vs. $\sqrt{s_2}$ distribution (Fig. 2.8b) as they are generated by `TAUOLA`. Both figures illustrate the known behaviour that both neutral $\pi\pi$ combinations tend to cluster in the intersection region of the ρ^0 bands. This qualitative criterion is provided by KS and IMR model. However, besides this qualitative aspect both models need to be checked for their ability to reproduce the kinematical properties quantitatively.

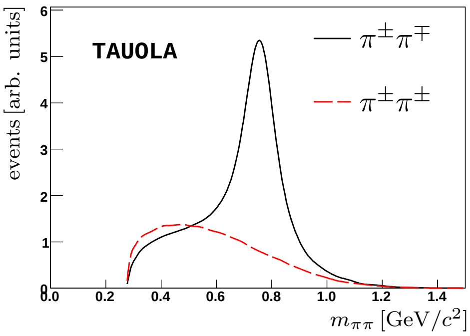


Figure 2.9: Illustration of $\rho^0 - \rho^0$ interference: Prediction for the invariant mass distributions of like-sign (dashed) and opposite-sign (solid) $\pi\pi$ combinations from the KS model. Every event enters twice the opposite-sign ($\sqrt{s_1}$ and $\sqrt{s_2}$), but only once the like-sign distribution.

A characteristic feature of the KS model is an interference between the two ρ^0 combinations reflected in the invariant mass distributions of like-sign and opposite-sign $\pi\pi$ combinations (Fig. 2.9). In the case of $\rho^0 - \rho^0$ interference the KS model predicts the like-sign mass spectrum to cross-over the opposite-sign spectrum in the low-mass region. This effect is related to the aforementioned behaviour that both neutral combinations are anxious to form a ρ^0 feeding up the ρ^0 peak from the lower mass region. This feature is absent in the case of incoherent superposition resulting in identical like-sign and opposite-sign spectra of $\pi\pi$ combinations not produced in ρ^0 decays. The existence of such an interference effect has been claimed by the MAC collaboration [65] and has been confirmed by the ARGUS collaboration [66].

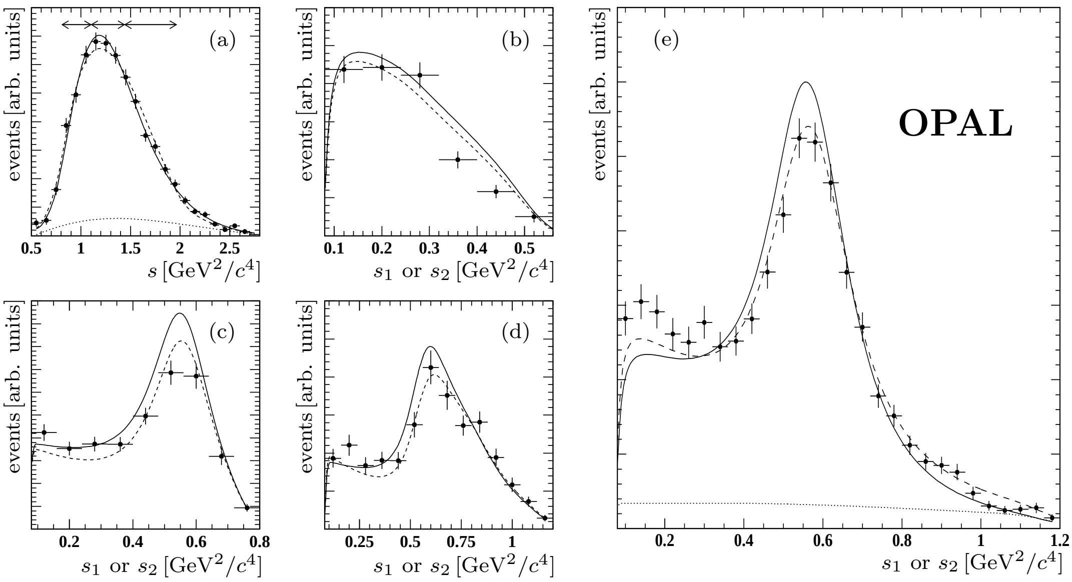


Figure 2.10: 3π mass squared spectrum and *Dalitz*-plot projections as measured by OPAL [56]: The arrows in the plot showing the 3π mass squared spectrum ((a)) indicate the s intervals used to generate the *Dalitz*-plot projections (b), (c), and (d). The projection plots illustrate the invariant mass squared of neutral $\pi\pi$ combinations s_j, where the second neutral combination has been restricted to the ρ^0 band $0.5 < s_k < 0.7\,\mathrm{GeV}^2/c^4$ with $j, k = 1, 2$ and $j \neq k$ (Fig. 2.8a). (e) illustrates the *Dalitz*-plot projection for the full s range. The spectra were fitted to the KS (solid line) and IMR (dashed line) model predictions, respectively. The dotted line indicates a polynomial background contribution of the IMR model taking into account the influence of the second time-ordered graph and non-pointlike hadrons (Sec. 2.5.2.2).

More reliable quantitative checks are represented by global fits to the 3π lineshape and *Dalitz*-plot densities as has been done by e.g. the OPAL collaboration. Fig. 2.10 briefly summarizes the results of the OPAL analysis [56]. A simultaneous fit has been performed to all four distributions displayed in Figs. 2.10a-d and both models have been found to be able to describe the 3π lineshape reasonably well on the given precision. On the other hand, Figs. 2.10b-d show deficits of both models in the simultaneous description of the *Dalitz*-plot projections. Both models are of compareable and rather poor agreement to the data as indicated by the values of χ^2/n_{dof} in Tab. 2.8, where n_{dof} denotes the number of degrees of freedom of the χ^2 fit. These disagreements are also visible in the *Dalitz*-plot projection for the full s

range (Fig. 2.10e). The rather high contribution of the polynomial background in the IMR model fit possibly indicates additional contributions to $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ besides the a_1 intermediate resonance. The possibility of such contributions has been investigated by e.g. ARGUS [67], DELPHI [64], and CLEO [63], which claim contributions from isoscalars or higher a_1 resonances.

Quantity	**KS Model**	**IMR Model**
$m_{a_1}/\,\mathrm{GeV}/c^2$	$1.262 \pm 0.009 \pm 0.007$	$1.219 \pm 0.007 \pm 0.002$
$\Gamma_{a_1}/\,\mathrm{GeV}/c^2$	$0.621 \pm 0.032 \pm 0.058$	$0.497 \pm 0.015 \pm 0.017$
D/S amplitude ratio	–	$-0.10 \pm 0.02 \pm 0.02$
χ^2/n_dof (global)	111/49	91.9/45
Polynomial background fraction	–	$(13.8 \pm 2.4)\,\%$

Table 2.8: Results of model fits to OPAL data [56]: Γ_{a_1} and the D/S amplitude ratio for the IMR model are calculated quantities, not parameters of the fit.

The discussion given should illustrate that to date there is no completely satisfactory description of the nature of the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. Hence, measurements relying on model descriptions of the decay dynamics of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ need to be done with caution. In this thesis this decay is used to search for the rare decay $B^- \to \tau^-\overline{\nu}_\tau$. The $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ kinematics has been simulated using the KS model, as it has been sketched in Sec. 2.5.2.1, with the `TAUOLA` parameter values listed in Tab. 2.7. However, at this point it should only be mentioned that the aforementioned imperfections of this model and the resulting model dependence have been taken into account in the analysis (Sec. 4.4).

Chapter 3

The *BaBar* Experiment

The data used to perform the analysis presented in this thesis have been collected with the *BaBar* detector located at the ***Positron-Electron-Project-II*** **(PEP-II)** storage rings at the ***Stanford Linear Accelerator Center*** **(SLAC)** on the west coast of the United States. This project has been planned and built as a high luminosity B meson factory and therefore not only precision tests can be accomplished, but also searches for rare decays on the B sector can be performed. The *BaBar* project consists of approximately 550 collaborators from 80 institutions in 10 countries and in the past few years it contributed many important results to our understanding of particle physics. This chapter gives a brief overview of the concept and the performance of the accelerator and the detector. A more detailed description can be found in Refs. [68, 69].

3.1 The PEP-II B Factory

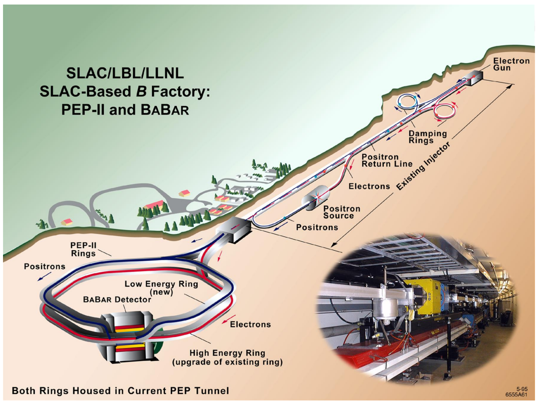


Figure 3.1: LINAC and PEP-II

In December 1993 the US congress decided to fund a B meson factory within the USA. After a short period of discussion SLAC with its already existing PEP tunnel has been chosen as location for this project. In order to satisfy the requirements on the new accelerator, PEP has been rebuilt by July 1998 to provide a peak luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Fig. 3.1 shows a schematic view of the machine. Electrons (e^-) and positrons (e^+) are accelerated in the ***Linear Accelerator*** **(LINAC)** before they are injected into the PEP-II storage rings. When they collide within the *BaBar* detector the e^- and e^+ beams have asymmetric energies of 9 GeV and 3.1 GeV, respectively, resulting in a relative boost of $\beta\gamma \approx 0.558$ between the laboratory and c.m. frame. The c.m. energy is approximately $\sqrt{s} \approx 10.58\,\text{GeV}$ and nearly corresponds to the mass of the $\Upsilon(4S)$ resonance. Since the mass of this bottomonium state, i.e. bound $b\bar{b}$ state, is about 20 MeV above the production threshold of B meson pairs, the $\Upsilon(4S)$ predominantly decays into two B mesons with a branching fraction of nearly 100 %.

The abovementioned asymmetry in the beam energies is the key to investigate the nature of $\not{CP}$ in the B system. The study of time-dependent CP asymmetries is the primary goal of the *BABAR* project. Secondary goals are measurements of bottom and charm meson decays with high precision and searches for rare decays. All of these scientific goals require a large amount of data.

Fig. 3.2 illustrates the time evolution of the integrated luminosity $\mathcal{L}$ since the project started in 1999. In almost eight years of data taking PEP-II delivered about $406\,\text{fb}^{-1}$. *BABAR* recorded nearly $391\,\text{fb}^{-1}$, where $37\,\text{fb}^{-1}$ have been taken about 40 MeV below the $\Upsilon(4S)$ resonance. This so-called Off Peak data sample is used by many analyses to study backgrounds from continuum events, i.e. $e^+e^- \to q\bar{q}$ processes, and $e^+e^- \to \tau^+\tau^-$. The run periods are labeled with Run1 to Run 5. The analysis presented in this thesis uses data accumulated within the first four run periods. These data sets with the integrated luminosities are summarized in Tab. 3.1.

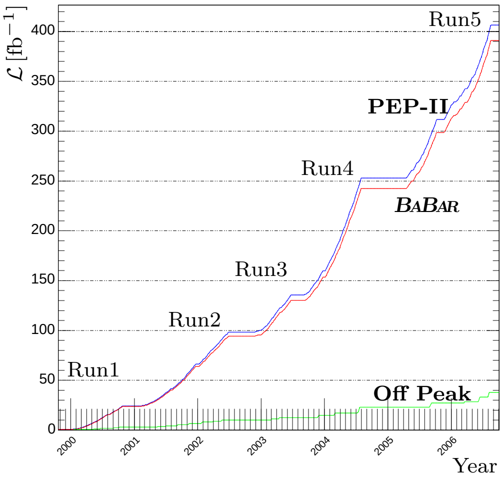


Figure 3.2: The upper curve shows the luminosity delivered by PEP-II. The middle curve illustrates the luminosity recorded by *BABAR* (On plus Off Peak). The Off Peak sample evolution is given by the lower curve.

In Tab. 3.2 the effective production cross sections of different e^+e^- reactions at the corresponding c.m. energy are given. With a cross section for $b\bar{b}$ production of

Run Period	$\mathcal{L}/\,\text{fb}^{-1}$	$\mathcal{L}_{\text{off}}/\,\text{fb}^{-1}$	$\mathcal{L}_{\text{on}}/\,\text{fb}^{-1}$	$N_{B\bar{B}}$
Run1	21.8	2.3	19.5	21181864
Run2	67.2	6.9	60.3	66441247
Run3	33.5	2.4	31.1	34076579
Run4	109.7	9.9	99.8	110107681
Run12	89.1	9.2	79.8	87623110
Run13	122.5	11.6	110.9	121699689
Run14	232.1	21.5	210.6	231807369

Table 3.1: *BABAR* data taking periods: The subscripts "on" and "off" denote the data samples taken on the $\Upsilon(4S)$ resonance (On Peak) and 40 MeV below (Off Peak), respectively. $N_{B\bar{B}}$ has been determined using the B counting procedure. The data sets denoted as Run1X are equal to the sum of Run1 to RunX (X=2,3,4).

$e^+e^- \rightarrow$	$b\bar{b}$	$c\bar{c}$	$s\bar{s}$	$u\bar{u}$	$d\bar{d}$	$\tau^+\tau^-$	$\mu^+\mu^-$	e^+e^-
σ/nb	1.10	1.30	0.35	1.39	0.35	0.94	1.16	≈ 40

Table 3.2: Cross sections of e^+e^- reactions at a c.m. energy of $\sqrt{s} = 10.58\,\text{GeV}$ [69].

$\sigma_{b\bar{b}} = 1.10\,\text{nb}$ and the On Peak luminosity listed in Tab. 3.1 the number of $B\overline{B}$ pairs included in the Run14 data set arises to $N_{B\overline{B}} \approx 230 \times 10^6$. Since many analyses, e.g. branching fraction measurements, rely on a well-known normalization to the total number of $\Upsilon(4S)$ events in a given data sample, the B *counting* technique has been developed [70]. The B counting uses the measured numbers of multihadron and dimuon events ($e^+e^- \rightarrow \mu^+\mu^-$) in the On and Off Peak samples to extract the number of $\Upsilon(4S) \rightarrow B\overline{B}$ events contained in the On Peak sample. Tab. 3.1 includes the number of $B\overline{B}$ pairs determined by the given procedure, where it has been assumed that the $\Upsilon(4S)$ exclusively decays into pairs of B mesons.[1] The high number of $B\overline{B}$ events allows precision measurements as well as searches for rare B decays.

3.2 The *BABAR* Detector

In order to fulfil the requirements, the *BABAR* detector (Fig. 3.3) consists of five main subsystems. From the beam pipe to outside of the detector the components are the ***Silicon Vertex Tracker*** **(SVT)**, the ***Drift Chamber*** **(DCH)**, the ***Detector of Internally Reflected Cherenkov light*** **(DIRC)**, an ***Electromagnetic Calorimeter*** **(EMC)**, and the ***Instrumented Flux Return*** **(IFR)**. The four innermost components are surrounded by a 1.5 T solenoid field. Since all components are of substantial importance for the analysis, they are described in the following in more detail.

3.2.1 Silicon Vertex Tracker

The SVT is the first part of the two-stage tracking system and provides precise reconstruction of charged particle trajectories. Furthermore, it has been explicitly designed to determine decay vertices near the interaction point with high resolution. For particles with transverse momenta less than $0.1\,\text{GeV}/c$ it provides the only momentum measurement since such particles cannot reach the drift chamber volume.

As visualized in Fig. 3.4, the SVT consists of five double-sided layers of silicon microstrip detectors. The inner three layers are mainly responsible for track and vertex reconstruction, while the outer ones provide information for matching SVT tracks with DCH tracks, if a charged particle reached the DCH. In order to measure time-dependent CP asymmetries mentioned to be the primary goal of the *BABAR* project, the mean vertex resolution along the z axis is required to be better than $80\,\mu\text{m}$. The design of the SVT fulfils this requirement and further ensures spatial

---

[1] In fact, this is an approximation, since non-$B\overline{B}$ $\Upsilon(4S)$ decays have not been completely excluded by measurements so far [13].

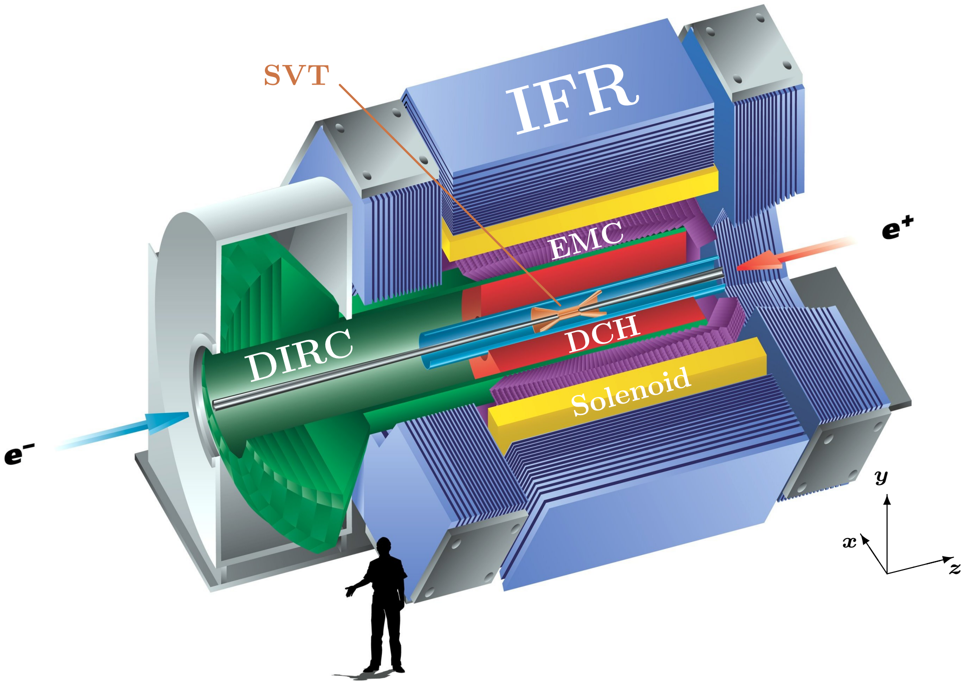


Figure 3.3: The *BABAR* Detector: The beam axis nearly corresponds to the z axis of the *BABAR* coordinate system.

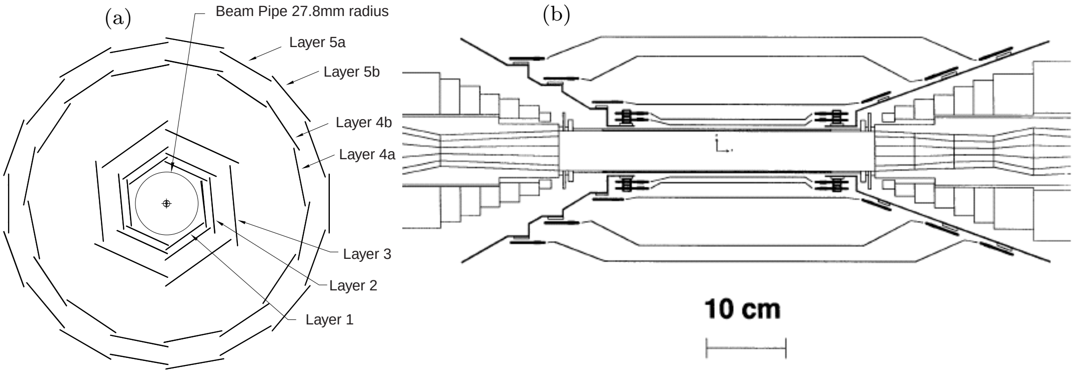


Figure 3.4: Schematic (a) front and (b) side view of the Silicon Vertex Tracker [68].

vertex resolutions of $10 - 15\,\mu$m in the three inner and $30 - 40\,\mu$m in the two outer layers, respectively.

The mean flight length of τ's produced in $B^- \to \tau^-\overline{\nu}_\tau$ decays has been calculated in the B rest frame to be about $115\,\mu$m (Sec. 2.5.1). Although one has to take into account the relative boost between the B rest frame and the laboratory frame, the secondary vertex of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay should be reconstructable and distinguishable from the decay vertex of the companion B meson produced in the $\Upsilon(4S)$ decay. It is expected that a possible vertex separation depends on the τ flight direction relative to the decay vertex of the second B meson (Sec. 4.5.3.6).

3.2.2 Drift Chamber

The DCH completes the two-stage *BABAR* tracking system and is therefore responsible for the precise momentum measurement of charged particles passing the SVT, i.e. particles with a transverse momentum above 0.1 GeV/c. It further provides ***Particle Identification*** **(PID)** by measuring the ionization loss dE/dx of low momentum charged particles passing the DCH volume.

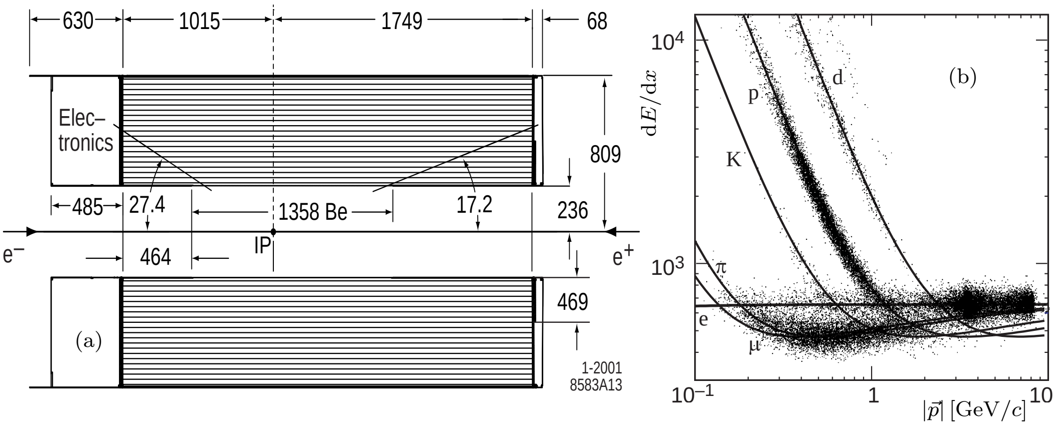


Figure 3.5: Drift Chamber and dE/dx: (a) shows a side view of the drift chamber with its dimensions and (b) illustrates the energy-loss per path for particles of different types traversing the DCH volume with three-momentum $\vec{p}$[68].

40 concentric layers of aluminium wires are spanned parallel to the beam axis (Fig. 3.5a) resulting in 7104 single hexagonal drift cells. The DCH is filled with a gas mixture of 80 % helium and 20 % isobutane, which has been optimized to achieve a good spatial and dE/dx resolution. A good measurement of the energy-loss per path length dE/dx is important for PID of charged particles with low momenta. Fig. 3.5b illustrates this quantity for different particle types as a function of momentum. A resolution of about 7 % allows $K^{\pm}/\pi^{\pm}$ separation up to 0.7 GeV/c. On the other hand, there is no PID potential using dE/dx above $\approx$ 0.7 GeV/c. But in order to reconstruct e.g. charm mesons decaying into kaons and pions, a good K/π separation is required over the entire momentum range. Therefore, an additional PID device has been designed and installed to provide K/π separation in the high momentum region; the detector of internally reflected *Cherenkov* light (DIRC).

3.2.3 *Cherenkov* Detector

The DIRC is a high-performance *Cherenkov* detector and provides PID for charged particles with momenta above 0.7 GeV/c. Fig. 3.6a illustrates the working principle. If a relativistic charged particle passes one of the bars of fused silica, *Cherenkov* light is emitted in a cone around the particle's three-momentum vector. The emission angle θ_C of the *Cherenkov* photons is related to the magnitude of the velocity v of the traversing charged particle and the refraction index n of the radiator medium.

The *Cherenkov* angle θ_C can be calculated as

$$\cos\theta_C = \frac{1}{\beta n} \quad \text{with} \quad \beta = \frac{v}{c} \tag{3.1}$$

and $n \approx 1.473$ for fused silica. Thus, a measurement of θ_C translates into the velocity v of the traversing particle. Using the momentum measurement provided by the tracking system the particle mass can be determined resulting in an identification of the particle type.

Apart from their functionality as radiators the bars serve as light pipes transporting the *Cherenkov* light via total internal reflection out of the inner detector volume into a water filled standoff box. The angle information is conserved during this process. The photons are then caught by an array of photomultipliers. The expected pattern on the surface of this array is essentially a conic section modified by the refraction at the exit of the *Cherenkov* light from the fused silica window (Fig. 3.6b). With this innovative design an excellent K/π separation over a wide momentum range from about 1 GeV up to 4 GeV has been achieved.

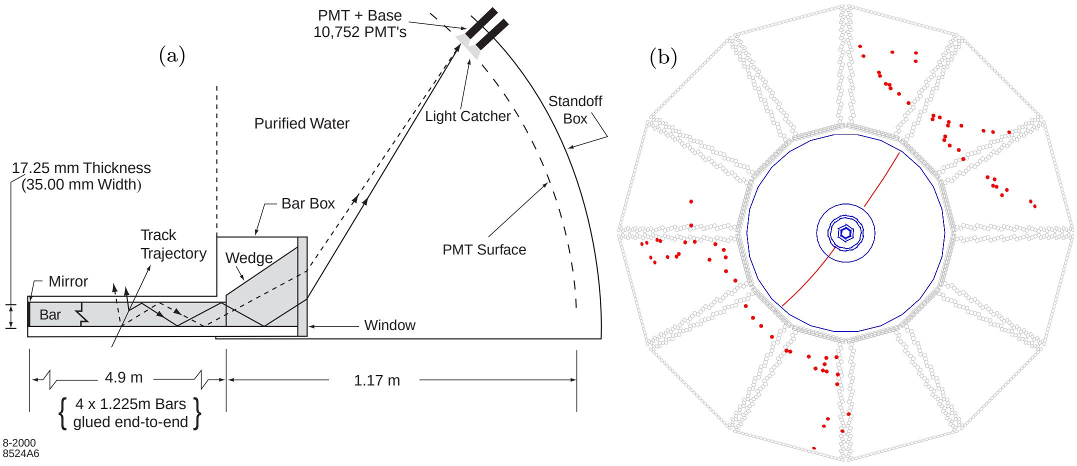


Figure 3.6: (a) principle functionality of the DIRC, (b) typical signal distribution in the standoff box [68]

3.2.4 Electromagnetic Calorimeter

The three inner components exclusively deal with charged particles, but for many important reconstruction channels, e.g. $B^0 \to \pi^0\pi^0$, it is indispensable to detect neutral particles and to measure their energies. Therefore, the *BABAR* detector contains an electromagnetic calorimeter consisting of 6580 thallium-doped CsI crystals, which are arranged in 56 rings in the azimuthal angle. Fig. 3.7 gives a longitudinal cross secion of the EMC.

Thallium-doped CsI has been chosen due to excellent angular and energy resolutions compared to other scintillator materials. With crystal lengths between 29.6 cm and 32.4 cm they cover about $16 - 17.5$ radiation lengths resulting in high light yield

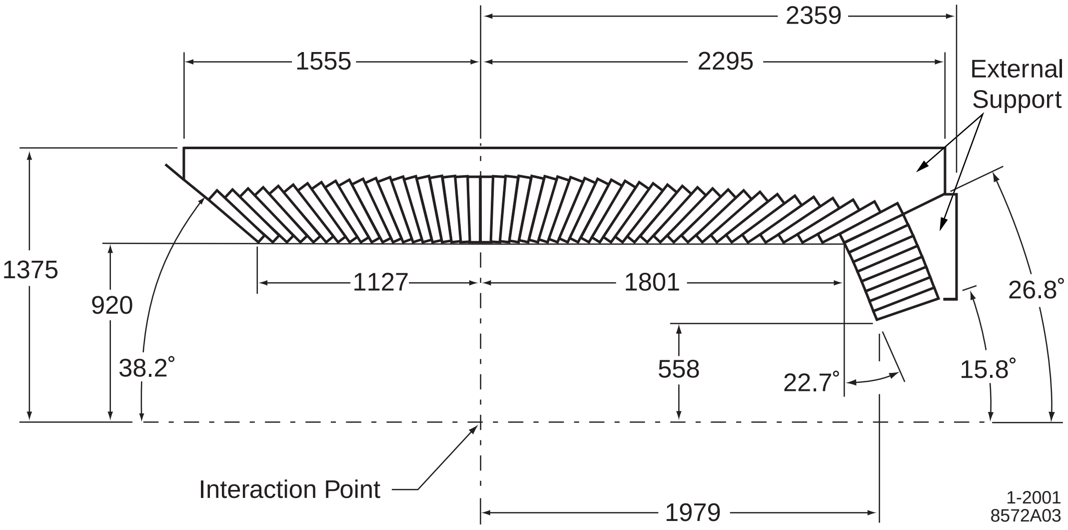


Figure 3.7: Longitudinal cross section of the upper half of the EMC: This figure shows the 48 barrel rings and 8 rings in the forward endcap section [68].

paired with a compact design. Initial photons or electrons going through the scintillator medium induce electromagnetic showers of secondary photons and electrons produced in alternating external bremsstrahlung and photon conversions. The energy deposition leads to an excitation of the medium. This excitation results in emission of photons in a wavelength range of visible light at the thallium atoms, which is then collected and read out by a front-end electronics placed on the rear face of the crystals.

Electrons can be identified by their energy deposition in the EMC E_{EMC} and the magnitude of the three-momentum $\vec{p}_{\mathrm{DCH}}$ as it has been reconstructed from the DCH hits. $e^{\pm}$ can be separated from π or other heavier hadrons, which are also able to induce showers in the EMC, by evaluating $E_{\mathrm{EMC}}/|\vec{p}_{\mathrm{DCH}}|$. Due to the negligible electron mass of $m_{e^{\pm}} = 0.511\,\mathrm{MeV}/c^2$ this ratio is expected to be $E_{\mathrm{EMC}}/|\vec{p}_{\mathrm{DCH}}| \approx 1$. apart from detector resolution effects for true $e^{\pm}$.

3.2.5 Instrumented Flux Return

Muons deposit a negligible fraction of their energy in the EMC since the energy emitted via bremsstrahlung by an accelerated particle with a rest mass m_0 is proportional to $1/m_0^2$. Hence, in contrast to electrons $\mu^{\pm}$ are not seen in the EMC. On the other hand, the detection and identification of $\mu^{\pm}$ is of similar importance, e.g. for the reconstruction of semileptonic $B \to X\mu\nu$ decays.

Therefore, the steel flux return of the magnet is segmented (Fig. 3.8). It is used as a $\mu^{\pm}$ filter and hadron absorber. By design the gaps between the 19 layers in the barrel region and the 18 layers of the end doors have been instrumented with ***Resistive Plate Chambers*** **(RPC's)**. Ionizing particles passing these RPC's induce streamers, which can then be detected. The instrumentation not only allows

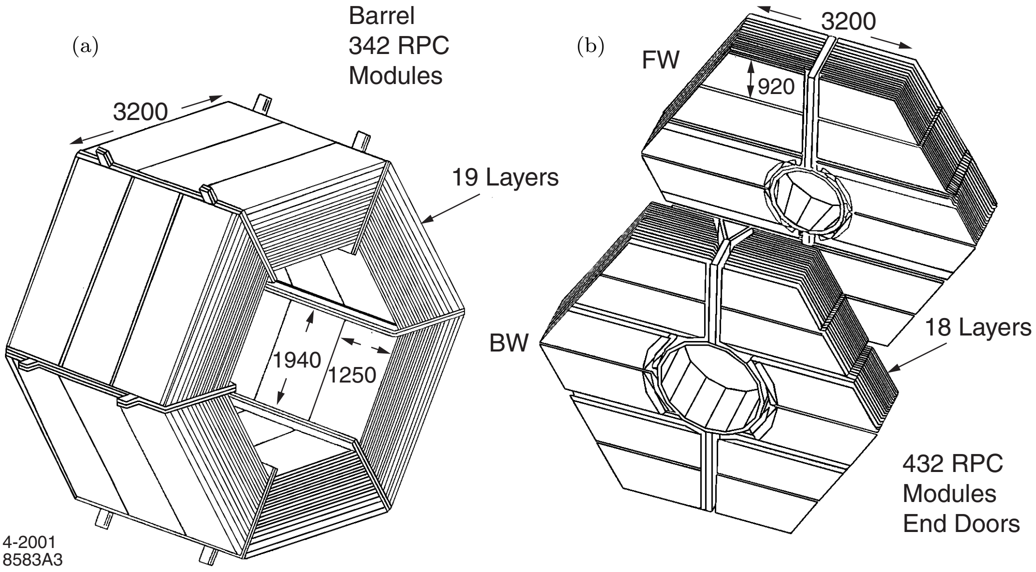


Figure 3.8: Overview of the Instrumented Flux Return: (a) barrel section and (b) the two end doors [68]

$\mu^\pm$ identification but also detection of the particles produced in hadron decays inside the steel plates, which provides detection of long-living neutral hadrons, primarily K_L^0 and neutrons. Due to performance problems of the RPC's, namely a significant decrease of detection efficiencies with time, it has been decided to replace them by ***Limited Streamer Tubes* (LST's)**. The replacement started in July 2004 and all sextants of the IFR barrel region have already been reinstrumented with LST's. Due to the reinstrumentation the IFR now provides an improved $\mu^\pm$ identification.

3.2.6 Trigger System and Data Acquisition

The two-level *BABAR* trigger system is responsible for a fast preselection of events of interest and consists of the ***Level 1* (L1)** hardware trigger followed by the ***Level 3* (L3)** trigger implemented in software. Thereby, it is necessary to guarantee stable and well-understood trigger efficiencies, while beam-induced background needs to be rejected sufficiently.

The L1 trigger decision is based on tracks in the DCH, EMC showers, and tracks detected in the IFR. The data are processed by specialized hardware processors, such as the DCH trigger, the EMC trigger, and the IFR trigger, which send the information to the ***Global L1 trigger* (GLT)**. The GLT combines the information, forms specific triggers, and delivers them to the fast control, which decides, if the event will be accepted or not. All these steps are accomplished within $12.8\,\mu$s after the e^+e^- collision. Typically, the L1 rates are about 2.5 kHz at a luminosity of $\mathcal{L} = 8 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Although the overall efficiency of $B\overline{B}$ events is $> 99.9\,\%$ the L1 trigger efficiency of $B^- \to \tau^-\overline{\nu}_\tau$ is about $99.7\,\%$ [68].

The L3 filter operates by refining the selection methods used in L1 and typically reduces the event rate by a factor of 10.

The trigger requirements have been chosen to reject background events, such as beam-wall or beam-gas interactions, but allow interesting physics processes to pass. For the purposes of calibration, efficiency studies, and B counting the selected data still contain a large amount of *Bhabha* ($e^+e^- \rightarrow e^+e^-$) and dimuon ($e^+e^- \rightarrow \mu^+\mu^-$) events.

Fig. 3.9 shows schematically the data flow from the detector to the mass storage. If the L1 trigger accepts an event and it passes the L3 filter, the complete information from the subsystems is dumped into the intermediate event store. The data are then converted into a `ROOT` based format [71] and are in principle ready for analysis. In order to accelerate analyses, the enormous data set is organized in so-called *Skims*, where every skim is a subset of the entire data sample after application of a predefined selection. Thus, analysts become able to run on preselected data sets of interest resulting in more efficient and less time-consuming analyses [72].

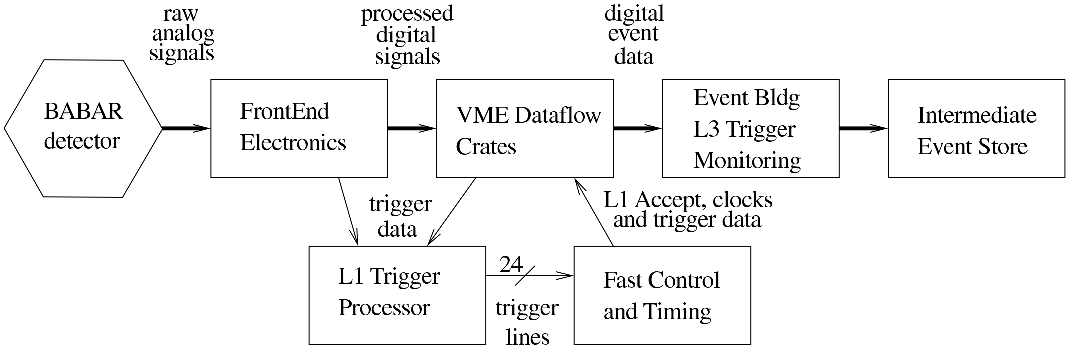


Figure 3.9: Schematic diagram of the data acquisition

Chapter 4

Description of the Analysis

Given the numbers of $B\overline{B}$ pairs listed in Tab. 3.1 and a presumed branching fraction of $\mathcal{B}_{\mathrm{exp}}(B^- \rightarrow \tau^-\overline{\nu}_\tau) = 10^{-4}$, one expects about 23000 $B^- \rightarrow \tau^-\overline{\nu}_\tau$ decays in the Run14 data set.[1] This number is remarkably high and indeed the B factories already measured branching fractions of exclusive B decay channels of the order 10^{-6} with high significance, e.g. $B^+ \rightarrow \eta K^+$ [73, 74]. However, the $B^- \rightarrow \tau^-\overline{\nu}_\tau$ channel suffers from the fact that due to the presence of at least two neutrinos in the final state the decay chain is not fully reconstructable. This makes the search for this purely leptonic decay challenging since high background from various sources needs to be suppressed.

Sec. 2.4.1 summarized the current status of searches for $B^- \rightarrow \tau^-\overline{\nu}_\tau$. As a consequence of the huge background arising from both other B decays and underlying continuum events ($e^+e^- \rightarrow q\overline{q}$, $q = u, d, s, c$) all listed measurements on the $\Upsilon(4S)$ resonance have been performed as ***Recoil analyses***. Such analyses are based on the reconstruction of one of the B mesons produced in the $\Upsilon(4S) \rightarrow B\overline{B}$ decay. All particles used to reconstruct this B_{tag}[2] are then removed from the list of particles detected in the entire event. After erasing B_{tag} the signal decay can be searched for in its so-called *Recoil* representing the remaining particles detected. The advantage of this recoil technique lies in a clean environment since almost all detected particles produced in the e^+e^- collision are assigned to the decay products. On the other hand, this method is restricted to exclusive B decay channels used for the B_{tag} reconstruction resulting in small overall reconstruction efficiencies and therefore in small significances.

The analysis described in this thesis aims to explore the potential of the 3-prong τ decay $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ for the search for $B^- \rightarrow \tau^-\overline{\nu}_\tau$ without an a priori B_{tag} reconstruction. Therefore, three different reconstruction techniques have been used to perform this search and have then been compared with respect to the expected significances and upper limits for $B^- \rightarrow \tau^-\overline{\nu}_\tau$:

---

[1] Given the assumption that the $\Upsilon(4S)$ always decays into $B\overline{B}$ pairs and $\mathcal{B}(\Upsilon(4S) \rightarrow B^0\overline{B}^0) = \mathcal{B}(\Upsilon(4S) \rightarrow B^+B^-) = 50\,\%$.

[2] In the following the B mesons of a $\Upsilon(4S)$ event are denoted as B_{sig} and B_{tag} for the B reconstructed into $\tau\nu_\tau$ and the companion B, respectively.

1. Inclusive Reconstruction (Sec. 4.5)
2. Semileptonic Reconstruction (Sec. 4.6)
3. Recoil of $B \to D^{(*)}\ell\nu_\ell$ (Sec. 4.7)

The differences of the procedures manifest in the treatment of the B_{tag}. But before the methods are described in detail, an overview of the ***Monte Carlo*** **(MC)** samples used in the analysis is given (Sec. 4.1).

4.1 Detector Simulation and Monte Carlo Samples

Since at the given c.m. energy different processes contribute to the total cross section (Tab. 3.2), various MC generators have been used to provide proper descriptions of the physics at $\sqrt{s} = 10.58\,\text{GeV}$.[3] The MC species, which are important for the analysis in terms of efficiency and background studies, are discussed in the following:

1. Continuum events are generated by `Jetset7.4` [75]. The fragmentation process, i.e. the breaking of the colour field between two primary quarks and the formation of hadrons, is simulated using the *Lund*-String model [76]. The `Jetset7.4` simulation includes the subsequent decays of the hadrons produced during this hadronization process.

2. τ-pair events with higher order radiative corrections are generated using the `KK2f` MC generator with the τ decays simulated with `TAUOLA` [62] as mentioned in Sec. 2.5.2. Final state radiation is inserted by the `PHOTOS` package [77].

3. Within *BABAR*, generic B decay chains are mainly simulated by a channel based MC generator called `EvtGen` [78]. This package uses decay amplitudes instead of probabilities for the simulation of decays. The exclusive channels are described by dedicated decay models, which handle either decays to specific sets of spin states, e.g. scalar to lepton plus neutrino, or more inclusive models, e.g. *Dalitz* decay models. Approximately 60 % of exclusive B decay chains are generated by `EvtGen`, whereas the remainder is handled by an interface to `Jetset7.4`.

4. For the purpose of selection optimization and efficiency evaluation signal MC is needed. Its composition is illustrated in Fig. 4.1: The B^+ is free to decay generically into any possible exclusive final state X, while B^- is forced to decay into $\tau^-\overline{\nu}_\tau$ with a subsequent $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ transition. The τ decay is generated by `EvtGen`, where the KS model with the `TAUOLA` parameter values discussed in Sec. 2.5.2.1 is used. The charge-conjugated configuration is included.

5. For later studies an additional MC sample has been generated including one B decaying into the signal channel, while the second B decays semileptonically into a cocktail of $D^{*0}\ell\nu_\ell$ and $D^0\ell\nu_\ell$. A more detailed description of this "signal vs. cocktail" sample is given in Sec. 4.7.6.3.

---

[3] *Bhabha* and dimuon events have not been used for MC studies.

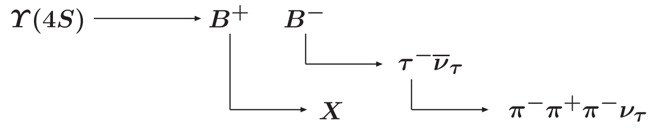


Figure 4.1: Decomposition of signal MC

The generation of physics processes is followed by a full simulation of the *BABAR* detector based on `GEANT4` [79] in order to simulate detection efficiencies and resolution effects. Final state photon radiation is realized by `PHOTOS`. Variations of the detector conditions with time have been taken into account and the signature of randomly triggered events has been mixed into the MC events to include beam backgrounds and to make the simulation as realistic as possible.

MC species	**Initial number of MC events $N_{\rm MC}/10^6$**						
	Run1	**Run2**	**Run3**	**Run4**	**Run12**	**Run13**	**Run14**
B^+B^- generic	24.25	119.36	67.97	322.71	143.60	211.58	534.28
$B^0\overline{B}^0$ generic	28.76	120.30	61.31	325.50	149.06	210.37	535.87
$e^+e^- \to c\bar{c}$	21.34	100.07	57.70	188.72	121.42	179.12	367.84
$e^+e^- \to u\bar{u}, d\bar{d}, s\bar{s}$	30.17	177.60	104.11	388.67	207.77	311.88	700.55
$e^+e^- \to \tau^+\tau^-$	22.10	99.79	69.85	211.53	121.89	191.74	403.27
signal	0.61	1.48	0.77	1.31	2.09	2.86	4.17
signal vs. cocktail	0.26	0.62	0.32	0.51	0.88	1.20	1.71

Table 4.1: Overview of the MC samples

Tab. 4.1 summarizes all relevant MC samples used in the analysis. It should be noted that the full generic $B\overline{B}$ MC sample (Run14) contains nearly 4.5 times the number of $B\overline{B}$ included in the corresponding On Peak data sample. From these numbers scaling factors have been calculated in order to scale the MC samples to the luminosities of the corresponding data sets. For the generic $B\overline{B}$ MC these factors are given by

$$f^{\rm sc}_{B\overline{B}} = \frac{N_{B\overline{B}}}{2 \cdot N_{\rm MC}} \tag{4.1}$$

with the number of $\Upsilon(4S) \to B\overline{B}$ events $N_{B\overline{B}}$ as derived from the B counting procedure (Tab. 3.1) under the assumption $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = \mathcal{B}(\Upsilon(4S) \to B^+B^-) = 50\,\%$. For continuum and $\tau^-\tau^+$ MC they are calculated from the luminosities and cross sections (Tab. 3.2)

$$f^{\rm sc}_{q\bar{q}} = \frac{\mathcal{L}_{\rm on} \cdot \sigma_{q\bar{q}}}{N_{\rm MC}}\,, \quad f^{\rm sc}_{\tau^+\tau^-} = \frac{\mathcal{L}_{\rm on} \cdot \sigma_{\tau^+\tau^-}}{N_{\rm MC}}\,. \tag{4.2}$$

For signal MC the scaling factors have been determined from

$$f^{\rm sc}_{\rm sig} = \frac{2 \cdot \frac{N_{B\overline{B}}}{2} \cdot \mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) \cdot \mathcal{B}_{\rm exp}(B^- \to \tau^-\overline{\nu}_\tau)}{N_{\rm MC}}\,. \tag{4.3}$$

MC species	**Scaling Factor f^{sc}**			
	Run1	**Run2**	**Run3**	**Run4**
B^+B^- generic	0.44	0.28	0.25	0.17
$B^0\overline{B}^0$ generic	0.37	0.28	0.28	0.17
$e^+e^- \to c\overline{c}$	1.19	0.78	0.70	0.69
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	1.35	0.71	0.62	0.54
$e^+e^- \to \tau^+\tau^-$	0.83	0.57	0.42	0.44
signal	3.2×10^{-4}	4.2×10^{-4}	4.1×10^{-4}	7.8×10^{-4}

Table 4.2: Overview of the MC scaling factors

Here, the term $N_{B\overline{B}}/2$ takes into account that only $\Upsilon(4S) \to B^+B^-$ events contain signal. The additional factor of two then accounts for the fact that both charged B mesons can decay via $B \to \tau\nu_\tau$ with a probability of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$. Naturally, the branching fractions need to be applied to calculate the number of signal events produced in a given $N_{B\overline{B}}$. For estimations from the signal MC an expected branching fraction of $\mathcal{B}_{\text{exp}}(B^- \to \tau^-\overline{\nu}_\tau) = 10^{-4}$ has been assumed. The scaling factors have been separately calculated for every run period, respectively (Tab. 4.2).

4.2 Charged Tracks and Neutral Clusters

The reconstruction of all relevant particles of the analysis is based on lists of ***Track*** and ***Cluster*** objects representing the event-by-event information provided by the BABAR detector.

The track object contains the information of a physical charged particle track fitted to the SVT and DCH hits. The track finding and fitting procedures uses the *Kalman* filter algorithm, where the detailed distribution of material and the full map of the magentic field is taken into account. The *Kalman* filter first fits a trajectory to the DCH hits. The resulting track is then extrapolated into the SVT and the SVT segments are added. If the segments match together, the *Kalman* fit is repeated using the full set of DCH and SVT hits [68].

Electromagnetic showers spread over many adjacent crystals and form a cluster of energy deposits. The neutral cluster object contains the full information of such a collection of energy deposits not associated to a charged track. Clusters are required to have at least one seed crystal with an energy deposit above 10 MeV. Energy deposits in the surrounding crystals of more than 1 MeV are considered as part of the cluster. Pattern recognition algorithms are used to find local energy maxima within the cluster. If more than one maximum has been found, the cluster is splitted into so-called ***Bumps***, where the cluster energy is portioned accordingly [68].

Charged tracks and neutral clusters are available as lists satisfying different quality criteria. In this analysis track objects from the ***Charged Tracks*** **(CT)**, ***GoodTracksVeryLoose*** **(GTVL)**, and ***GoodTracksLoose*** **(GTL)** lists defined in App. A.1 have been used. The definitions of neutral objects, such as ***CalorNeutral*** **(CN)** and ***GoodPhotonLoose*** **(GPL)**, are given in App. A.2.

4.3 Particle Identification

The next sections clarify that PID information is indispensable to perform a search for the decay $B^- \to \tau^-\overline{\nu}_\tau$. The PID information is available as lists containing reconstructed candidates, which pass the corresponding PID criteria. In this analysis lepton ($\ell^\pm = e^\pm, \mu^\pm$) as well as kaon ($K^\pm$) identification have been used. The $e^\pm$ and $K^\pm$ identification is performed by likelihood-based selectors, while the $\mu^\pm$-ID is based on a neural network. The PID selectors combine information provided by the detector subsystems. A detailed description of the PID criteria and performances can be found in App. A.3.

In general, CT candidates can fulfil the requirements of more than one PID selector. Hence, in this analysis the identification of reconstructed candidates is applied in the following order based on the PID efficiencies and mis-ID rates (App. A.3): If a CT candidate passes the $e^\pm$ selection, it is treated as $e^\pm$. If not, the candidate is checked to be included in the μ-ID list. If the candidate passed the muon selection, it is marked as $\mu^\pm$, else the K-ID selection is applied. The candidate is called a $K^\pm$ if it passes this selection. If the candidate is rejected by all three selectors, it is treated as $\pi^\pm$. This procedure guarantees that the particle type of a CT is set to exactly one species.

The CT four-momentum vector p has been modified by

$$p = \begin{pmatrix} \sqrt{|\vec{p}|^2 + m_\xi^2} \\ \vec{p} \end{pmatrix} \quad \text{with} \quad \xi \in \{e^\pm; \mu^\pm; K^\pm; \pi^\pm\} \tag{4.4}$$

in order to assign the nominal particle masses m_ξ of the identified particle type ξ, where $\vec{p}$ denotes the three-momentum vector measured in the tracking system.

4.4 *Dalitz*-Plot Correction

As claimed in Sec. 2.5.2, the *Dalitz*-plot of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ shows a clear pattern induced by the peculiar kinematics of this process and therefore it can be exploited to suppress combinatorial 3π background. On the other hand, it has also been mentioned in Sec. 2.5.3 that neither the KS nor the IMR model reproduce the *Dalitz*-plot and its projections sufficiently well. Hence, before the $B^- \to \tau^-\overline{\nu}_\tau$ reconstruction techniques are discussed, it follows an explanation how this insufficiency of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ *Dalitz*-plot description is taken into account.

A powerful way to test and correct MC descriptions is to select a well defined control sample and compare the properties of interest to the corresponding MC expectation. Since in our case the properties of the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ should be investigated, a large τ sample is needed, which is accessible and cleanly reconstructable from the *BABAR* data. $e^+e^- \to \tau^+\tau^-$ events fulfil all these requirements. τ-pair events are produced with large cross section (Tab. 3.2). Furthermore, due to their unique topology such events can be reconstructed with low background.

In particular, we used τ-pair events with one τ decaying into a leptonic final state, while the companion τ lepton decayed via $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. Events of such an 1-3 topology[4] can be reconstructed very cleanly and backgrounds from underlying $\Upsilon(4S)$ and $e^+e^- \rightarrow q\bar{q}$ events are negligible. In the following the selection of a τ control sample is discussed in detail.

4.4.1 $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ Control Sample Selection

The 1-3 $\tau^+\tau^-$ event selection is based on the *Tau1N* skim [80], which has been developed to efficiently preselect $e^+e^- \rightarrow \tau^+\tau^-$ events, where one τ decays into a final state with one charged particle, while the other one produces three or more charged particles. In principle, this preselection divides the event into two hemispheres, which are then required to include one and at least three GTVL candidates, respectively. This requirement exploits the special topology of τ-pairs and already reduces non-τ backgrounds considerably. τ-pair events represent an enormous portion of both the On Peak as well as the Off Peak sample. Thus, the entire Run14 data sample has been used corresponding to an integrated luminosity of $\mathcal{L} = \mathcal{L}_{\text{on}} + \mathcal{L}_{\text{off}} = 232.1\,\text{fb}^{-1}$ (Tab. 3.1). The procedure described in the following aims at a clean selection of $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ decays, where the second τ lepton is required to decay into a leptonic final state $\tau^+ \rightarrow \ell^+\nu_\ell\bar{\nu}_\tau$. Hence, only events with exactly four GTL candidates and a total charge c_{tot} of zero are accepted (Fig. 4.2), where c_{tot} is defined as

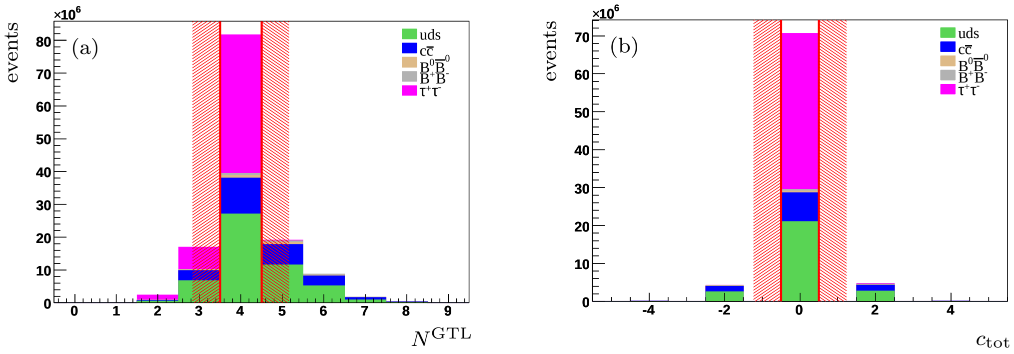


Figure 4.2: τ-pair selection: (a) the GTL multiplicity, and (b) the total event charge after the requirement $N^{\text{GTL}} = 4$. The term "*uds*" in the legends denotes continuum events of type $e^+e^- \rightarrow u\bar{u}, d\bar{d}, s\bar{s}$. The red lines illustrate the cuts applied.

$$c_{\text{tot}} = \sum_i c_i \quad \text{with} \quad i \in \text{CT} \tag{4.5}$$

and the track charge c_i. Furthermore, at least one identified lepton passing the GTL selection is required and a cut on the sum of the EMC deposits of GPL candidates of $E_{\text{sum}}^{\text{GPL}} < 0.5\,\text{GeV}$ (Fig. 4.3) has been applied.

---

[4] One of the τ leptons decays into a mode with one charged track (1-prong). The second τ decayed into three charged tracks (3-prong).

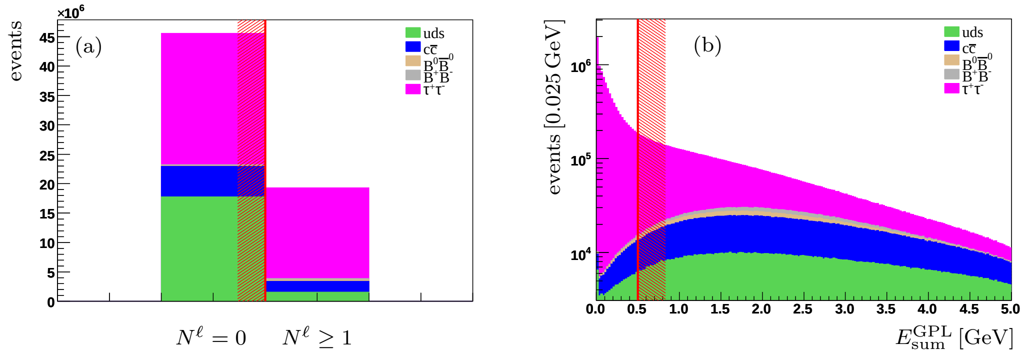


Figure 4.3: τ-pair selection: (a) classifies events with ($N^\ell \geq 1$) and without ($N^\ell = 0$) an identified lepton candidate, and (b) shows the energy sum of GPL candidates. The red lines illustrate the cuts applied.

After this selection the sample is by far dominated by τ-pair events. Backgrounds from continuum and $\Upsilon(4S)$ events are negligible as can be seen in Fig. 4.4. In order to reject outliers in the 3π invariant mass $m_{3\pi}$, we accept events with $0.5 < m_{3\pi} < 2.0\,\text{GeV}/c^2$, where 3π candidates are defined as combinations of the three remaining GTL after exclusion of the identified $\ell^\pm$ candidate.

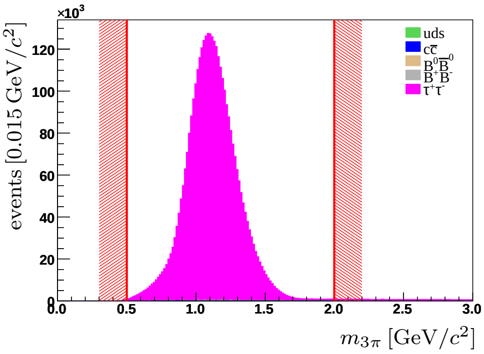


Figure 4.4: Invariant 3π mass after $E_{\text{sum}}^{\text{GPL}}$ requirement: The red lines illustrate the cuts applied.

Besides the reduction of non-τ backgrounds it is important to understand the decomposition of the remaining sample. Since we want to correct the kinematics of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ in the MC, it is indispensable to reject other τ decays, which might also be poorly known and therefore not correctly modeled in the MC. One component, which could possibly introduce additional deviations into the *Dalitz*-plot, is represented by the five-body decay $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ described by a simple phase space distribution in the MC simulation. On the other hand, the structure of this decay is not clear and resonances might contribute. Such unmodeled resonances would introduce differences between data and MC spoiling the correction of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ kinematics. On the other hand, the branching fraction of this five-body decay is well measured (Tab. 2.6). Thus, if this process could be strongly suppressed in the MC, this component should also be negligible in data. A suppression of the $\tau \to 3\pi\pi^0\nu$ process from the control sample can be achieved by a tightening of the requirement on neutral energy in the entire event.

Fig. 4.5 illustrates the number of GPL candidates per event N^{GPL} separately for correctly reconstructed $\tau \to 3\pi\nu$, $\tau \to 3\pi\pi^0\nu$, and combinatorial backgrounds in τ-pair events. Besides the fact that the mean of the data distribution differs significantly from the MC expectation, the possibility of suppression of $\tau \to 3\pi\pi^0\nu$ as well as combinatorial τ-background becomes visible. A cut on $N^{\mathrm{GPL}} = 0$ considerably improves the pureness of the $\tau \to 3\pi\nu$ control sample, even though this selection results in a discrepancy in efficiency between data and MC.

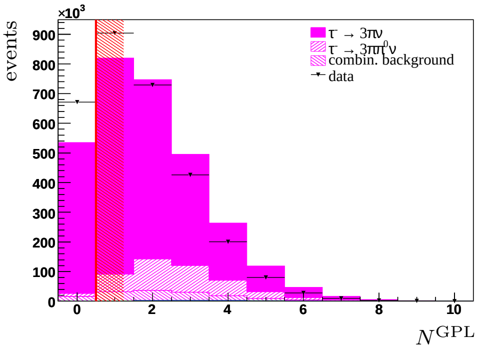


Figure 4.5: Multiplicity of GPL candidates after $m_{3\pi}$ requirement

At this stage of the control sample selection the MC indicates an almost background free $\tau^+ \to \ell^+\nu_\ell\overline{\nu}_\tau$ vs. $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ sample. However, it is necessary to check if other processes not included in the MC simulation survive the selection. Such a possible background source is represented by radiative *Bhabha* events with a subsequent conversion of the radiated photon ($e^+e^- \to e^+e^-\gamma \to e^+e^-e^+e^-$). These events are characterized by four tracks, no neutral energy, a total charge of zero, and electron tracks. Hence, in principle such events are selected by the described procedure.

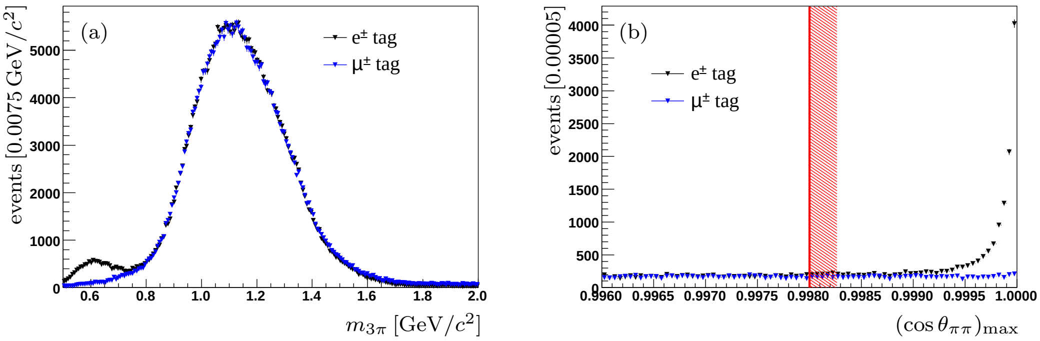


Figure 4.6: (a) illustrates $m_{3\pi}$ extracted from the data sample for $e^\pm$ and $\mu^\pm$ tagged events, respectively. The difference in the region $m_{3\pi} < 0.8\,\mathrm{GeV}/c^2$ is induced by radiative *Bhabha* events. This is confirmed by (b) the distribution of $(\cos\theta_{\pi\pi})_{\mathrm{max}}$.

In order to check for such a contribution, $m_{3\pi}$ has been plotted separately from data events including an identified $e^\pm$ ($e^\pm$ tag) or $\mu^\pm$ ($\mu^\pm$ tag), respectively. The distributions have been scaled to the same peak height at $m_{3\pi} \approx 1.1\,\mathrm{GeV}/c^2$ (Fig. 4.6a). The comparison uncovers an enhancement of $e^\pm$ tagged events supporting the assumption of surviving radiative *Bhabha* events. Typically, the γ is radiated within a narrow cone around the three-momentum vector of the emitting $e^\pm$. Due to kinematical reasons the e^+e^--pair generated in the γ conversion are expected to enclose a small angle. Therefore, the three tracks are all included within one narrow cone and angular correlations are exploitable to reject these events. We define

$$(\cos\theta_{\pi\pi})_{\mathrm{max}} = \max\left(\cos\theta_{\pi_i\pi_j}\right) \quad \text{with} \quad i,j \in \{1;2;3\} \wedge i \neq j\,, \tag{4.6}$$

where $\theta_{\pi_i\pi_j}$ denotes the angle between the three-momentum vectors of the i-th and j-th GTL candidate used to reconstruct the 3π candidate. Fig. 4.6b gives the distribution of this variable again for $e^{\pm}$ and $\mu^{\pm}$ tagged events, where the same scaling factor has been used. Indeed, an enhancement at high $(\cos\theta_{\pi\pi})_{\max}$ corresponding to small maximal angles between the tracks becomes visible. This confirms the assumption of radiative *Bhabha* events and a cut at $(\cos\theta_{\pi\pi})_{\max} < 0.998$ has been introduced.

A test of the selected control sample for additional unmodeled background components can be achieved by exploiting the topology of τ-pairs using event shape variables. In particular, the angular distribution of detected particles produced in the e^+e^- collision is useful to separate events of different topologies. The event shape variable R_2 reflects the opening angles between the flight directions of detected particles and is defined as

$$R_2 = \frac{H_2}{H_0} \quad \text{with} \quad H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{E_{\text{tot}}^2} P_l(\cos\theta_{ij}) \quad \text{and} \quad l = 0, 1, 2\dots . \tag{4.7}$$

This quantity is calculated as the ratio of two angular moments H_2 and H_0, which have been proposed by *Fox* and *Wolfram* in 1978 [81] to separate two-jet and three-jet events in e^+e^- annihilation. The l-th *Fox-Wolfram* moment H_l is calculated as a double sum over a set of candidates i with three-momenta $\vec{p}_i$ in the c.m. frame.[5] The angular dependence is introduced by the *Legendre* polynomials $P_l(\cos\theta_{ij})$ with the opening angle $\cos\theta_{ij}$ between the momentum vectors of candidate i and j. E_{tot} denotes the sum of energies measured in the c.m. frame of all candidates. R_2 takes a value between zero and one, where small values correspond to isotropic events of spherical geometry, while non-spherical particle distributions are characterized by high R_2.

Since the τ leptons are produced with high momenta in $e^+e^- \to \tau^+\tau^-$ at the given c.m. energy, the decay products are expected to be emitted in narrow cones

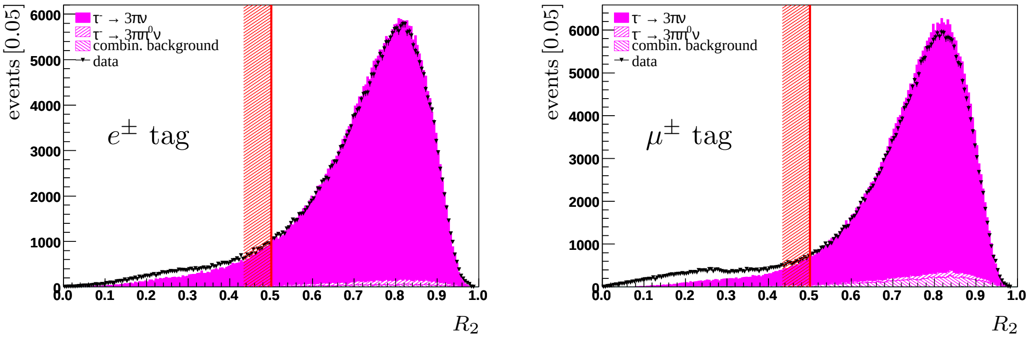


Figure 4.7: Unmodeled contributions in τ control sample: The plots show R_2 for $e^{\pm}$ and $\mu^{\pm}$ tagged events, respectively. The red lines illustrate the cuts applied.

---

[5] Throughout this thesis energies and momentum vectors are given in the c.m. frame. If a different frame is used it is explicitly mentioned.

around the τ three-momenta resulting in non-spherical particle distributions corresponding to high R_2 values. This behaviour is visualized in Fig. 4.7, where R_2 has been calculated from all CT and CN candidates. Furthermore, the distributions indeed indicate unmodeled contributions in the data, which cluster at low R_2 for both $\ell^{\pm}$ types. The origin of this enhancement has not been investigated, but the selection has been extended by $R_2 > 0.5$ to remove these unknown background components.

B^+B^-	$B^0\overline{B}^0$	$c\overline{c}$	$u\overline{u}, d\overline{d}, s\overline{s}$	$\tau^+\tau^-$	**Off Peak**	**On Peak**	**Data**
5	17	52	184	976501	53156	526118	578873

Table 4.3: Number of events after τ control sample selection: The number of data events is equal to the sum of Off Peak and On Peak events, where the number of Off Peak events has been scaled down with $s_{\text{off}}/s_{\text{on}} \approx 10.54^2/10.58^2 \approx 0.9925$ in order to account for the $1/s$ dependence of the cross section for e^+e^- annihilation. The MC numbers have not been scaled to Run14 luminosity.

Tab. 4.3 summarizes the decomposition of the control sample after the entire τ-pair selection. This high-purity (95 %) $\tau^+ \to \ell^+\nu_\ell\overline{\nu}_\tau$ vs. $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ control sample can be used to test the MC description of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ kinematics.

4.4.2 *Dalitz*-Plot Reweighting

The procedure described in the following aims at the extraction of weighting factors $w_{3\pi}(\sqrt{s_1}, \sqrt{s_2})$ depending on the invariant masses of both neutral $\pi\pi$ combinations (Sec. 2.5.2) in order to reweight 3π candidates in the MC simulation. Fig. 4.8 illus-

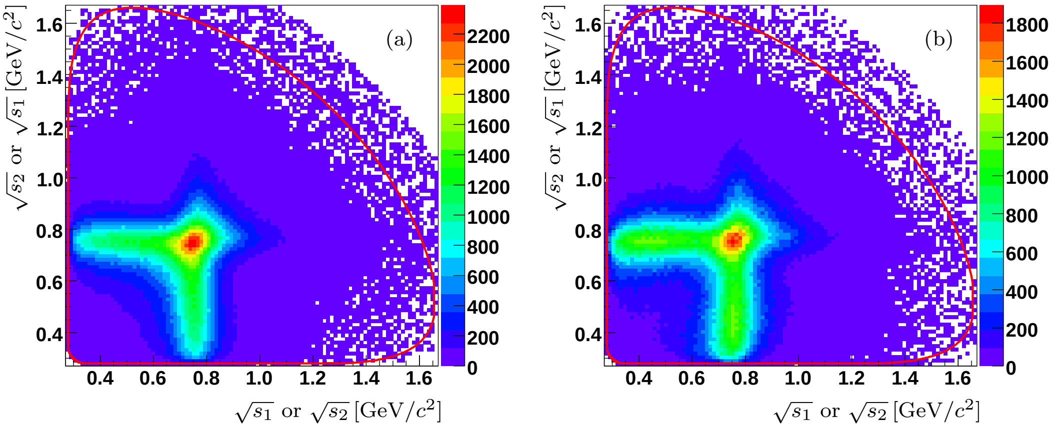


Figure 4.8: *Dalitz*-plots of τ control sample from (a) combined MC and (b) On Peak plus Off Peak data: Assuming symmetric *Dalitz*-plots, the distributions have been symmetrized, i.e. every event entered twice, once at ($\sqrt{s_1}$,$\sqrt{s_2}$) and once at ($\sqrt{s_2}$,$\sqrt{s_1}$). The MC distribution has been scaled to the number of entries in the data sample. The red lines illustrate the kinematically allowed region for $0.5 < m_{3\pi} < 1.8\,\text{GeV}/c^2$.

trates the *Dalitz*-plots for the entire kinematically allowed range

$$2m_{\pi^\pm} < \sqrt{s_{1,2}} < m_{\tau^\pm} - m_{\pi^\pm} \approx 0.28 < \sqrt{s_{1,2}} < 1.66\ \mathrm{GeV}/c^2 \tag{4.8}$$

extracted from the τ control sample. The deviations of the MC expectation and the *Dalitz*-plot observed in the data sample are confirmed by the projections (Fig. 4.9a) and the invariant 3π mass spectrum (Fig. 4.9b).

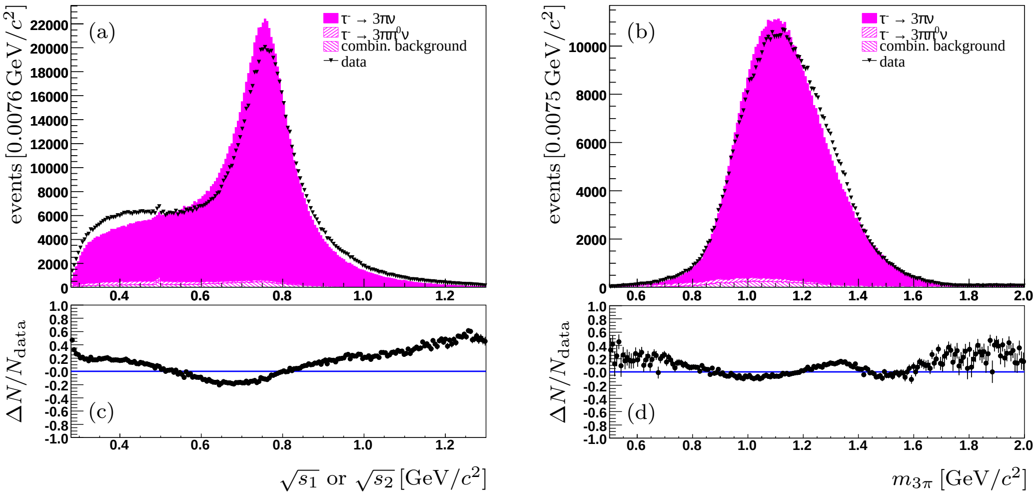


Figure 4.9: (a) *Dalitz*-plot projection and (b) invariant 3π mass as obtained from control sample (Non-τ backgrounds are included in the MC distributions). (c,d) illustrate the bin-by-bin deviations between data and MC extracted from (a) and (b), respectively ($\Delta N = N_{\mathrm{data}} - N_{\mathrm{MC}}$).

The reweighting factor for a given 3π combination can be directly obtained from the *Dalitz*-plots. Depending on its position within the *Dalitz*-plane the weight is calculated as the ratio of the actual bin content found in the data sample and the corresponding bin content in the combined MC expectation, where the MC distribution has been rescaled to the same integral as the data distribution (Fig. 4.10a). In low statistic regions bins have been summed up in order to obtain more reliable weighting factors (Fig. 4.10b).

In order to validate this technique, the weights have been applied to the MC control sample itself, which should automatically transform the MC *Dalitz*-plot into the data distribution. Fig. 4.11 shows the projections and the 3π mass spectrum after reweighting of correctly reconstructed $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ candidates. While the projection almost perfectly agrees to the data points (Fig. 4.11a), there are still deviations in the 3π mass spectrum above $1.1\,\mathrm{GeV}/c^2$ (Fig. 4.11b). This residual deviation has been taken into account as a source of systematic uncertainties (Sec. 4.7.9.5). However, there is a strong improvement.

Due to the rescaling of the MC sample to the integral of the corresponding data distribution the extracted weights only account for the difference of the two-dimensional shape in the *Dalitz*-plane. Absolute deviations arising from differences

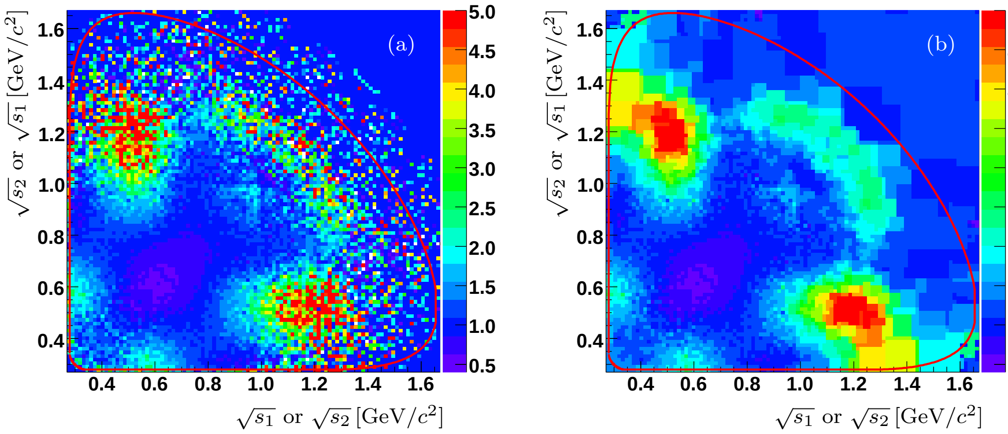


Figure 4.10: *Dalitz*-plot correction factors extracted from the τ control sample: (a) gives the ratio of the symmetrized *Dalitz*-plots. (b) shows the same quantity, where bins in low-statistics regions have been added to get more reliable weights.

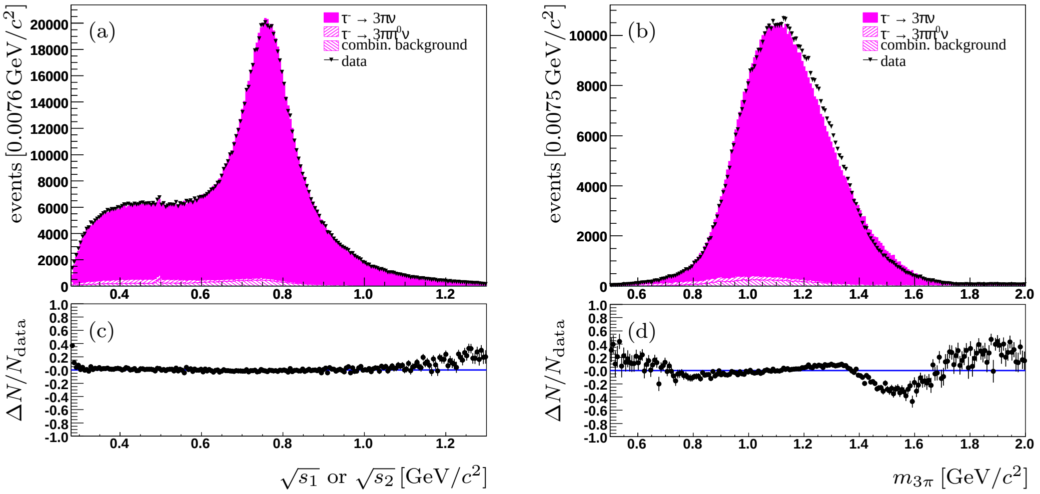


Figure 4.11: (a) *Dalitz*-plot projection and (b) invariant 3π mass as obtained from the control sample after the reweighting.

in the reconstruction efficiency have not been taken into account. It is assumed that such effects will not affect the relative differences of the two *Dalitz*-plots but only result in a relative overall normalization factor. However, such a momentum dependence should be visible as a modification of the π momentum spectra as well as the 3π momentum spectrum after the reweighting has been applied. It has been verified that neither the π momentum spectra nor the 3π momentum spectrum have been modified by the reweighting technique. As it was expected, the reweighting procedure only modifies the *Dalitz*-plot and the invariant 3π mass.

One should note that the weighting factors $w_{3\pi}$ have been determined from *Lorentz*-invariant quantities ($\sqrt{s_1}$, $\sqrt{s_2}$). This is important since the *Dalitz*-plot correction factors are later applied to signal MC, where the τ leptons are produced in B meson decays and therefore in a different kinematical range. Naturally, this results in lower momenta of the 3π candidates originated in the τ decay. While the 3π candidates in $e^+e^- \to \tau^+\tau^-$ events are produced with momenta up to $|\vec{p}_{3\pi}| \approx 5\,\mathrm{GeV}/c$, the endpoint of the $|\vec{p}_{3\pi}|$ spectrum in $B^- \to \tau^-\overline{\nu}_\tau$ events has been found to be at $|\vec{p}_{3\pi}| \approx 2.7\,\mathrm{GeV}/c$, as it will be shown in Sec. 4.5.2. However, since the shape of the *Dalitz*-plot in principle does not depend on the kinematical region of the τ decay, the weights extracted from the given control sample are applicable for $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decays on the lower momentum scale, whereas detector resolution effects are supposed to be negligible.

In summary, using the above procedure weights $w_{3\pi}(\sqrt{s_1}, \sqrt{s_2})$ have been calculated from the data and MC *Dalitz*-plots of a high-purity $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ control sample. Since the 3π candidate selections used within the $B^- \to \tau^-\overline{\nu}_\tau$ reconstruction techniques explicitly exploit kinematical properties, such as the invariant 3π mass $m_{3\pi}$ as well as $\sqrt{s_1}$ and $\sqrt{s_2}$, the application of these weighting factors results in a modification of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ and therefore of the $B^- \to \tau^-\overline{\nu}_\tau$ reconstruction efficiency. Correctly reconstructed 3π candidates in signal MC have been reweighted by these weights, where the reconstructed invariant masses of the neutral $\pi\pi$ combinations have been used to extract them from the ratio of the two *Dalitz*-plots (Fig. 4.10b). All figures corresponding to 3π candidates in signal MC shown in this chapter include the *Dalitz*-plot correction factors. In the following the three different reconstruction techniques are discussed in detail and the potential of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay mode for a $B^- \to \tau^-\overline{\nu}_\tau$ search is clarified.

4.5 Inclusive Reconstruction

As mentioned before the $B^- \to \tau^-\overline{\nu}_\tau$ search presented in this thesis does not start with an a priori B_{tag} reconstruction in order to select signal events with high efficiency. Fig. 4.12 sketches the strategy to realize this goal. For the inclusive search the entire Run12 data and MC samples have been used.[6]

Sec. 2.5.1 and Sec. 3.2.1 mentioned that the main advantage of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ vs. the 1-prong τ decays lies in the possibility of vertexing of the three π tracks in the SVT. Although the B_{tag} is not explicitly reconstructed, the tracks not belonging to the 3π candidate can be used to reconstruct a B_{tag} vertex. The quality of separation of these two vertices depends on the flight direction of the τ relative to the B_{tag} decay region. If the τ three-momentum vector points to the B_{tag} decay vertex, one expects indistinguishable vertices. On the other hand, if the τ strides away from the B_{tag} decay, the vertices should be separable.

---

[6] As long as the MC and data sets used are not explicitly mentioned in the captions, all plots shown in this section correspond to the entire Run12 samples, where all distributions have been rescaled to the On Peak Run12 luminosity.

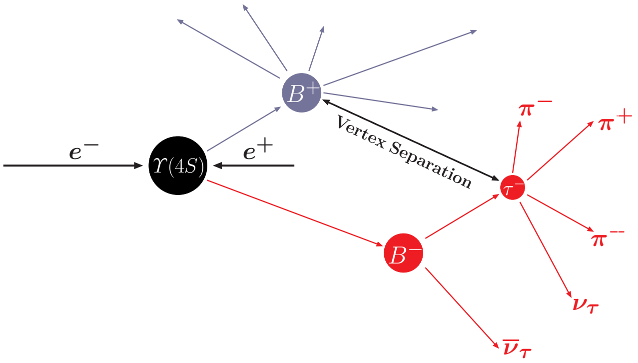


Figure 4.12: Potential of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ for a search for $B^- \to \tau^-\overline{\nu}_\tau$: The non-detectable neutrinos on the B_{sig} side [🟥] induce a high missing momentum. The π tracks can be vertexed to a common 3π decay vertex, which is approximately equal to the $\tau \to 3\pi\nu_\tau$ decay vertex. The remaining tracks detected can be used to reconstruct the B_{tag} [🟪] decay vertex and to determine the B_{tag} flavour.

The presence of at least two neutrinos in the final state represents the main difficulty in the reconstruction. On the other hand, the invisibility of the neutrinos can be used for event selection requiring high missing momentum. Furthermore, the properties of the reconstructable $a_1 \to 3\pi$ candidate[7] can be exploited and correlations of the B_{tag} and B_{sig} side might be useful for background suppression.

Since *BABAR* possesses a high experience in *Flavour Tagging* from time-dependent CP studies, i.e. the determination of the B flavour from its decay products, this technique can further be applied to reject background.

4.5.1 Event Preselection

The event selection starts with a loose preselection based on the topology of signal events. Fig. 4.13 shows distributions of variables characterizing the topologies of signal and background events.

The number of GTL candidates per event N^{GTL} is very useful to suppress τ-pair background illustrated in Fig. 4.13b. Events with $5 \leq N^{\text{GTL}} \leq 10$ are selected. In principle, one might expect the lower cut to be too restrictive (Fig. 4.13a), but besides at least three good tracks needed for the a_1 reconstruction additional tracks are required to access information of the B_{tag} side.

As visualized in Fig. 4.13b, an enormous portion of background arises from continuum events ($e^+e^- \to q\overline{q}$). Typically, such events differ from $\Upsilon(4S) \to B\overline{B}$ by their angular particle distributions. Since the primary $q\overline{q}$-pair is produced with high momentum transfer, hadrons are predominantly generated in fragmentation processes. Depending on the energy scale the resulting particles are more or less bundled in jets since they are produced with low transverse momenta relative to the jet axis. This property of $q\overline{q}$ events is reflected in the event shape variable R_2, which has already been used within the τ control sample selection (Eq. 4.7). Figs. 4.13e,f illustrate the

---

[7] In the following 3π candidates are also denoted as a_1 candidates.

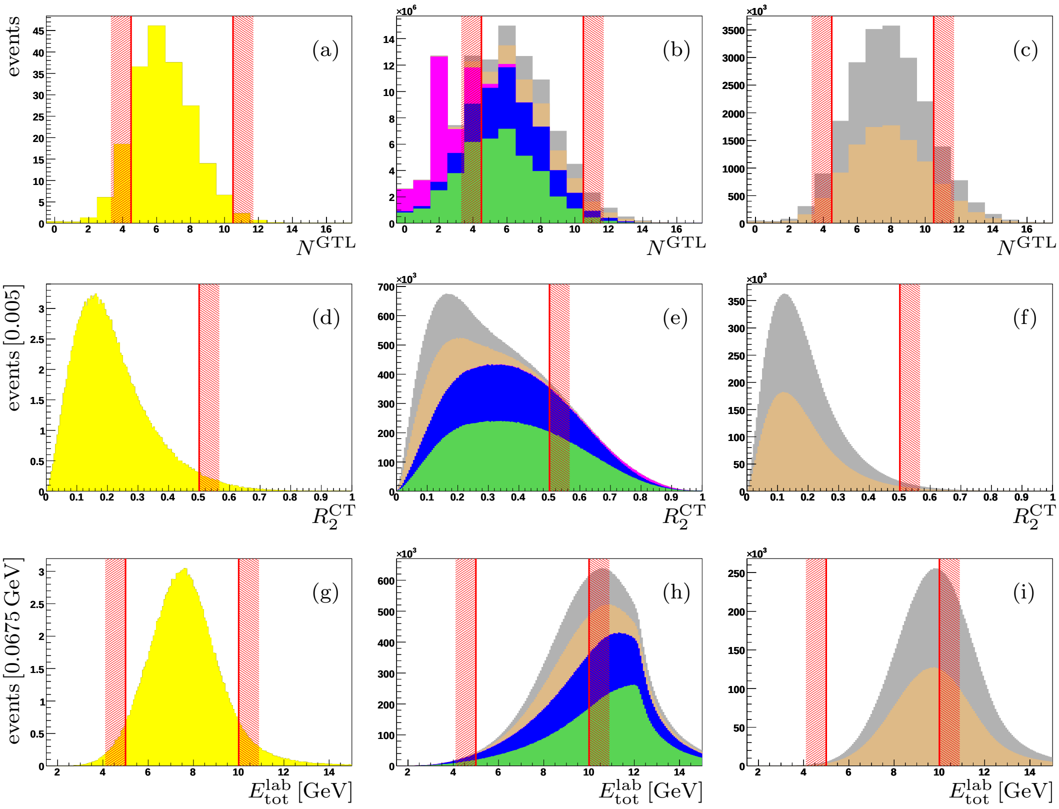


Figure 4.13: Distributions of (a-c) N^{GTL}, (d-f) R_2^{CT} after N^{GTL} selection, and (g-i) $E_{\mathrm{tot}}^{\mathrm{lab}}$ after N^{GTL} and R_2^{CT} selection. The plots have been created using a subset of the Run1 MC sets and the samples are scaled to the On Peak Run1 luminosity. The red lines indicate the cuts applied. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜])

described behaviour. Here, R_2 has been calculated from CT candidates only and is therefore denoted as R_2^{CT}. In average, $q\bar{q}$ events cluster at higher R_2^{CT}. However, the opening angle of the jet-cones depends on the energy scale and the jet structure becomes more distinct at higher c.m. energies. Thus, after the selection of events with $R_2^{\mathrm{CT}} < 0.5$ a large amount of $q\bar{q}$-pairs remains in the selected sample. On the other hand, a tighter R_2^{CT} requirement results in loss of signal efficiency as can be seen in Fig. 4.13d.

The unseen neutrinos in the final state induce a signature of missing energy for signal events. A comparison of the total visible energy $E_{\mathrm{tot}}^{\mathrm{lab}}$ measured in the laboratory frame between signal MC and backgrounds (Figs. 4.13g-i) establishes the opportunity of background suppression by selecting events with $E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$. A lower cut ($E_{\mathrm{tot}}^{\mathrm{lab}} > 5\,\mathrm{GeV}$) is motivated by rejection of two-photon events ($e^+e^- \to \gamma\gamma$) in the data sample not included in the MC. Such events tend to cluster below 5 GeV.

Besides the requirements on N^{GTL}, R_2^{CT}, and $E_{\mathrm{tot}}^{\mathrm{lab}}$ events are required to contain at least one "good" 3π candidate defined in Sec. 4.5.2.

4.5.2 a_1 Candidate Preselection

The kinematics of the reconstructable 3π candidate has been introduced as one important handle to search for $B^- \to \tau^- \overline{\nu}_\tau$. Indeed, the 3π properties are helpful to preselect events.

a_1 candidates are reconstructed as combinations of three GTL candidates with charge $|c_{3\pi}| = |c_{\pi_1} + c_{\pi_2} + c_{\pi_3}| = 1$ and

- invariant 3π mass of $0.8 < m_{3\pi} < 1.6\,\mathrm{GeV}/c^2$,
- magnitude of 3π three-momentum of $1.5 < |\vec{p}_{3\pi}| < 2.7\,\mathrm{GeV}/c$, and
- cosine of angle between $\vec{p}_{3\pi}$ and $\vec{p}_{\mathrm{miss}}$ (Eq. 4.9) $\cos\Theta_{3\pi,\mathrm{miss}} < -0.8$.

Fig. 4.14 illustrates the corresponding distributions for correctly reconstructed a_1 candidates in signal MC and backgrounds. The term "correctly reconstructed a_1" denotes 3π candidates found to be generated in a true $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay using MC information, where the τ has been produced in a $B^- \to \tau^- \overline{\nu}_\tau$ decay.

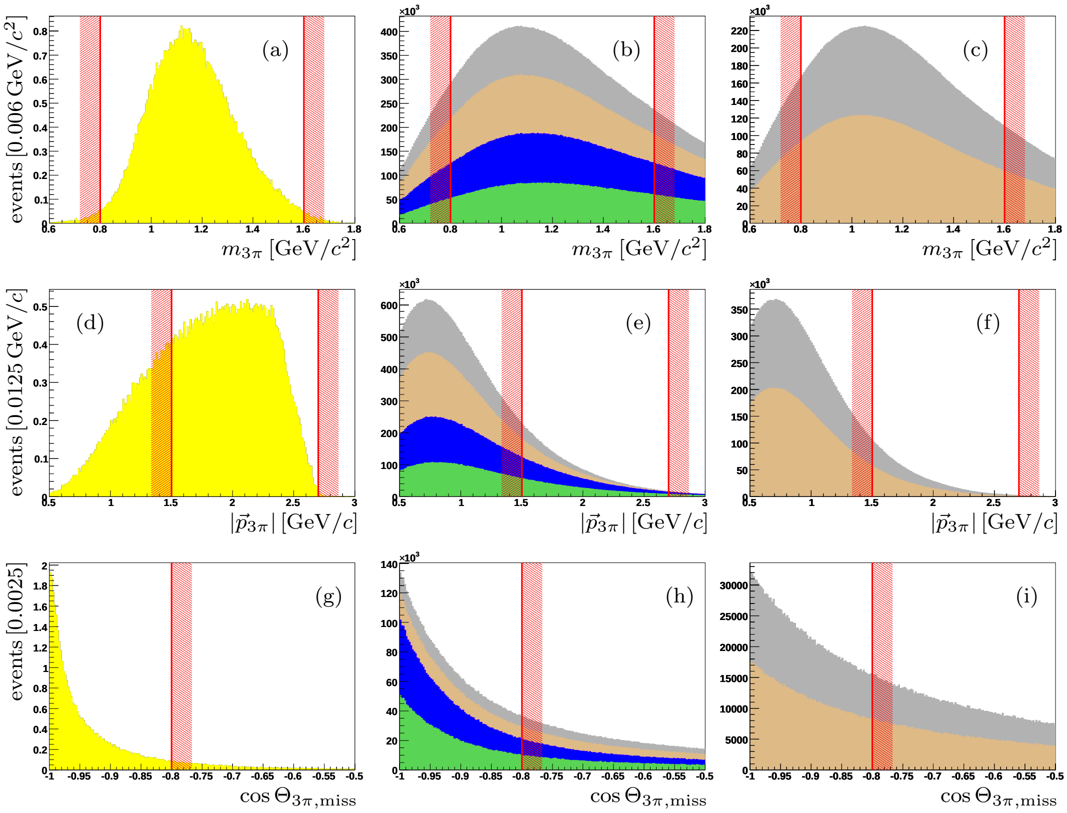


Figure 4.14: Distributions of (a-c) $m_{3\pi}$, (d-f) $|\vec{p}_{3\pi}|$ after $m_{3\pi}$ selection, and (g-i) $\cos\Theta_{3\pi,\mathrm{miss}}$ after $m_{3\pi}$ and $|\vec{p}_{3\pi}|$ selection. The plots have been created using a subset of the Run1 MC sets and the samples are scaled to the On Peak Run1 luminosity. The red lines illustrate the cuts applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [■])

As already mentioned in Sec. 2.5.2 the a_1 is a broad resonance. Therefore, one would not expect a high separation power between signal and background in its invariant mass spectrum. Figs. 4.14b,c illustrate the high background contribution within the 3π mass range.

In contrast, $|\vec{p}_{3\pi}|$ offers the possibility of background suppression (Figs. 4.14d-f). Although the relatively strong requirement reduces the signal efficiency by a factor of approximately 30 %, an enormous portion of background is rejected. The upper cut is motivated by the kinematical limit shown in Fig. 4.14d.

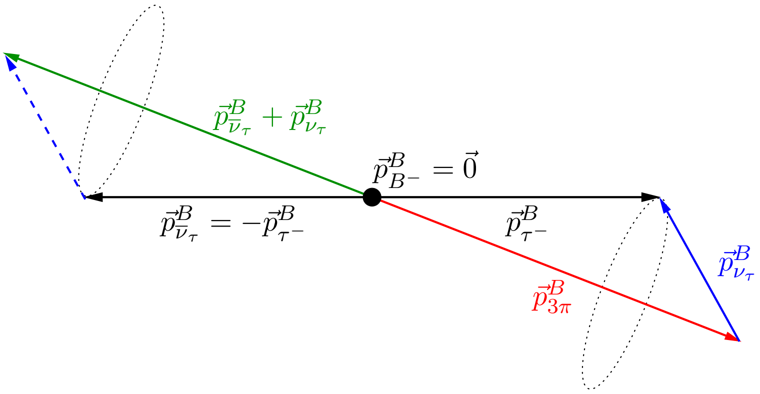


Figure 4.15: Vector diagram of $B^- \to \tau^-\overline{\nu}_\tau$: The plot shows the three-momentum vectors of the particles involved in $B^- \to \tau^-\overline{\nu}_\tau$ with the subsequent $\tau^- \to (3\pi)^-\nu_\tau$ decay. The combined neutrino three-momentum vector is antiparallel and equal in magnitude to the 3π three-momentum vector. The ellipses illustrate the only fixed direction to be given by $\vec{p}^B_{3\pi}$. The superscript "B" indicates that all three-momenta are given in the B rest frame.

Fig. 4.15 illustrates that the 3π three-momentum vector $\vec{p}^B_{3\pi}$ and the combined momentum of the neutrinos $\vec{p}^B_{\overline{\nu}\nu} \equiv \vec{p}^B_{\overline{\nu}_\tau} + \vec{p}^B_{\nu_\tau}$ in the B rest frame are back-to-back and equal in magnitude. Due to the B motion relative to the c.m. frame with an average momentum of 320 MeV/c the B rest frame slightly differs from the c.m. frame. However, correlations of $\vec{p}_{3\pi}$ and the missing momentum $\vec{p}_{\text{miss}}$ are exploitable, where the missing four-momentum is defined as

$$p_{\text{miss}} = \begin{pmatrix} E_{\text{miss}} \\ \vec{p}_{\text{miss}} \end{pmatrix} = \begin{pmatrix} \sqrt{s} \\ \vec{0} \end{pmatrix} - \sum_i \begin{pmatrix} E_i \\ \vec{p}_i \end{pmatrix} \quad \text{with} \quad i \in \text{CT} \cup \text{CN}\,. \tag{4.9}$$

In Eq. 4.9 the four-momenta of all reconstructed CT and CN objects are added. If all particles produced on the B_{tag} side have been detected and the a_1 candidate has been correctly reconstructed, p_{miss} is equal to the combined two-neutrino four-momentum vector $p_{\overline{\nu}\nu}$. Apart from detector resolution effects and reconstruction issues the cosine of the angle between $\vec{p}_{3\pi}$ and $\vec{p}_{\text{miss}}$ ($\cos\Theta_{3\pi,\text{miss}}$) is expected to cluster at -1. As visible in Fig. 4.14g signal events show the expected behaviour and mainly $B\overline{B}$ background can be suppressed (Figs. 4.14h-i).

4.5.3 Event Selection

The previous sections explained the event preselection depending on event topology and a_1 candidate requirements. In this section a variety of discriminating variables are presented, which can be combined to perform an signal selection. Instead of simple cutting on single quantities, as it has been done within the event preselection, there are more sophisticated procedures to separate event types efficiently in the n-dimensional variable hyperspace. Such methods account for correlations among the variables and are able to select a given event type with higher efficiencies, while other event species are strongly suppressed.

4.5.3.1 Multivariate Procedures

Multivariate methods are useful to select event samples by using highly correlated quantities. Two such procedures are explained in detail.

A popular multivariate procedure was developed by *R. A. Fisher* [82] in 1936 in order to sort species of plants based on taxonomic observations. However, the algorithm has already been successfully applied by many experiments for classification of physical events. The ***Fisher Discriminant*** **(FD)** is defined as a linear combination of a set of n input variables x_i

$$\mathcal{F} = \alpha_0 + \sum_{i=1}^{n} \alpha_i x_i \,, \tag{4.10}$$

where the offset α_0 just centers the sample mean of $\mathcal{F}$ at zero. The *Fisher* coefficients α_i are determined by analyzing a so-called training sample of events belonging to different event species, i.e. signal and background. The coefficients are calculated from the covariance matrices of the input variables for both species U^{sig} and U^{bkg} and the mean values of the input variables μ_j^{sig} and μ_j^{bkg} extracted from the training samples as

$$\alpha_i = \sum_{j=1}^{n} (U_{ij}^{\text{sig}} + U_{ij}^{\text{bkg}})^{-1} \times (\mu_j^{\text{sig}} - \mu_j^{\text{bkg}}) \,. \tag{4.11}$$

In the case of uncorrelated input variables, i.e. diagonal covariance matrices U^{sig} and U^{bkg}, one obtains

$$\alpha_i = \frac{\mu_i^{\text{sig}} - \mu_i^{\text{bkg}}}{(\sigma_i^{\text{sig}})^2 + (\sigma_i^{\text{bkg}})^2} \,. \tag{4.12}$$

Eq. 4.12 clarifies that the coefficients α_i maximize the between-class separation (numerator) relative to the within-class variance (denominator). Simply speaking, the FD analysis determines an axis in the n-dimensional variable hyperspace $\mathbb{R}^n$. When projecting the event classes upon this axis, they are pushed as far as possible away from each other, while events of the same class are confined in a close vicinity. The performance of this algorithm depends on the shapes and the relative position of the n-dimensional distributions of the two event classes (Fig. 4.16). Non-linearly

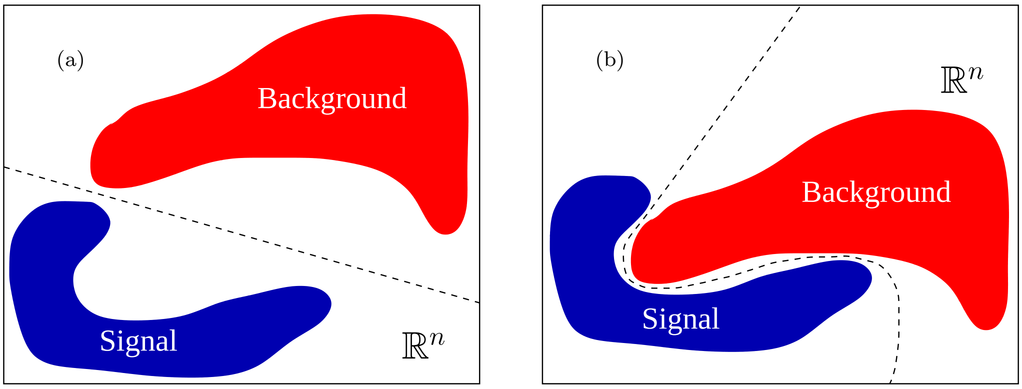


Figure 4.16: Linear separability: Examples for signal and background distributions in $\mathbb{R}^n$, which are (a) linearly separable and (b) not linearly separable.

separable event classes illustrated in Fig. 4.16b require other methods, which are able to take into account higher order correlations between the input variables.

An attractive non-linear algorithm is represented by a certain kind of neural networks known as *Feed-forward Multilayer Perceptron*. Fig. 4.17 illustrates a possible propagation scheme of such an ***Artificial Neural Network* (ANN)**. In general, the ANN consists of one input layer, k hidden layers, and one output layer, where every layer contains a certain number of neurons. All inputs to a neuron of a given layer are linear combinations of the neuron outputs of the previous layer. The gen-

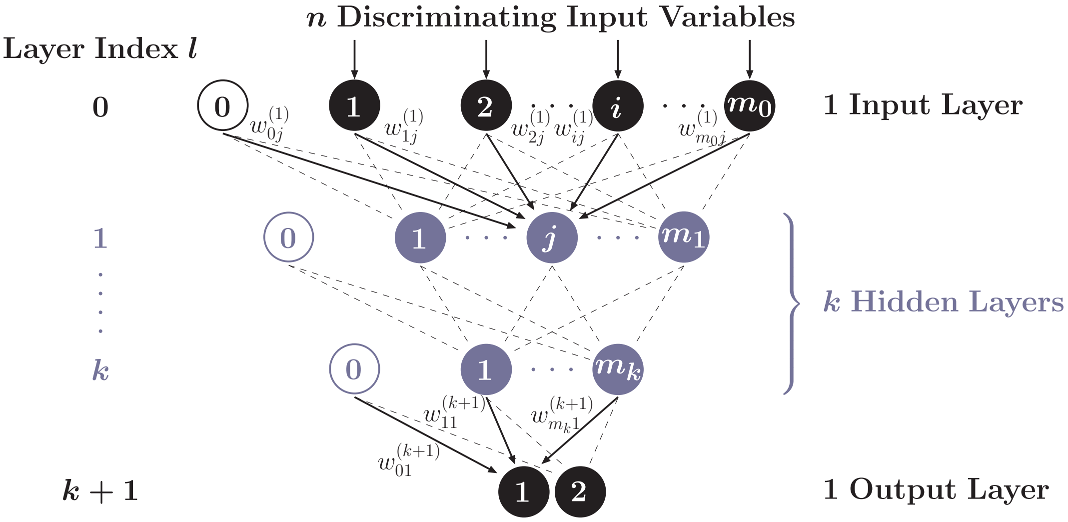


Figure 4.17: Propagation scheme of Artificial Neural Network: The input layer contains as many active neurons as input variables ($m_0 = n$). The output layer consists of two neurons, where one of them returns the signal event weight, while the other one returns the background event weight. In between, there is a variable number of k hidden layers with arbitrary numbers of active neurons. The neurons indicated by the open circles are the bias neurons.

eral mathematic definition of an artificial neuron of the hidden layer l at position j is given as

$$y_j^{(l)} = g\left(w_{0j}^{(l)} + \sum_{i=1}^{m_{l-1}} w_{ij}^{(l)} y_i^{(l-1)}\right) \quad \text{with} \quad l \in [1, 2, ...k], j \in [1, 2, ...m_l]\,. \tag{4.13}$$

g is an activation function that weights how powerful the output from this neuron should be, based on the sum of the input. Here, a sigmoid activation function has been chosen

$$g(y) = (1 + e^{-y})^{-1}\,. \tag{4.14}$$

Additionally to the active neurons of each layer, there are bias neurons as illustrated by the open circles in Fig. 4.17. These neurons have no input from the previous layer, but they are connected to each active neuron of the next layer. These special neurons always emit $y_0^{(l)} = 1$ leading to an additional input weight $w_{0j}^{(l)}$ for the active neurons (Eq. 4.13). The introduction of such bias neurons guarantee that g only depends on y, which speeds up the iteration process substantially. Within this iterative procedure the neurons are trained using a training sample, i.e. the input weights $w_{ij}^{(l)}$ of the neuron-neuron connections are optimized to separate the two event species included in the training sample. Afterwards, the trained ANN can be used to analyze a given data sample. It is important to note that the weights are trained using a statistically limited sample and therefore rely on statistical fluctuations. Hence, the number of training events needs to be high enough to represent a good approximation of the properties of the data sample, which should be analyzed. Furthermore, in order to achieve an optimal separation the number of iteration steps (cycles) have to be chosen accordingly.

Both aforementioned methods represent useful tools to optimize event selections since the n-dimensional parameter space is mapped to the linear scale $\mathbb{R}^n \rightarrow \mathbb{R}$. The FD algorithm is easy to handle, but might not be able to provide optimal separation. In contrast, ANN's are very complex and require an extensive computational framework. Nevertheless, this non-linear algorithm could possibly give better separation between signal and background.

4.5.3.2 Multivariate Event Selection

A multivariate event selection has been performed using a `Fortran` package, which provides both the FD algorithm as well as an interface to an ANN as introduced in the previous section. The ANN algorithm was developed by *P. Gay et al.* [83]. The network has been constructed with two hidden layers ($k = 2$) with $m_0 = n$, $m_1 = n-1$, and $m_2 = n-2$. The number of cycles has been chosen to be $n_{\text{cyc}} = 10000$, which has been found to be reasonable.

A set of nine discriminating variables has been used to suppress background on event basis. Essentially, these quantities describe the event shape, the missing momentum, and particle multiplicities. The complete list with the corresponding signal and background distributions is given in App. B.1. Exemplarily, three of the

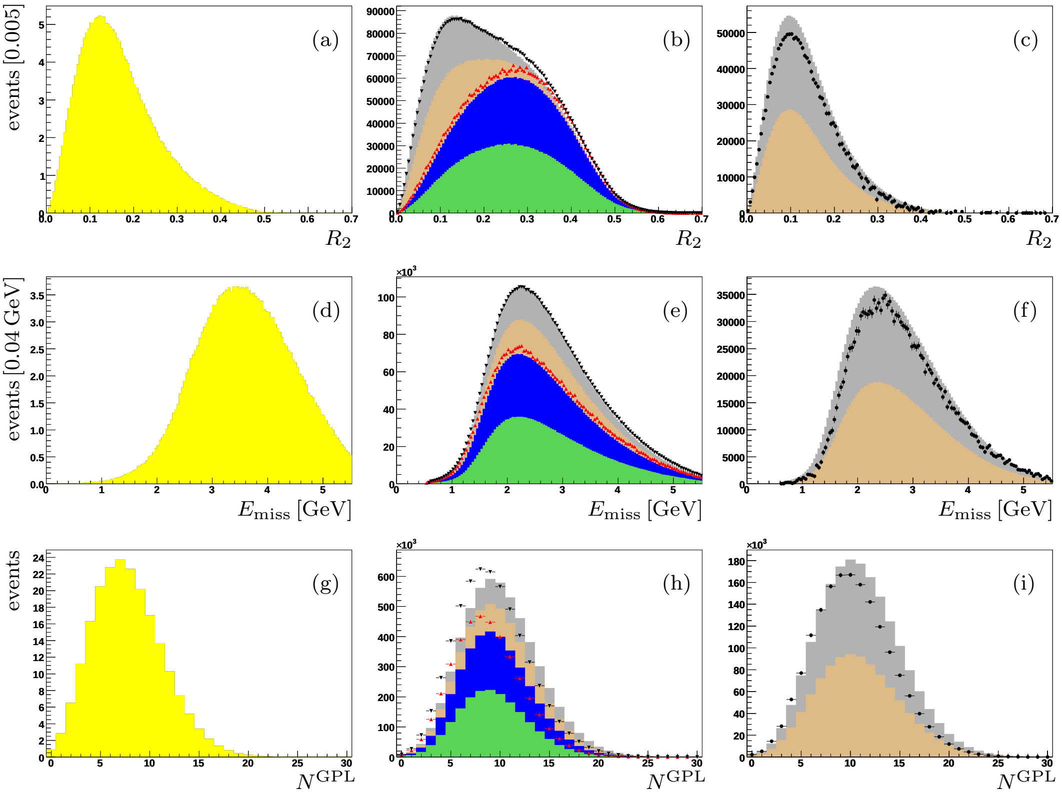


Figure 4.18: Distributions of (a-c) R_2, (d-f) $E_{\rm miss}$, and (g-i) $N^{\rm GPL}$ after event preselection. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

input variables are illustrated in Fig. 4.18, such as R_2, the missing energy $E_{\rm miss}$ both calculated from all CT and CN candidates, and the GPL multiplicity $N^{\rm GPL}$.

These figures show the MC predictions with the corresponding data points overlaid. The quality of the MC description of non-$\Upsilon(4S)$ events can be judged by comparing the Off Peak data points with the combined $q\overline{q}/\tau^+\tau^-$ distributions in Figs. 4.18b,e,h. In contrast, the Off Peak subtracted On Peak data points represent $\Upsilon(4S) \to B\overline{B}$ events and have to be compared with the combined $B^0\overline{B}^0/B^+B^-$ MC expectation (Figs. 4.18c,f,i). Apart from $N^{\rm GPL}$ the distributions of the input variables show good agreement in shape.[8] On the other hand, a significant deviation in the normalization in all quantities is seen. The origin of these differences remains unclear, but this issue needs to be addressed in order to estimate background contributions (Sec. 4.5.3.4).

In order to use multivariate methods a training sample has been chosen. The signal sample consists of nearly 33000 signal MC events passing the event preselection. The background sample is represented by the same number of On Peak data

---

[8] The huge disagreement in $N^{\rm GPL}$ (Figs. 4.18h,i) becomes important within the recoil analysis for background estimation and signal efficiency determination (Sec. 4.7.5.2).

events. This seems to be surprising since data events possibly include signal events now declared as background. On the other hand, at this level of the analysis the contribution of signal in the selected data sample is negligible and will not be visible in any of the input variables. Thus, neither the FD nor the ANN notices the presence of signal events in the background sample. The advantage of using On Peak data lies in the proper decomposition of continuum, τ-pair, and $B\overline{B}$ events and there is no need to create a MC admixture by hand. The signal and background sample have been created as an appropriate admixture of Run1 and Run2 events in order to account for possible changes of the detector conditions and background levels from one run period to another.

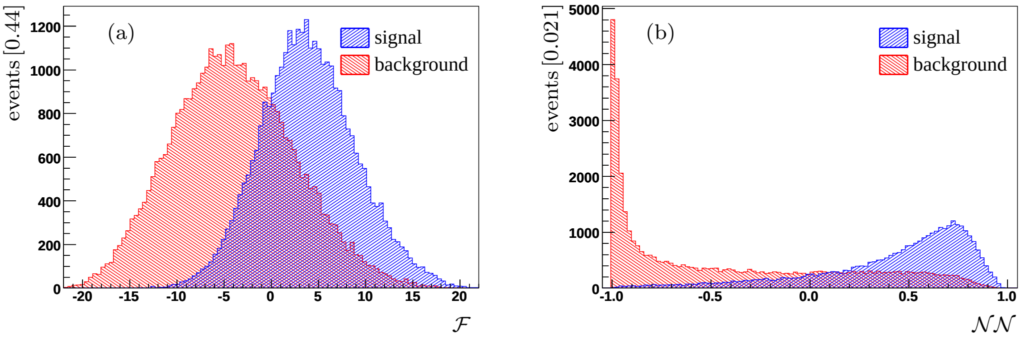


Figure 4.19: Signal-Background separation of event based ANN: (a) *Fisher* discriminant and (b) ANN output determined from the training sample

After training of the *Fisher* coefficients α_i and the ANN weights $w_{ij}^{(l)}$ the training sample has been analyzed for the performance of both algorithms. Fig. 4.19 compares the *Fisher* discriminant $\mathcal{F}$ as well as the ANN output $\mathcal{NN}$. In order to judge the performances of both algorithms the signal and background efficiencies have been determined for different selection requirements on these quantities. Fig. 4.20 visualizes the evolution of the efficiencies with restrictions on $\mathcal{F}$ and $\mathcal{NN}$. Every point gives the background efficiency $\varepsilon_{\mathrm{bkg}}^{\mathrm{train}}$ vs. signal efficiency $\varepsilon_{\mathrm{sig}}^{\mathrm{train}}$ for a given cut on $\mathcal{F}$ or $\mathcal{NN}$. The higher the ratio $\varepsilon_{\mathrm{sig}}^{\mathrm{train}}/\varepsilon_{\mathrm{bkg}}^{\mathrm{train}}$ the better the event classes can be separated. For example, while a cut on $\mathcal{NN}$ which keeps 80 % of the signal will retain about 22 % of the background, a cut on $\mathcal{F}$ rejects only 72 % of the background for the same signal efficiency. Hence, for the given set of input variables the ANN provides a better separation power than $\mathcal{F}$, i.e. higher order correlations can be exploited for background fighting.

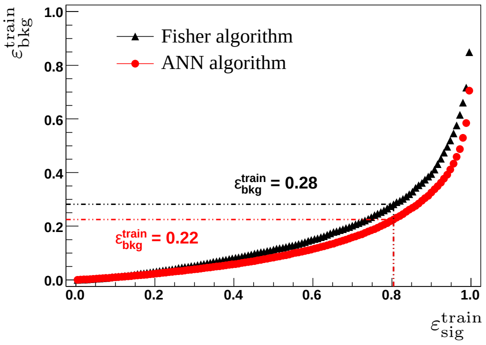


Figure 4.20: Separation power of $\mathcal{F}$ and $\mathcal{NN}$ for multivariate event selection

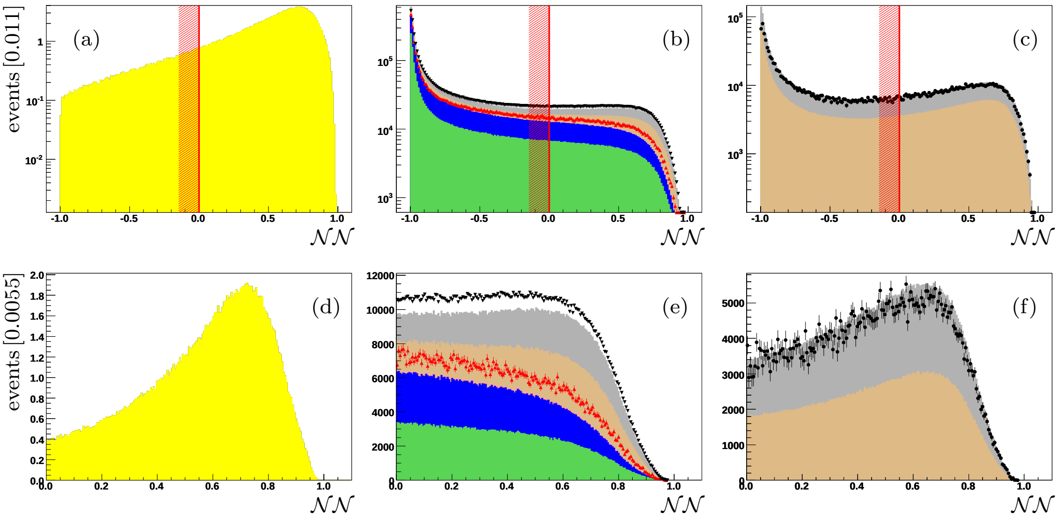


Figure 4.21: Output of the event based ANN after event preselection: (d-f) show the $\mathcal{NN}$ distributions in the range between zero and one. The red lines illustrate the selection applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

The trained ANN has been applied to the Run12 samples. The resulting $\mathcal{NN}$ distributions are illustrated in Fig. 4.21. In order to compare the $\mathcal{NN}$ distributions in data and MC the bin-by-bin differences $\Delta N = N_{\text{data}} - N_{\text{MC}}$ of the scaled numbers of events per bin have been calculated for Off Peak data vs. MC continuum (Fig. 4.22a) as well as for Off Peak subtracted On Peak data vs. generic $B\overline{B}$ MC (Fig. 4.22b). Perfect data-MC agreement results in a flat distribution within the entire range, where the bin-by-bin deviations should be consistent with zero within statistical errors. While the $\mathcal{NN}$ shape of the $B\overline{B}$ background is well described as indicated by the relatively flat distribution (Fig. 4.22b), Fig. 4.22a uncovers a disagreement of the continuum/τ-pair description. Besides this disagreement in shape the already mentioned normalization discrepancy persists in the signal region. As illustrated in Fig. 4.22, the $B\overline{B}$ background simulation agrees with the data on the 5 % level, while the MC continuum expectation strongly differs from Off Peak data. However, events passing the multivariate event selection are required to satisfy $\mathcal{NN} > 0$.

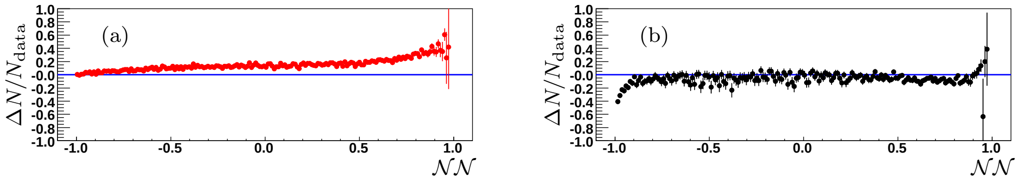


Figure 4.22: Data-MC comparison of $\mathcal{NN}$: (a) Deviations of the scaled numbers of Off Peak events and MC continuum/τ-pairs, and (b) of the scaled numbers of Off Peak subtracted On Peak events and generic $B\overline{B}$ MC.

4.5.3.3 Multivariate a_1 Candidate Selection

Besides the variables of the loose a_1 preselection (Fig. 4.14) a variety of quantities has been investigated for their separation potential and many variables have been found, which potentially could contribute to separate signal from background. However, none of them is dominating the others with respect to its separation power. Thus, a second ANN for multivariate a_1 selection has been constructed. The complete list of the twelve input variables is given in App. B.2.

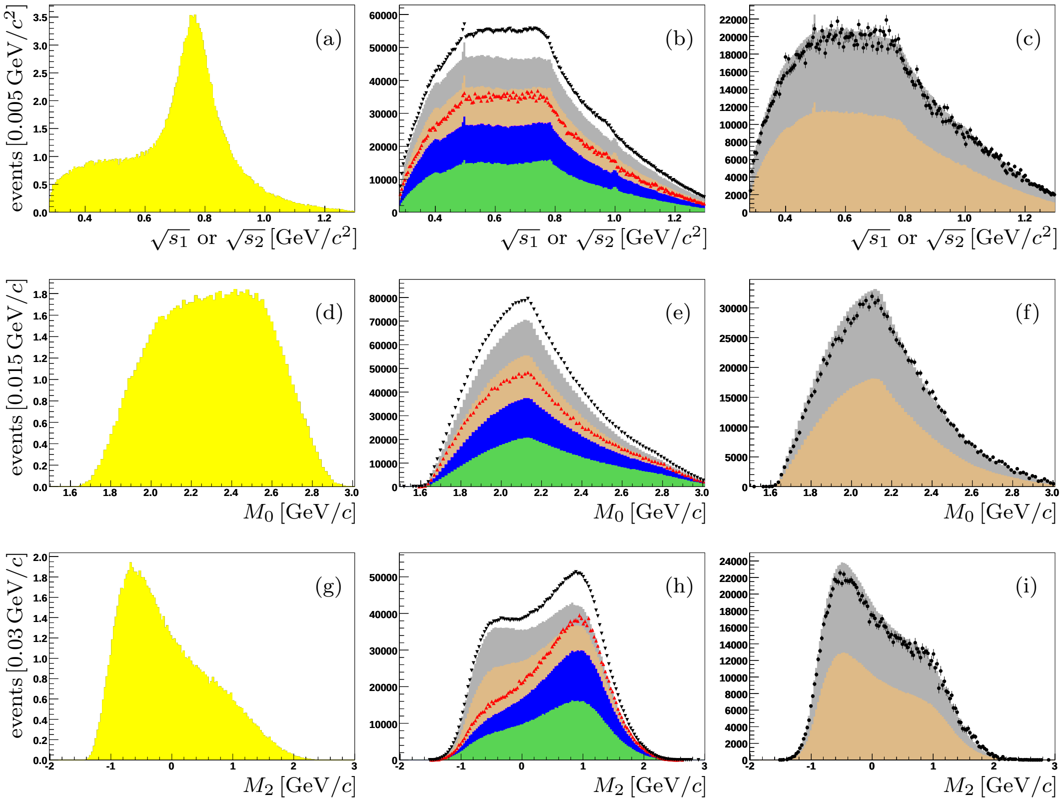


Figure 4.23: (a-c) *Dalitz*-plot projections, (d-f) M_0, and (g-i) M_2 after $\mathcal{NN} > 0$. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Although $m_{3\pi}$, $|\vec{p}_{3\pi}|$, and $\cos\Theta_{3\pi,\text{miss}}$ have already been used for the a_1 preselection, these variables are included in the a_1 candidate based ANN in order to further exploit the characteristic signal shapes. Another strong separation arises from the *Dalitz*-plot since combinatorial background does not exhibit the typical $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ structure. Figs. 4.23a-c illustrate the *Dalitz*-plot projections as derived from correctly reconstructed a_1 in signal MC and combinatorial background, respectively. The same figure shows angular moments M_0 (Figs. 4.23d-f) and M_2

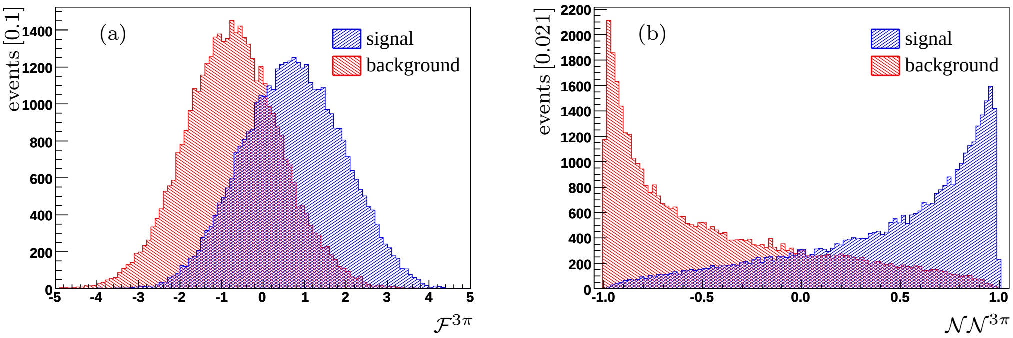


Figure 4.24: Signal-Background separation of a_1 candidate based ANN: (a) *Fisher* discriminant and (b) ANN output determined from the training sample

(Figs. 4.23g-i). These quantities are defined as

$$M_l = \sum_{i=1}^{3} |\vec{p}_{\pi_i}| \times |\cos\theta_{\pi_i}|^l \,, \tag{4.15}$$

where the index i runs over the GTL candidates of the given 3π combination. The angle θ_{π_i} is measured between the i-th track three-momentum vector and the thrust axis (App. B.1, Eq. B.1) calculated from all CT and CN not belonging to the a_1 candidate. Eq. 4.15 implies $M_0 = |\vec{p}_{\pi_1}| + |\vec{p}_{\pi_2}| + |\vec{p}_{\pi_3}|$, while M_2 includes angular correlations between B_{sig} and B_{tag}.

About 38000 correctly reconstructed a_1 and the same number of background candidates represent the training sample (Fig. 4.24). Again, events from different run periods have been mixed to take into account different detector and beam background conditions. The performance test (Fig. 4.25) clearly favours the ANN method resulting in significantly lower background efficiencies. Fig. 4.26 illustrates $\mathcal{NN}^{3\pi}$ for the entire Run12 samples, where the signal MC distribution (Fig. 4.26a) only includes correctly reconstructed a_1 candidates. On the other hand, there is a large fraction of combinatorial background in signal MC events, which need to be decreased. Furthermore, at this stage of the selection the events still include multiple a_1 candidates and one has to choose one of them as the best candidate. Sec. 4.5.3.4 explains how this decision is made and discusses the treatment of combinatorial background in signal MC events.

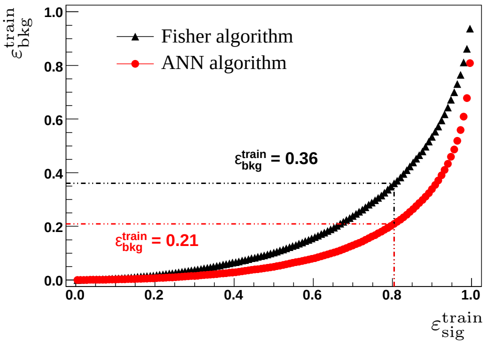


Figure 4.25: Separation Power of $\mathcal{F}^{3\pi}$ and $\mathcal{NN}^{3\pi}$ for multivariate a_1 selection

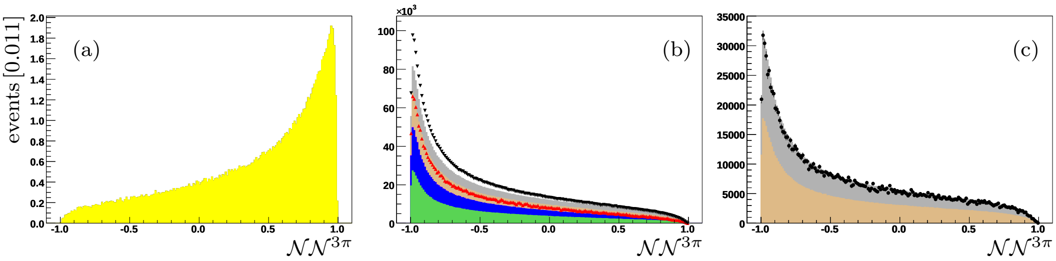


Figure 4.26: Output of a_1 candidate based ANN after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

4.5.3.4 Best a_1 Candidate Selection

Since the B_{tag} is not explicitly reconstructed, tracks originated in its decay chain contribute to wrong a_1 combinations in signal MC passing the preselection. Fig. 4.27a illustrates $\mathcal{NN}^{3\pi}$ for all reconstructed 3π candidates in signal MC. The fraction of combinatorial background at this level of reconstruction is about 79 % and is expected to depend on the B_{tag} decay channel. The higher the number of charged particles generated on the companion B side, the higher the probability to pick up wrong 3π combinations. In order to minimize such a mode dependence, one best a_1 candidate from the pool of preselected candidates per event has been chosen.

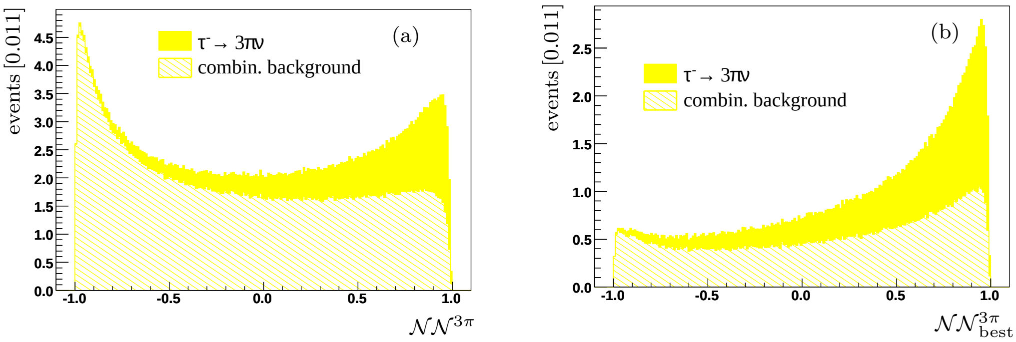


Figure 4.27: Combinatorial background in signal events: (a) shows $\mathcal{NN}^{3\pi}$ of all a_1 candidates passing the preselection. (b) gives the same quantity for best a_1 candidates.

Since $\mathcal{NN}^{3\pi}$ can be interpreted as some sort of a "likelihood" of a given a_1 candidate to be a true one, the candidate with the highest $\mathcal{NN}^{3\pi}$ is selected as the best one (Fig. 4.27b). This simple selection reduces the combinatorial background fraction to about 57 %, while about 78 % of correctly reconstructed 3π candidates passes this selection.

The ANN output of the best candidate per event $\mathcal{NN}^{3\pi}_{\text{best}}$ is shown in Fig. 4.28 and the data and MC distributions are in rather good agreement in shape (Fig. 4.29). In contrast to the event based ANN, this is also true for the continuum description.

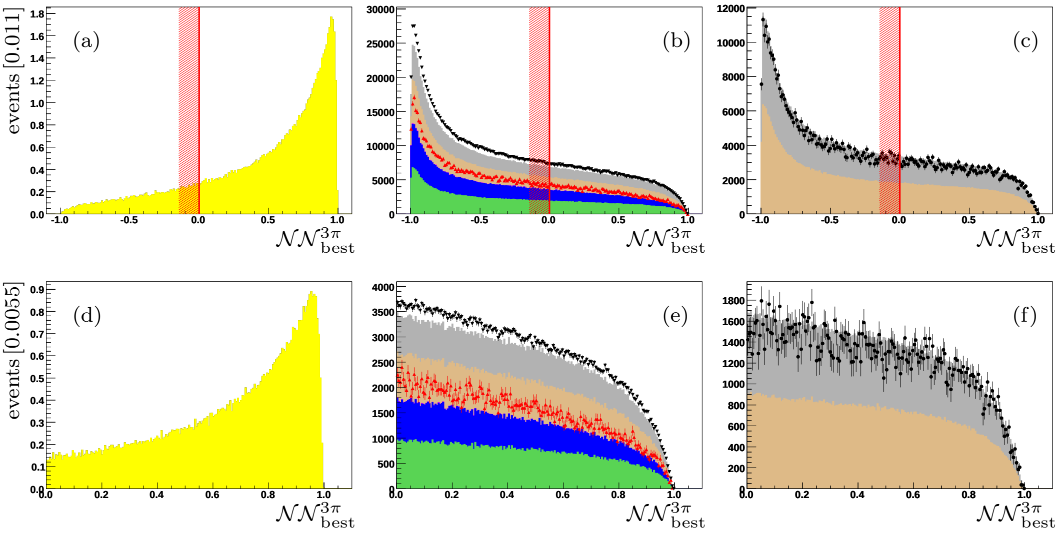


Figure 4.28: ANN output of best a_1 candidates after $\mathcal{NN} > 0$. The red lines illustrate the selection applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Events passing the best a_1 candidate selection are required to contain a best 3π candidate with an ANN output of $\mathcal{NN}^{3\pi}_{\text{best}} > 0$, which further decreases the combinatorial background fraction in signal MC to 49 %.

A cut flow table summarizing the scaled numbers of events passing the step-by-step selection requirements is given in Tab. 4.4. As illustrated in Fig. 4.29, the deviations in normalization of the MC expectations relative to the corresponding data points after the multivariate event selection (Sec. 4.5.3.2) persist after the best a_1 candidate selection. The differences in normalization have been calculated separately for continuum/τ-pair and $\varUpsilon(4S) \to B\overline{B}$ events to be

$$R_{\text{off}} = \frac{N^{\text{sc}}_{\text{off}}}{N^{\text{sc}}_{q\overline{q}} + N^{\text{sc}}_{\tau^+\tau^-}} = 1.189 \pm 0.007\,, \quad R_{\varUpsilon(4S)} = \frac{N^{\text{sc}}_{\text{on}} - N^{\text{sc}}_{\text{off}}}{N^{\text{sc}}_{B^+B^-} + N^{\text{sc}}_{B^0\overline{B}^0}} = 0.949 \pm 0.007\,. \tag{4.16}$$

N^{sc} denotes the scaled number of events passing $\mathcal{NN}^{3\pi}_{\text{best}} > 0$ (Tab. 4.4). These factors need to be taken into account for background estimations from the MC samples.

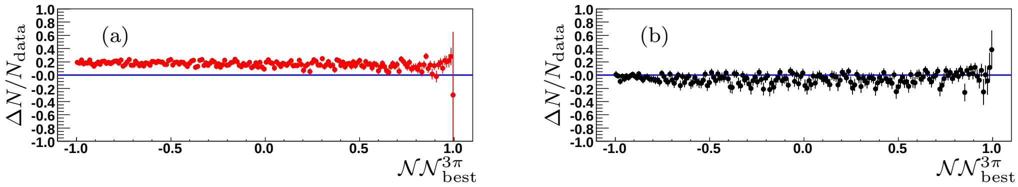


Figure 4.29: Data-MC comparison of $\mathcal{NN}^{3\pi}_{\text{best}}$: Deviations of the scaled numbers of (a) Off Peak events and MC continuum/τ-pairs, and (b) Off Peak subtracted On Peak events and generic $B\overline{B}$ MC.

Selection	**signal**	$\boldsymbol{B^+B^-}$	$\boldsymbol{B^0\overline{B}^0}$	$\boldsymbol{c\overline{c}}$
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	596.2	17838327.6	19303130.5	25748626.9
$N^{3\pi} > 0$	209.4	908657.2	997397.1	1786659.2
$\mathcal{NN} > 0$	177.8	318372.3	378547.9	336431.2
$\mathcal{NN}^{3\pi}_{\mathrm{best}} > 0$	123.3	107454.8	126528.8	96172.6
Selection	$\boldsymbol{u\overline{u}, d\overline{d}, s\overline{s}}$	$\boldsymbol{\tau^+\tau^-}$	**Off Peak**	**On Peak**
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	24994574.8	179729.5	58828347.7	90912971.4
$N^{3\pi} > 0$	1949724.6	6085.3	4068107.1	5745488.4
$\mathcal{NN} > 0$	394852.0	3078.2	889746.2	1551910.9
$\mathcal{NN}^{3\pi}_{\mathrm{best}} > 0$	128627.7	788.9	268223.5	490118.1

Table 4.4: Cut Flow Table: The table summarizes the scaled numbers of events of the Run12 samples passing the given cuts for MC and data, respectively. $N^{3\pi} > 0$ stands for the requirement on events to include at least one reconstructed 3π candidate passing the a_1 preselection. All numbers have been scaled to the Run12 luminosity.

Selection	**signal**	$\boldsymbol{B^+B^-}$	$\boldsymbol{B^0\overline{B}^0}$	$\boldsymbol{c\overline{c}}$
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	72.94	40.57	43.95	24.83
$N^{3\pi} > 0$	25.63 (35.14)	2.08 (5.12)	2.28 (5.19)	1.72 (6.94)
$\mathcal{NN} > 0$	21.77 (84.92)	0.73 (35.00)	0.86 (37.94)	0.32 (18.81)
$\mathcal{NN}^{3\pi}_{\mathrm{best}} > 0$	15.45 (70.96)	0.25 (33.78)	0.29 (33.44)	0.09 (28.62)
Selection	$\boldsymbol{u\overline{u}, d\overline{d}, s\overline{s}}$	$\boldsymbol{\tau^+\tau^-}$	**Off Peak**	**On Peak**
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	14.99	0.24	15.49	19.33
$N^{3\pi} > 0$	1.17 (7.80)	0.01 (3.38)	1.07 (6.92)	1.22 (6.32)
$\mathcal{NN} > 0$	0.24 (20.21)	0.00 (50.49)	0.23 (21.87)	0.33 (27.01)
$\mathcal{NN}^{3\pi}_{\mathrm{best}} > 0$	0.08 (32.57)	0.00 (25.54)	0.07 (30.14)	0.10 (31.58)

Table 4.5: Efficiency Table: The reconstruction efficiencies in % for MC and data, respectively. The numbers in brackets are efficiencies with respect to the previous cut.

Tab. 4.5 lists the cumulative and step-by-step reconstruction efficiencies in order to illustrate the effect of the single selection steps. About 15 % of signal MC pass the entire selection, where about 50 % of these events have not been correctly reconstructed. The requirement of at least one good 3π candidate rejects 2/3 of signal reflecting the detector acceptance and the selection requirement on $|\vec{p}_{3\pi}|$. The background efficiencies are of the order 1 ‰. In order to perform a search on the 10^{-4} level, the signal to background ratio needs to be increased, which can be achieved by an optimization of the $\mathcal{NN}$ and $\mathcal{NN}^{3\pi}_{\text{best}}$ selection requirements with respect to a quality criterion. The choice of such a quantity and the optimization procedure is explained in the next section.

4.5.3.5 Optimization of Multivariate Selection

Any cut optimization relies on a quantity to optimize on. This analysis searches for a rare decay and in principle there are two criteria, which represent different "moods" to optimize. The expected significance

$$S_{\text{exp}} = \frac{N_{\text{sig}}^2}{N_{\text{sig}} + N_{\text{bkg}}} \tag{4.17}$$

can be easily calculated, where N_{sig} and N_{bkg} denote the expected numbers of signal and background events after application of a set of cuts. Since N_{sig} enters in higher order than N_{bkg}, a maximization of this criterion corresponds to "discovery mood".

In contrast, if one wants to exclude the existence of the searched process in a given sample ("exclusion mood"), the expected upper limit UL_{exp} with $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) < \text{UL}_{\text{exp}}$ should be minimized. The extraction of an upper limit UL is more complicated than the calculation of S_{exp} and there is no commonly used procedure for calculation. The exact definition of the term "Upper Limit" used in this thesis and its determination procedure is discussed in the results section (Sec. 5.2). At this point of the analysis this quantity has only been used as a quality criterion for optimization. Therefore, it should only be noted that UL depends on the reconstruction efficiency of signal events ε_{sig} and the already declared N_{bkg} in a given sample. UL at a given confidence level (C.L.) is calculated with respect to a number of observed events N_{obs}.

The subscript "exp" is introduced to clarify that S_{exp} and UL_{exp} are determined from the expected number of signal events N_{sig} corresponding to a branching fraction of $\mathcal{B}_{\text{exp}}(B^- \to \tau^-\overline{\nu}_\tau) = 10^{-4}$ and N_{obs} is set to $N_{\text{obs}} = N_{\text{sig}} + N_{\text{bkg}}$.

In order to find an optimal combination of cuts on both ANN outputs, the cut values $\mathcal{NN}_{\text{cut}}$ and $\mathcal{NN}^{3\pi}_{\text{cut}}$ have been varied between zero and one. For every cut combination the expected significance and upper limit have been calculated, where only events with a correctly reconstructed best a_1 candidate are included in ε_{sig} and N_{sig}. Otherwise, the procedure would optimize for combinatorial background candidates and the aforementioned B_{tag} decay channel dependence. Furthermore, the normalization differences R_{off} and $R_{\Upsilon(4S)}$ (Eq. 4.16) have been taken into account for the N_{bkg} calculation.

The resulting S^{12}_{exp} and $\text{UL}^{12}_{\text{exp}}$ values for different cut combinations are illustrated in Fig. 4.30, where the superscript "12" denotes the Run12 sample N_{sig} and N_{bkg} correspond to. The optimal cut combinations have been found to be $\mathcal{NN} > 0.25$, $\mathcal{NN}^{3\pi}_{\text{best}} > 0.8$ for S^{12}_{exp} and $\mathcal{NN} > 0.3$, $\mathcal{NN}^{3\pi}_{\text{best}} > 0.8$ for $\text{UL}^{12}_{\text{exp}}$, respectively.

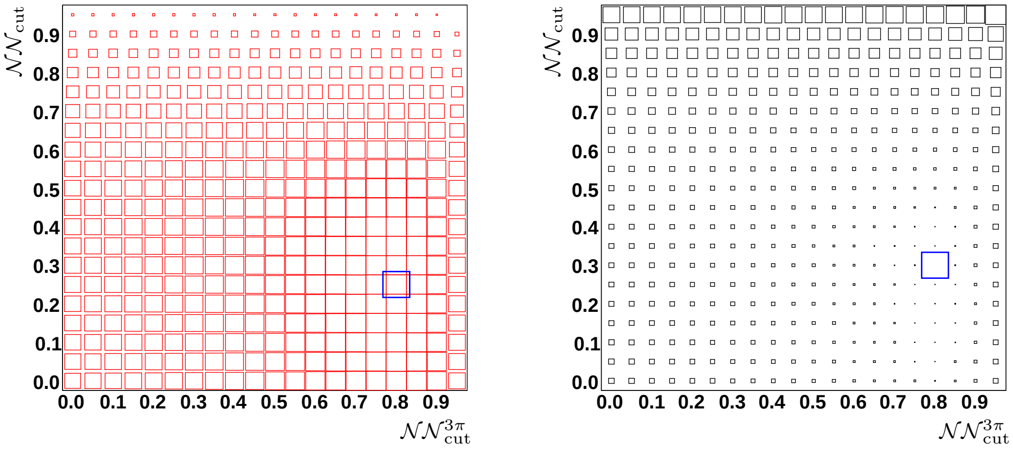


Figure 4.30: Optimization of the multivariate selection: The plot illustrates the development of $S^{12}_{\rm exp}$ (left) and $\mathrm{UL}^{12}_{\rm exp}$ of 90 % C.L. (right) for different combinations of cuts ($\mathcal{NN} > \mathcal{NN}_{\rm cut}$, $\mathcal{NN}^{3\pi}_{\rm best} > \mathcal{NN}^{3\pi}_{\rm cut}$) on the event and candidate based ANN outputs. The optima found for both criteria are indicated by the blue-bordered squares.

After application of the best cut combination the significance takes a value of $S^{12}_{\rm exp} = 0.013$. The upper limit has been calculated to $\mathrm{UL}^{12}_{\rm exp} = 12.4 \times 10^{-4}$ at 90 % C.L.. No systematic uncertainties of $\varepsilon_{\rm sig}$, $N_{\rm sig}$, or $N_{\rm bkg}$ have been taken into account, which are expected to increase $\mathrm{UL}^{12}_{\rm exp}$. However, due to the low significance it has been decided to optimize the cuts by minimizing $\mathrm{UL}^{12}_{\rm exp}$ corresponding to "exclusion mood".

4.5.3.6 Vertex Separation

Besides the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ kinematics the separation of the τ and $B_{\rm tag}$ decay vertices has been introduced as one tool to reject backgrounds from various sources. Vertexing of a set of charged tracks is performed by least χ^2 fitting algorithms with different constraints and many *BABAR* analyses use such algorithms to determine productions vertices of particles. A detailed description of the *BABAR* vertexing algorithms is given in Ref. [84].

The three GTL of a given 3π candidate have been fitted to a common vertex using the `Cascade` algorithm under geometrical constraint. The operation of `Cascade` is similar to the `GeoKin` algorithm [84], but it shows a better performance. Since the a_1 resonance is very broad, it immediately decays after it has been produced. Hence, the τ decay vertex can be assumed to be identical to the 3π production vertex.

While the signal side vertexing of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay is fairly simple, the $B_{\rm tag}$ vertex reconstruction requires a more complex algorithm since tracks directly originated in the $B_{\rm tag}$ decay are not distinguishable from tracks induced by secondary decays, e.g. $K^0_S \to \pi^+\pi^-$. The `FvtClusterer` algorithm is a vertexer that clusterizes a given set of charged tracks to a best common vertex. Since the $B_{\rm tag}$

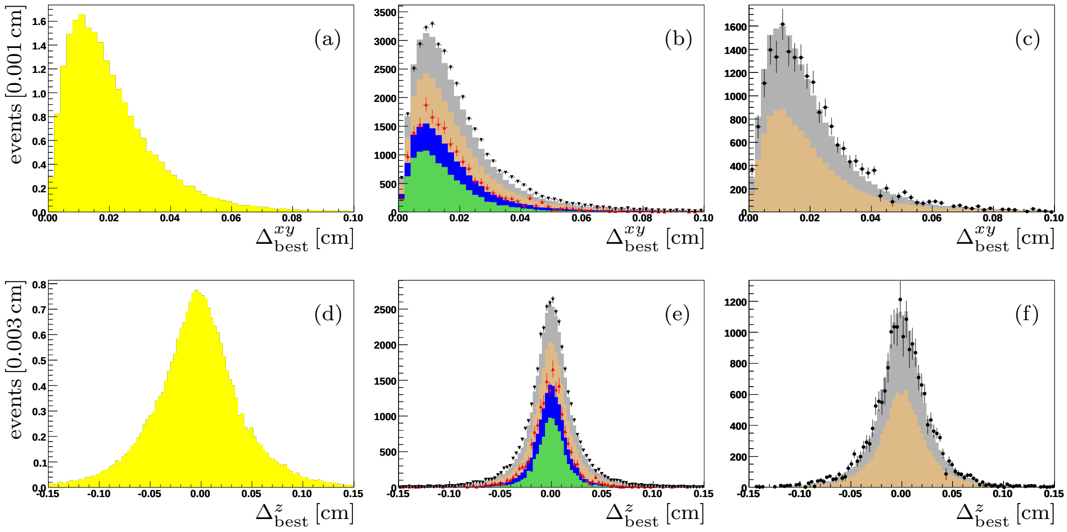


Figure 4.31: Vertex Separation of the best a_1 candidate and $B_{\rm tag}$: Separation in (a-c) xy and (d-f) z after $\mathcal{NN} > 0.3$ and $\mathcal{NN}^{3\pi}_{\rm best} > 0.8$. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

vertex corresponding to a given 3π candidate should be reconstructed, the set of tracks consists of all CT candidates not belonging to the actual 3π candidate. The algorithm starts with a fit of all of these tracks. If the global vertex fit probability $\mathcal{P}(\chi^2, n_{\rm dof})$ is less than a given cutoff value ($\mathcal{P}(\chi^2, n_{\rm dof}) < 0.001$), the track with the highest contribution to the χ^2 is removed from the input set. This procedure is repeated iteratively until the minimal value of $\mathcal{P}(\chi^2, n_{\rm dof})$ allowed has been exceeded. Besides a geometrical constraint the $B_{\rm tag}$ vertex fit is performed using a second constraint. The fitted vertex is forced to be placed in the vicinity of the spot of the primary e^+e^- collision since due to the low B lifetime the $B_{\rm tag}$ decay occurs near the e^+e^- interaction region. Fig. 4.31 illustrates the separation in the transverse plane (Δ^{xy}) and along the beam axis (Δ^z) of the vertices corresponding to best a_1 candidates passing the optimized multivariate selection. In particular, the $\Delta^z_{\rm best}$ distribution of correctly reconstructed a_1 (Fig. 4.31d) is broader compared to the background components (Figs. 4.31e,f), which offers the possibility of background suppression by cutting out the central region around zero.

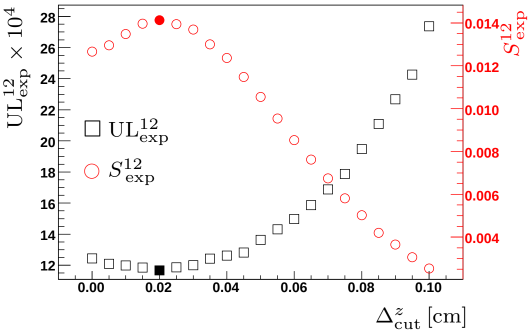


Figure 4.32: Optimization of the vertex separation cut: The plot illustrates $S^{12}_{\rm exp}$ and $\mathrm{UL}^{12}_{\rm exp}$ for different cuts $|\Delta^z_{\rm best}| > \Delta^z_{\rm cut}$. The filled markers indicate the optima found.

Indeed, as visible in Fig. 4.32 a selection of candidates satisfying $|\Delta^z_{\text{best}}| > 0.02\,\text{cm}$ slightly improves $\text{UL}^{12}_{\text{exp}}$. However, it seems that the vertexing does not considerably improve the selection of signal events.

Besides the possibility of vertexing of charged tracks to reconstruct the B_{tag} decay vertex one can exploit the fact that the B_{tag} flavour can be determined from its decay products. Therefore, for a given 3π candidate the remaining CT and CN objects are usable to access information about the B_{tag} quark content and this information might be helpful for signal selection.

4.5.3.7 *B* Flavour Tagging

B flavour tagging is indispensable for measurements of CP asymmetries. Therefore, a standard tagging algorithm has been developed for the CP analyses within BABAR [85, 86]. This algorithm determines the B meson flavour from a given set of reconstructed particles assumed to belong to its decay chain. A detailed explanation of this standard tool is given in Refs. [87, 88].

An example for the correlation of the B quark content to its decay products are semileptonic decays $B \to X\ell\nu_\ell$ characterized by a high-energetic $\ell^\pm$ in the final state. Given a true semileptonic B decay, the $\ell^\pm$ charge is fixed by the B flavour; B mesons including a b quark (B^-, $\overline{B}^0$) produce negatively charged leptons since the b quark transition proceeds via $b \to W^-Q$ with $Q = \{c, u\}$. The negatively charged W boson decays leptonically $W^- \to \ell^-\overline{\nu}_\ell$ and vice versa.

Category $\mathcal{C}$	Name	Definition
63	Lepton	$\lvert\mathcal{P}\rvert > 0.8$ and ($\lvert\mathcal{P}_e\rvert > 0.8$ or $\lvert\mathcal{P}_\mu\rvert > 0.8$)
64	Kaon I	$\lvert\mathcal{P}\rvert > 0.8$ and ($\lvert\mathcal{P}_e\rvert < 0.8$ and $\lvert\mathcal{P}_\mu\rvert < 0.8$)
65	Kaon II	$0.6 < \lvert\mathcal{P}\rvert < 0.8$
66	Kaon-Pion	$0.4 < \lvert\mathcal{P}\rvert < 0.6$
67	Pion	$0.2 < \lvert\mathcal{P}\rvert < 0.4$
68	Other	$0.1 < \lvert\mathcal{P}\rvert < 0.2$
0	Untagged	$\lvert\mathcal{P}\rvert < 0.1$

Table 4.6: Definition of tagging categories returned by tagging algorithm: $\mathcal{P}_e$ and $\mathcal{P}_\mu$ denote the neural net outputs returned by the lepton sub-taggers. Details are given in Refs. [85–88].

Besides $B \to X\ell\nu_\ell$ decays there are more processes, which carry information about the B flavour. The standard tagging algorithm uses various sub-taggers, where every one of them performs tagging using a flavour-sensitive process based on a neural network. The outputs of these sub-taggers are then used as inputs of an overlying neural network trained to separate b and $\overline{b}$. Besides the final output referred as tagging probabilities $\mathcal{P}$ six hierarchical and mutually exclusive tagging categories $\mathcal{C}$ are returned by the tagging algorithm (Tab. 4.6).

Using the aforementioned standard algorithm tagging probabilities $\mathcal{P}_{\text{best}}$ and categories $\mathcal{C}_{\text{best}}$ have been extracted from all CT and CN not belonging to an actual best a_1 candidate (Fig. 4.33). Neither $\mathcal{C}_{\text{best}}$ nor $\mathcal{P}_{\text{best}}$ seem to be useful for signal selection,

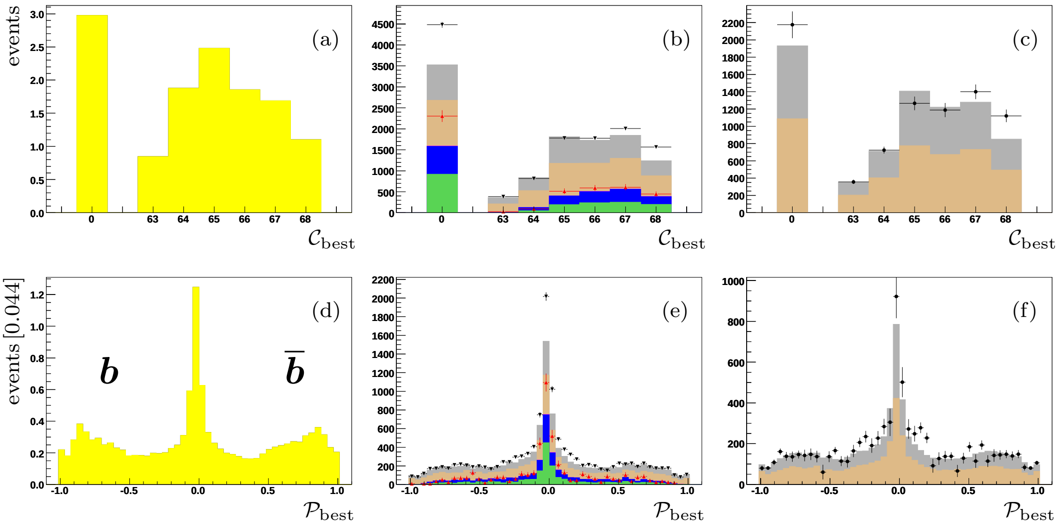


Figure 4.33: (a-c) Tagging categories and (d-f) probabilities of best a_1 candidates after $|\Delta^z_{\rm best}| > 0.02\,$cm. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

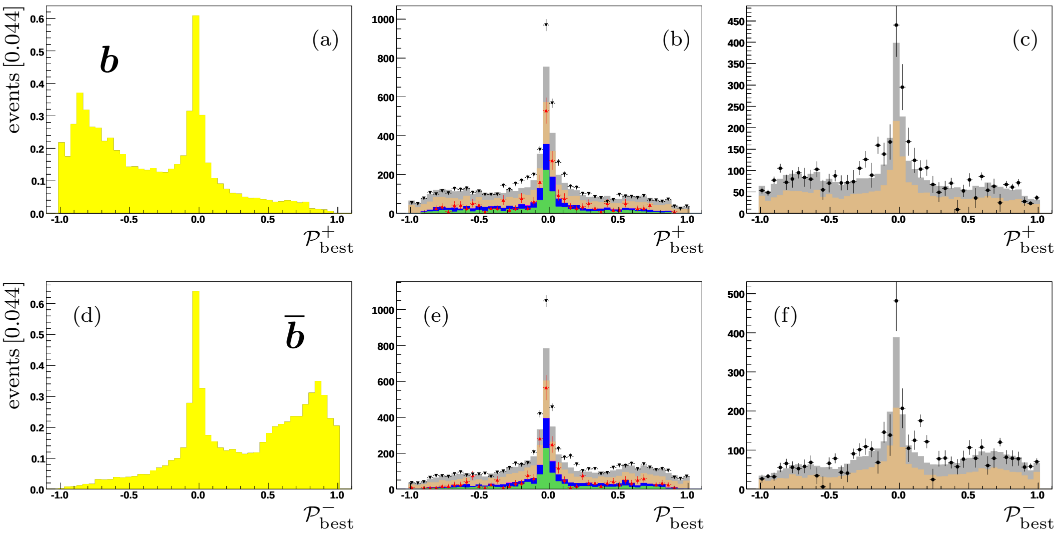


Figure 4.34: Correlation of the tagging probabilities to the charge of the best a_1 candidate. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

but the tagging probability is strongly correlated to the charge of the 3π candidate $c_{3\pi}$ as illustrated in Fig. 4.34. $\mathcal{P}^{\pm}_{\rm best}$ denotes the tagging probability corresponding to positively or negatively charged best a_1 candidates, respectively. Since the $B_{\rm sig}$ charge is conserved in signal events ($B^{\pm} \to \tau^{\pm} \to (3\pi)^{\pm}$), the $B_{\rm tag}$ charge is of opposite sign to $c_{3\pi}$ for correctly reconstructed 3π candidates. B^- mesons contain b quarks and therefore cluster at low $\mathcal{P}^{+}_{\mathrm{best}}$ and vice versa (Figs. 4.34a,d). Of course this behaviour is not visible in background distributions due to the absence of a charge correlation for 3π candidates not originated in signal decays. Hence, cuts on $\mathcal{P}^{\pm}_{\mathrm{best}}$ depending on the a_1 charge can possibly increase S^{12}_{exp} and lower $\mathrm{UL}^{12}_{\mathrm{exp}}$. Fig. 4.35 shows the development of both criteria with cuts on $\mathcal{P}^{\pm}_{\mathrm{best}}$. A cut of $\mathcal{P}^{+}_{\mathrm{best}} < 0.3$ and $\mathcal{P}^{-}_{\mathrm{best}} > -0.3$ results in the best expected upper limit.

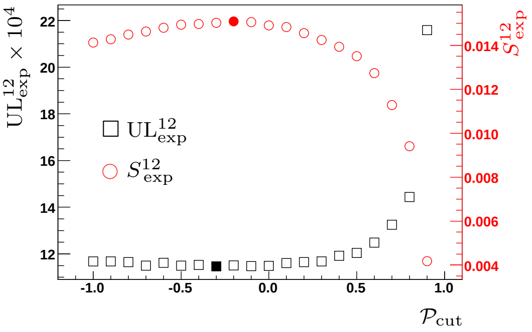


Figure 4.35: Optimization of $\mathcal{P}_{\mathrm{best}}$ selection: The plot illustrates S^{12}_{exp} and $\mathrm{UL}^{12}_{\mathrm{exp}}$ for different cuts ($\mathcal{P}^{+}_{\mathrm{best}} < -\mathcal{P}_{\mathrm{cut}}$, $\mathcal{P}^{-}_{\mathrm{best}} > \mathcal{P}_{\mathrm{cut}}$). The filled markers indicate the optima found for both criteria.

It has been shown that cuts on vertex separations and flavour tagging quantities slightly improve the signal selection. On the other hand, the best expected upper limit at this point is $\mathrm{UL}^{12}_{\mathrm{exp}} = 11.5 \times 10^{-4}$, i.e. these criteria result in a very small improvement compared to $\mathrm{UL}^{12}_{\mathrm{exp}} = 12.4 \times 10^{-4}$ without these selection requirements (Sec. 4.5.3.5). This leads to the question: Is there a better way to exploit vertex and tagging information?

4.5.3.8 Modification of Multivariate a_1 Selection

In order to investigate if vertexing and tagging provide higher separation power than it was visible by directly cutting, the corresponding quantities Δ^{xy} and Δ^{z} (App. B.3) as well as $\mathcal{C}$, $\mathcal{P}$, and the 3π charge $c_{3\pi}$ (App. B.4) have been appended to the list of the kinematical input variables of the a_1 candidate based ANN. The neural network have been retrained using the same training sample and the performances have been compared. The inclusion of vertex and tagging quantities indeed considerably improves the separation power of the ANN as shown in Fig. 4.36. It illustrates a comparison of $\varepsilon^{\mathrm{train}}_{\mathrm{sig}}$ vs. $\varepsilon^{\mathrm{train}}_{\mathrm{bkg}}$ extracted from the 3π candidate training sample for three different input variable sets:

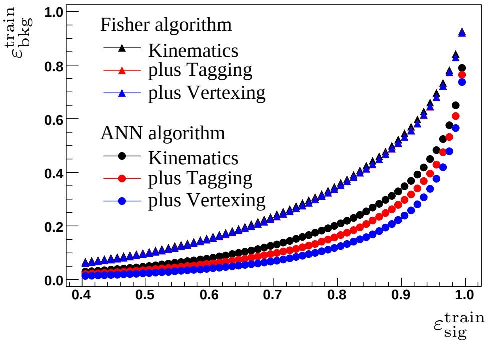


Figure 4.36: Separation Power of *Fisher* discriminant and ANN for different input variable sets

1. $n = 12$: Kinematical properties as discussed before
2. $n = 15$: Kinematics plus $\mathcal{C}$, $\mathcal{P}$, and $c_{3\pi}$ representing B_{tag} flavour tagging
3. $n = 17$: Kinematics plus B_{tag} flavour tagging plus Δ^{xy} and Δ^{z} (vertexing)

While the separation power of the FD is not affected, the ANN algorithm exploits the new variables and vertexing as well as tagging have an impact on the signal and background separation.

Since the additional input quantities modify the output of the a_1 candidate based ANN (Fig. 4.37), the optimization procedure previously described in Sec. 4.5.3.5 has been repeated resulting in new optimal selection requirements ($\mathcal{NN} > 0.25$, $\mathcal{NN}^{3\pi}_{\text{best}} > 0.95$). The expected significance and upper limit after application of this new optimized multivariate selection are

$$S^{12}_{\text{exp}} = 0.031\,, \quad \text{UL}^{12}_{\text{exp}} = 8.4 \times 10^{-4} \quad (90\,\%\,\text{C.L.})\,. \tag{4.18}$$

Therefore, the application of vertexing and flavour tagging quantities as input variables of the a_1 candidate based ANN considerably improved S^{12}_{exp} and $\text{UL}^{12}_{\text{exp}}$ compared to a cut based selection.

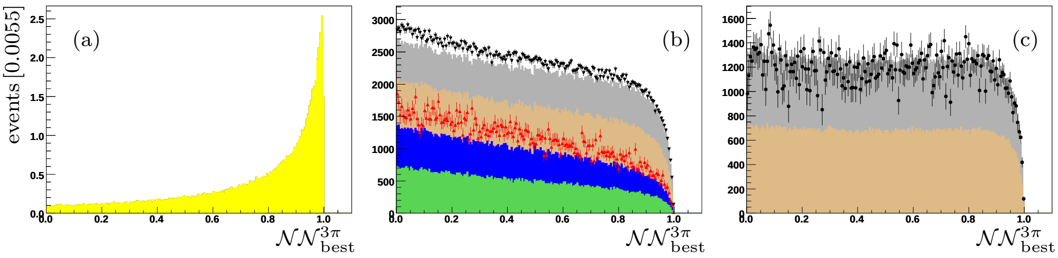


Figure 4.37: Output of modified a_1 candidate based ANN: $\mathcal{NN}^{3\pi}_{\text{best}}$ after $\mathcal{NN} > 0$ in the region $\mathcal{NN}^{3\pi}_{\text{best}} > 0$. The peak at high $\mathcal{NN}^{3\pi}_{\text{best}}$ in (a) is sharpened compared to the corresponding distribution extracted from signal MC with kinematical input variables only (Fig. 4.28d). (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [■], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

4.5.4 Result and Conclusion

The numbers of events passing the entire "inclusive" selection[9] and the reconstruction efficiencies after the aforementioned modification are summarized in Tab. 4.7.

Corresponding to an efficiency of about 2 %, we expect 15.8 reconstructed signal events in the entire Run12 On Peak sample. This number corresponds to events with a correctly reconstructed 3π candidate. The fraction of combinatorial background events at this stage of reconstruction is still about 25 %. As mentioned in Sec. 4.5.3.4 this fraction depends on the decomposition of the selected sample, i.e. B_{tag} decay modes. Since this decomposition is not necessarily well reproduced by the MC description, such events including a wrong best 3π candidate are not accounted for as signal for calculations of S_{exp} and UL_{exp}. Furthermore, the $\mathcal{NN}$ shape of MC continuum/τ-pair events has been found to considerably differ from the observation

---

[9] The term "inclusive" clarifies that an inclusive tagging represented by the standard algorithm has been applied to determine the B_{tag} flavour. No explicit reconstruction of the B_{tag} decay products has been performed.

Mode	$\boldsymbol{N_{\mathrm{sel}}}$	$\boldsymbol{N_{\mathrm{sel}}^{\mathrm{sc}}}$	$\varepsilon/\%$
signal	40541.7 ± 201.3	15.8 ± 0.1	1.9361 ± 0.0096
B^+B^-	10260.0 ± 101.3	2571.2 ± 25.4	0.0059 ± 0.0001
$B^0\overline{B}{}^0$	10995.0 ± 104.9	3185.9 ± 30.4	0.0073 ± 0.0001
$c\overline{c}$	1158.0 ± 34.0	986.1 ± 29.0	0.0010 ± 0.0000
$u\overline{u}$, $d\overline{d}$, $s\overline{s}$	1289.0 ± 35.9	1036.6 ± 28.9	0.0006 ± 0.0000
$\tau^+\tau^-$	22.0 ± 4.7	15.5 ± 3.3	0.0000 ± 0.0000

Table 4.7: Raw (N_{sel}) and luminosity scaled numbers ($N_{\mathrm{sel}}^{\mathrm{sc}}$) of selected MC events and reconstruction efficiencies (ε) after $\mathcal{NN} > 0.25$ and $\mathcal{NN}_{\mathrm{best}}^{3\pi} > 0.95$: The signal numbers correspond to events containing a correctly reconstructed best a_1 candidate and include the *Dalitz*-plot weights. All numbers correspond to the Run12 MC samples (Tab. 4.1).

in Off Peak data (Fig. 4.22a). Hence, the expected number of $q\overline{q}/\tau^+\tau^-$ events has been estimated from the Off Peak sample. We find 305.0 ± 17.5 Off Peak events. Taken into account the luminosity ratios $\mathcal{L}_{\mathrm{on}}/\mathcal{L}_{\mathrm{off}}$ for Run1 and Run2 the rescaled continuum/τ-pair expectation for the Run12 sample arises to

$$N_{q\overline{q}}^{\mathrm{sc}} + N_{\tau^+\tau^-}^{\mathrm{sc}} \equiv N_{\mathrm{off}}^{\mathrm{sc}} = 2605.0 \pm 149.2\,. \tag{4.19}$$

This value again strongly differs from the MC expectation of 2038.2 ± 41.1 events (Tab. 4.7) believed to be induced by the aforementioned deviations of the $\mathcal{NN}$ shape as well as the normalization discrepancy. In contrast, both the $\mathcal{NN}$ and $\mathcal{NN}_{\mathrm{best}}^{3\pi}$ shapes for $\Upsilon(4S)$ events are well described by the generic $B\overline{B}$ MC (Figs. 4.22b and 4.29b). The residual normalization difference can be taken into account by the previously defined factor $R_{\Upsilon(4S)}$ determined from the Off Peak subtracted On Peak sample after $\mathcal{NN}_{\mathrm{best}}^{3\pi} > 0$, where the modification of the multivariate a_1 selection changed this factor to $R_{\Upsilon(4S)} = 0.951 \pm 0.007$ compared to the corresponding factor obtained without vertexing or tagging quantities as input variables of the a_1 candidate based ANN (Eq. 4.16).

The total number of expected background events has been calculated using $N_{\mathrm{off}}^{\mathrm{sc}}$ (Eq. 4.19) and $N_{B\overline{B}}^{\mathrm{sc}}$ from Tab. 4.7 to be

$$N_{\mathrm{bkg}}^{\mathrm{sc}} = R_{\Upsilon(4S)} \cdot (N_{B^+B^-}^{\mathrm{sc}} + N_{B^0\overline{B}{}^0}^{\mathrm{sc}}) + N_{\mathrm{off}}^{\mathrm{sc}} = 8080.0 \pm 163.4\,. \tag{4.20}$$

With the expected number of correctly reconstructed signal events $N_{\mathrm{sig}}^{\mathrm{sc}} = 15.8$, the signal efficiency $\varepsilon_{\mathrm{sig}} = 1.94\,\%$ (Tab. 4.7), and the MC expectation (Eq. 4.20) one obtains the expected significance S_{exp}^{12} and upper limit $\mathrm{UL}_{\mathrm{exp}}^{12}$ corresponding to the Run12 sample as already given in the previous section (Eq. 4.18).

Both quantities S_{exp} as well as $\mathrm{UL}_{\mathrm{exp}}$ depend on the luminosity. Hence, $N_{\mathrm{sig}}^{\mathrm{sc}}$ and $N_{\mathrm{bkg}}^{\mathrm{sc}}$ have been extrapolated to the entire Run14 sample assuming the same reconstruction efficiencies as calculated for the Run12 period. The event numbers are multiplied by the luminosity ratio $\mathcal{L}_{\mathrm{on}}^{14}/\mathcal{L}_{\mathrm{on}}^{12} = 2.64$ (Tab 3.1) and one obtains

$$S_{\mathrm{exp}}^{14} = 0.081\,, \qquad \mathrm{UL}_{\mathrm{exp}}^{14} = 5.6 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{4.21}$$

The upper limit does not include any systematic corrections or uncertainties, e.g. the uncertainty of the background expectation. The precision of $N_{\mathrm{bkg}}^{\mathrm{sc}}$ is of substantial meaning and the error on this number has been incorporated as one systematic uncertainty to recalculate $\mathrm{UL}_{\mathrm{exp}}^{14}$ to[10]

$$\mathrm{UL}_{\mathrm{exp}}^{14} = 10.3 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{4.22}$$

The uncertainty of the background expection $N_{\mathrm{bkg}}^{\mathrm{sc}}$ for the recalculation of $\mathrm{UL}_{\mathrm{exp}}^{14}$ has been determined from the errors on $N_{B^+B^-}^{\mathrm{sc}}$, $N_{B^0\bar{B}^0}^{\mathrm{sc}}$, and $N_{\mathrm{off}}^{\mathrm{sc}}$ as expected for the full Run14 sample. The expected upper limit increases by a factor of two, which illustrates the high impact of the background uncertainty.

It is concluded that although the inclusive reconstruction has uncovered the possibility to use kinematical properties of the reconstructed $a_1 \to 3\pi$ candidate, the a_1 decay vertex separation to the B_{tag} decay vertex, as well as B_{tag} flavour tagging, this technique is not able to sufficiently suppress background. It has to be investigated if other techniques provide equal or even better significances and simultaneously result in a lower background level and therefore in smaller absolute uncertainties of the expected number of background events.

4.6 Semileptonic Reconstruction

Semileptonic decays $B^+ \to X^0\ell^+\nu_\ell$ amount for $(10.9 \pm 0.4)\,\%$ of the total charged B meson decay rate [13][11]. Due to the high B meson mass the $\ell^\pm$ is produced with high momentum depending on the mass of the X^0 system. Therefore, the selection of a high-energetic lepton may be useful for background rejection, e.g. continuum background. In continuum events leptons of compareable energy are mainly generated in semileptonic charmed meson decays (D, D^*), whose production is strongly suppressed compared to light meson states, like K or π. The semileptonic

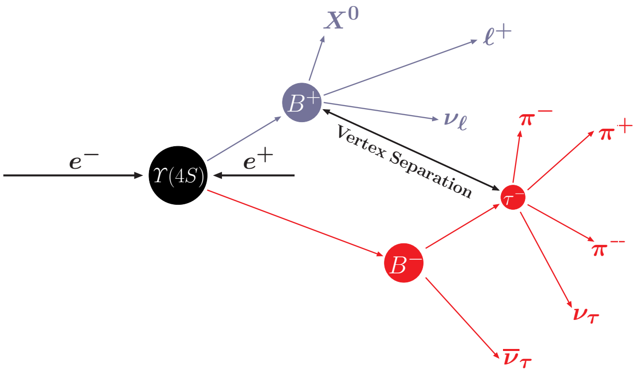


Figure 4.38: Analysis strategy of the Semileptonic Reconstruction: The B_{tag} [🟪] charge is fixed by the charge of the high-energetic lepton and can be used to reject wrong 3π combinations on the B_{sig} side [🟥] by exploiting charge correlation.

---

[10] The incorporation of systematic uncertainties in the upper limit extraction is discussed in Sec. 5.2.

[11] The charge-conjugated process is included.

selection strategy is sketched in Fig. 4.38. As already stated during the description of the inclusive tagging algorithm (Sec. 4.5.3.7), the B_{tag} flavour is determined by the lepton charge c_ℓ. Thus, true a_1 are forced to be of opposite charge to c_ℓ, which further rejects wrong 3π combinations. In analogy to the inclusive technique, the B_{tag} decay vertex can be determined from the remaining tracks represented by the $X^0\ell^+$ system.

The semileptonic reconstruction has been performed on the Run12 samples.[12]

4.6.1 Event and a_1 Candidate Preselection

Basically, the event preselection is similar to the method described for the inclusive reconstruction in Sec. 4.5.1 apart from two additional requirements.

At first, events are required to include at least one identified high-energetic lepton. Fig. 4.39 illustrates the momentum spectra of the highest-energetic identified $\ell^\pm$ per event determined for $e^\pm$ and $\mu^\pm$, respectively. In general, leptons from $B \to X\ell\nu$ decays are harder than $\ell^\pm$ from other sources, e.g. secondary charmed meson decays as visualized for signal (Figs. 4.39a,d) and $\Upsilon(4S)$ background events (Figs. 4.39c,f). Since we search for events of type $B^+ \to X^0\ell^+\nu_\ell$ vs. $B^- \to \tau^-\overline{\nu}_\tau$, an identified lepton with magnitude of its three-momentum of $|\vec{p}_\ell| > 1\,\mathrm{GeV}/c$ is required. Simultaneously,

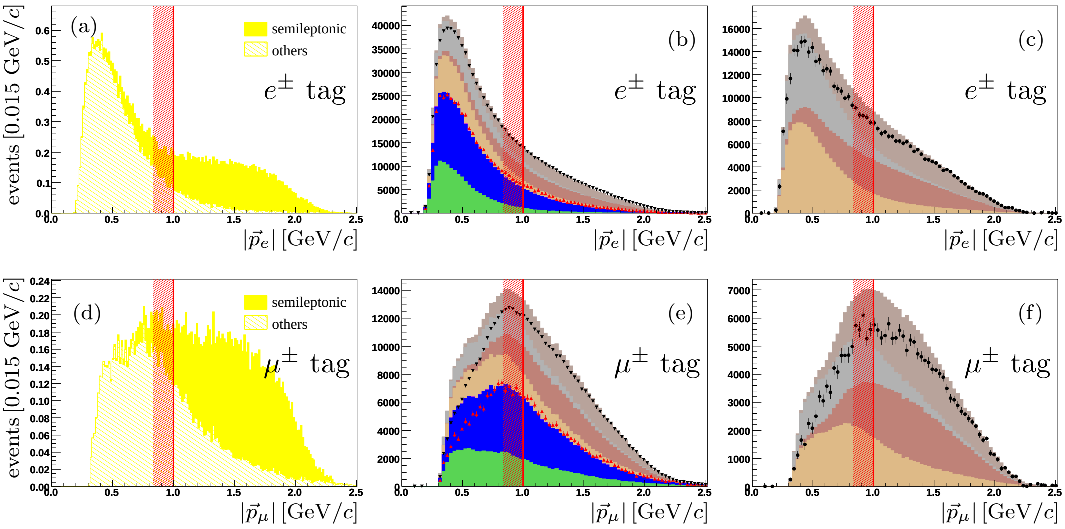


Figure 4.39: Momentum distribution of highest-energetic lepton per event identified as either $e^\pm$ (a-c) or $\mu^\pm$ (d-f). The hatched areas in the signal plots (a,d) illustrate the spectrum for $\ell^\pm$ candidates not produced in semileptonic B decays. The red lines illustrate the cuts applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to X^-\ell^+\nu_\ell$ [🟫], other $B^0\overline{B}^0$ [🟧], $B^+ \to X^0\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

---

[12] As long as the MC and data sets used are not explicitly mentioned in the captions, all plots shown in this section correspond to the entire Run12 samples, where all distributions have been rescaled to the On Peak Run12 luminosity.

events including a second lepton above the given momentum cut are rejected in order to suppress $\Upsilon(4S)$ events with two semileptonic B decays. In fact, the momentum spectra are not perfectly modeled. There are high discrepancies in the low $|\vec{p}_\mu|$ region in the continuum description. Moreover, the spectra found for Off Peak subtracted On Peak events are harder compared to the corresponding MC prediction. This introduces differences of the reconstruction efficiencies in MC and data.

However, once an event has been selected 3π candidates are required to satisfy the same criteria as discussed in Sec. 4.5.2. The aforementioned charge correlation of the high-energetic lepton and true $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ candidates is usable for background suppression as illustrated in Fig. 4.40. In most cases correctly reconstructed a_1 candidates are oppositely charged to the highest-energetic $\ell^\pm$ candidate (Fig. 4.40a). Due to mis-ID and the fact that a portion of lepton candidates are not produced in semileptonic B_{tag} decays there are some entries at $c_{3\pi} + c_\ell = \pm 2$. On the other hand, the fraction of such candidates is small and thus 3π candidates are required to meet $c_{3\pi} + c_\ell = 0$.

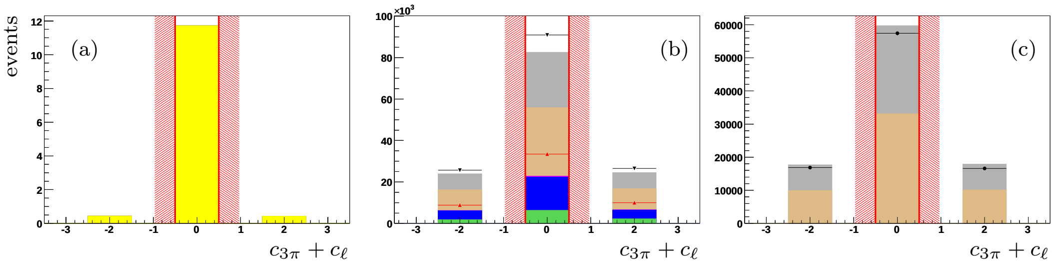


Figure 4.40: The sum of charges of the high-energetic $\ell^\pm$ candidate and 3π candidates after the a_1 candidate preselection. The red lines illustrate the cuts applied. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Again, events are required to contain at least one 3π candidate satisfying the a_1 preselection criteria (Sec. 4.5.2) and the aforementioned requirement on $c_{3\pi}$.

4.6.2 Multivariate Event and a_1 Candidate Selection

Although the inclusive reconstruction has been found to be not sufficient in suppressing background, the concept of multivariate techniques has been found to be useful to efficiently select events of certain type. Thus, the multivariate event and a_1 candidate selections have been adopted.

The list of input variables of the event based ANN has not been modified. Nevertheless, due to the enhancement of semileptonic B decays compared to the inclusive reconstruction the average topology has been changed in the selected sample possibly resulting in different performances of the ANN's. For example, all quantities related to the missing four-momentum p_{miss} "feel" the presence of the neutrino produced in $B \to X\ell\nu_\ell$. The E_{miss} modification of selected signal MC events is illustrated in Fig. 4.41a. The non-detectable ν_ℓ increases the missing energy considerably. The

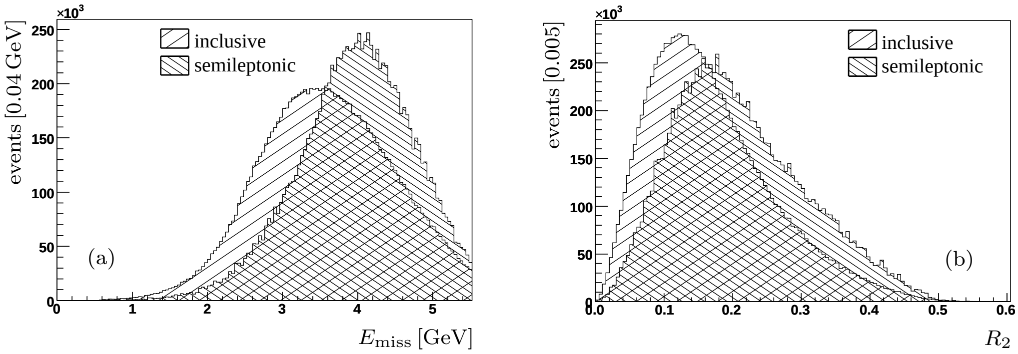


Figure 4.41: Comparison of (a) E_{miss} and (b) R_2 after inclusive and semileptonic event preselection: The distributions have been scaled to the same integral.

change of the event topology is illustrated by the R_2 distributions (Fig. 4.41b). The complete set of the input variables is given in App. C.1 and Fig. 4.42 shows the event based ANN output after training with about 10000 signal MC and the same amount of On Peak events passing the semileptonic event preselection.

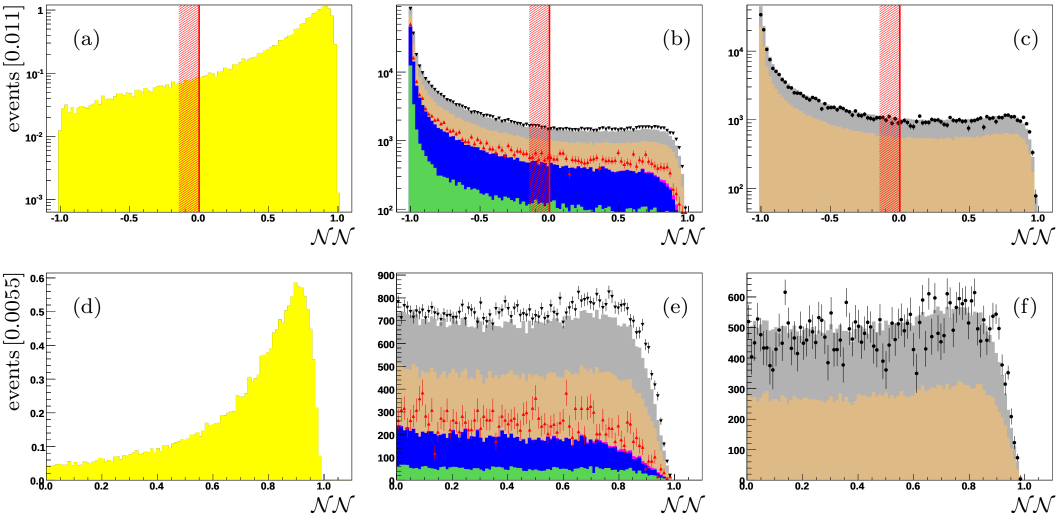


Figure 4.42: (a-c) illustrate $\mathcal{NN}$ after the semileptonic preselection, (d-f) show the $\mathcal{NN}$ distributions in the range above zero. The red lines illustrate the selection applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

The retrained ANN provides better separation compared to the event based ANN of the inclusive reconstruction induced by the E_{miss} modification. The data-MC comparison of $\mathcal{NN}$ (Fig. 4.43) uncovers a similar behaviour. While the $\Upsilon(4S)$ background is well described in shape and normalization, the Off Peak data differ from the MC $q\overline{q}/\tau$-pair expectation. However, similarly to the inclusive technique events with $\mathcal{NN} < 0$ have been rejected before the a_1 candidate based ANN has been retrained.

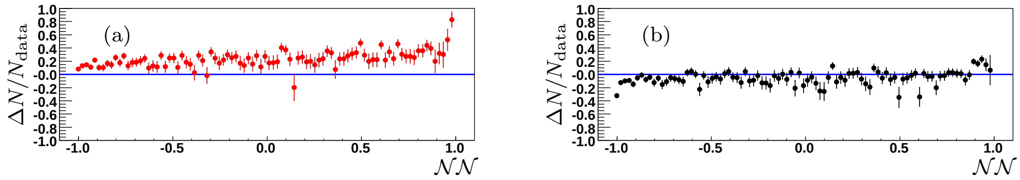


Figure 4.43: Data-MC Comparison of $\mathcal{NN}$: Deviations of the scaled numbers of (a) Off Peak events and MC continuum/τ-pairs, and (b) Off Peak subtracted On Peak events and generic $B\overline{B}$ MC.

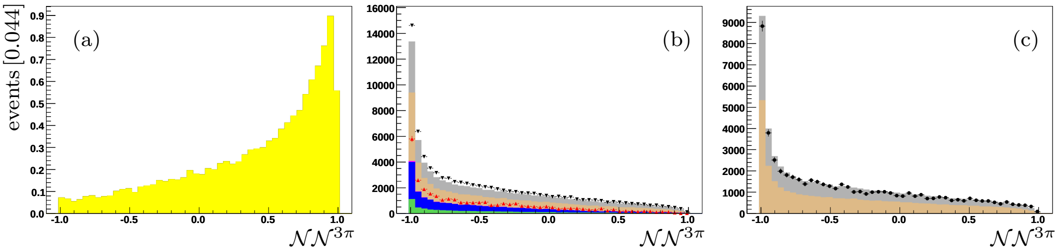


Figure 4.44: Output of the a_1 based ANN (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

In contrast to the inclusive reconstruction, the B flavour tagging is performed by the lepton selection, which obsoletes the inclusive tagging algorithm. Thus, the tagging variables $\mathcal{P}$, $\mathcal{C}$, and $c_{3\pi}$ have been removed from the list of input variables of the a_1 candidate based ANN. Hence, the ANN has been trained using the kinematical variables (App. C.2) plus the vertexing variables Δ^{xy} and Δ^z (App. C.3). The resulting ANN output (Fig. 4.44) is less separating compared to the inclusive reconstruction. Again, this is a consequence of the neutrino produced on the B_{tag} side. The additional contribution to the missing four-momentum spoils the rela-

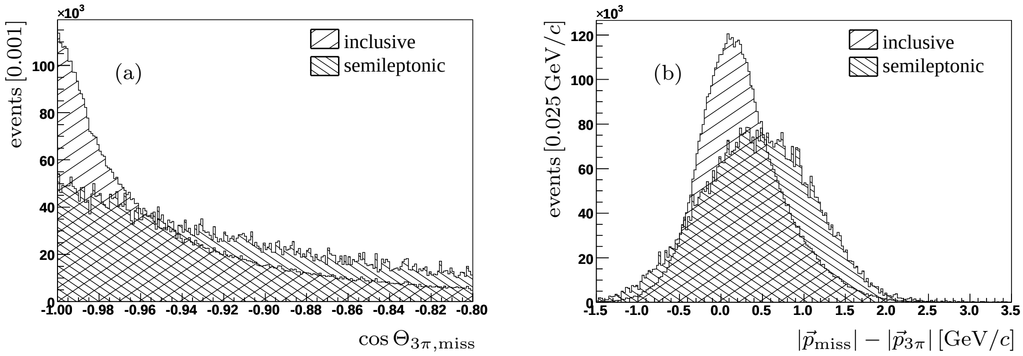


Figure 4.45: Impact of Semileptonic Preselection on (a) $\cos\Theta_{3\pi,\mathrm{miss}}$ and (b) $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$ of correctly reconstructed a_1 candidates: The distributions have been scaled to the same integral.

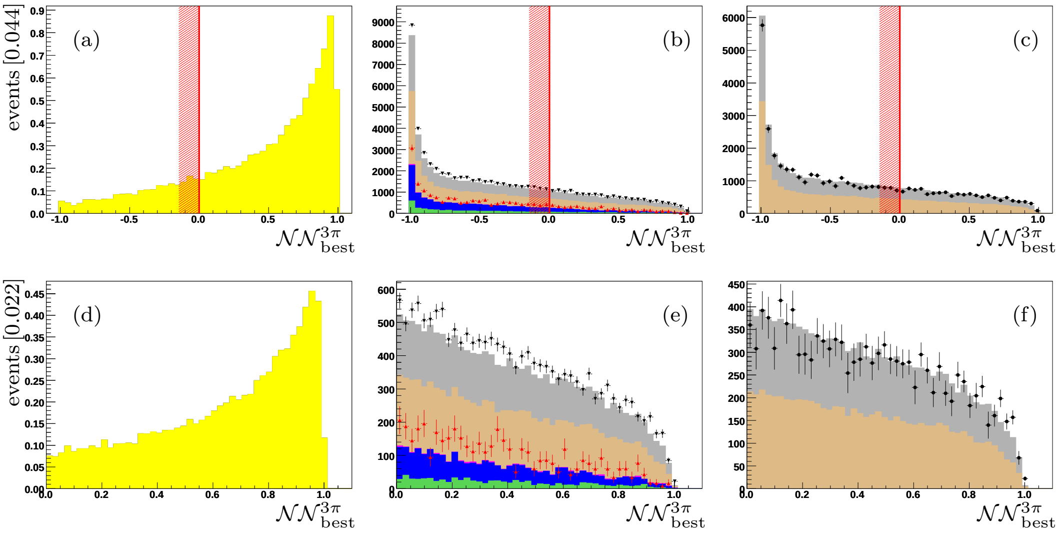


Figure 4.46: ANN output of best a_1 candidates: $\mathcal{NN}^{3\pi}_{\text{best}}$ after $\mathcal{NN} > 0$. The red lines illustrate the selection applied. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

tions between $\vec{p}_{3\pi}$ and $\vec{p}_{\text{miss}}$ (Sec. 4.5.2) since $\vec{p}_{\text{miss}}$ is not longer approximately equal to the sum of the three-momenta of the two signal side neutrinos. The impact of the semileptonic preselection on $\cos\Theta_{3\pi,\text{miss}}$ as well as $|\vec{p}_{\text{miss}}| - |\vec{p}_{3\pi}|$ for correctly reconstructed a_1 candidates in signal MC is illustrated in Fig. 4.45.

However, the data-MC comparison of $\mathcal{NN}^{3\pi}_{\text{best}}$ (Figs. 4.46 and 4.47) verify a good agreement in shape, while the normalization issue is still visible for continuum/τ-pair events. Analogous to the inclusive technique, normalization correction factors have been calculated from the scaled numbers of events satisfying $\mathcal{NN}^{3\pi}_{\text{best}} > 0$ to

$$R_{\text{off}} = \frac{N^{\text{sc}}_{\text{off}}}{N^{\text{sc}}_{q\bar{q}} + N^{\text{sc}}_{\tau^+\tau^-}} = 1.535 \pm 0.072\,, \quad R_{\Upsilon(4S)} = \frac{N^{\text{sc}}_{\text{on}} - N^{\text{sc}}_{\text{off}}}{N^{\text{sc}}_{B^+B^-} + N^{\text{sc}}_{B^0\bar{B}^0}} = 0.974 \pm 0.019\,. \tag{4.23}$$

Again, these factors are applied for the extraction of S^{12}_{exp} as well as $\text{UL}^{12}_{\text{exp}}$ for different cut combinations ($\mathcal{NN}_{\text{cut}}$, $\mathcal{NN}^{3\pi}_{\text{cut}}$) to determine the best selection requirements.

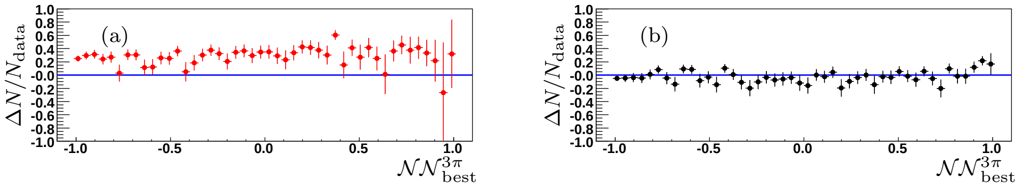


Figure 4.47: Data-MC Comparison of $\mathcal{NN}^{3\pi}_{\text{best}}$: Deviations of the scaled numbers of (a) Off Peak events and MC continuum/τ-pairs, and (b) Off Peak subtracted On Peak events and generic $B\bar{B}$ MC.

4.6.3 Result and Conclusion

The optimization procedure has been accomplished the same way as it has been done within the inclusive reconstruction (Sec. 4.5.3.5). $\mathcal{NN}_{\rm cut}$ and $\mathcal{NN}^{3\pi}_{\rm cut}$ have been varied between zero and one and the resulting scaled numbers of events in signal MC and the background samples have been translated into $S^{12}_{\rm exp}$ and $\mathrm{UL}^{12}_{\rm exp}$, where the correction factors $R_{\rm off}$ and $R_{\Upsilon(4S)}$ from Eq. 4.23 have been taken into account. Both quality criteria suggested a best cut combination of $\mathcal{NN} > 0.75$ and $\mathcal{NN}^{3\pi}_{\rm best} > 0.75$ confirming the improved separation in $\mathcal{NN}$ as well as the worse separation in $\mathcal{NN}^{3\pi}_{\rm best}$ compared to the inclusive reconstruction induced by the additional neutrino.

Mode	$N_{\rm sel}$	$N^{\rm sc}_{\rm sel}$	$\varepsilon/\%$
signal	6078.3 ± 78.0	2.4 ± 0.0	0.2903 ± 0.0037
B^+B^-	765.0 ± 27.7	194.0 ± 7.0	0.0004 ± 0.0000
$B^0\overline{B}{}^0$	914.0 ± 30.2	265.8 ± 8.8	0.0006 ± 0.0000
$c\overline{c}$	31.0 ± 5.6	25.8 ± 4.6	0.0000 ± 0.0000
$u\overline{u}$, $d\overline{d}$, $s\overline{s}$	33.0 ± 5.7	24.8 ± 4.3	0.0000 ± 0.0000
$\tau^+\tau^-$	24.0 ± 4.9	16.0 ± 3.3	0.0000 ± 0.0000

Table 4.8: Raw ($N_{\rm sel}$) and luminosity scaled numbers ($N^{\rm sc}_{\rm sel}$) of selected MC events and reconstruction efficiencies (ε) after $\mathcal{NN} > 0.75$ and $\mathcal{NN}^{3\pi}_{\rm best} > 0.75$: The signal numbers correspond to events containing a correctly reconstructed best a_1 candidate and include the *Dalitz*-plot weights. The numbers of events correspond to the entire Run12 MC samples as summarized in Tab. 4.1.

The entire expected number of background events can be calculated from the scaled numbers of events as summarized in Tab. 4.8 to be

$$N^{\rm sc}_{\rm bkg} = R_{\Upsilon(4S)} \cdot (N^{\rm sc}_{B^+B^-} + N^{\rm sc}_{B^0\overline{B}{}^0}) + R_{\rm off} \cdot (N^{\rm sc}_{q\overline{q}} + N^{\rm sc}_{\tau^+\tau^-}) = 550.1 \pm 18.4\,, \quad (4.24)$$

where in contrast to Sec. 4.5.4 the non-$\Upsilon(4S)$ background expectation is not taken from Off Peak data since only 9 ± 3 Off Peak events passed the entire selection leading to high extrapolation uncertainties. We expect $N^{\rm sc}_{\rm sig} = 2.4$ indicating an improvement of the signal to background ratio. On the other hand, $S^{12}_{\rm exp}$ and $\mathrm{UL}^{12}_{\rm exp}$ have been found to be not competitive to the corresponding values found for the best cut combination in the inclusive reconstruction (Eq. 4.18). We find

$$S^{12}_{\rm exp} = 0.010\,, \quad \mathrm{UL}^{12}_{\rm exp} = 14.0 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.}) \qquad (4.25)$$

and after extrapolation to the Run14 sample using $\mathcal{L}^{14}_{\rm on}/\mathcal{L}^{12}_{\rm on} = 2.64$ (Tab 3.1)

$$S^{14}_{\rm exp} = 0.027\,, \quad \mathrm{UL}^{14}_{\rm exp} = 8.8 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \qquad (4.26)$$

After incorporation of the $N^{\rm sc}_{\rm bkg}$ uncertainty induced by the limited MC statistics and the correction factors $R_{\rm off}$ and $R_{\Upsilon(4S)}$ $\mathrm{UL}^{14}_{\rm exp}$ increases to

$$\mathrm{UL}^{14}_{\rm exp} = 11.2 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \qquad (4.27)$$

This again illustrates the importance to decrease the background level. After inclusion of the $N^{\text{sc}}_{\text{bkg}}$ uncertainty $\text{UL}^{14}_{\text{exp}}$ is compareable to the result found within the inclusive reconstruction (Eq. 4.22), which is a consequence of the lower background level, even though the relative uncertainty of the background estimation $\Delta N^{\text{sc}}_{\text{bkg}}/N^{\text{sc}}_{\text{bkg}} \approx 3.3\,\%$ (Eq. 4.24) increased compared to the inclusive reconstruction.

It is concluded that such a simple semileptonic selection without any requirements on the remaining X^0 system is not sufficient to search for $B^- \to \tau^-\overline{\nu}_\tau$. Instead, an explicit reconstruction of the X^0 system in combination with the gained experiences how to exploit 3π properties might result in an improved selection compared to the inclusive reconstruction. As mentioned in the introduction of this chapter, such an exclusive B_{tag} reconstruction suffers from low reconstruction efficiency. On the other hand, it has also been noted that such a technique results in a cleaner environment since the signal decay can be searched for in the recoil of the reconstructed B_{tag}. Thus, it could improve the significance as well as the understanding of the origin of data-MC discrepancies already visualized in e.g. neutral multiplicity distributions (N^{GPL}) or absolute normalizations.

4.7 Search in the Recoil of $B \to D^{(*)}\ell\nu_\ell$

Two techniques to search for the decay $B^- \to \tau^-\overline{\nu}_\tau$ have been presented in Secs. 4.5 and 4.6. Although the multivariate algorithms have been established as important tools to efficiently select signal events, neither the inclusive nor the semileptonic technique has been found to be sufficient in rejecting background by combining information on properties of the B_{sig} and B_{tag} side.

Thus, it has been decided to apply an exclusive reconstruction of the second B meson with a subsequent search for the signal decay in the recoil of the selected B_{tag} candidate. The choice of the decay channels used for a B_{tag} reconstruction is motivated by two criteria. The decay modes need to cover a high contribution of the total B decay rate, i.e. the decay channels used for B_{tag} reconstruction own high branching fractions. Furthermore, the B_{tag} decay chain has to be easily reconstructable with high efficiency since due to the restriction to exclusive B decay modes the event sample to search for $B^- \to \tau^-\overline{\nu}_\tau$ strongly decreases depending on the total branching fraction covered by the B_{tag} reconstruction channels.

Decay Mode	$\mathcal{B}$/%
$B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$	5.73 ± 0.22
$B^+ \to \overline{D}^{0}\ell^+\nu_\ell$	2.15 ± 0.22
$B^0 \to D^{*-}\ell^+\nu_\ell$	5.35 ± 0.20
$B^0 \to D^{-}\ell^+\nu_\ell$	2.12 ± 0.20
$\overline{D}^{*0} \to \overline{D}^0\pi^0$	61.90 ± 2.90
$\overline{D}^{*0} \to \overline{D}^0\gamma$	38.10 ± 2.90
$D^{*-} \to \overline{D}^0\pi^-$	67.70 ± 0.50
$D^{*-} \to D^-\pi^0$	30.70 ± 0.50
$\overline{D}^0 \to K^+\pi^-$	3.80 ± 0.07
$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	7.72 ± 0.28
$\overline{D}^0 \to K^+\pi^-\pi^0$	14.10 ± 0.50
$\overline{D}^0 \to K^0_S\pi^-\pi^+$	2.90 ± 0.19
$D^- \to K^+\pi^-\pi^-$	9.51 ± 0.12
$D^- \to K^0_S\pi^-$	1.47 ± 0.06
$K^0_S \to \pi^+\pi^-$	69.20 ± 0.05
$\pi^0 \to \gamma\gamma$	98.80 ± 0.03

Table 4.9: B_{tag} reconstruction channels and corresponding branching fractions used in the recoil analysis

Semileptonic decays $B \to X_c \ell \nu$ proceed via a $b \to c$ transition and amount about 10 % of all B decays. The largest contributions are covered by decays into $X_c = D, D^*$ mesons. Tab. 4.9 summarizes possible reconstruction modes of charged and neutral $B \to D^{(*)}\ell\nu_\ell$ decay chains with the corresponding branching fractions. The branching fractions of the listed modes add to about 2.2 % (1.4 %) of all charged (neutral) B decays. These channels have been used for the B_{tag} reconstruction.

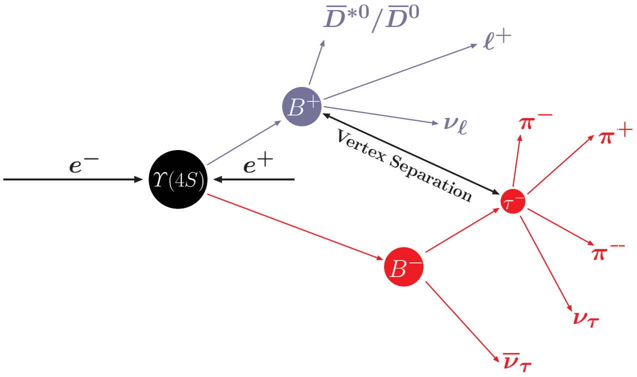


Figure 4.48: Search in the Recoil of $B \to D^{(*)}\ell\nu_\ell$: One B meson [🟪] is exclusively reconstructed in $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ or $B^+ \to \overline{D}^{0}\ell^+\nu_\ell$. All tracks and neutral candidates used for the reconstruction are removed from the entire event. Afterwards, the signal is searched for in the remaining event [🟥]. In contrast to the inclusive and semileptonic selection, the B_{tag} decay vertex can be directly reconstructed from its daughters. Charge correlation of B_{tag} and 3π candidates can be exploited.

The analysis strategy is illustrated in Fig. 4.48. On the first look it seems to be surprising that decays of neutral B mesons are also reconstructed, even though $B^- \to \tau^-\overline{\nu}_\tau$ can only be found in the recoil of charged ones. In fact, the reconstruction of neutral B mesons is of particular importance for background rejection from $\Upsilon(4S) \to B^0\overline{B}^0$ events shown in Sec. 4.7.3.5.

The search in the recoil of $B \to D^{(*)}\ell\nu_\ell$ has been performed on the full Run14 data and MC sets.

4.7.1 Branching Fraction Reweighting

All branching fractions given in Tab. 4.9 are taken from Ref. [13] apart from the value of $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$. The current world average of this branching fraction is totally dominated by a simultaneous measurement of $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ and $B^0 \to D^{*-}\ell^+\nu_\ell$ published by the CLEO collaboration in 2003 [89]

$$\mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) = (6.50 \pm 0.20 \pm 0.43)\,\% \tag{4.28}$$

with a total relative error of 7.3 %. On the other hand, the branching fraction of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ has been measured by various experiments with much smaller uncertainties and Ref. [13] quotes a world average of

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = (5.35 \pm 0.20)\,\% \,. \tag{4.29}$$

These two branching fractions can be related by the ratio of the B mean lifetimes $\tau_{B^+}/\tau_{B^0} = 1.071 \pm 0.009$ [13] assuming isospin symmetry (App. F)

$$\mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) = \frac{\tau_{B^+}}{\tau_{B^0}}\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) \tag{4.30}$$

and one easily obtains

$$\mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) = (5.73 \pm 0.22)\,\% \tag{4.31}$$

with a relative error of about 3.8 %. Although isospin symmetry is broken due to the non-vanishing difference of the u and d quark masses, the influence of this symmetry violation on $B \to D^*\ell\nu_\ell$ decay rates is expected to be negligible. The light spectator quark (u or d) is barely noticed by the heavy b quark decaying into c and can be substituted without changing the decay rate in leading order (Fig. 4.49).

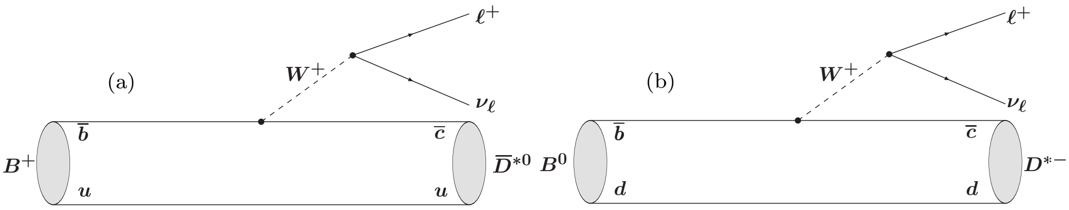


Figure 4.49: *Feynman* diagrams for $B \to D^*\ell\nu_\ell$: (a) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ and (b) $B^0 \to D^{*-}\ell^+\nu_\ell$ are symmetric in substitution of the spectator quarks u and d.

The branching fractions of the given processes assumed in the MC simulation differ from the values of Tab. 4.9. Hence, a dedicated correction factor equal to the ratio of the measured (Tab. 4.9) and the assumed branching fraction has been determined for every mode. All MC events have then been scanned on truth-level for the decay modes listed and an event-by-event weighting factor has been calculated as the product of the correction factors of all aforementioned processes found in a given MC event.

4.7.2 Event Preselection

Contrary to the inclusive and semileptonic reconstruction, where no preselected sample has been used, this analysis is based on the *BToDlnu* skim [90] representing an enriched sample of the abovementioned reconstruction modes. This skim has been explicitly developed to provide a high statistics sample to search for rare decays.

Analogous to the inclusive and semileptonic technique, the selection starts with a loose event preselection using the same quantities as before (Secs. 4.5.1 and 4.6.1):

- $5 \leq N^{\mathrm{GTL}} \leq 10$
- $R_2^{\mathrm{CT}} < 0.5$
- $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$

Moreover, the lepton momentum cut introduced for the semileptonic selection has been applied (Sec. 4.6.1):

- at least one identified lepton with $|\vec{p}_\ell| > 1\,\mathrm{GeV}/c$
- no second identified lepton with $|\vec{p}_\ell| > 1\,\mathrm{GeV}/c$

Again, the second requirement rejects events with two semileptonic B decays. It is shown in Sec. 4.7.4.2 that this requirement needs to be dropped for some efficiency studies.

Since the emphasis of the reconstruction in the recoil technique has been moved from the B_{sig} to the B_{tag} side, events at this stage of the selection are not longer required to include at least one reconstructed 3π candidate. This is one important difference to the two other techniques. The B_{tag} selection is therefore completely decoupled from the signal side selection and the B_{tag} and B_{sig} reconstruction efficiencies can be determined separately. However, the meaning of this remark is clarified in Sec. 4.7.4.2.

An overview of the B_{tag} reconstruction is given in Sec. 4.7.3. All plots and numbers correspond to the entire Run14 sets as long as no explicit data set is mentioned.

4.7.3 B_{tag} Reconstruction

Within the *BToDlnu* skim selection the requirements on the D, D^*, and $D\ell$ candidates[13] have been chosen to be not too restrictive in order to provide high efficiency. Thus, this preselected sample still contains high combinatorial backgrounds from non-semileptonic decays and it is indispensable to reduce the average number of $D\ell$ candidates per event.

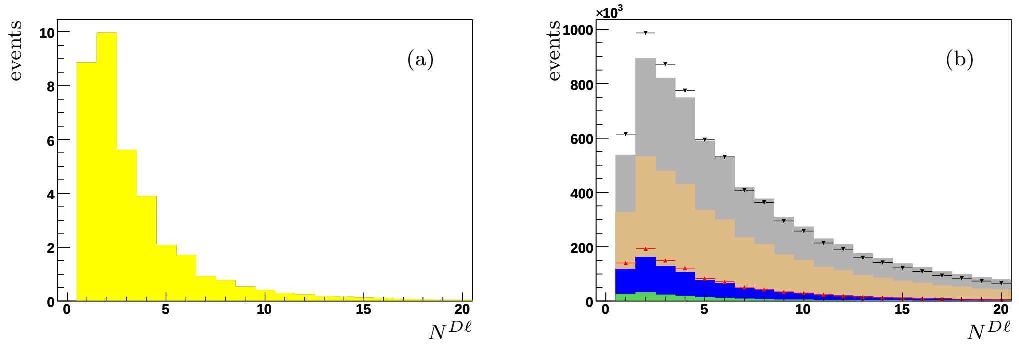


Figure 4.50: Number of reconstructed $D\ell$ candidates per event included in *BToDlnu* skim: (a) Signal MC shows a significantly lower average $D\ell$ candidate multiplicity compared to (b) background events. This is induced by lower track and neutral multiplicities resulting in lower combinatorial background in signal events. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [■], Off Peak [▲], On Peak [▼])

---

[13] Combinations of D or D^* candidates and a lepton candidate are denoted as $D\ell$ combinations.

Fig. 4.50 illustrates the number of reconstructed $D\ell$ candidates $N^{D\ell}$ for signal and background events. The average $D\ell$ multiplicity of the combined background distribution has been found to be $\langle N^{D\ell} \rangle \approx 7$. Thus, selection criteria given in Ref. [90] have been tightened in the actual reconstruction. However, the D, D^*, and $D\ell$ candidate preselections explained in the following still aim at a high B_{tag} reconstruction efficiency. Therefore, only loose cuts are applied and a multivariate selection of $D\ell$ candidates is used to select "good" B_{tag} candidates.

4.7.3.1 *D* Candidate Preselection

As already given in Tab. 4.9, D candidates are reconstructed in six decay channels; four neutral and two charged D decay modes. Some of these modes include π^0 as well as K^0_S. Their definitions are given in Apps. A.4 (π^0_{all}) and A.5 (K^0_S).

For the combination of pions and kaons to D candidates $\pi^\pm$ tracks are required to fulfil the GTVL criteria, whereas the identified $K^\pm$ are taken from the pool of GTL candidates. π^0s needed to reconstruct $\overline{D}^0 \to K^+\pi^-\pi^0$ have been selected as π^0_{all} (App. A.4). $\pi^\pm$ and $K^\pm$ candidates of a given D candidate are fitted to a common vertex using the `Cascade` algorithm under geometrical constraint as introduced in Sec. 4.5.3.6. While $\pi^0 \to \gamma\gamma$ information can not be used in the vertex fit, the $K^0_S \to \pi^+\pi^-$ decay vertex information is taken into account.

The invariant mass distributions of D candidates found to belong to a true $B \to D^{(*)}\ell\nu_\ell$ decay for all six modes as derived from generic B^+B^- and $B^0\overline{B}^0$ MC are summarized in Fig. 4.51. The presence of π^0s in $\overline{D}^0 \to K^+\pi^-\pi^0$ results in a broader invariant D mass distribution compared to the other D modes (Fig. 4.51c).

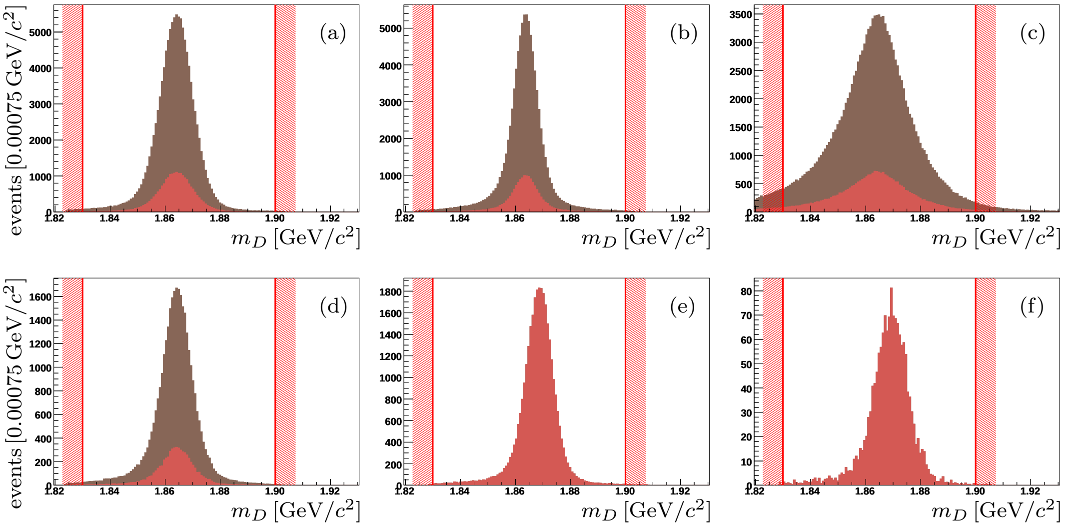


Figure 4.51: Distributions of invariant mass of correctly reconstructed D candidates: (a) $\overline{D}^0 \to K^+\pi^-$, (b) $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, (c) $\overline{D}^0 \to K^+\pi^-\pi^0$, (d) $\overline{D}^0 \to K^0_S\pi^-\pi^+$, (e) $D^- \to K^+\pi^-\pi^-$, and (f) $D^- \to K^0_S\pi^-$. The red lines illustrate the cuts applied. ($B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫])

Therefore, a loose cut of $1.83 < m_D < 1.90\,\mathrm{GeV}/c^2$ has been applied to select D candidates. Although this requirement could be tightened for the non-π^0 D modes, it has been decided to use the same mass cut for all modes. Later on, m_D acts as one of the input variables of the $D\ell$ candidate based ANN and the differences in shape are exploited by the algorithm (Sec. 4.7.3.4). One has to remark that naturally no signal contribution from $\Upsilon(4S) \to B^+B^-$ is visible in the charged D mass distributions (Fig. 4.51e,f) since D^{*0} always decay into D^0 mesons (Tab. 4.9).

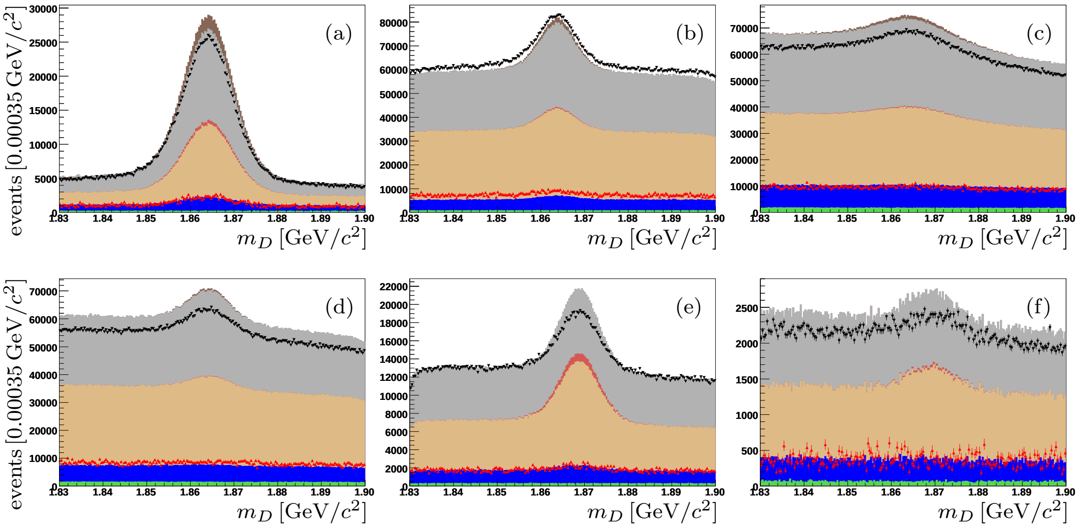


Figure 4.52: Distributions of invariant mass of D candidates: (a) $\overline{D}^0 \to K^+\pi^-$, (b) $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, (c) $\overline{D}^0 \to K^+\pi^-\pi^0$, (d) $\overline{D}^0 \to K^0_S\pi^-\pi^+$, (e) $D^- \to K^+\pi^-\pi^-$, and (f) $D^- \to K^0_S\pi^-$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

A comparison of the combined MC and the corresponding data distributions (Fig. 4.52) uncovers deviations in the absolute normalization. Furthermore, while the peak positions and widths are well-described, the relative peak heights differ between data and MC indicating different D production rates. Fig. 4.52 further clarifies the definition of "correctly reconstructed" candidates. The peaks are represented by true D mesons reconstructed in the given channels, but only a portion of them has been produced in true $B \to D^{(*)}\ell\nu_\ell$ decays.

4.7.3.2 D^* Candidate Preselection

Before the selection criteria are discussed in detail, some explanations on the kinematical properties of the $D^* \to D\pi/\gamma$ decays are given.

Tab. 4.10 lists the D^* reconstruction modes and the corresponding differences of the invariant D^* and D meson masses Δm. The nominal Δm values range between $140\,\mathrm{MeV}/c^2$ and $145\,\mathrm{MeV}/c^2$. Hence, in $D^* \to D\pi$ decays almost the entire D^* energy is converted into the masses of the decay products. With pion masses of $135\,\mathrm{MeV}/c^2$ (π^0) or $140\,\mathrm{MeV}/c^2$ ($\pi^\pm$) the magnitude of the π momentum in the D^*

Decay Mode	m_{D^*}	m_D	$\Delta m = m_{D^*} - m_D$
$\overline{D}^{*0} \to \overline{D}^0\pi^0/\gamma$	2006.7 ± 0.4	1864.5 ± 0.4	142.12 ± 0.07
$D^{*-} \to \overline{D}^0\pi^-$	2010.0 ± 0.4	1864.5 ± 0.4	145.42 ± 0.01
$D^{*-} \to D^-\pi^0$	2010.0 ± 0.4	1869.3 ± 0.4	140.64 ± 0.10

Table 4.10: Nominal $D^* - D$ mass differences [13]: All values are given in MeV/c^2.

rest frame can be calculated to be $|\vec{p}_\pi^{\,D^*}| \approx 45\,\mathrm{MeV}/c$. Due to the relative boost between the D^* rest and the laboratory frame the magnitude of the π three-momentum measured in the laboratory frame $|\vec{p}_\pi^{\,\mathrm{lab}}|$ depends on the flight direction of the π and the D^* momentum. However, for D^* originated in $B \to D^*\ell\nu_\ell$ the pions are produced with low momentum and they are in the following referred to as "soft" π.

In order to reconstruct $D^{*\pm}$ candidates, $\pi^\pm$ taken from the GTVL list have been combined with the selected D^0, where with respect to the aforementioned kinematics the $\pi^\pm$ have to satisfy $|\vec{p}_\pi^{\,\mathrm{lab}}| < 0.45\,\mathrm{GeV}/c$. The same requirement has been used to select π^0 candidates for $D^{*-} \to D^-\pi^0$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$. The exact definition of these "π^0_{soft}" candidates can be found in App. A.4. For the $\overline{D}^{*0} \to \overline{D}^0\gamma$ channel GPL candidates with a raw bump energy of $E > 0.1\,\mathrm{GeV}$ have been used.

The differences of the invariant masses Δm of the reconstructed D^* and corresponding D candidates for all four D^* modes are shown in Fig. 4.53. The variable Δm is very useful to reject combinatorial $D\pi$ backgrounds. Since all uncertainties of the D momentum reconstruction effectively cancel in Δm, the width is given by the soft π momentum resolution. Although the relative momentum resolution of the tracking devices and the EMC increases towards lower momenta, the absolute momentum resolution of the soft $\pi^\pm$ and π^0 are at the level of a few MeV/c since the pions are produced with very low momenta. In contrast, γ from $\overline{D}^{*0} \to \overline{D}^0\pi^0$ are originated with higher momenta resulting in a worse Δm resolution compared to the π channels (Fig. 4.53c).

The red lines in Fig. 4.53 illustrate the cuts applied depending on the reconstruction mode. D^* candidates reconstructed in $D^* \to D\pi$ are required to have a Δm within $\pm 3\,\mathrm{MeV}/c^2$ around the nominal values (Tab. 4.10). For $\overline{D}^{*0} \to \overline{D}^0\gamma$ this criterion is extended to $\pm 20\,\mathrm{MeV}/c^2$. As already stated, $D\ell$ candidates are later on selected by an ANN and in principle Δm could be included as an input variable instead of applying the given selection. On the other hand, this variable can only be calculated for the $B \to D^*\ell\nu_\ell$ modes. Thus, Δm is not qualified to act as input of a multivariate algorithm trained to select candidates of all four $B \to D^{(*)}\ell\nu_\ell$ modes.

4.7.3.3 $D\ell$ Candidate Preselection

The previously described hadron candidates (D, D^*) have been combined with a identified lepton track $(\ell = e^\pm, \mu^\pm)$ to form $D\ell$ candidates with

- $|\vec{p}_\ell| > 1\,\mathrm{GeV}/c$ and
- $0.5 < |\vec{p}_h| < 2.5\,\mathrm{GeV}/c$ with $h = D, D^*$.

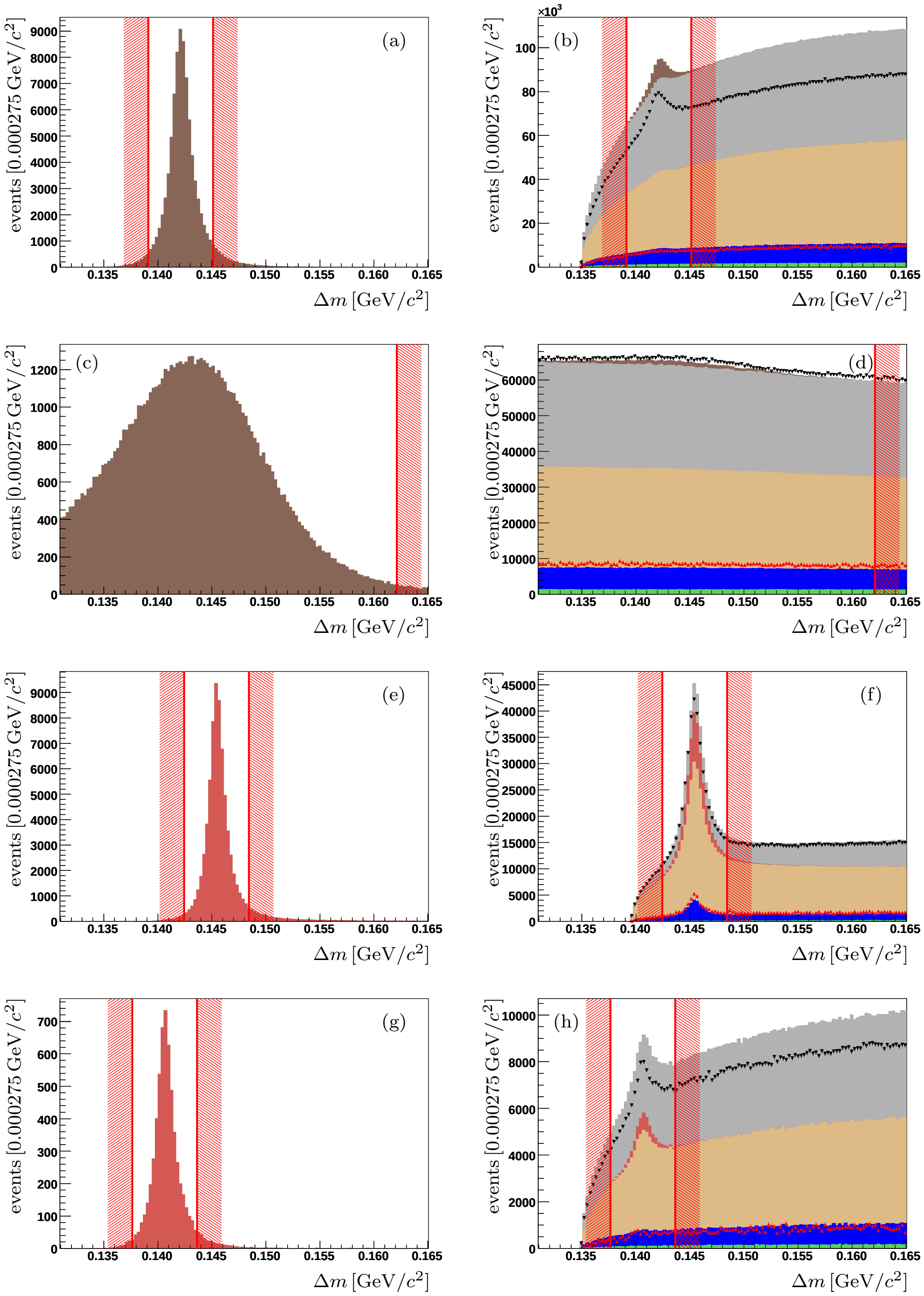


Figure 4.53: Δm distributions of D^* candidates found to be produced in a true $B \to D^*\ell\nu_\ell$ decay (left) and all D^* candidates (right) reconstructed in (a,b) $\overline{D}^{*0} \to \overline{D}^0\pi^0$, (c,d) $\overline{D}^{*0} \to \overline{D}^0\gamma$, (e,f) $D^{*-} \to \overline{D}^0\pi^-$, and (g,h) $D^{*-} \to D^-\pi^0$. The red lines illustrate the cuts applied. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟧], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

It has been shown within the semileptonic reconstruction (Sec. 4.6.1) that leptons originated in semileptonic B decays possess a harder momentum spectrum than leptons from other sources. Thus, a cut at $1\,\mathrm{GeV}/c$ appropriately rejects non-semileptonic processes and enhances $B \to D^{(*)}\ell\nu_\ell$. Exemplarily, the spectra of the magnitude of the hadron three-momentum $\vec{p}_h$ of correctly reconstructed and all $D\ell$ candidates prior the $|\vec{p}_h|$ selection for both $B \to D\ell\nu_\ell$ modes are given in Fig. 4.54. Again, the selection is focused on high B_{tag} reconstruction efficiency. However, $|\vec{p}_h|$ provides high separation potential for the selection of true $B \to D^{(*)}\ell\nu_\ell$ and in fact more variables have been found to be exploitable for separation. These quantities have been combined for an efficient $D\ell$ selection.

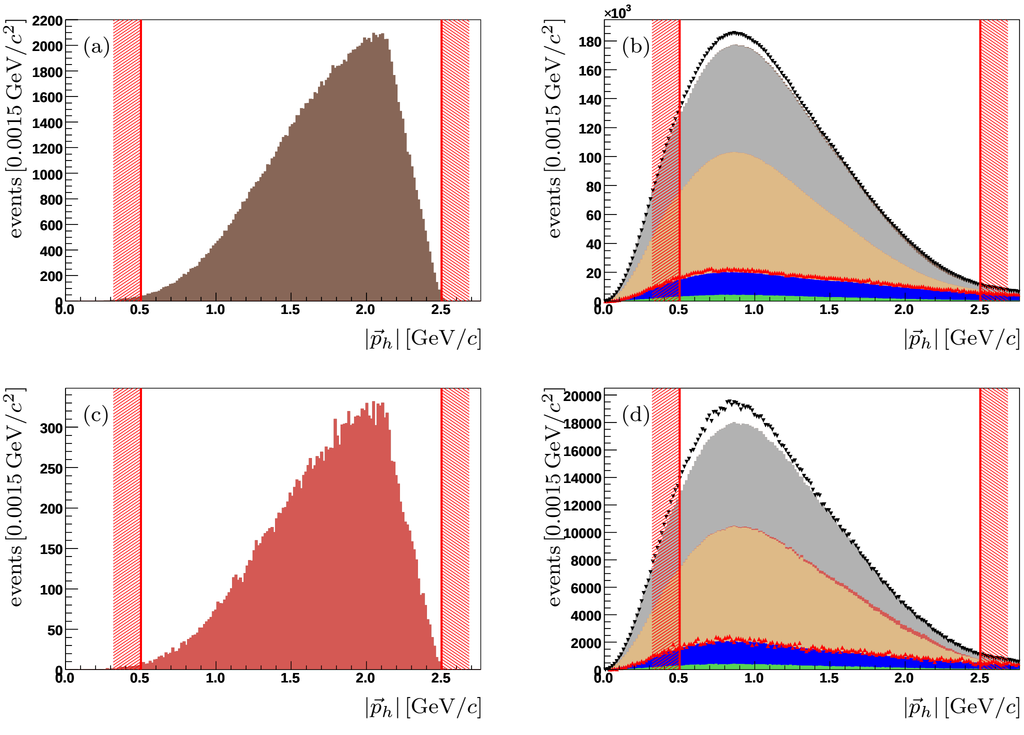


Figure 4.54: Magnitude of hadron momentum of correctly reconstructed $D\ell$ candidates (left) and all $D\ell$ candidates (right) for (a,b) $B^+ \to \overline{D}^0\ell^+\nu_\ell$ and (c,d) $B^0 \to D^-\ell^+\nu_\ell$. The red lines illustrate the cuts applied. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

4.7.3.4 Multivariate $D\ell$ Candidate Selection

Even though the D, D^*, and $D\ell$ preselection cuts have been chosen to retain true $B \to D^{(*)}\ell\nu_\ell$ candidates resulting in high background, the average number of $D\ell$ candidates in background events drops from $\langle N^{D\ell}\rangle \approx 7$ to $\langle N^{D\ell}\rangle \approx 3$ after the $D\ell$ preselection. However, in order to search for $B^- \to \tau^-\overline{\nu}_\tau$ in the recoil of $B \to D^{(*)}\ell\nu_\ell$, one has to choose one particular $D\ell$ candidate from the pool of selected candidates in a given event. A method to select best candidates has already been used for 3π candidates within the inclusive and semileptonic technique (Sec. 4.5.3.4). Here, the decision was based on the output of an ANN. Hence, again an ANN has been constructed to further suppress combinatorial background and to find one best B_{tag} most likely to be a true $B \to D^{(*)}\ell\nu_\ell$. The input variables of this $D\ell$ candidate based ANN are

1. the invariant mass of the reconstructed D candidate m_D,
2. the magnitude of the lepton momentum $|\vec{p}_\ell|$,
3. the magnitude of the hadron momentum $|\vec{p}_h|$,
4. the cosine of the angle between the three-momenta of the hadron and the lepton candidate $\cos\Theta_{h,\ell}$, and
5. the cosine of the angle between the three-momenta of the B meson and the $D\ell$ candidate $\cos\Theta_{B,D\ell}$.

While m_D, $\vec{p}_\ell$, $\vec{p}_h$, as well as $\cos\Theta_{h,\ell}$ have already been defined, the meaning of $\cos\Theta_{B,D\ell}$ demands a more detailed discussion.

Due to the non-reconstructable neutrino the B_{tag} is not fully reconstructed and its four-momentum vector $p_B = (E_B, \vec{p}_B)$ remains unknown. On the other hand, assuming the reconstructed $D\ell$ candidate to be part of a true $B \to D^{(*)}\ell\nu_\ell$ decay and the neutrino to be massless one can calculate $\cos\Theta_{B,D\ell}$ to be (App. G)

$$\cos\Theta_{B,D\ell} = \frac{2E_B E_{D\ell} - m_B^2 - m_{D\ell}^2}{2|\vec{p}_B||\vec{p}_{D\ell}|} \tag{4.32}$$

using the reconstructed four-momentum vector of the $D\ell$ candidate $p_{D\ell} = (E_{D\ell}, \vec{p}_{D\ell})$ and its mass $m_{D\ell} = \sqrt{E_{D\ell}^2 - |\vec{p}_{D\ell}|^2}$. Even though p_B has not been reconstructed, the B_{tag} energy and the magnitude of its three-momentum is fixed in the c.m. frame by the initial e^+e^- state

$$E_B = \frac{\sqrt{s_{\text{on}}}}{2}, \qquad |\vec{p}_B| = \sqrt{E_B^2 - m_B^2} \tag{4.33}$$

with the c.m. energy $\sqrt{s_{\text{on}}}$ and the nominal B meson mass m_B of charged or neutral B mesons depending on the reconstruction mode. To be more concrete, the variable $\cos\Theta_{B,D\ell}$ has the meaning of a hypothesis test. For true $B \to D^{(*)}\ell\nu_\ell$ decays $\cos\Theta_{B,D\ell}$ is equal to the cosine of the angle between the B meson and the $D\ell$ system and the calculated value is expected to be within the mathematical range ±1 apart from detector resolution effects. Instead, combinatorial background candidates are not restricted to the mathematical range since the four-momenta are not balanced in that case ($p_B \neq p_{D\ell} + p_\nu$).

Thus, $\cos\Theta_{B,D\ell}$ provides a high separation potential as it is visualized for the $B \to D^*\ell\nu_\ell$ modes in Fig. 4.55. Furthermore, the $\cos\Theta_{B,D\ell}$ background shape particularly depends on the $D\ell$ reconstruction mode. In fact, such mode-by-mode dependencies of the signal and background shapes are visible in all those input variables. In Sec. 4.7.3.1 the m_D distributions of the different D modes have already

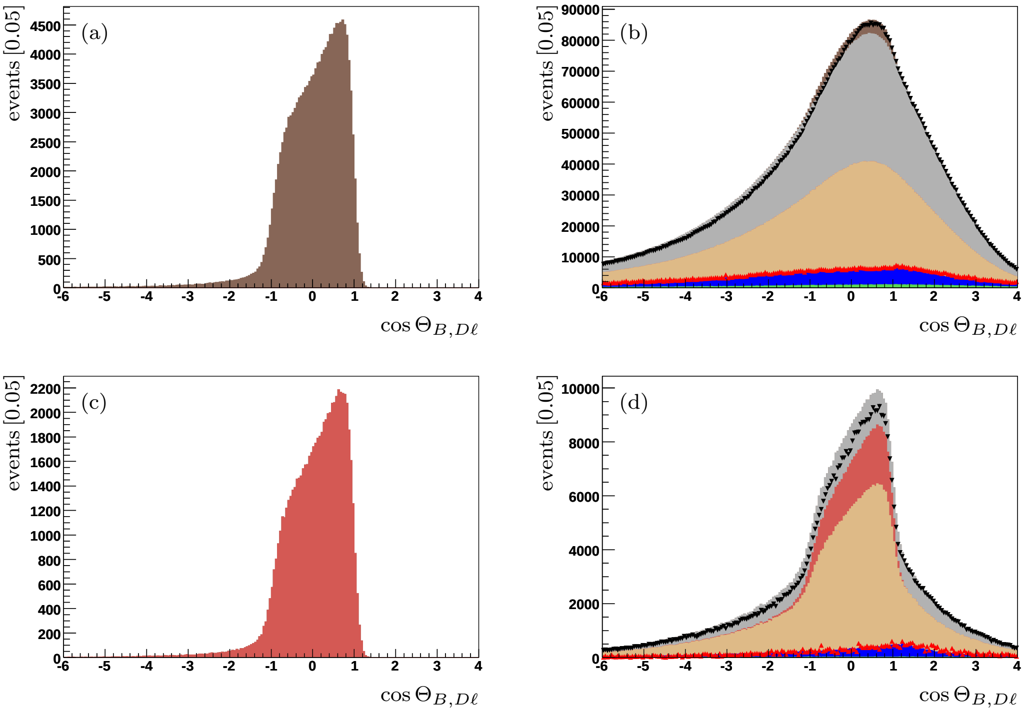


Figure 4.55: $\cos\Theta_{B,D\ell}$ of correctly reconstructed $D\ell$ candidates (left) and all $D\ell$ candidates (right) for (a,b) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, and (c,d) $B^0 \to D^{*-}\ell^+\nu_\ell$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

been shown and the $\overline{D}^0 \to K^+\pi^-\pi^0$ mass distribution has been mentioned to significantly differ from the distributions of the non-π^0 modes. In order to take into account such mode-by-mode dependencies, indices for the D, D^*, and $D\ell$ reconstruction modes have been introduced and included as three additional input variables. The completed list of input variables and the corresponding reconstruction mode-dependent distributions after the $D\ell$ preselection for correctly reconstructed $D\ell$ as well as all combinatorial backgrounds are summarized in App. D.1. These eight variables have been combined using the ANN algorithm already introduced in Sec. 4.5.3.1. In order to account for differences in reconstruction efficiencies and background levels from one run period to another, the ANN has been trained separately for every run period.

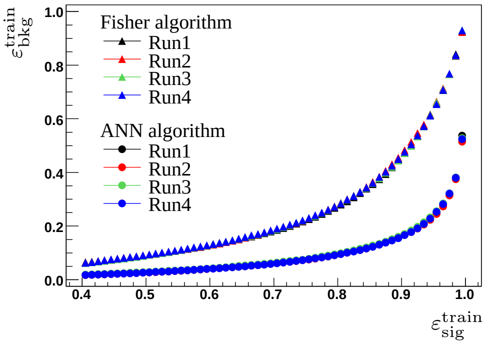


Figure 4.56: Separation Power of the $D\ell$ candidate based ANN and FD shown for the four run periods, respectively.

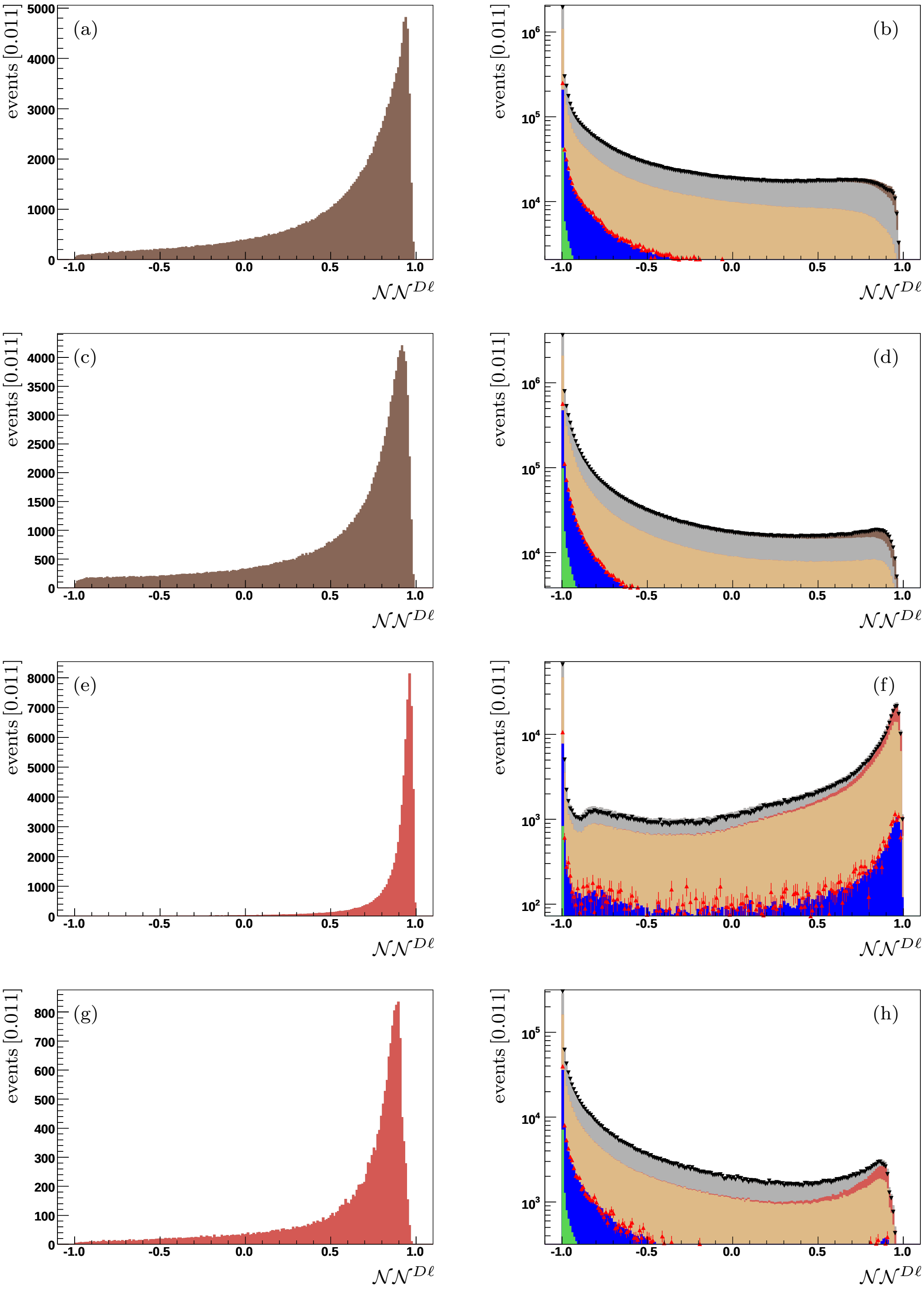


Figure 4.57: ANN output of correctly reconstructed $D\ell$ candidates (left) and all $D\ell$ candidates for (a,b) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (c,d) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (e,f) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (g,h) $B^0 \to D^-\ell^+\nu_\ell$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

Since the main background components arise from combinatorics in $\Upsilon(4S)$ events, the training sample consist of an admixture of generic B^+B^- and $B^0\overline{B}^0$ events. Signal is defined as $D\ell$ candidates found to be produced in true $B \to D^{(*)}\ell\nu_\ell$ decays. All other $D\ell$ combinations are defined as background, even if this sample includes true D or D^* mesons. The training samples consist of about 40000 signal and the same amount of background candidates passing the $D\ell$ preselection for every run period. Thus, the entire training sample includes about 160000 signal and background candidates. This relatively large training sample has been chosen since the inclusion of the reconstruction modes by dedicated indices effectively splits the sample into combinations of these indices. In order to avoid training on statistical fluctuations, the mode combinations have to contribute an adequate number of $D\ell$ candidates. Fig. 4.56 illustrates the evolution of $\varepsilon_{\text{sig}}^{\text{train}}$ vs. $\varepsilon_{\text{bkg}}^{\text{train}}$ with restrictions on the ANN output and the *Fisher* discriminant. Again, the ANN algorithm provides better separation. Furthermore, this figure shows that apart from some small deviations in the region $\varepsilon_{\text{sig}}^{\text{train}} > 0.9$ all four run periods follow the same curve, i.e. the influence of statistical fluctuations on the ANN output is negligible.

The resulting ANN outputs $\mathcal{NN}^{D\ell}$ of the $D\ell$ candidate based network for all four B_{tag} reconstruction modes are shown in Fig. 4.57. Although the distributions of some input variables uncovered deviations between MC and data, e.g. a harder $|\vec{p}_\ell|$ spectrum in data for all four modes (App. D.1), the $\mathcal{NN}^{D\ell}$ shapes agree on the 5 % level. However, there are still differences in the absolute normalization in all channels, which need to be addressed (Sec. 4.7.4.2).

As it has been mentioned at the beginning of this section, at this stage of the B_{tag} selection events possibly include multiple $D\ell$ candidates and one has to decide, which one should be chosen. $\mathcal{NN}^{D\ell}$ acts as the criterion to make this decision.

4.7.3.5 Best $D\ell$ Candidate Selection

If multiple $D\ell$ candidates are present in a given event, the one with the highest $\mathcal{NN}^{D\ell}$ is defined as the best $D\ell$ candidate and all others are rejected. This selection influences the shape of the ANN output as can be seen in Fig. 4.58 since higher $\mathcal{NN}^{D\ell}$ are preferred, but the comparison of MC and data still shows a good agreement in all distributions over the entire range. Fig. 4.58 further indicates that events containing a best $D\ell$ candidate satisfying $\mathcal{NN}^{D\ell}_{\text{best}} > 0$ have been selected. Furthermore, at this analysis stage the meaning of the reconstruction of neutral B meson decays is enlightened.

Since $\Upsilon(4S) \to B^0\overline{B}^0$ events contain true $B^0 \to D^{(*)-}\ell^+\nu_\ell$ decays, one expects a high probability to select neutral $D\ell$ candidates as the best ones. In contrast, in $\Upsilon(4S) \to B^+B^-$ events charged best $D\ell$ combinations should be preferred. Indeed, the $B^0 \to D^{(*)-}\ell^+\nu_\ell$ contributions to the entire number of best $D\ell$ candidates satisfying $\mathcal{NN}^{D\ell}_{\text{best}} > 0$ have been found to strongly differ between generic $B^0\overline{B}^0$ and B^+B^- MC. This behaviour is shown in Fig. 4.59. In generic B^+B^- MC the neutral B_{tag} reconstruction modes amount for about 8.4 % of all reconstructed best $D\ell$ combinations. In contrast, the neutral $D\ell$ contributions in $B^0\overline{B}^0$ events are approximately 37 %.

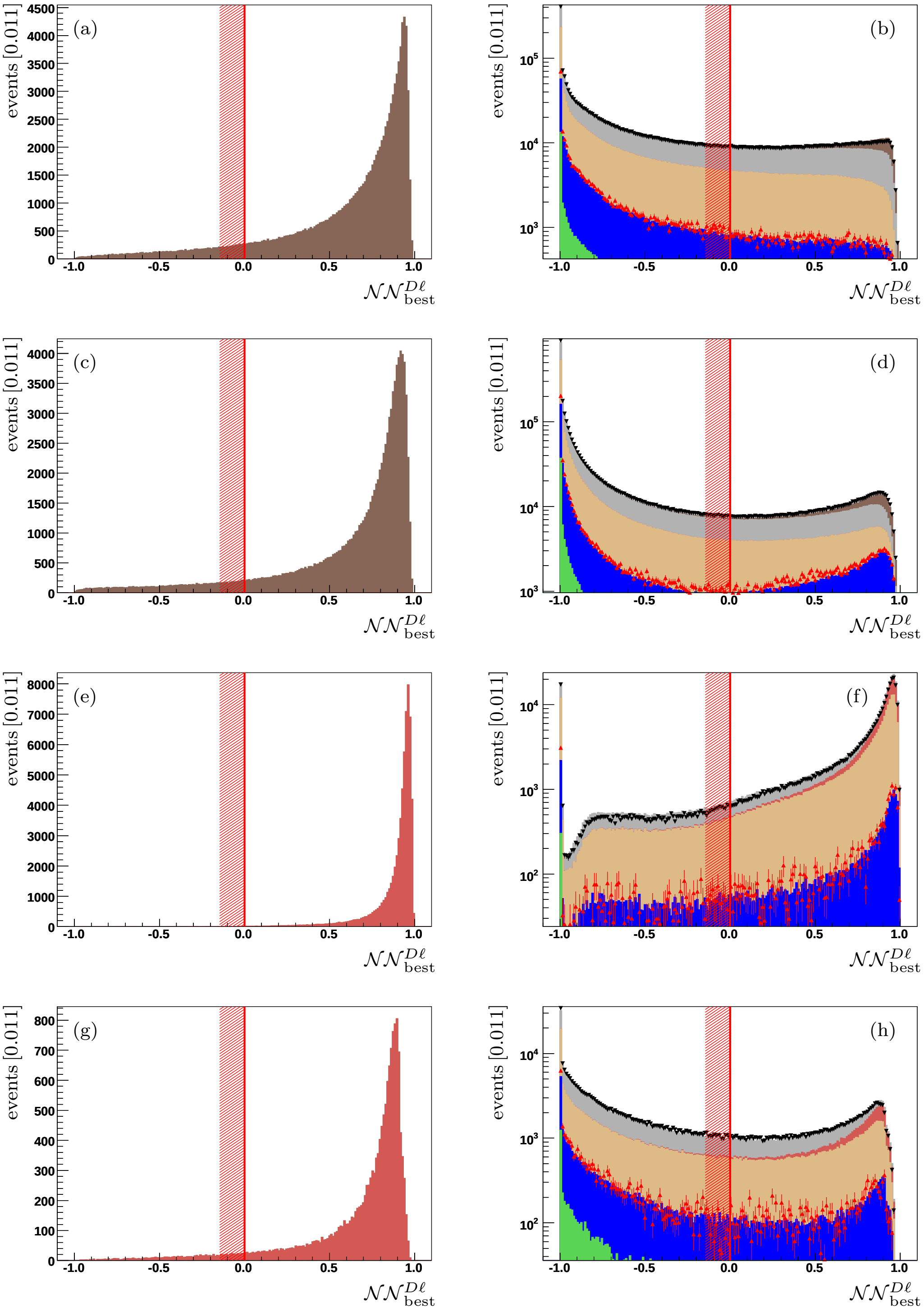


Figure 4.58: ANN output of correctly reconstructed best $D\ell$ candidates (left) and all best $D\ell$ candidates for (a,b) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (c,d) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (e,f) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (g,h) $B^0 \to D^-\ell^+\nu_\ell$. The red lines illustrate the cuts applied. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

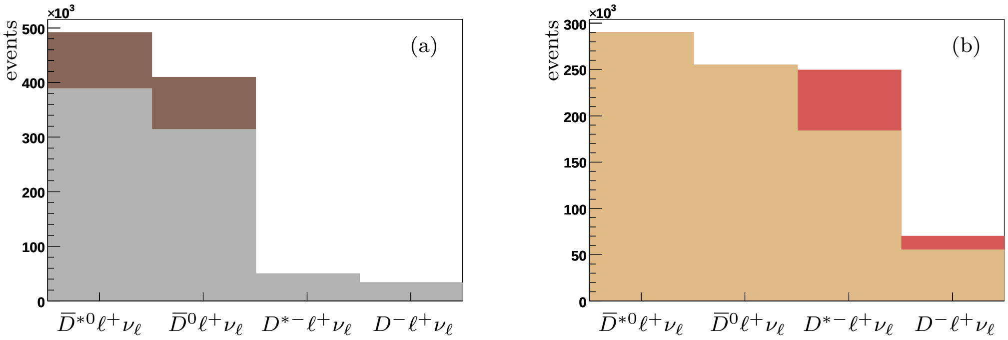


Figure 4.59: Best $D\ell$ candidate reconstruction modes after $\mathcal{NN}^{D\ell}_{\text{best}} > 0$ ($B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜])

In general, due to differences of track and neutral multiplicities the relative contributions of the four B_{tag} reconstruction modes in $B^- \to \tau^-\overline{\nu}_\tau$ events may differ from the distribution extracted from generic B^+B^- MC events (Fig. 4.59a). For example, a lower number of CN candidates results in less π^0_{soft} candidates and affect the D^* occurrence harming the $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ channel. Indeed, Fig. 4.60 illustrates the aforementioned behaviour of signal MC, where the hatched area indicates not correctly reconstructed best $D\ell$ candidates. In contrast to B^+B^- events, the $B^+ \to \overline{D}^0\ell^+\nu_\ell$ mode has been found to be more populated than the $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ mode. This difference is believed to be due to lower π^0_{soft} multiplicities in signal MC, which not only decreases the number of $\overline{D}^{*0}\ell^+$ candidates but simultaneously raises the number of $\overline{D}^0\ell^+$ since $\overline{D}^0$ candidates not combined with a π^0_{soft} may pass the $B^+ \to \overline{D}^0\ell^+\nu_\ell$ selection. However, both $B^0 \to D^{(*)-}\ell^+\nu_\ell$ modes are barely populated. For signal MC the neutral B_{tag} modes amount for only 6.6 % of all best $D\ell$ candidates. Therefore, the neutral reconstruction modes can be used for an efficient veto of $\Upsilon(4S) \to B^0\overline{B}^0$. Thus, in the following events are required to include a best $D\ell$ candidate reconstructed in one of the two $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ modes. Events with a neutral best B_{tag} candidate are explicitly excluded.

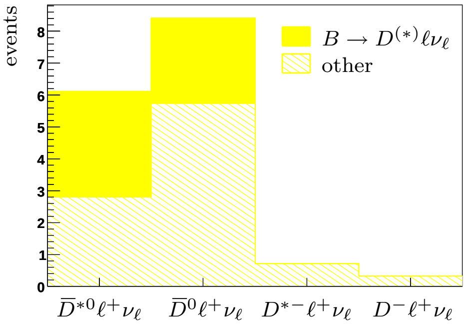


Figure 4.60: Best $D\ell$ candidate reconstruction modes after $\mathcal{NN}^{D\ell}_{\text{best}} > 0$ in $B^- \to \tau^-\overline{\nu}_\tau$ signal MC.

4.7.4 Tagging Efficiency

In this section the B_{tag} yields and reconstruction efficiencies ε_{tag} are given and compared between MC and data. Moreover, a tagging efficiency correction procedure is presented accounting for data-MC normalization discrepancies.

4.7.4.1 Determination of Tagging Efficiency

In Tab. 4.11 the numbers of events passing the selection steps are listed for all MC and data samples, where these event numbers are neither luminosity scaled nor reweighted by the branching fraction ratios (Sec. 4.7.1). The raw B_{tag} reconstruction efficiencies determined from these numbers are summarized in Tab. 4.12, where the cumulative values have been calculated using the initial numbers of events (Tab. 4.1). The *BToDlnu* skim selection selects about 1/4 of all $\Upsilon(4S)$ events, but only 1/8 of signal events. In fact, this factor of two is exactly, what one would expect since the *BToDlnu* sample is still dominated by combinatorial background. Signal MC events contain only three tracks on the B_{sig} side. Hence, the number of selected $D\ell$ candidates is dominated by the generic decay of the second B. Since B^+B^- and $B^0\overline{B}^0$ events include two B mesons both decaying into a generic final state, the probability to find at least one $D\ell$ candidate is expected to be nearly two compared to signal MC.

Selection	**signal**	$\boldsymbol{B^+B^-}$	$\boldsymbol{B^0\overline{B}^0}$	$\boldsymbol{c\overline{c}}$
BToDlnu skim	515160.0	136515495.0	139262490.0	35853490.0
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	413672.0	68063286.0	73900644.0	9034843.0
$\lvert\vec{p}_\ell\rvert > 1\,\mathrm{GeV}/c$	312474.0	54408195.0	58742646.0	5582492.0
$N^{D\ell} > 0$	69748.0	13549947.0	14146437.0	1027613.0
$\mathcal{NN}_{\mathrm{best}}^{D\ell} > 0$	30375.0	4464344.0	4102138.0	260007.0
$\lvert c_{\mathrm{best}}^{D\ell}\rvert = 1$	28368.0	4085948.0	2566045.0	228220.0
Selection	$\boldsymbol{u\overline{u}, d\overline{d}, s\overline{s}}$	$\boldsymbol{\tau^+\tau^-}$	**Off Peak**	**On Peak**
BToDlnu skim	27952833.0	3241917.0	5277843.0	109934676.0
$5 \leq N^{\mathrm{GTL}} \leq 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	3769173.0	90021.0	1077029.0	39213012.0
$\lvert\vec{p}_\ell\rvert > 1\,\mathrm{GeV}/c$	2023445.0	62974.0	664124.0	29602914.0
$N^{D\ell} > 0$	258712.0	1114.0	107030.0	6949221.0
$\mathcal{NN}_{\mathrm{best}}^{D\ell} > 0$	41639.0	154.0	24359.0	2044319.0
$\lvert c_{\mathrm{best}}^{D\ell}\rvert = 1$	38160.0	145.0	21573.0	1644046.0

Table 4.11: Cut Flow Table: The table summarizes the raw (unscaled, unweighted) numbers of events passing the given selection criteria. $N^{D\ell} > 0$ denotes the requirement of at least one reconstructed $D\ell$ candidate passing the preselection. Events passing the entire B_{tag} selection include a charged best $D\ell$ candidate ($\lvert c_{\mathrm{best}}^{D\ell}\rvert = 1$).

After the requirement of a charged best $D\ell$ candidate the B_{tag} reconstruction efficiency in signal MC differ from generic B^+B^- MC, which again is a consequence of different average track and neutral multiplicities. Therefore, the tagging efficiency $\varepsilon_{\mathrm{tag}}$ is derived from signal MC as the number of reconstructed events passing the B_{tag} selection divided by the initial number of signal events. For the entire Run14

Selection	**signal**	$\boldsymbol{B^+B^-}$	$\boldsymbol{B^0\overline{B}^0}$	$\boldsymbol{c\overline{c}}$
BToDlnu skim	12.35	25.55	25.99	9.75
$5 \le N^{\mathrm{GTL}} \le 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	9.91 (80.30)	12.74 (49.86)	13.79 (53.07)	2.46 (25.20)
$\lvert\vec{p}_\ell\rvert > 1\,\mathrm{GeV}/c$	7.49 (75.54)	10.18 (79.94)	10.96 (79.49)	1.52 (61.79)
$N^{D\ell} > 0$	1.67 (22.32)	2.54 (24.90)	2.64 (24.08)	0.28 (18.41)
$\mathcal{NN}_{\mathrm{best}}^{D\ell} > 0$	0.73 (43.55)	0.84 (32.95)	0.77 (29.00)	0.07 (25.30)
$\lvert c_{\mathrm{best}}^{D\ell}\rvert = 1$	0.68 (93.39)	0.76 (91.52)	0.48 (62.55)	0.06 (87.77)
Selection	$\boldsymbol{u\overline{u}, d\overline{d}, s\overline{s}}$	$\boldsymbol{\tau^+\tau^-}$	**Off Peak**	**On Peak**
BToDlnu skim	3.99	0.80	1.84	3.55
$5 \le N^{\mathrm{GTL}} \le 10$ $R_2^{\mathrm{CT}} < 0.5$ $5 < E_{\mathrm{tot}}^{\mathrm{lab}} < 10\,\mathrm{GeV}$	0.54 (13.48)	0.02 (2.78)	0.37 (20.41)	1.27 (35.67)
$\lvert\vec{p}_\ell\rvert > 1\,\mathrm{GeV}/c$	0.29 (53.68)	0.02 (69.95)	0.23 (61.66)	0.96 (75.49)
$N^{D\ell} > 0$	0.04 (12.79)	0.00 (1.77)	0.04 (16.12)	0.22 (23.47)
$\mathcal{NN}_{\mathrm{best}}^{D\ell} > 0$	0.01 (16.09)	0.00 (13.82)	0.01 (22.76)	0.07 (29.42)
$\lvert c_{\mathrm{best}}^{D\ell}\rvert = 1$	0.01 (91.64)	0.00 (94.16)	0.01 (88.56)	0.05 (80.42)

Table 4.12: Tagging Efficiency Table: The raw B_{tag} reconstruction efficiencies in %. The numbers given in brackets are the efficiencies with respect to the previous cut.

sample one obtains

$$\varepsilon_{\mathrm{tag}} = \frac{28368}{4173000} = (0.6800 \pm 0.0040)\,\%\,, \tag{4.34}$$

which significantly differs from the corresponding B^+B^- value of $(0.7648 \pm 0.0004)\,\%$. As already mentioned, $\varepsilon_{\mathrm{tag}}$ is influenced by track and neutral multiplicities. The N^{CT} and N^{CN} distributions have been found to vary comparing the Run13 and Run4 data sets induced by higher peak luminosities during the Run4 data taking period resulting in a higher beam-background level. Thus, $\varepsilon_{\mathrm{tag}}$ has been determined separately for Run13 and Run4 as summarized in Tab. 4.13, whereas the MC events have been reweighted by the branching fraction correction factors. The table indicates higher

Run Period	$\boldsymbol{B^+ \to \overline{D}^{*0}\ell^+\nu_\ell}$	$\boldsymbol{B^+ \to \overline{D}^{0}\ell^+\nu_\ell}$	$\boldsymbol{B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell}$
Run13	0.2849 ± 0.0032	0.3923 ± 0.0037	0.6772 ± 0.0049
Run4	0.2802 ± 0.0046	0.3863 ± 0.0054	0.6665 ± 0.0071
Run14	0.2834 ± 0.0026	0.3904 ± 0.0031	0.6738 ± 0.0040

Table 4.13: The table summarizes the B_{tag} reconstruction efficiencies in %, where the branching fraction weighting factors have been taken into account.

$\varepsilon_{\rm tag}$ for Run13 compared to the Run4 period, but the values are consistent within errors. Thus, the run periods have not been treated separately.

The $B_{\rm tag}$ efficiencies summarized in Tab. 4.13 do not include any systematic corrections. On the other hand, many distributions shown in the last sections, e.g. m_D of $\overline{D}^0 \to K^0_S\pi^-\pi^+$ candidates (Fig. 4.52d) or Δm of $\overline{D}^{*0} \to \overline{D}^0\pi^0$ (Fig. 4.53b), uncovered normalization deviations between MC and data. Such differences influence the tagging efficiency as derived from the MC and require further investigations.

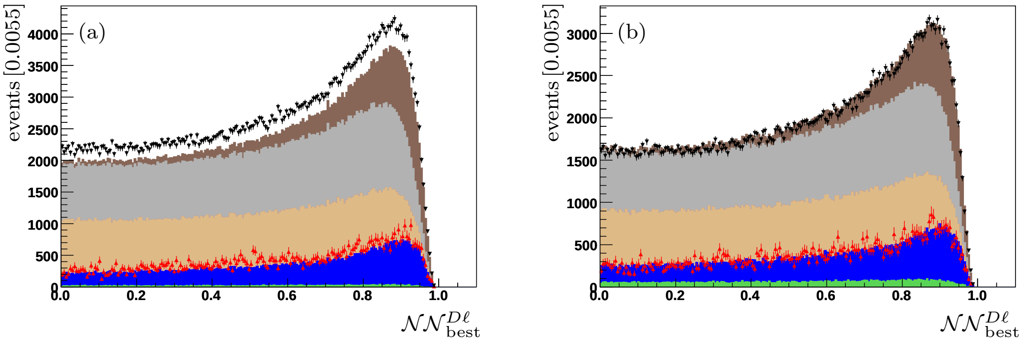


Figure 4.61: ANN output of best $B^+ \to \overline{D}^0\ell^+\nu_\ell$ candidates after $|c^{D\ell}_{\rm best}| = 1$: $\mathcal{NN}^{D\ell}_{\rm best}$ is shown for (a) $B^+ \to \overline{D}^0 e^+\nu_e$ and (b) $B^+ \to \overline{D}^0\mu^+\nu_\mu$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

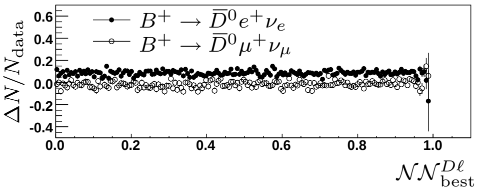


Figure 4.62: Data-MC comparison of $\mathcal{NN}^{D\ell}_{\rm best}$ of $B^+ \to \overline{D}^0\ell^+\nu_\ell$ candidates: Deviations of the scaled numbers of On Peak events and full MC.

The sample has been divided into De and $D\mu$ candidates in order to find systematic deviations induced by the lepton identification. The ANN outputs of $B^+ \to \overline{D}^0 e^+\nu_e$ and $B^+ \to \overline{D}^0\mu^+\nu_\mu$ candidates after $B_{\rm tag}$ selection are compared between data and MC in Fig. 4.61. While the MC underestimates the corresponding data distribution in $B^+ \to \overline{D}^0 e^+\nu_e$, the $B^+ \to \overline{D}^0\mu^+\nu_\mu$ channel shows good agreement. Since the absolute number of selected $D\ell$ candidates depends on more than the $\ell^\pm$ identification efficiency (K-ID rates, mis-ID rates, track and neutral reconstruction efficiencies, K^0_S reconstruction efficiency), one cannot conclude the μ-ID efficiency to be well reproduced in the MC and the e-ID efficiency to be worse. On the other hand, deviations induced by the D reconstruction enter both lepton modes to the same amount. Thus, it can be concluded from Fig. 4.61 that at least one of the two lepton identification rates is not sufficiently reproduced in the MC. The data-MC comparison (Fig. 4.62) indicates a difference on the 10 % level in normalization of the two lepton species.

Another source of tagging efficiency differences in data and MC arises from the D^* reconstruction. The $\overline{D}^{*0} \to \overline{D}^0\pi^0$ selection suffers from uncertainties of the soft π^0 reconstruction efficiency. As already visible in the abovementioned Δm distribution (Fig. 4.53b), the MC tends to overestimate the π^0_{soft} efficiency, which enriches the pool of $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ candidates with a D^* reconstructed in $\overline{D}^{*0} \to \overline{D}^0\pi^0$. The bin-by-bin comparison of the ANN outputs for best $D^{*0}e$ and $D^{*0}\mu$ candidates is illustrated in Fig. 4.63, where the D^{*0} are reconstructed in $D^0\pi^0$. The difference of the two lepton species is still visible, but $\Delta N/N_{\text{data}}$ moved down by about 20 % in both cases. This confirms a data-MC discrepancy of the π^0_{soft} reconstruction efficiency. The corresponding $\mathcal{NN}^{D\ell}_{\text{best}}$ distributions are displayed in Fig. 4.64.

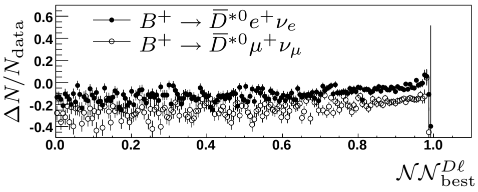


Figure 4.63: Data-MC comparison of $\mathcal{NN}^{D\ell}_{\text{best}}$ of $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ ($\overline{D}^{*0} \to \overline{D}^0\pi^0$) candidates: Deviations of the scaled numbers of On Peak events and full MC.

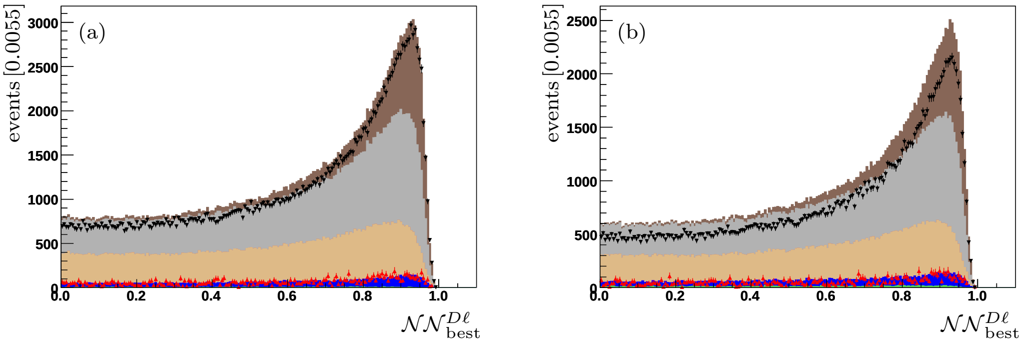


Figure 4.64: ANN output of best $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ ($\overline{D}^{*0} \to \overline{D}^0\pi^0$) candidates after $|c^{D\ell}_{\text{best}}| = 1$: $\mathcal{NN}^{D\ell}_{\text{best}}$ is shown for (a) $B^+ \to \overline{D}^{*0}e^+\nu_e$, and (b) $B^+ \to \overline{D}^{*0}\mu^+\nu_\mu$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

The above discussion clarifies the challenge to determine exactly the B_{tag} reconstruction efficiency. There are many sources of data-MC disagreements arising from the $B \to D^{(*)}\ell\nu_\ell$ reconstruction chain. In fact, within *BABAR* dedicated analyses have been performed to study deviations of track reconstruction [91], PID efficiency [92], and π^0 selection efficiencies [93]. However, this analysis uses a different strategy to correct ε_{tag} from a signal-free control sample.

4.7.4.2 Tagging Efficiency Correction using Double-Tagged Events

As discussed in the previous section, data-MC disagreements in the B_{tag} reconstruction efficiency have been found, which are believed to be due to differences in PID efficiencies, tracking efficiencies, and π^0 selection efficiencies. Their influences can be determined from a control sample.

This sample consists of events with two reconstructed $D\ell$ candidates, where both of them passed the entire $D\ell$ preselection. Naturally, this implies the requirement of ”no second identified lepton with $|\vec{p}_\ell| > 1\,\text{GeV}/c$” as part of the event preselection

(Sec. 4.7.2) to be dropped since both $D\ell$ candidates are required to include a lepton with $|\vec{p}_\ell| > 1\,\mathrm{GeV}/c$. Furthermore, the $D\ell$'s are required to be oppositely charged and are not allowed to share any track or neutral candidate since this procedure aims at the selection of events of type $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ vs. $B^- \to D^{(*)0}\ell^-\overline{\nu}_\ell$. If more than two such $B_{\rm tag}$ candidates have been found, the best and the second best candidates are chosen based on their ANN outputs.

$\varepsilon_{\rm tag}$ correction factors can be determined from the numbers of selected double-tagged events $N_{\rm DT}$ in data and MC. Assuming $N_{\rm DT} = \varepsilon_{\rm tag}^2 N_{B^+B^-}$ one obtains

$$\frac{N_{\rm DT}^{\rm Data}}{N_{\rm DT}^{\rm MC}} = \left(\frac{\varepsilon_{\rm tag}^{\rm Data}}{\varepsilon_{\rm tag}^{\rm MC}}\right)^2 \frac{N_{B^+B^-}^{\rm Data}}{N_{B^+B^-}^{\rm MC}} \tag{4.35}$$

with the initial number of $\Upsilon(4S) \to B^+B^-$ events $N_{B^+B^-}$. The generic B^+B^- MC sample has been explicitly scaled to the same number of $\Upsilon(4S)$ events embedded in the On Peak sample (Sec. 4.1). This leads to $N_{B^+B^-}^{\rm Data} = N_{B^+B^-}^{\rm MC}$ and the $B_{\rm tag}$ reconstruction efficiency correction is given by

$$C_{\rm tag} = \frac{\varepsilon_{\rm tag}^{\rm Data}}{\varepsilon_{\rm tag}^{\rm MC}} = \sqrt{\frac{N_{\rm DT}^{\rm Data}}{N_{\rm DT}^{\rm MC}}}\,. \tag{4.36}$$

Strictly speaking, Eq. 4.36 only holds if the double-tagged sample does not include any non-B^+B^- events. Fig. 4.65 shows the ANN output of the best $D\ell$ candidate reconstructed in double-tagged events, where the second best $D\ell$ candidate passed exactly the same selection. As already claimed, the $B_{\rm tag}$ selection efficiency depends on the reconstruction mode, which requires mode-dependent correction factors. Thus, Fig. 4.65a illustrates $\mathcal{NN}_{\rm best}^{D\ell}$ of double-tagged events reconstructed as $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ vs. $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$, whereas Fig. 4.65b shows the second configuration ($B^+ \to \overline{D}^0\ell^+\nu_\ell$ vs. $B^- \to D^0\ell^-\overline{\nu}_\ell$). Mixed events of type $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ vs. $B^- \to D^0\ell^-\overline{\nu}_\ell$ or $B^+ \to \overline{D}^0\ell^+\nu_\ell$ vs. $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ have not been used to determine

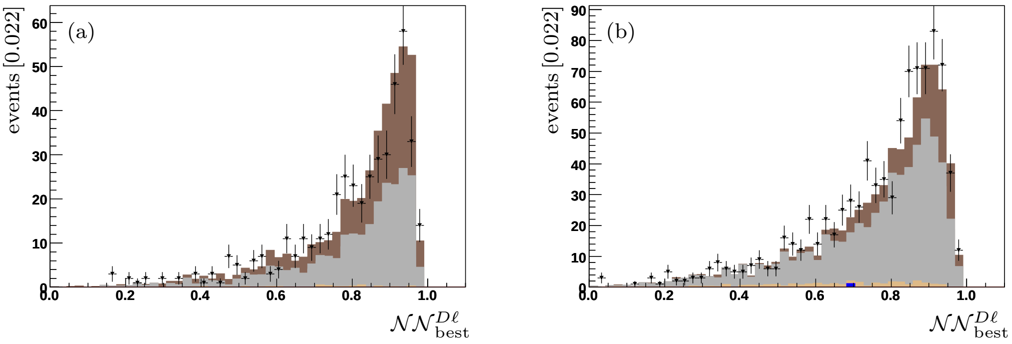


Figure 4.65: ANN output of best $D\ell$ candidates after $B_{\rm tag}$ selection in double-tagged events: (a) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ vs. $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ and (b) $B^+ \to \overline{D}^0\ell^+\nu_\ell$ vs. $B^- \to D^0\ell^-\overline{\nu}_\ell$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ vs. $B^- \to D^{(*)0}\ell^-\overline{\nu}_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

the correction factors. The purities of B^+B^- events in the selected double-tagged sample have been found to be 99.1 % and 96.2 % for the $\overline{D}^{*0}\ell^+$ vs. $D^{*0}\ell^-$ and $\overline{D}^0\ell^+$ vs. $D^0\ell^-$ configurations, respectively.

Quantity	$B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$	$B^+ \to \overline{D}^{0}\ell^+\nu_\ell$	$B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$
$\varepsilon_{\text{tag}}/\%$	0.2834 ± 0.0026	0.3904 ± 0.0031	0.6738 ± 0.0040
$N_{\text{DT}}^{\text{Data}}$	438.0 ± 20.9	885.0 ± 29.7	–
$N_{\text{DT}}^{\text{MC}}$	445.4 ± 9.8	784.1 ± 13.1	–
$\sqrt{N_{\text{DT}}^{\text{Data}}/N_{\text{DT}}^{\text{MC}}}$	0.992 ± 0.026	1.062 ± 0.019	–
$N_{\text{ST}}^{\text{Data}}$	811553.0 ± 900.9	832493.0 ± 912.4	
$N_{\text{ST}}^{\text{MC}}$	841553.0 ± 439.5	796091.6 ± 441.4	–
$N_{\text{ST}}^{\text{Data}}/N_{\text{ST}}^{\text{MC}}$	0.964 ± 0.001	1.046 ± 0.001	–
$\varepsilon_{\text{tag}}^{\text{corr}}/\%$	0.2810 ± 0.0079	0.4147 ± 0.0080	0.6958 ± 0.0112

Table 4.14: The table includes the uncorrected B_{tag} reconstruction efficiencies ε_{tag}, the numbers of double and single-tagged events found in On Peak data and full MC, and the corrected B_{tag} efficiency $\varepsilon_{\text{tag}}^{\text{corr}}$.

Tab. 4.14 summarizes the mode-dependent numbers of double-tagged events for data and MC, respectively. As a cross-check the numbers of single-tagged events N_{ST} and their data-MC ratios after the B_{tag} selection are also listed for both modes. In principle, a correction could be determined from the difference in normalization of single-tagged events (Fig. 4.66). On the other hand, this technique assumes that the data-MC efficiency ratios are equal for all background components. The correction extracted from the double-tagged sample is valid for B^+B^- events and can therefore be used to correct ε_{tag} for signal MC. However, the corrections from double and single-tagged events are consistent within 1.2σ (Tab. 4.14). The uncertainties of $\varepsilon_{\text{tag}}^{\text{corr}}$ are mainly driven by the limited statistics of the double-tagged sample.

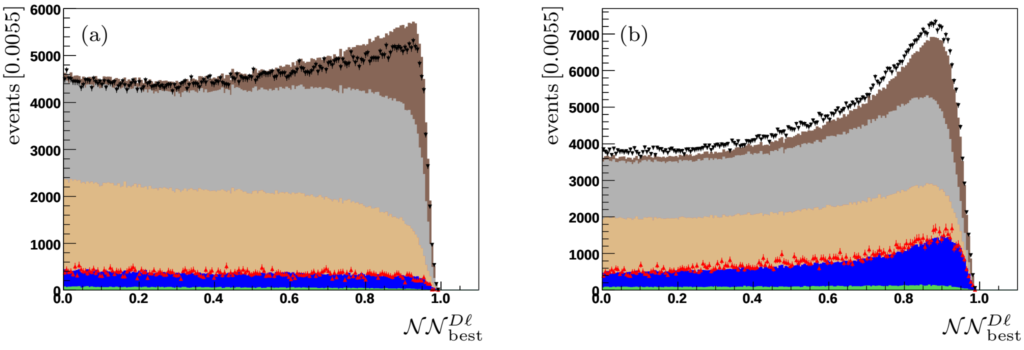


Figure 4.66: ANN output of best (a) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ and (b) $B^+ \to \overline{D}^0\ell^+\nu_\ell$ candidates after $|c_{\text{best}}^{D\ell}| = 1$ in single-tagged events. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [■], Off Peak [▲], On Peak [▼])

4.7.5 Restrictions on the Recoil

The determination and correction of ε_{tag} closes the discussion of the B_{tag} selection. This section mainly deals with the remaining neutral energy $E^{\text{CN}}_{\text{rem}}$ and its importance for the signal side selection.

4.7.5.1 $N^{\text{GTL}}_{\text{rem}}$ and $E^{\text{CN}}_{\text{rem}}$

Once a best $D\ell$ candidate has been selected, all tracks and neutrals used for its reconstruction have been removed from the event. Afterwards, the remaining event has been searched for patterns expected for $B^- \to \tau^-\overline{\nu}_\tau$ with the subsequent decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. Ideally, signal events should be visible as exactly three tracks reconstructed in the tracking system and no EMC bumps not associated to charged tracks since only charged particles are produced on the B_{sig} side. In fact, detector effects spoil this pattern of the remaining event for true $B^- \to \tau^-\overline{\nu}_\tau$ decays.

Later on, 3π candidates are reconstructed from three GTL candidates not belonging to the reconstructed B_{tag} (Sec. 4.7.6.1). The number of remaining GTL

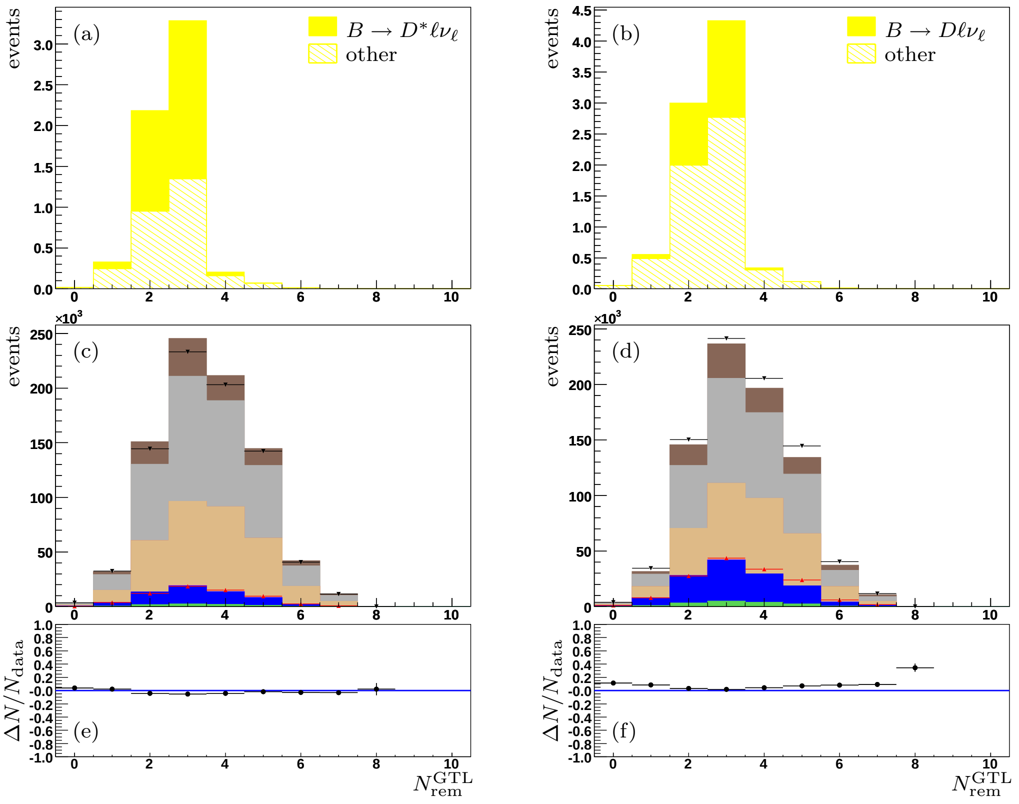


Figure 4.67: Number of remaining GTL candidates after $|c^{D\ell}_{\text{best}}| = 1$ for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ (left) and $B^+ \to \overline{D}^0\ell^+\nu_\ell$ tagged events (right). (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟧], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

candidates $N_{\mathrm{rem}}^{\mathrm{GTL}}$ after subtraction of all B_{tag} daughters is illustrated in Fig. 4.67. Figs. 4.67a,b indicate the expected pattern for signal MC. Due to the limited acceptance of the BABAR detector tracks are able to escape without being detected. Furthermore, tracks initiated by charged pions originated in $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ may not fulfil the GTL requirements. Both effects result in less than three GTL candidates in the recoil of the B_{tag} candidate for true signal events. The shape of the background distributions (Figs. 4.67c,d) are well-described by the MC as visible in the comparison plots of the On Peak data points and the full MC expectation (Figs. 4.67e,f).

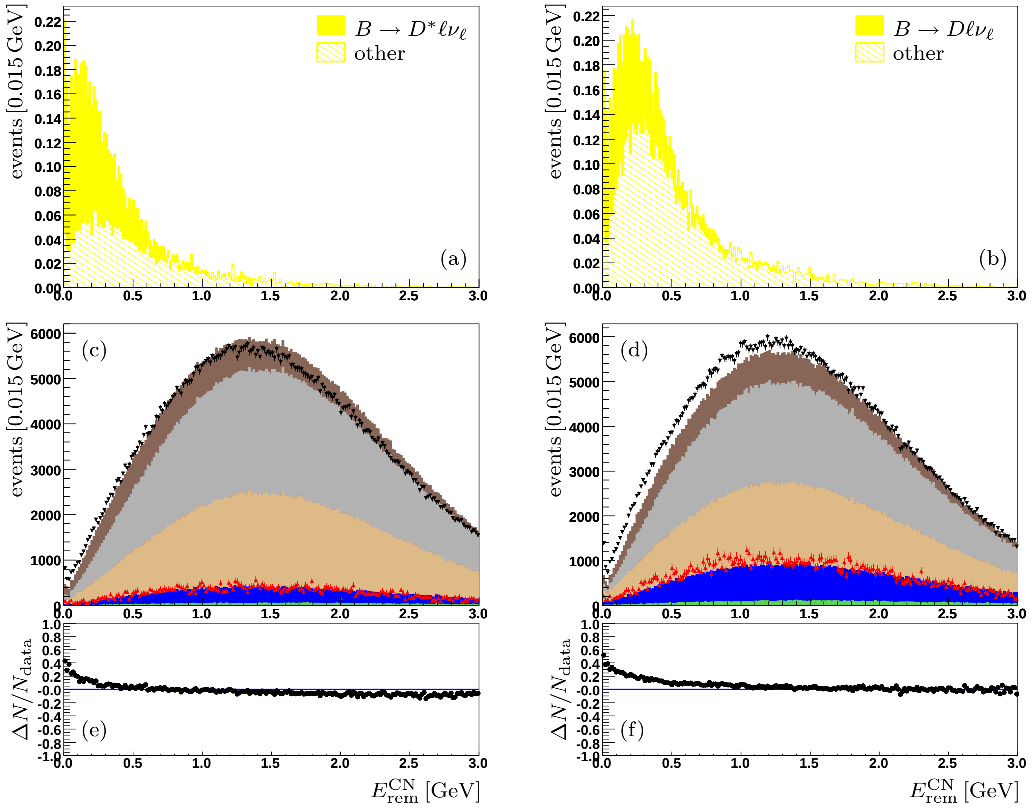


Figure 4.68: Remaining neutral energy after $|c_{\mathrm{best}}^{D\ell}| = 1$ for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ (left), and $B^+ \to \overline{D}^0\ell^+\nu_\ell$ tagged events (right). (e,f) illustrate the On Peak data vs. full MC bin-by-bin deviations. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟧], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

Signal events are expected to show no neutral energy in the EMC after erasing the B_{tag} candidate. The variable $E_{\mathrm{rem}}^{\mathrm{CN}}$ is defined as the sum of the energies of all CN candidates not belonging to B_{tag}, i.e. γ candidates used to reconstruct $\overline{D}^{*0} \to \overline{D}^0\pi^0$ ($\pi^0 \to \gamma\gamma$), $\overline{D}^{*0} \to \overline{D}^0\gamma$, and $\overline{D}^0 \to K^+\pi^-\pi^0$ ($\pi^0 \to \gamma\gamma$) have been removed from the CN list. The energy of all remaining CN candidates measured in the EMC (laboratory frame) have been summed up. The resulting signal MC distributions

(Figs. 4.68a,b) display enhancements at low $E_{\rm rem}^{\rm CN}$ values. Deviations from $E_{\rm rem}^{\rm CN} = 0$ are caused by various effects, such as beam-background, wrong $B_{\rm tag}$ reconstruction, and hadronic splitoffs. Photons radiated by the beams enter the EMC and can be selected as CN candidates leading to additional energy in the EMC. Moreover, if the $B_{\rm tag}$ has not been correctly reconstructed, γ produced in the $B_{\rm tag}$ decay chain are assigned to the signal side. For example, if a true $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ decay is reconstructed as $B^+ \to \overline{D}^0\ell^+\nu_\ell$, the soft D^* daughter (π^0, γ) falsifies the $E_{\rm rem}^{\rm CN}$ distribution. Such contributions are illustrated by the hatched areas in Figs. 4.68a,b.

However, correctly reconstructed $B_{\rm tag}$ (filled areas) contribute a large part in the range $E_{\rm rem}^{\rm CN} > 0$ believed to be induced by beam-background photons as well as interactions of the $\pi^\pm$ tracks with the detector material. If a $\pi^\pm$ traverses the EMC material, a shower of secondary particles from the initial hadronic interaction is created. These particles, such as π^0s or nuclear spallation byproducts, can travel laterally through neighbouring crystals and deposit their energy in a certain distance to the point of impact of the $\pi^\pm$. In this case a separated "splitoff" bump not associated to the initial $\pi^\pm$ track could be selected by the cluster algorithm and enters the CN list. The probability to select such hadronic splitoffs depends on the track multiplicity. However, the three $\pi^\pm$ tracks originated in the τ decay of signal events could initiate such extra EMC bumps and therefore modify $E_{\rm rem}^{\rm CN}$.

Although $E_{\rm rem}^{\rm CN}$ is washed-out due to the aforementioned effects, this variable provides high separation potential between signal and background and in fact $E_{\rm rem}^{\rm CN}$ is one of the variables the final selection has been optimized on (Sec. 4.7.7). Unfortunately, the data-MC comparison shown in Figs. 4.68e,f exhibits a strong disagreement in shape right in the signal region $E_{\rm rem}^{\rm CN} < 0.5\,{\rm GeV}$. This disagreement needs to be addressed. Otherwise, the resulting underestimation of background in the given region leads to fake signal in data.

4.7.5.2 Neutral Multiplicity

A huge disagreement in the number of neutral candidates per event between data and MC has been mentioned several times (Figs. 4.18h,i). At this stage of the recoil analysis this disagreement becomes important since it has been identified as the source of the previously mentioned deviations of the $E_{\rm rem}^{\rm CN}$ distributions.

Fig. 4.69 illustrates the data-MC differences in the number of remaining CN candidates $N_{\rm rem}^{\rm CN}$. The higher $N_{\rm rem}^{\rm CN}$ the higher their energy sum $E_{\rm rem}^{\rm CN}$ resulting in a shift of the $E_{\rm rem}^{\rm CN}$ distributions in MC towards higher values. In order to antagonize this effect, the MC description of $N_{\rm rem}^{\rm CN}$ has to be corrected.

The $E_{\rm rem}^{\rm CN}$ deviations have been assumed to be entirely due to the disagreement in $N_{\rm rem}^{\rm CN}$. No differences in the EMC energy response have been taken into account. The correction procedure compares $N^{\rm CN}$ after the $B_{\rm tag}$ selection in bins of energy, i.e. event-by-event the numbers of CN candidates with an energy in a given range are counted. Therefore, 14 energy ranges have been defined. The MC and data samples have been further divided into six subsamples ($N^{\rm GTL} = 5, 6, 7, 8, 9, 10$) with respect to the GTL multiplicity motivated by the aforementioned correlation of hadronic splitoffs and track multiplicities. In fact, a large

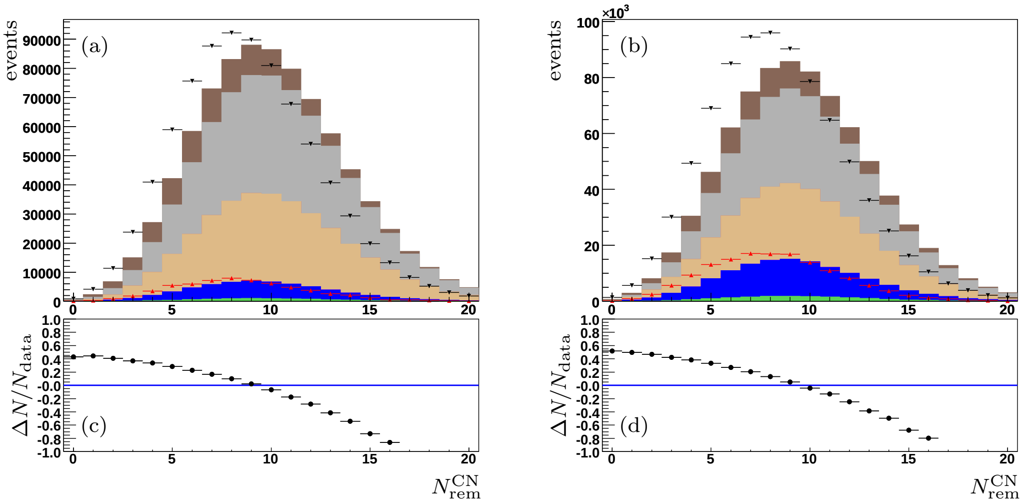


Figure 4.69: Number of remaining CN candidates after $|c_{\mathrm{best}}^{D\ell}| = 1$ for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ (left) and $B^+ \to \overline{D}^0\ell^+\nu_\ell$ tagged events (right). ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

amount of the N^{CN} deviations is believed to be induced by deficits in the simulation of hadronic splitoffs in the EMC. Exemplarily, Fig. 4.70 displays the multiplicity $N^{\mathrm{CN}}(E, N^{\mathrm{GTL}})$ of CN candidates with energies of $0.02 < E < 0.03$ GeV in events with $N^{\mathrm{GTL}} = 8$, where the the CN energy E has been measured in the laboratory frame. The mean values of the full MC and On Peak data distributions μ_{MC} and μ_{Data} with

$$\mu = \overline{N^{\mathrm{CN}}(E, N^{\mathrm{GTL}})} \tag{4.37}$$

strongly differ as indicated by the lines in Fig. 4.70. The deviation of their ratio from $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}} = 1$ acts as a measure of the level of disagreement. Fig. 4.71 summarizes these ratios for all 14 CN energy bins and six N^{GTL} subsamples, where the ratios have been determined separately for the Run13 and Run4 sets. All these ratios are listed in App. E for the Run13 and Run4 periods, respectively. In general, significantly lower factors have been found in Run4 in the low energy region $E < 0.2$ GeV (Figs. 4.71c,d) indicating higher discrepancies compared to the Run13 set. Fig. 4.71 further illustrates that additional low-energetic EMC bumps in the MC are mainly responsible for higher average CN multiplicities. This supports the assumption of hadronic splitoffs to act as the leading source of these deviations. Since splitoff bumps are produced as a part of hadronic showers, they are

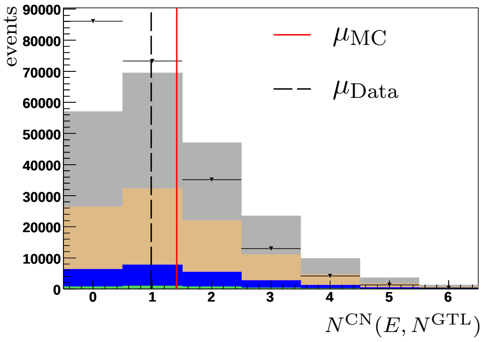


Figure 4.70: Number of CN candidates with energies of $0.02 < E < 0.03$ GeV in events with $N^{\mathrm{GTL}} = 8$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], B^+B^- [⬜], On Peak [▼])

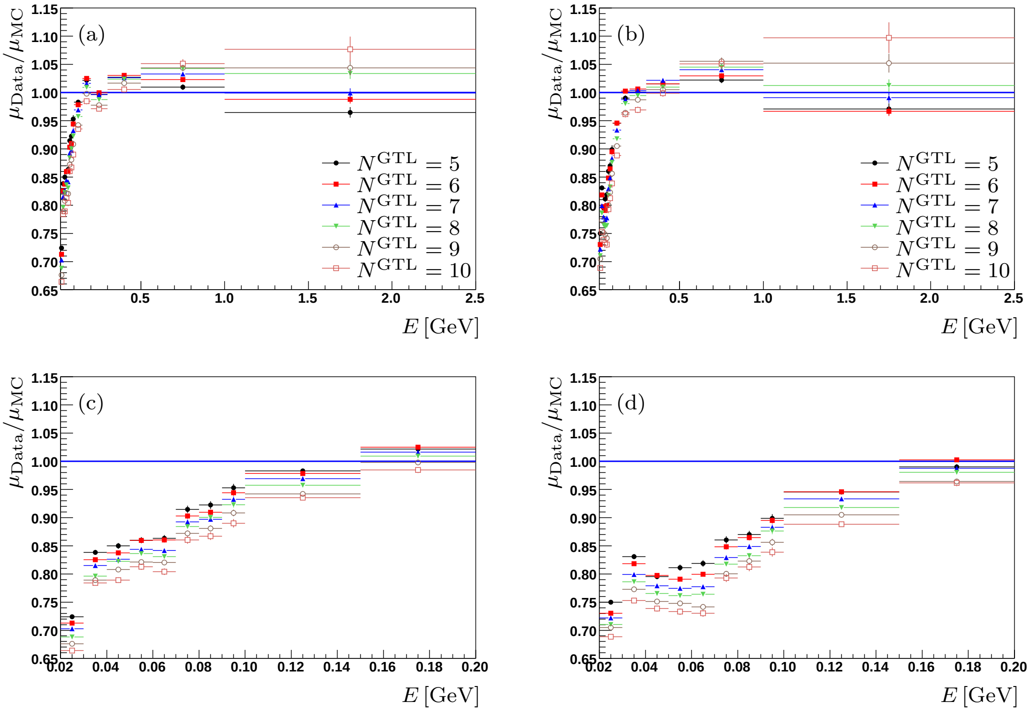


Figure 4.71: N^{GTL} correction factors determined from (a,c) Run13 and (b,d) Run4 samples. The blue lines indicate $\mu_{\text{Data}}/\mu_{\text{MC}} = 1$.

expected to contain a small portion of the entire cluster energy. Moreover, the ratios determined in the low-energy region (Figs. 4.71c,d) tend to decrease with increasing N^{GTL}, i.e. stronger data-MC disagreement for high track multiplicities. Again, this is expected for hadronic splitoffs as discussed before. The high-energy region $E > 0.3\,\text{GeV}$ is not less interesting. As can be seen in Figs. 4.71a,b, the MC simulation is characterized by a lack of neutral candidates in this region ($\mu_{\text{Data}}/\mu_{\text{MC}} > 1$) shifting $E^{\text{CN}}_{\text{rem}}$ towards lower values. However, the ratios given in Fig. 4.71 can be used as correction factors to improve the modeling of $N^{\text{CN}}_{\text{rem}}$ and therefore $E^{\text{CN}}_{\text{rem}}$ in the MC simulation.

4.7.5.3 Neutral Multiplicity Correction

In order to correct $N^{\text{CN}}(E, N^{\text{GTL}})$ in the MC, the ratios of the means are interpreted as probabilities to remove ($\mu_{\text{Data}} < \mu_{\text{MC}}$) or to add ($\mu_{\text{Data}} > \mu_{\text{MC}}$) CN candidates. In the case of "neutral killing" the probability is given by

$$\mathcal{P}_{\text{kill}} = 1 - \frac{\mu_{\text{Data}}}{\mu_{\text{MC}}}\,. \tag{4.38}$$

Depending on N^{GTL} and their EMC energies CN candidates have been killed randomly with probability $\mathcal{P}_{\text{kill}}$ if $\mu_{\text{Data}}/\mu_{\text{MC}} < 1$ is valid for the corresponding energy range. Such candidates have not longer been used for the calculation of $E^{\text{CN}}_{\text{rem}}$. In contrast, bins with $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}} > 1$ require a more complicated procedure since additional candidates have to be "invented" and added to the CN list. The adding probability

$$\mathcal{P}_{\mathrm{add}} = p_{\mathrm{Data}} \left(\frac{\mu_{\mathrm{Data}}}{\mu_{\mathrm{MC}}} - 1 \right) \tag{4.39}$$

depends on an a priori probability p_{Data} to find a CN candidate in the given energy range determined from On Peak data. Once a candidate has been added, a pseudoenergy has been generated within the limits of the actual energy bin assuming an uniform energy distribution.[14] Afterwards, this pseudoenergy has been taken into account in the $E_{\mathrm{rem}}^{\mathrm{CN}}$ calculation.

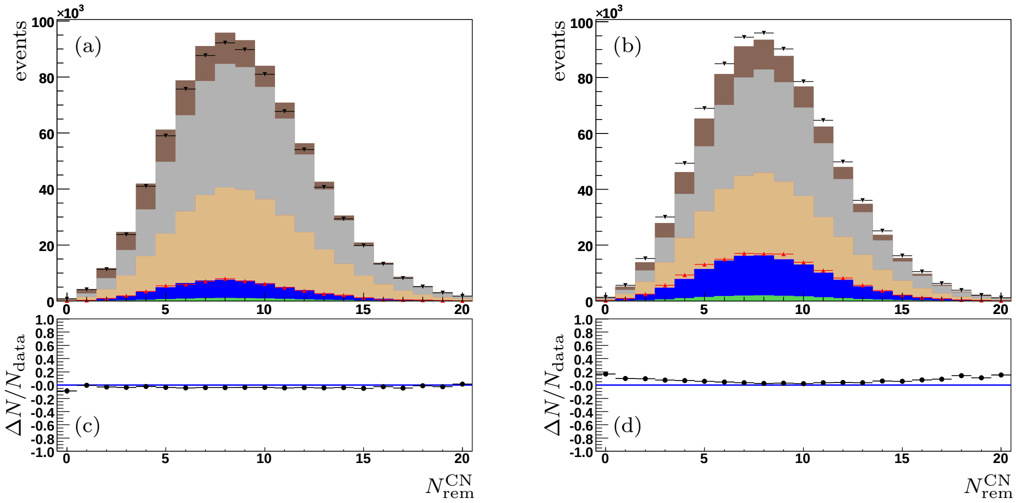


Figure 4.72: Number of remaining CN candidates after N^{CN} correction for $B^+ \to \bar{D}^{*0}\ell^+\nu_\ell$ (left) and $B^+ \to \bar{D}^0\ell^+\nu_\ell$ tagged events (right). ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟨], $B^+ \to \bar{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

After application of the given procedure the $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$ ratios have been found to be consistent with one. It is notable that the correction factors have been extracted using all CN candidates including neutral candidates of the B_{tag} side of events passing the B_{tag} selection. Nevertheless, after correction of all MC samples the distributions of the number of remaining neutrals agree well in shape in data and corrected MC (Fig. 4.72). This is true for non-$\Upsilon(4S)$ as well as $\Upsilon(4S)$ events, even though the corrections have not been determined separately from comparisons of Off Peak vs. MC continuum/τ-pair and On minus Off Peak vs. $\Upsilon(4S)$ MC. Thus, the correction factors account for detector effects and are assumed to be not affected by different event topologies of continuum and $\Upsilon(4S)$ events.

Although no energy correction of CN candidates has been applied, the $E_{\mathrm{rem}}^{\mathrm{CN}}$ distributions of tagged events are corrected to high extent as shown in the next section.

---

[14] The CN energy distribution is known to be non-uniform, but due to the low adding factors the effect of the "neutral adding" has been found to be small.

4.7.5.4 Validation of $E_{\text{rem}}^{\text{CN}}$ and Recoil Selection

Fig. 4.73 illustrates the impact of the N^{CN} correction on the MC prediction of the $E_{\text{rem}}^{\text{CN}}$ distributions. Indeed, the distributions agree well after the correction has been applied (Figs. 4.73e,f). Naturally, signal MC events have also been corrected since the same deficits of the MC description of the aforementioned detector effects need to be taken into account.

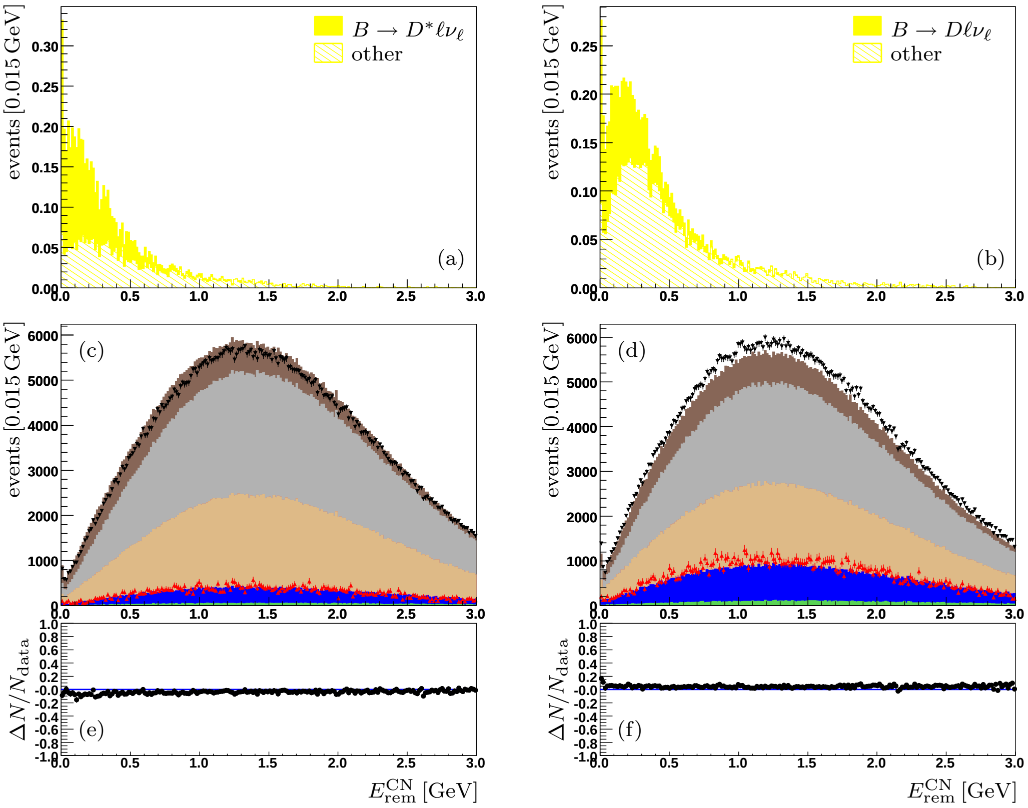


Figure 4.73: Remaining neutral energy after N^{CN} correction for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ (left) and $B^+ \to \overline{D}^0\ell^+\nu_\ell$ tagged events (right). (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [▨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

The $E_{\text{rem}}^{\text{CN}}$ modeling after the N^{CN} correction can be validated using the double-tagged sample introduced in Sec. 4.7.4.2. $E_{\text{rem}}^{\text{CN}}$ in double-tagged events is calculated as the sum of the EMC energies measured in the laboratory frame of all remaining CN candidates after subtraction of both reconstructed $D\ell$ candidates. In principle, the resulting distributions correspond to the expected $E_{\text{rem}}^{\text{CN}}$ spectra in signal events if both $D\ell$ candidates have been correctly reconstructed. However, the data-MC comparison plots before N^{CN} correction for double-tagged events (Figs. 4.74a-d)[15]

---

[15] Contrary to the ε_{tag} correction plots of Fig. 4.65, these figures include mixed events of type $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ vs. $B^- \to D^0\ell^-\overline{\nu}_\ell$ or $B^+ \to \overline{D}^0\ell^+\nu_\ell$ vs. $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ to increase statistics.

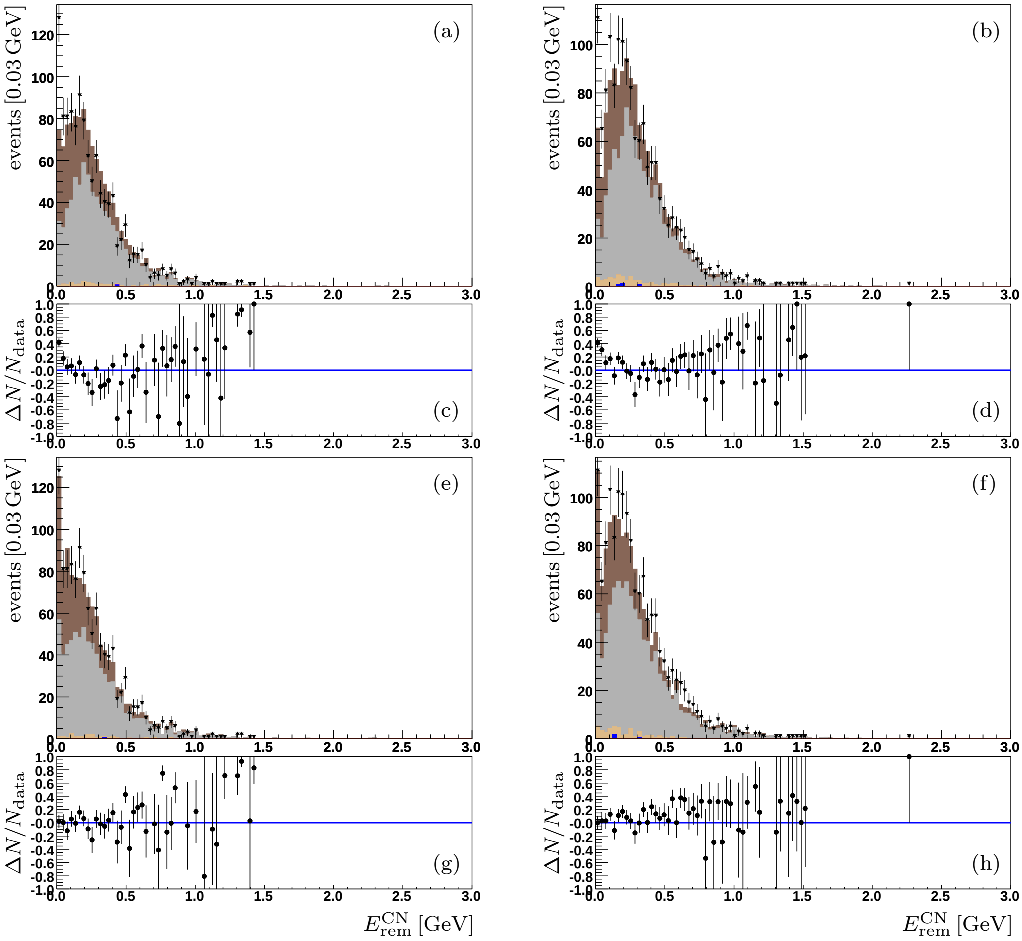


Figure 4.74: Remaining neutral energy (a,b) before and (e,f) after $N^{\rm CN}$ correction for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ (left) and $B^+ \to \overline{D}^0\ell^+\nu_\ell$ (right) double-tagged events. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ vs. $B^- \to D^{(*)0}\ell^-\overline{\nu}_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

confirm the disagreement in the range $E_{\rm rem}^{\rm CN} < 0.5\,{\rm GeV}$ already mentioned for the single-tagged events (Fig. 4.68c-f). The $N^{\rm CN}$ correction procedure results in an improved $E_{\rm rem}^{\rm CN}$ modeling in the MC double-tagged sample as illustrated in Figs. 4.74e,f. In the region $E_{\rm rem}^{\rm CN} < 0.5\,{\rm GeV}$ the shape is well-described by the corrected MC. In fact, this good agreement of double-tagged events validate the $E_{\rm rem}^{\rm CN}$ description in the corrected signal MC (Figs 4.73a,b).

After the $N^{\rm CN}$ correction $N_{\rm rem}^{\rm GTL}$ and $E_{\rm rem}^{\rm CN}$ have been used to restrict the recoil to the pattern expected for signal events before 3π candidates are reconstructed in the remaining event. Events passing the recoil selection are required to satisfy

- $N_{\rm rem}^{\rm GTL} = 3$ and
- $E_{\rm rem}^{\rm CN} < 1\,{\rm GeV}$.

In fact, $E_{\text{rem}}^{\text{CN}}$ implies a very high separation potential and allows a tighter selection. However, an optimization of the selection criterion has been accomplished after the reconstruction of a_1 candidates in the recoil of the best $D\ell$ candidate.

4.7.6 Signal Side Reconstruction

After the recoil selection $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ candidates can be reconstructed from the three remaining GTL candidates the recoil of the B_{tag} has been restricted on. The experiences on the kinematical properties of this τ decay are applicable. Again, after a loose preselection an a_1 candidate based ANN is constructed to efficiently select good candidates.

4.7.6.1 a_1 Candidate Preselection

Within the recoil analysis the same quantities as given in Sec. 4.5.2 have been used for preselection of 3π candidates, but since the emphasis of the reconstruction is provided by the B_{tag} side, the requirements on $|\vec{p}_{3\pi}|$ and $\cos\Theta_{3\pi,\text{miss}}$ have been modified. With respect to B_{sig} selection efficiencies a given combination of three remaining GTL candidates is accepted if it satisfies:

- $0.8 < m_{3\pi} < 1.6\,\text{GeV}/c^2$,
- $0.5 < |\vec{p}_{3\pi}| < 2.7\,\text{GeV}/c$, and
- $c_{3\pi} + c_{\text{best}}^{D\ell} = 0$.

The requirement on $|\vec{p}_{3\pi}|$ has been relaxed compared to the tight criterion of the inclusive reconstruction ($1.5 < |\vec{p}_{3\pi}| < 2.7\,\text{GeV}/c$). 3π combination are forced to be oppositely charged to the best $D\ell$ candidate. No restrictions have been applied on $\cos\Theta_{3\pi,\text{miss}}$. As already discussed (Sec. 4.6.2), a semileptonic reconstruction $B \to X\ell\nu$ spoils the correlations of the missing and the 3π three-momentum vectors due to the presence of an additional neutrino produced on the B_{tag} side, i.e. $\vec{p}_{\text{miss}}$ and $\vec{p}_{3\pi}$ are not expected to be back-to-back and equal in magnitude (Fig. 4.45). Thus, no $\cos\Theta_{3\pi,\text{miss}}$ selection has been used to preselect 3π candidates. But in contrast to the $B \to X\ell\nu$ technique of Sec. 4.6, the four-momentum of the $X\ell$ system is explicitly reconstructed in the recoil analysis represented by the $D\ell$ candidate. Hence, assuming a correctly reconstructed B_{tag} candidate in rest the three-momentum vector of the neutrino produced within the semileptonic decay is simply given by

$$\vec{p}_\nu = \vec{p}_B - \vec{p}_{D\ell} = \vec{0} - \vec{p}_{D\ell} = -\vec{p}_{D\ell} \tag{4.40}$$

with the three-momentum vector $\vec{p}_{D\ell}$ of the $D\ell$ candidate. Hence, one can recalculate the missing three-momentum vector corresponding to the B_{sig} side by subtraction of $\vec{p}_\nu$. Even though the non-zero B_{tag} momentum has been neglected and a large fraction of the B_{tag} sample has not been correctly reconstructed, this simple $\vec{p}_{\text{miss}}$ modification results in the expected behaviour ($\cos\Theta_{3\pi,\text{miss}} \approx -1$, $|\vec{p}_{\text{miss}}| - |\vec{p}_{3\pi}| \approx 0$) for correctly reconstructed a_1 candidates in signal MC (Fig. 4.75). Both distribu-

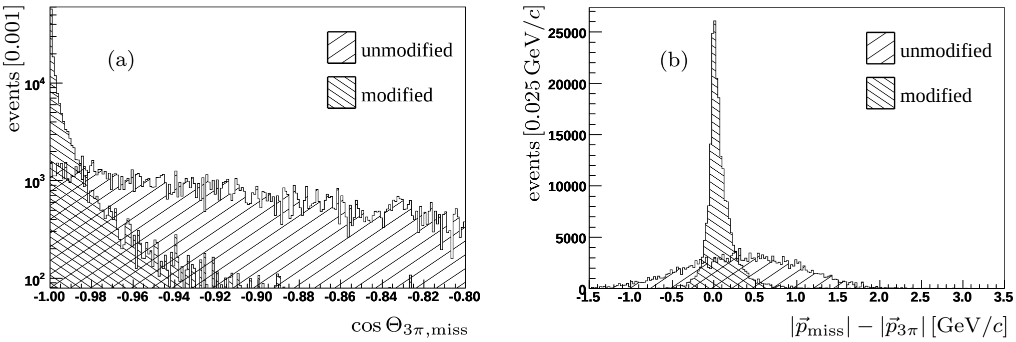


Figure 4.75: Modification of missing momentum: (a) $\cos\Theta_{3\pi,\mathrm{miss}}$ and (b) $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$ of correctly reconstructed a_1 candidates in signal MC with and without $\vec{p}_{\mathrm{miss}}$ modification. The distributions have been scaled to the same integral.

tions show sharp peaks, even though not correctly reconstructed B_{tag} candidates are included. Again, these variables have been combined with other 3π quantities using an a_1 candidate based ANN. But before the multivariate a_1 selection is explained, a brief review of the vertexing is given in Sec. 4.7.6.2.

4.7.6.2 Vertex Separation

The reconstruction of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay vertex has been introduced as a motivation to favour this 3-prong τ decay to perform a search for the decay $B^- \to \tau^-\overline{\nu}_\tau$. Indeed, within the inclusive reconstruction (Sec. 4.5.3.8) an inclusion of the separations of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ and the B_{tag} decay vertices along the z axis (Δ^z) and in the transverse plane (Δ^{xy}) improved the separation power of $\mathcal{NN}^{3\pi}_{\mathrm{best}}$ considerably (Fig. 4.36). Since the companion B meson has not been reconstructed within the inclusive technique, its decay vertex has been extracted from tracks found not to belong to a given 3π candidate (Sec. 4.5.3.6). However, within the recoil analysis the B_{tag} is built by a well-defined set of tracks and neutrals and its decay vertex has been directly obtained from these candidates using the `Cascade` algorithm. Non-negligible mean flight lengths of long-living particles, e.g. K^0_S from $\overline{D}^0 \to K^0_S\pi^-\pi^+$, have been taken into account.

Fig. 4.76 compares Δ^z extracted by the vertexing procedures of the inclusive and the recoil techniques. While the width of the Δ^z distribution of correctly reconstructed a_1 candidates is not affected by the $B \to D^{(*)}\ell\nu_\ell$ selection (Figs. 4.76a,d), the background distributions have been found to be broadened. This is true for combinatorial $B\overline{B}$ (Figs. 4.76b,e) as well as continuum background (Figs. 4.76c,f). The r.m.s. of the background distributions increase by a factor of about 1.3 resulting in a negligible potential to suppress $B\overline{B}$ background using Δ^z. $\Upsilon(4S) \to B^+B^-$ events exhibit the strongest broadening, which is the result of a clear vertex separation of random three-GTL combinations recoiling against correctly reconstructed B_{tag} candidates. Furthermore, $e^+e^- \to c\overline{c}$ events contain charmed mesons like D or D^* formed by the primary c quarks and light quarks originated in the fragmen-

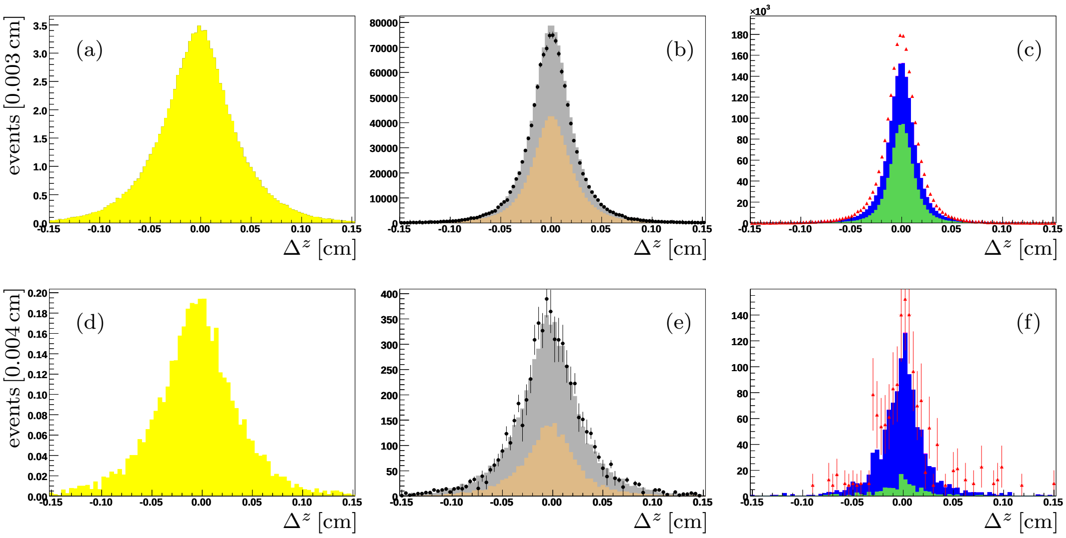


Figure 4.76: Δ^z after a_1 candidate preselections of (a-c) inclusive reconstruction, and (d-f) $B \to D^{(*)}\ell\nu_\ell$ recoil technique: The inclusive selection plots correspond to the entire Run12 data set, while the plots derived from the recoil technique correspond to the Run14 sample. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On minus Off Peak [▼−▲=●])

tation process. Correctly reconstructed D or D^* candidates combined with random $\ell^\pm$ candidates might be spatially separated coincidently from the remaining tracks in the event causing a broadening of Δ^z.

Although Δ^z after the B_{tag} selection no longer provides suppression potential of $B\overline{B}$ background, the small residual separation power of signal and continuum events is exploited by including Δ^z and Δ^{xy} in an a_1 based ANN (Sec. 4.7.6.3). However, one has to conclude that the potential of vertex reconstruction for background suppression is not sufficient; neither with nor without an explicit B_{tag} reconstruction.

4.7.6.3 Multivariate a_1 Candidate Selection

Even though vertex separation is not the key to perform a successful search for the process $B^- \to \tau^-\overline{\nu}_\tau$, the kinematics of the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ still provides high potential of background suppression.

For example, the separation power of $\cos\Theta_{3\pi,\text{miss}}$ and $|\vec{p}_{\text{miss}}| - |\vec{p}_{3\pi}|$ is illustrated in Fig. 4.77. Here, the signal distributions (Fig. 4.77a,d) include correctly (filled areas) and not correctly (hatched areas) reconstructed a_1 candidates. At this stage of the analysis not correctly reconstructed a_1 candidates amount for about 11 % of the selected a_1 sample in signal MC, i.e. as a result of the exclusive B_{tag} reconstruction the combinatorial background decreased considerably compared to the inclusive technique. This is the consequence of the clean environment of the signal side achieved by the recoil technique since combinations of GTL candidates originated on the B_{tag} side and tracks from the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay are strongly suppressed.

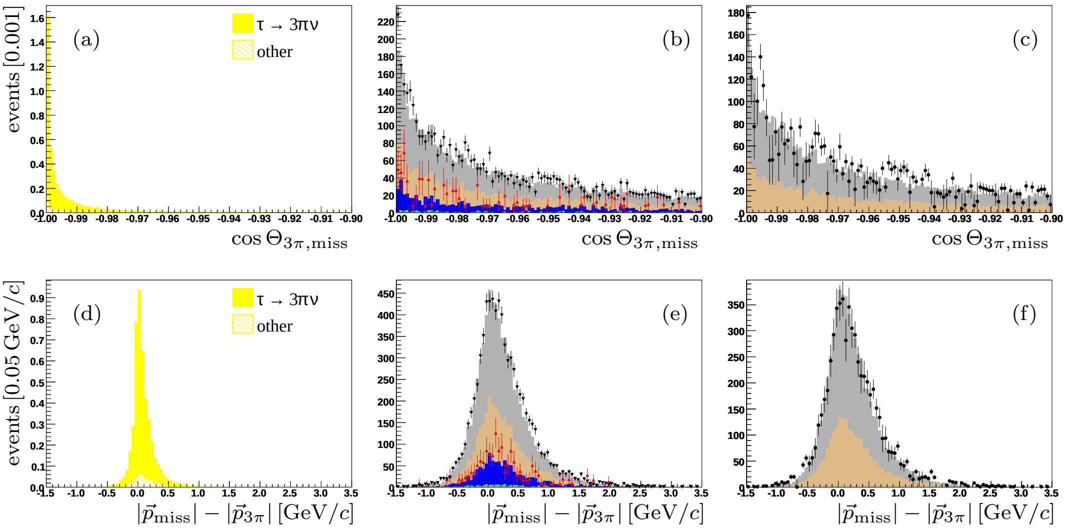


Figure 4.77: Distributions of (a-c) $\cos\Theta_{3\pi,\mathrm{miss}}$ and (d-f) $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$ after a_1 preselection. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

$\cos\Theta_{3\pi,\mathrm{miss}}$ and $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$ have been combined with ten additional variables. The distributions of these twelve input variables are summarized in Apps. D.2 (kinematical quantities) and D.3 (vertex separation).

Since $B\overline{B}$ events dominate the sample after the recoil restrictions, the background part of the training sample contains an appropriate admixture of tagged $\Upsilon(4S) \to B^+B^-$ and $\Upsilon(4S) \to B^0\overline{B}^0$ MC events passing the a_1 preselection. The signal part is represented by the signal vs. cocktail MC mentioned in Sec. 4.1. The exact decomposition of this MC species is illustrated in Fig. 4.78. While one B meson decays via the signal channel, the other B meson is forced to decay into $D^*\ell\nu$ or $D\ell\nu$ with subsequent D^* and D decays corresponding to the reconstruction modes. The B and D branching fractions of these modes have been rescaled to sum up to 100 %, while the relative fractions of the specific modes to the sum of all these modes

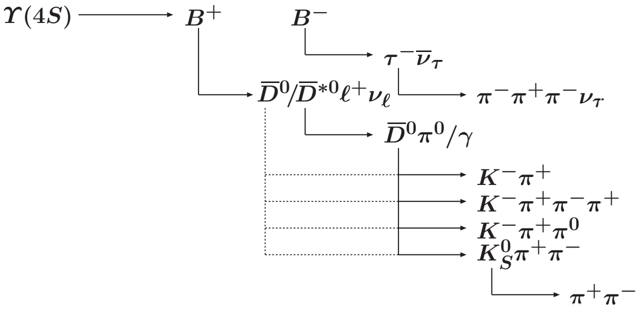


Figure 4.78: Decomposition of signal vs. cocktail MC: The dotted lines show the decay tree of $B^+ \to \overline{D}^0\ell^+\nu_\ell$.

have been kept. The entire training sample is composed of about 14000 correctly reconstructed a_1 candidates from signal vs. cocktail MC and the same amount of combinatorial $B\overline{B}$ background candidates passing the a_1 preselection. Again, the ANN is strongly favoured compared to the FD algorithm.

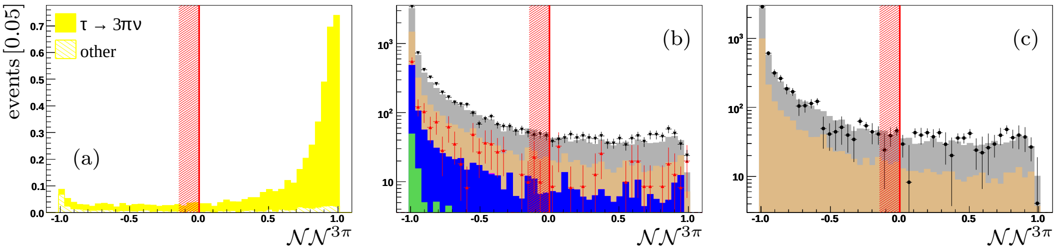


Figure 4.79: Output of a_1 candidate based ANN for (a) signal MC, (b) full background, and (c) $B\overline{B}$ background. The red lines illustrate the cuts applied. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

The resulting ANN output is given in Fig. 4.79. This figure illustrates the high separation between signal and backgrounds in $\mathcal{NN}^{3\pi}$. One should note that no best a_1 selection is needed since the recoil of the B_{tag} candidate has been restricted to exactly three GTL candidates resulting in at most one reconstructed a_1 candidate passing the a_1 preselection. Events passing the multivariate a_1 selection are required to satisfy $\mathcal{NN}^{3\pi} > 0$.

The contribution of signal events containing a not correctly reconstructed 3π candidate drops from 11 % to 6.5 % after application of the aforementioned criterion. This fraction further decreases with harder requirements on the ANN output. Moreover, due to the B_{tag} reconstruction and the restrictions on its recoil this small part is not longer expected to depend on the B_{tag} decay channel. Thus, in the following such combinatorial background candidates in signal MC events are accounted for as signal.

4.7.7 Final Cut Optimization

The quantities $E_{\text{rem}}^{\text{CN}}$ and $\mathcal{NN}^{3\pi}$ provide the highest potentials to separate signal from background events and therefore these variables have been used to optimize the expected upper limit $\text{UL}_{\text{exp}}^{14}$. In principle, one could simply calculate $\text{UL}_{\text{exp}}^{14}$ using the counted numbers of surviving signal and background events satisfying different selection criteria on both quantities. On the other hand, due to the limited MC statistics some regions of the $E_{\text{rem}}^{\text{CN}} - \mathcal{NN}^{3\pi}$ plane contain only a handful of events satisfying $\mathcal{NN}^{3\pi} > 0$ and such an optimization procedure becomes sensitive to statistical fluctuations. Alternatively, in order to obtain reliable estimations of the numbers of signal and background events passing a given cut combination, analytical functions can be used to approximate both distributions and to extract fractional integrals included in a given $E_{\text{rem}}^{\text{CN}} - \mathcal{NN}^{3\pi}$ range.

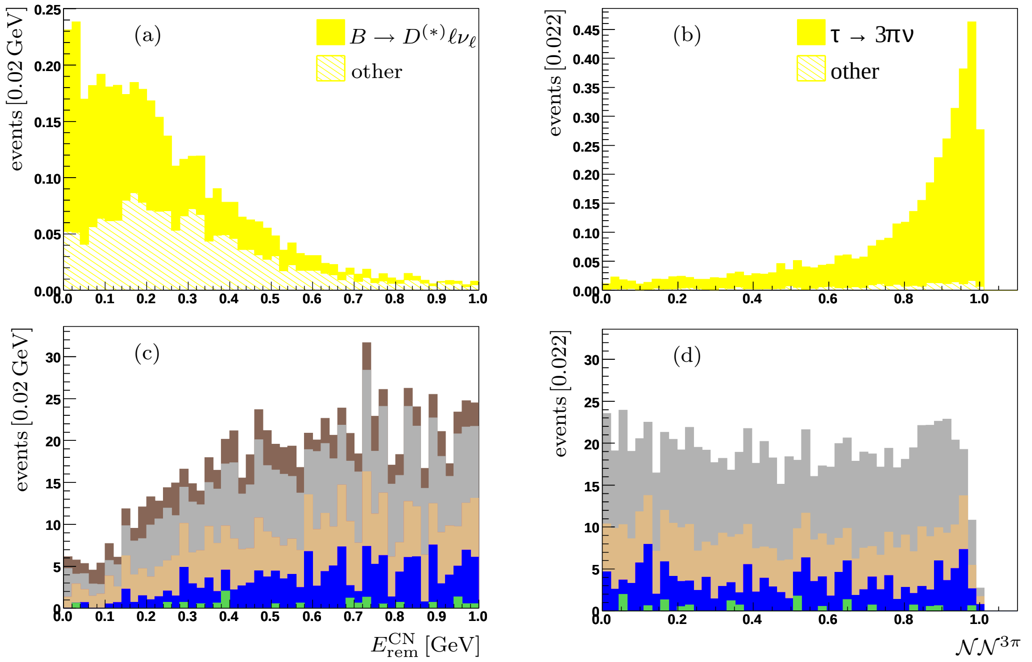


Figure 4.80: (a,c) $E_{\rm rem}^{\rm CN}$ and (b,d) $\mathcal{NN}^{3\pi}$ after $\mathcal{NN}^{3\pi} > 0$ for (a,b) signal MC and (c,d) full MC background. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟧], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜])

The MC distributions of both quantities are illustrated in Fig. 4.80. Both distributions contain correctly and not correctly reconstructed a_1 as well as $D\ell$ candidates indicated by the filled and hatched parts in the signal distributions (Figs. 4.80a,b), respectively. Furthermore, Fig. 4.80 includes both $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ reconstruction modes. In principle, these distributions could be described by analytical functions, but the shape of $\mathcal{NN}^{3\pi}$ of signal candidates (Fig. 4.80b) introduces a complication.

4.7.7.1 Transformation of $\mathcal{NN}^{3\pi}$

The ANN algorithm returns outputs between minus one and one. Thus, the number of entries above $\mathcal{NN}^{3\pi} = 1$ is forced to be zero resulting in a very sharp edge paired with a long left-side tail visible in Fig. 4.80b. This behaviour of $\mathcal{NN}^{3\pi}$ is difficult to be described by an analytical function and requires a transformation.

The transformed a_1 based ANN output $\mathcal{XNN}^{3\pi}$ is defined as

$$\mathcal{XNN}^{3\pi} = a_0 + a_1 \cdot \ln\left[\frac{\mathcal{NN}^{3\pi}_{\rm max} - \mathcal{NN}^{3\pi}_{\rm min}}{\mathcal{NN}^{3\pi} - \mathcal{NN}^{3\pi}_{\rm min}} - 1\right] \tag{4.41}$$

with

$$\mathcal{NN}^{3\pi} \in \left[\mathcal{NN}^{3\pi}_{\rm min}, \mathcal{NN}^{3\pi}_{\rm max}\right] = [0, 1]\,. \tag{4.42}$$

The relation of $\mathcal{XNN}^{3\pi}$ and $\mathcal{NN}^{3\pi}$ given in Eq. 4.41 is illustrated in Fig. 4.81, where the parameters a_0 and a_1 take the values $a_0 = 0.5$ and $a_1 = 0.06$. The advantage of such an one-to-one transformation of the interval [0,1] onto itself is shown in Fig. 4.82. The resulting $\mathcal{XNN}^{3\pi}$ distributions are gaussian-like and can therefore be easily approximated. It should be noted that the separation potential of $\mathcal{NN}^{3\pi}$ is not harmed by this procedure.

The final cut optimization procedure (Sec. 4.7.7.2) uses fits to the $E^{\mathrm{CN}}_{\mathrm{rem}}$ and $\mathcal{XNN}^{3\pi}$ distributions and determines $\mathrm{UL}^{14}_{\mathrm{exp}}$ for different cut combinations from the fitted functions.

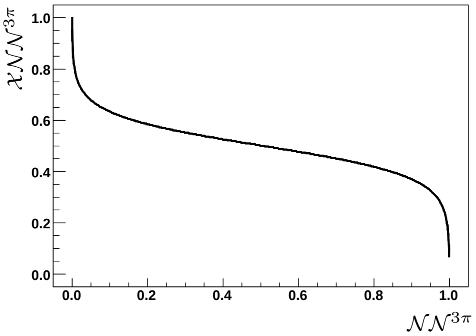


Figure 4.81: $\mathcal{NN}^{3\pi}$ Transformation Function

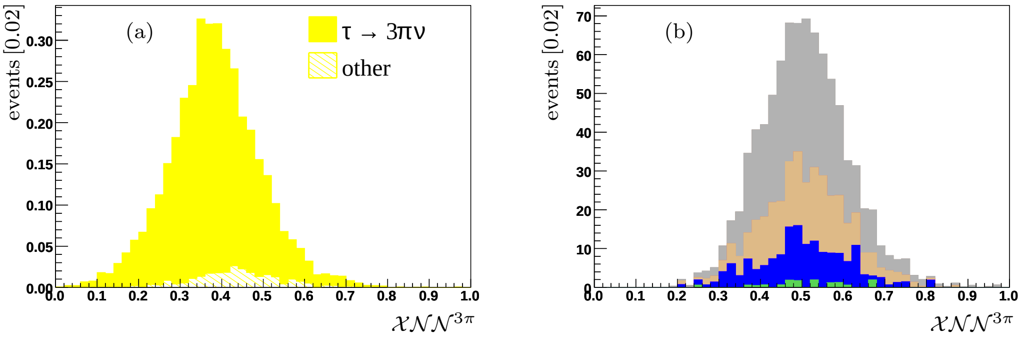


Figure 4.82: Transformed output of a_1 candidate based ANN after $\mathcal{NN}^{3\pi} > 0$ (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [◼])

4.7.7.2 Final Cut Optimization Procedure

The final cut optimization procedure is based on analytical functions fitted to the $E^{\mathrm{CN}}_{\mathrm{rem}}$ and $\mathcal{XNN}^{3\pi}$ distributions extracted from signal MC and the full MC background samples. The fitted functions $f(x)$ are illustrated in Fig. 4.83, where the distributions have been scaled to the Run14 luminosity and further include the branching fraction as well as the *Dalitz*-plot weighting factors. The details of the parametrizations are given in App. H.

The functions $f(x)$ can be used to estimate the number of events passing a given cut combination $\left([E^{\mathrm{CN}}_{\mathrm{rem}}]_{\mathrm{cut}}, [\mathcal{XNN}^{3\pi}]_{\mathrm{cut}}\right)$. The fractional integral $F(x_{\mathrm{cut}})$ included in the region $[x_{\mathrm{min}}, x_{\mathrm{cut}}]$ is given as

$$F(x_{\mathrm{cut}}) = \frac{\int_{x_{\mathrm{min}}}^{x_{\mathrm{cut}}} f(x)\mathrm{d}x}{\int_{x_{\mathrm{xmin}}}^{x_{\mathrm{max}}} f(x)\mathrm{d}x} \quad \text{with} \quad x \in \{E^{\mathrm{CN}}_{\mathrm{rem}}, \mathcal{XNN}^{3\pi}\} \tag{4.43}$$

in the fit range $[x_{\mathrm{min}}, x_{\mathrm{max}}]$. The number of events passing both cuts on $E^{\mathrm{CN}}_{\mathrm{rem}}$ and

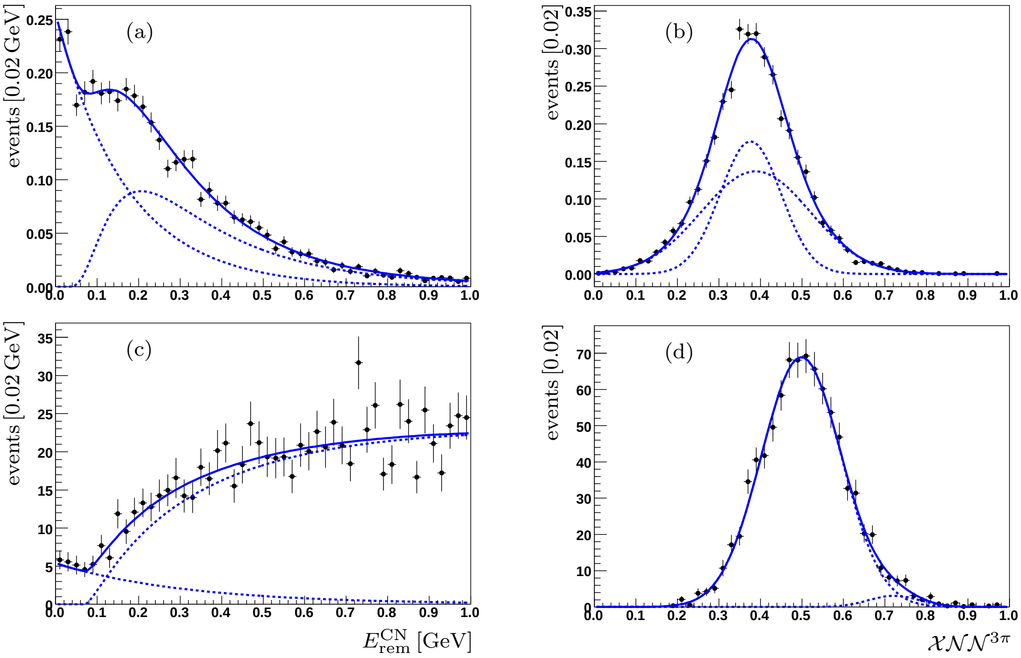


Figure 4.83: Fits to the distributions of (a,c) $E_{\mathrm{rem}}^{\mathrm{CN}}$ and (b,d) $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ after $\mathcal{N}\mathcal{N}^{3\pi} > 0$ for (a,b) signal MC and (c,d) full MC background.

$\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ is approximated by

$$N\left(\left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}}, \left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}\right) = F\left(\left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}}\right) \cdot F\left(\left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}\right) \cdot N_{\mathrm{sel}}\,, \tag{4.44}$$

where N_{sel} denotes the luminosity scaled and reweighted (branching fraction and *Dalitz*-plot weights) number of events after the multivariate a_1 selection. The product ansatz of Eq. 4.44 neglects correlations between $E_{\mathrm{rem}}^{\mathrm{CN}}$ and $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$. On the other hand, these quantities have been found to be correlated by about 18 % in signal MC and 16 % in the full background sample. However, the described procedure to estimate $N\left(\left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}}, \left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}\right)$ is only used for cut optimization. The actual determination of the expected number of signal and background events is discussed in Sec. 4.7.8.

Since the estimation of $N\left(\left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}}, \left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}\right)$ assumes $E_{\mathrm{rem}}^{\mathrm{CN}}$ and $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ to be uncorrelated, it is not necessary to scan the entire $E_{\mathrm{rem}}^{\mathrm{CN}} - \mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ plane for the best cut combination as it has been done within the inclusive reconstruction using $\mathcal{N}\mathcal{N}$ and $\mathcal{N}\mathcal{N}_{\mathrm{best}}^{3\pi}$ (Sec. 4.5.3.5). Instead, the cut on $E_{\mathrm{rem}}^{\mathrm{CN}}$ has been optimized using $\mathrm{UL}_{\mathrm{exp}}^{14}$ before $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ has been scanned for the best requirement after application of the optimal $E_{\mathrm{rem}}^{\mathrm{CN}}$ criterion.

Fig. 4.84a shows the evolution of S_{exp}^{14} and $\mathrm{UL}_{\mathrm{exp}}^{14}$ with different requirements $E_{\mathrm{rem}}^{\mathrm{CN}} < \left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}}$. The minimal $\mathrm{UL}_{\mathrm{exp}}^{14}$ has been found at $\left[E_{\mathrm{rem}}^{\mathrm{CN}}\right]_{\mathrm{cut}} = 0.25\,\mathrm{GeV}$ and nearly corresponds to the maximum of the expected significance S_{exp}^{14}. Fig. 4.84b illustrates the same quantities as a function of $\left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}$ ($\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < \left[\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}\right]_{\mathrm{cut}}$)

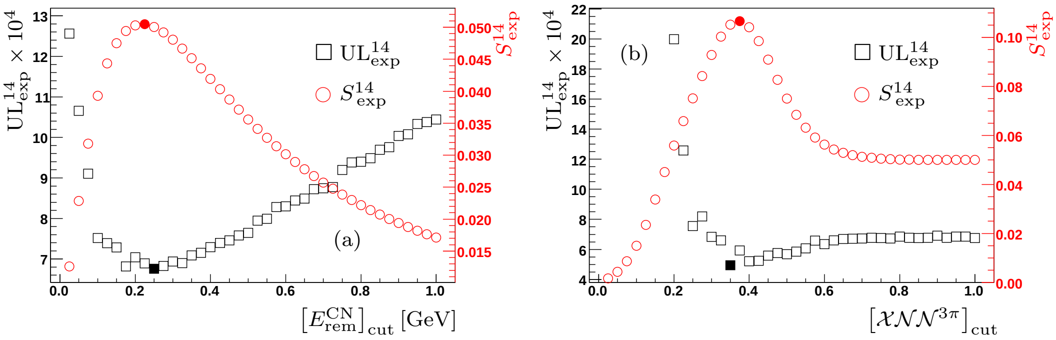


Figure 4.84: Expected upper limit and significance for different cuts on (a) $E^{\mathrm{CN}}_{\mathrm{rem}}$ after $\mathcal{NN}^{3\pi} > 0$ and (b) $\mathcal{XNN}^{3\pi}$ after the optimal cut on $E^{\mathrm{CN}}_{\mathrm{rem}}$ shown in (a). The optima are illustrated by the filled markers.

derived from $f(\mathcal{XNN}^{3\pi})$ and the fractional integral $F\left([E^{\mathrm{CN}}_{\mathrm{rem}}]_{\mathrm{cut}}\right) = F(0.25\,\mathrm{GeV})$ (Eq. 4.44). Again, the optima of $\mathrm{UL}^{14}_{\mathrm{exp}}$ and S^{14}_{exp} are located at nearly the same position and $\mathcal{XNN}^{3\pi} < 0.35$ has been found as the optimal $\mathcal{XNN}^{3\pi}$ requirement.

As already mentioned, the number of events extracted from Eq. 4.44 is estimated supposing uncorrelated selection variables. Furthermore, it assumes the $E^{\mathrm{CN}}_{\mathrm{rem}}$ and $\mathcal{XNN}^{3\pi}$ distributions in signal MC and combined background MC to be perfectly modeled by the analytical functions. Possible differences in the shapes of both quantities between data and MC are not taken into account. Hence, finally the expected numbers of signal and background events have not been determined from the fitted functions. However, the optimal cut combination $E^{\mathrm{CN}}_{\mathrm{rem}} < 0.25\,\mathrm{GeV}$ and $\mathcal{XNN}^{3\pi} < 0.35$ defines the signal region in the $E^{\mathrm{CN}}_{\mathrm{rem}} - \mathcal{XNN}^{3\pi}$ plane to search for signal events.

4.7.8 Signal Efficiency and Expected Number of Events

Once the final selection requirements have been defined one has to determine the expected number of background events and the signal efficiency. These quantities are needed to extract the upper limit from the observed number of On Peak data events in the signal region after all selection steps.

The selection steps of the signal side selection are given in Tab. 4.15. It summarizes the raw as well as the scaled/weighted event numbers passing the signal side selection requirements. Again, the scaled numbers of signal events correspond to an expected branching fraction of $\mathcal{B}_{\mathrm{exp}}(B^- \to \tau^-\overline{\nu}_\tau) = 10^{-4}$. The corresponding raw selection efficiencies with respect to the number of events after the B_{tag} selection ($|c^{D\ell}_{\mathrm{best}}| = 1$) are listed in Tab. 4.16.

After the final selection the raw signal efficiency is

$$\varepsilon_{\mathrm{sig}} = \frac{2091}{28368} = (7.371 \pm 0.155)\,\%\,. \tag{4.45}$$

$\varepsilon_{\mathrm{sig}}$ from Eq. 4.45 neither includes the branching fraction weighting factors (Sec. 4.7.1)

Selection	**signal**		B^+B^-		$B^0\overline{B}^0$	
B_{tag} selection	28368	(14.52)	4085948	(899836.74)	2566045	(544385.46)
$N^{\text{GTL}}_{\text{rem}} = 3$ $E^{\text{CN}}_{\text{rem}} < 1\,\text{GeV}$	13935	(7.08)	331929	(73127.96)	155158	(32532.01)
$N^{3\pi} > 0$	9131	(4.64)	17651	(3865.07)	10614	(2220.74)
$\mathcal{N}\mathcal{N}^{3\pi} > 0$	7750	(3.87)	2113	(463.79)	1174	(240.99)
$E^{\text{CN}}_{\text{rem}} < 0.25\,\text{GeV}$	4719	(2.31)	282	(62.88)	161	(32.10)
$\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < 0.35$	2091	(0.95)	25	(5.43)	10	(2.39)
Selection	$c\overline{c}$		$u\overline{u}, d\overline{d}, s\overline{s}$		$\tau^+\tau^-$	
B_{tag} selection	228220	(168901.15)	38160	(24450.60)	145	(70.94)
$N^{\text{GTL}}_{\text{rem}} = 3$ $E^{\text{CN}}_{\text{rem}} < 1\,\text{GeV}$	17022	(12561.36)	2300	(1485.27)	28	(13.46)
$N^{3\pi} > 0$	1394	(1017.77)	175	(116.99)	2	(1.01)
$\mathcal{N}\mathcal{N}^{3\pi} > 0$	209	(150.89)	19	(13.97)	0	(0.00)
$E^{\text{CN}}_{\text{rem}} < 0.25\,\text{GeV}$	15	(9.49)	1	(0.62)	0	(0.00)
$\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < 0.35$	1	(0.70)	1	(0.62)	0	(0.00)

Table 4.15: Cut Flow Table: This table gives the raw and in brackets the scaled (luminosity scaling, branching fraction weighting, *Dalitz*-plot weighting) numbers of events passing the given cuts. $N^{3\pi} > 0$ stands for the requirement of a reconstructed 3π candidate passing the a_1 preselection.

Selection	**signal**		B^+B^-		$B^0\overline{B}^0$	
$N^{\text{GTL}}_{\text{rem}} = 3$ $E^{\text{CN}}_{\text{rem}} < 1\,\text{GeV}$	49.122		8.124		6.047	
$N^{3\pi} > 0$	32.188	(65.5)	0.432	(5.3)	0.414	(6.8)
$\mathcal{N}\mathcal{N}^{3\pi} > 0$	27.320	(84.9)	0.052	(12.0)	0.046	(11.1)
$E^{\text{CN}}_{\text{rem}} < 0.25\,\text{GeV}$	16.635	(60.9)	0.007	(13.3)	0.006	(13.7)
$\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < 0.35$	7.371	(44.3)	0.001	(8.9)	0.000	(6.2)
Selection	$c\overline{c}$		$u\overline{u}, d\overline{d}, s\overline{s}$		$\tau^+\tau^-$	
$N^{\text{GTL}}_{\text{rem}} = 3$ $E^{\text{CN}}_{\text{rem}} < 1\,\text{GeV}$	7.459		6.027		19.310	
$N^{3\pi} > 0$	0.611	(8.2)	0.459	(7.6)	1.379	(7.1)
$\mathcal{N}\mathcal{N}^{3\pi} > 0$	0.092	(15.0)	0.050	(10.9)	0.000	(0.0)
$E^{\text{CN}}_{\text{rem}} < 0.25\,\text{GeV}$	0.007	(7.2)	0.003	(5.3)	0.000	(0.0)
$\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < 0.35$	0.000	(6.7)	0.003	(100.0)	0.000	(0.0)

Table 4.16: Signal Efficiency Table: This table gives raw B_{sig} selection efficiencies in % determined using the MC samples with respect to the number of events passing the B_{tag} selection ($|c^{D\ell}_{\text{best}}| = 1$). The values given in brackets are the step-by-step efficiencies.

nor the *Dalitz*-plot weights determined from the τ control sample (Sec. 4.4.2). If one takes into account both reweighting factors the signal efficiency drops by about 9 % to

$$\varepsilon_{\text{sig}} = (6.775 \pm 0.150)\,\% \tag{4.46}$$

mainly driven by the *Dalitz*-plot weights. Once more, this clarifies the importance of the *Dalitz*-plot correction. a_1 candidates satisfying the aforementioned final selection criteria mainly populate the intersection region of the ρ^0 bands within the $\sqrt{s_1} - \sqrt{s_2}$ plane ($0.6 < \sqrt{s_{1,2}} < 0.9\,\text{GeV}/c^2$) and therefore own *Dalitz*-plot weights below one (Fig. 4.10) resulting in the decrease of ε_{sig}.

4.7.8.1 Signal Efficiency Correction

Although the value of ε_{sig} given in Eq. 4.46 contains relative corrections of the branching fractions of the B_{tag} reconstruction modes and the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ *Dalitz*-plot weighting factors, absolute data-MC deviations influencing the signal side selection have not been taken into account. On the other hand, possible deviations of the GTL reconstruction efficiencies between data and MC affect ε_{sig}. Within *BABAR* such effects have been studied by an independent analysis of various control samples [91]. It has been found that in average the MC overestimates the track reconstruction efficiency. An application of an overall correction factor of −0.8 % per GTL candidate to the MC expectation has been recommended by this analysis. Since three GTL candidates have been combined to form an a_1 candidate, one has to account for a correction factor of

$$C_{\text{sig}} = (1 - 0.008)^3 = 0.9762 \tag{4.47}$$

per 3π combination and from Eq. 4.46 the corrected signal efficiency arises to

$$\varepsilon_{\text{sig}}^{\text{corr}} = C_{\text{sig}} \cdot \varepsilon_{\text{sig}} = (6.614 \pm 0.146)\,\%\,. \tag{4.48}$$

Since the B_{tag} and B_{sig} reconstructions are totally decoupled, $\varepsilon_{\text{sig}}^{\text{corr}}$ and the corrected B_{tag} reconstruction efficiency $\varepsilon_{\text{tag}}^{\text{corr}}$ (Tab. 4.14) have been multiplied to the total reconstruction efficiency of $B^- \to \tau^-\overline{\nu}_\tau$ events with the subsequent decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$

$$\varepsilon_{\text{tot}}^{\text{corr}} = \varepsilon_{\text{tag}}^{\text{corr}} \cdot \varepsilon_{\text{sig}}^{\text{corr}} = (4.60 \pm 0.13) \times 10^{-4}\,. \tag{4.49}$$

With $\varepsilon_{\text{tot}}^{\text{corr}}$ and $\mathcal{B}_{\text{exp}}(B^- \to \tau^-\overline{\nu}_\tau) = 10^{-4}$ the expected number of signal events reconstructed in the entire Run14 On Peak data sample can be calculated to be $N_{\text{sig}} = 0.99$. This value is meaningless as long as the expected number of background events N_{bkg} has bot been determined.

4.7.8.2 Expected Number of Background Events

Even though the numbers of events passing the entire selection have already been summarized in Tab. 4.15, a more compact overview of the background expectations is given in Tab. 4.17. The full background expectation after the final selection has been determined to $N_{\text{bkg}} = 9.14$, where the largest contribution arises from $\Upsilon(4S) \to B^+B^-$ events. The dominant part of this background source (60 %) has been identified as events of type $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ vs. $B^- \to X_c^0\ell^-\overline{\nu}_\ell$, where one lepton has not been selected by the PID selectors. Here, X_c^0 denotes higher excited charm states like D^{**} as well as non-resonant $D^{(*)}\pi$ states.

Mode	$\boldsymbol{N_{\text{sel}}}$	$\boldsymbol{N_{\text{sel}}^{\text{wght}}}$	$\boldsymbol{N_{\text{sel}}^{\text{sc}}}$
B^+B^-	25.0 ± 5.0	24.80 ± 4.98	5.43 ± 1.09
$B^0\overline{B}^0$	10.0 ± 3.2	9.88 ± 3.14	2.39 ± 0.76
$c\overline{c}$	1.0 ± 1.0	1.00 ± 1.00	0.70 ± 0.70
$u\overline{u}$, $d\overline{d}$, $s\overline{s}$	1.0 ± 1.0	1.00 ± 1.00	0.62 ± 0.62
$\tau^+\tau^-$	0.0 ± 0.0	0.00 ± 0.00	0.00 ± 0.00
combined	–	–	9.14 ± 1.62

Table 4.17: The raw (N_{sel}), branching fraction reweighted ($N_{\text{sel}}^{\text{wght}}$), and luminosity scaled ($N_{\text{sel}}^{\text{sc}}$) numbers of background events after the final selection. All values correspond to the entire Run14 MC samples as summarized in Tab. 4.1. The errors of $N_{\text{sel}}^{\text{sc}}$ have been added in quadrature to obtain the error of the combined background expectation.

However, this background expectation assumes the $E_{\text{rem}}^{\text{CN}}$ and $\mathcal{XNN}^{3\pi}$ MC predictions to agree perfectly to the data, i.e. in shape and absolute normalization. Although the $E_{\text{rem}}^{\text{CN}}$ shape has been corrected by the N^{CN} correction procedure (Sec. 4.7.5.3) at the stage of the best $D\ell$ candidate selection and has been verified for signal using the double-tagged sample, the background prediction needs to be tested at this analysis stage. Furthermore, deviations in the $\mathcal{XNN}^{3\pi}$ shape would also falsify the background expectation. Therefore, data-MC discrepancies after the multivariate a_1 selection have been investigated.

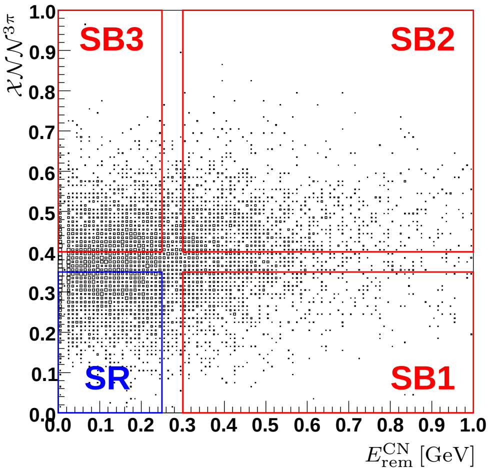


Figure 4.85: $E_{\text{rem}}^{\text{CN}}$ vs. $\mathcal{XNN}^{3\pi}$ of signal events after $\mathcal{NN}^{3\pi} > 0$: The signal region (SR) is defined by the selection criteria extracted by the optimization procedure. SB1, SB2, and SB3 denote the sidebands.

The $E_{\text{rem}}^{\text{CN}} - \mathcal{XNN}^{3\pi}$ plane after the multivariate a_1 selection is illustrated for signal MC in Fig. 4.85 with the signal region (**SR**) delimited by the cut values found by the optimization procedure ($E_{\text{rem}}^{\text{CN}} < 0.25\,\text{GeV}$, $\mathcal{XNN}^{3\pi} < 0.35$). In order to study data-MC disagreements, three sideband regions have been defined:

SB1: $E_{\text{rem}}^{\text{CN}} > 0.30\,\text{GeV}, \mathcal{XNN}^{3\pi} < 0.35$

SB2: $E_{\text{rem}}^{\text{CN}} > 0.30\,\text{GeV}, \mathcal{XNN}^{3\pi} > 0.40$

SB3: $E_{\text{rem}}^{\text{CN}} < 0.25\,\text{GeV}, \mathcal{XNN}^{3\pi} > 0.40$

Although these sideband regions are not signal free as visible in Fig. 4.85 it is shown later that the expected numbers of signal events in these regions are negligible compared to the background contributions. The corresponding distributions determined from the combined MC and On Peak data are given in Fig. 4.86. In order to judge how the MC compares to On Peak data, the projections of these distributions in the given sideband regions are shown in Fig. 4.87. The data-MC comparison plots (Figs. 4.87e,f) indicate a well-described $\mathcal{XNN}^{3\pi}$ background shape. On the other hand, a difference in normalization is visible over the entire range. This is not surprising since a lack of MC events has already been mentioned at the stage of the B_{tag} reconstruction (Sec. 4.7.4). In contrast to $\mathcal{XNN}^{3\pi}$, Fig. 4.87e uncovers a residual tendency of the MC to underestimate the number of events at low $E_{\text{rem}}^{\text{CN}}$. This residual deviation of the $E_{\text{rem}}^{\text{CN}}$ shape after the multivariate a_1 selection needs to be addressed.

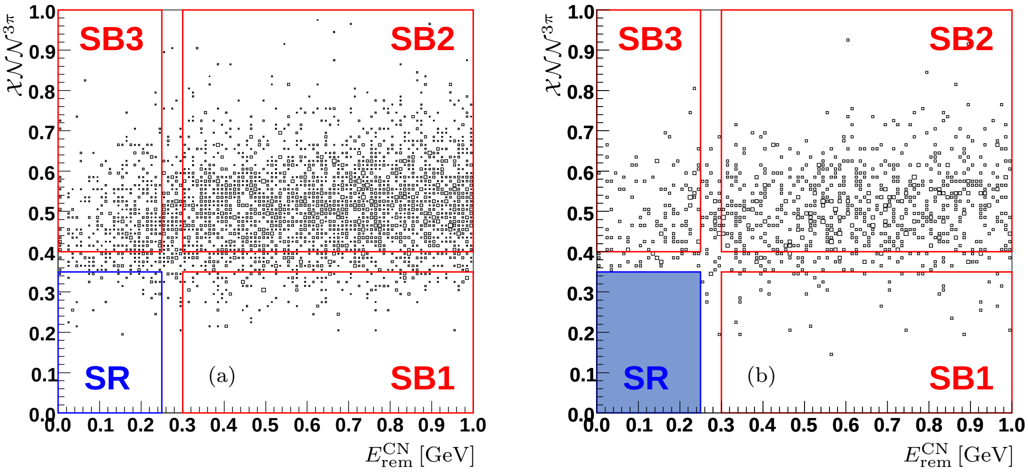


Figure 4.86: $E_{\text{rem}}^{\text{CN}}$ vs. $\mathcal{XNN}^{3\pi}$ of (a) combined MC background and (b) On Peak data after $\mathcal{NN}^{3\pi} > 0$: In order to avoid possible bias, the signal region (SR) has been covered in the On Peak data distribution.

The luminosity scaled and reweighted numbers of events contained in the sideband regions and the relative data-MC deviations are listed in Tab. 4.18. As already

Quantity	SB1	SB2	SB3
N_{sig}	0.38 ± 0.01	0.66 ± 0.02	0.84 ± 0.02
N_{MC}	40.41 ± 4.01	621.42 ± 13.99	77.66 ± 4.45
N_{data}	40.00 ± 6.32	675.00 ± 25.98	104.00 ± 10.20
$N_{\text{data}}/N_{\text{MC}}$	0.990 ± 0.185	1.086 ± 0.048	1.339 ± 0.152

Table 4.18: Scaled numbers of events after $\mathcal{NN}^{3\pi} > 0$ and ratios of the observed number of On Peak data events and the combined MC background expectation in the three sideband regions, respectively.

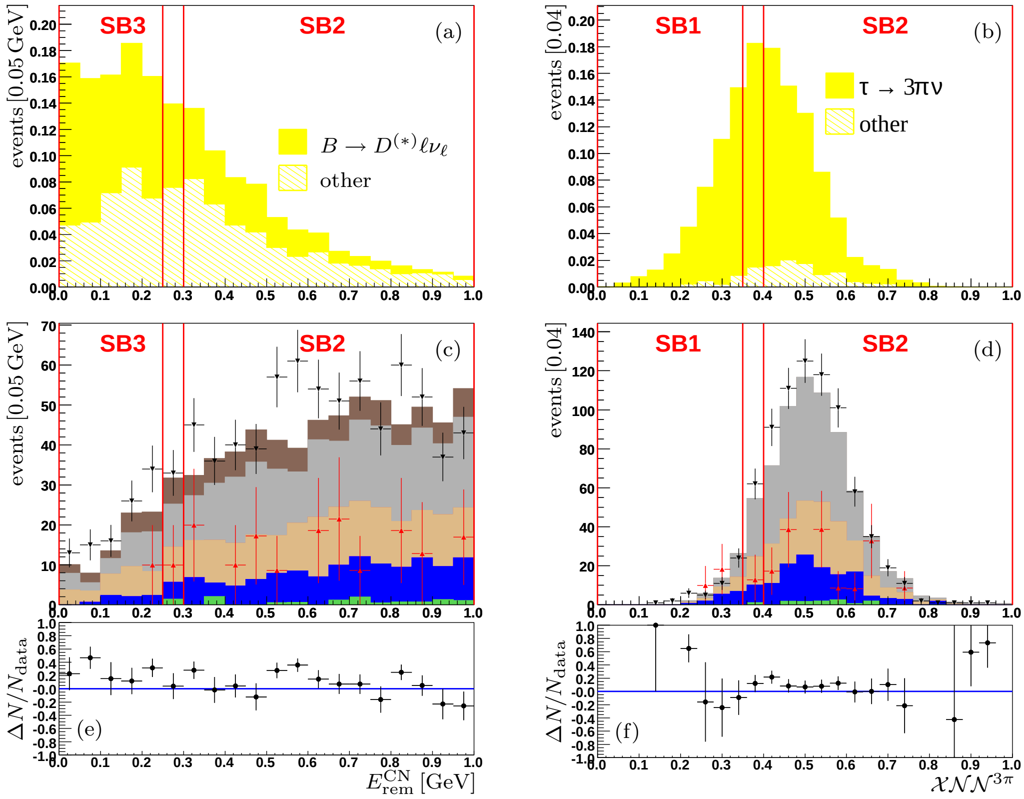


Figure 4.87: $E_{\rm rem}^{\rm CN}$ and $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ projections after $\mathcal{N}\mathcal{N}^{3\pi} > 0$ in the sideband regions for (a,b) signal MC and (c,d) full background. (a,c) show $E_{\rm rem}^{\rm CN}$ of events satisfying $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} > 0.4$. (b,d) illustrate $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ of events passing $E_{\rm rem}^{\rm CN} > 0.3\,\text{GeV}$. The bin-by-bin deviations of data and MC are given in (e,f). (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}{}^0$ [🟧], $B^+ \to \overline{D}{}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼])

noted, the expected signal contributions in all of these regions are negligible compared to the full MC background expectation and therefore all On Peak data events included in the sideband regions are supposed to be background.

The ratio $N_{\rm data}^{\rm SB2}/N_{\rm MC}^{\rm SB2}$ of the observed number of On Peak data events and the luminosity scaled/reweighted number of MC events in SB2 is different compared to the corresponding ratio $N_{\rm data}^{\rm SB3}/N_{\rm MC}^{\rm SB3}$ calculated for the region SB3. This fact reflects the tendency already seen in Fig. 4.87e. Assuming the relative difference of these ratios to be the same comparing SB1 and SR, one can extrapolate the observed disagreement in the region $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} > 0.4$ to the region $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi} < 0.35$, where the tendency seen in $\mathcal{X}\mathcal{N}\mathcal{N}^{3\pi}$ for $E_{\rm rem}^{\rm CN} > 0.3\,\text{GeV}$ (Fig. 4.87f) has to be taken into account in the extrapolation. Based on this assumption a correction factor for the expected number of background events in the signal region is calculable as

$$C_{\rm bkg} = \left(\frac{N_{\rm data}^{\rm SB3}}{N_{\rm MC}^{\rm SB3}} \Bigg/ \frac{N_{\rm data}^{\rm SB2}}{N_{\rm MC}^{\rm SB2}} \right) \cdot \frac{N_{\rm data}^{\rm SB1}}{N_{\rm MC}^{\rm SB1}} = 1.220 \pm 0.272 \tag{4.50}$$

with the ratios $N_{\text{data}}/N_{\text{MC}}$ as given in Tab. 4.18 for the three sideband regions, respectively. The largest uncertainty of C_{bkg} is introduced by the statistical uncertainty of the event numbers included in SB1 and is accounted for as a systematic uncertainty (Sec. 4.7.9). However, finally the expected number of background events in the signal region arises to

$$N_{\text{bkg}}^{\text{corr}} = C_{\text{bkg}} \cdot N_{\text{bkg}} = 11.16 \pm 3.18 \tag{4.51}$$

with the uncorrected background expectation $N_{\text{bkg}} = (9.14 \pm 1.62)$ from Tab. 4.17 and an uncertainty of about 28 % mainly induced by the correction factor.

After the final selection we expect approximately one signal event over a total background of roughly eleven background events with all systematic corrections applied. From these numbers the expected significance and the expected upper limit corresponding to the entire Run14 sample have been calculated to

$$S_{\text{exp}}^{14} = 0.081\,, \qquad \text{UL}_{\text{exp}}^{14} = 6.7 \times 10^{-4} \quad (90\,\%\,\text{C.L.})\,. \tag{4.52}$$

These values differ from the corresponding values of Fig. 4.84b shown in Sec. 4.7.7.2 for the best cut combination since for the optimization procedure used the uncorrected signal and background expectations extracted from the fractional integrals of the analytical functions. Furthermore, the optimization procedure did not account for the correlation of $E_{\text{rem}}^{\text{CN}}$ and $\mathcal{XNN}^{3\pi}$.

In comparison with $\text{UL}_{\text{exp}}^{14}$ without any systematic uncertainties extracted within the inclusive reconstruction (Eq. 4.21), the recoil analysis does not seem to be competitive. On the other hand, the influence of systematic uncertainties on $\text{UL}_{\text{exp}}^{14}$ has already been discussed and in order to judge the qualities of the different techniques one has to determine the systematic uncertainties and their impact on $\text{UL}_{\text{exp}}^{14}$.

4.7.9 Systematic Uncertainties

The quantities needed to calculate the upper limit for the branching fraction of the decay $B^- \to \tau^-\overline{\nu}_\tau$, such as

1. the corrected number of background events in the signal region $N_{\text{bkg}}^{\text{corr}}$,
2. the corrected B_{tag} reconstruction efficiency $\varepsilon_{\text{tag}}^{\text{corr}}$,
3. the corrected B_{sig} reconstruction efficiency $\varepsilon_{\text{sig}}^{\text{corr}}$,
4. the initial number of $\Upsilon(4S) \to B\overline{B}$ events included in the On Peak data sample $N_{B\overline{B}}$, and
5. the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ branching fraction $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$

are affected by different sources of systematic uncertainties. This section summarizes their contributions and explains the techniques to extract them.

4.7.9.1 Limited MC Statistics

Fundamental uncertainties of $N_{\text{bkg}}^{\text{corr}}$, $\varepsilon_{\text{tag}}^{\text{corr}}$, and $\varepsilon_{\text{sig}}^{\text{corr}}$ are induced by the limited MC statistics used to determine these quantities. Thus, the relative statistical errors of the corresponding uncorrected quantities extracted from the signal and background MC samples are assigned as systematic uncertainties. The relative systematic uncertainties are summarized in Tab. 4.19.

Quantity	**Central Value**	**Error**	**Relative Error**
ε_{tag}	0.680 %	0.004 %	±0.6 %
ε_{sig}	7.371 %	0.155 %	±2.1 %
N_{bkg}	9.14	1.62	±17.7 %

Table 4.19: Systematic uncertainties due to limited MC statistics

Besides these statistical errors of the uncorrected quantities additional uncertainties are related to the systematic corrections applied to determine $N_{\text{bkg}}^{\text{corr}}$, $\varepsilon_{\text{tag}}^{\text{corr}}$, as well as $\varepsilon_{\text{sig}}^{\text{corr}}$.

4.7.9.2 Systematic Corrections

The factors C_{bkg} and C_{tag}, which have been determined to correct N_{bkg} and ε_{tag} for data-MC disagreements, are associated with the statistical errors of the MC and data samples used for extraction. In the case of $N_{\text{bkg}}^{\text{corr}}$ the uncertainty is given by the statistical errors of the sideband regions, while the limited statistics of the double-tagged sample causes systematic uncertainties of $\varepsilon_{\text{tag}}^{\text{corr}}$.

As mentioned in Sec. 4.7.8.1, the ε_{sig} correction factor has been determined by an independent analysis of various control samples accounting for data-MC disagreements of the track reconstruction efficiencies. As a result, the GTL reconstruction efficiency has been corrected by −0.8 %. The uncertainty of this factor is given as 1.4 % [91]. Since the a_1 candidates are formed by three GTL candidates, the absolute uncertainty of the correction factor C_{sig} arises to 4.2 %.

Quantity	**Central Value**	**Error**	**Relative Error**
C_{tag}	1.033	0.015	±1.5 %
C_{sig}	0.976	0.042	±4.3 %
C_{bkg}	1.220	0.272	±22.3 %

Table 4.20: Systematic uncertainties of correction factors

The relative errors of the correction factors listed in Tab. 4.20 have been used as systematic uncertainties inserted by the systematic corrections. The high uncertainty of C_{bkg} is mainly induced by the small number of entries in sideband region SB1 in the combined MC as well as On Peak data sample.

4.7.9.3 N^{CN} Correction

$\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ as well as $N_{\mathrm{bkg}}^{\mathrm{corr}}$ are affected by the N^{CN} correction (Sec. 4.7.5.3) since this technique modified the $E_{\mathrm{rem}}^{\mathrm{CN}}$ distributions of and therefore influences the numbers of events contained in the signal region.

Quantity	Central Value	$+2\sigma$	-2σ
$N_{\mathrm{MC}}^{\mathrm{SB1}}$	40.41 ± 4.01	41.14 ± 4.09	41.48 ± 4.10
$N_{\mathrm{MC}}^{\mathrm{SB2}}$	621.42 ± 13.99	613.87 ± 13.91	623.49 ± 13.99
$N_{\mathrm{MC}}^{\mathrm{SB3}}$	77.66 ± 4.45	76.72 ± 4.42	82.95 ± 4.66
C_{bkg}	1.220 ± 0.272	1.199 ± 0.268	1.117 ± 0.249
N_{bkg}	9.14 ± 1.62	9.14 ± 1.62	9.31 ± 1.63
$N_{\mathrm{bkg}}^{\mathrm{corr}}$	11.16 ± 3.18	10.97 ± 3.13	10.40 ± 2.94
$\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}$	–	-0.19	-0.76
$\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}/N_{\mathrm{bkg}}^{\mathrm{corr}}$	–	$-1.7\,\%$	$-6.8\,\%$
$\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	0.06614 ± 0.00146	0.06477 ± 0.00145	0.06712 ± 0.00147
$\Delta\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	–	-0.00137	$+0.00098$
$\Delta\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}/\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	–	$-2.1\,\%$	$+1.5\,\%$

Table 4.21: Comparison of the luminosity scaled and branching fraction reweighted numbers of MC background events and corrected signal efficiency extracted with the central $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$ values, minimal killing/maximal adding ($+2\sigma$) and maximal killing/minimal adding (-2σ).

As shown in Fig. 4.71, the neutral multiplicity correction factors determined from all CN candidates of events passing the B_{tag} selection suffer from statistical uncertainties of the mean values μ_{Data} and μ_{MC} extracted from the $N^{\mathrm{CN}}(E, N^{\mathrm{GTL}})$ distributions. In order to extract systematic uncertainties induced by the statistical errors of $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$, all factors have been varied (correlated) by $+2\sigma$ (minimal killing/maximal adding) and -2σ (maximal killing/minimal adding).

Tab. 4.21 summarizes the effects of the variations on the number of background events in the $E_{\mathrm{rem}}^{\mathrm{CN}} - \mathcal{XNN}^{3\pi}$ sideband regions, the resulting corrected number of background events in the signal region , as well as the corrected signal reconstruction efficiency. The relative shifts $\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}/N_{\mathrm{bkg}}^{\mathrm{corr}}$ ($+0$, $-6.8\,\%$) and $\Delta\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}/\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ ($+1.5\,\%$, $-2.1\,\%$) have been assigned as systematic uncertainties. One should mention that the correlated variation of all N^{CN} correction factors results in maximal variation of $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ and $N_{\mathrm{bkg}}^{\mathrm{corr}}$. Thus, these N^{CN} correction uncertainties have to be considered as conservative.

4.7.9.4 *Dalitz*-Plot Reweighting

Besides the N^{CN} correction $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ depends on the weighting factors extracted from the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ control sample as discussed in Sec. 4.4.2. In analogy to the N^{CN} correction factors, the statistical uncertainties of these weighting factors have to be taken into account. The weights $w_{3\pi}(\sqrt{s_1}, \sqrt{s_2})$ with the invariant masses $\sqrt{s_{1,2}}$ of the two neutral $\pi\pi$ combinations of a given 3π candidate have been varied within $\pm 2\sigma$ and $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ have been recalculated. Again, since all weighting factors have been modified in one direction, this procedure corresponds to maximal variation.

Quantity	Central Value	$+2\sigma$	-2σ
$\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	0.06614 ± 0.00146	0.06839 ± 0.00149	0.06389 ± 0.00144
$\Delta\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	–	$+0.00225$	-0.00225
$\Delta\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}/\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$	–	$+3.4\,\%$	$-3.4\,\%$

Table 4.22: Comparison of the corrected signal efficiency extracted with the central $w_{3\pi}(\sqrt{s_1}, \sqrt{s_2})$ values and after variation within $\pm 2\sigma$.

The relative systematic uncertainty on $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ induced by this source has been found to be $\pm 3.4\,\%$ (Tab. 4.22).

4.7.9.5 3π Lineshape

An additional contribution to the systematic uncertainties of the corrected signal reconstruction efficiency arises from deviations of the 3π lineshape. Such a residual discrepancy has been found in the τ control sample after application of the *Dalitz*-plot reweighting (Fig. 4.11b,d).

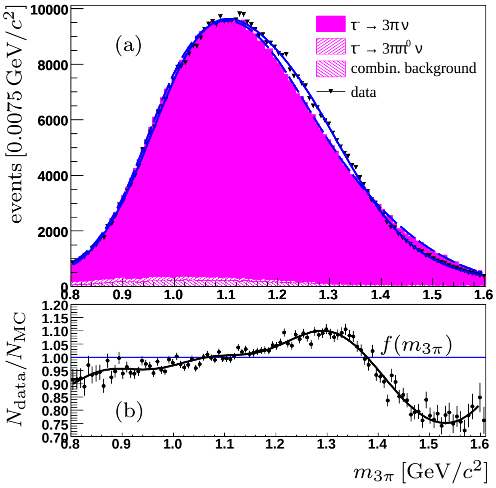


Figure 4.88: (a) illustrates $m_{3\pi}$ of the τ control samples after *Dalitz*-plot reweighting in the range $0.8 < m_{3\pi} < 1.6$ GeV/c^2 together with the fitted functions for MC (dashed line) and data (solid line), respectively. (b) shows the corresponding bin-by-bin ratios with the ratio of the fitted functions overlaid.

In order to determine the effect of this residual data-MC disagreement on the signal reconstruction efficiency, an additional correction function has been extracted from the τ control sample by fits to the invariant 3π masses in data and MC after the *Dalitz*-plot weights have been applied.

The $m_{3\pi}$ distributions (Fig. 4.88a) have been parametrized by the sum of three gaussians, respectively. Although this choice is not physically motivated, both shapes are well described by these functions. Once the $m_{3\pi}$ distributions have been parametrized, one can compute a $m_{3\pi}$ weighting function $f(m_{3\pi})$ as the ratio of the data and the MC functions. In Fig. 4.88b the resulting function has been overlaid to the bin-by-bin ratios $N_{\mathrm{data}}/N_{\mathrm{MC}}$ of the number of entries per bin. This parametrization of the residual data-MC disagreement found in the τ control sample has been used to extract an additional weighting factor $w(m_{3\pi})$ for the 3π candidates included in the $E_{\mathrm{rem}}^{\mathrm{CN}} - \mathcal{XNN}^{3\pi}$ signal region based on their invariant mass. The recalculation of $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ yields

$$\varepsilon_{\mathrm{sig}}^{\mathrm{corr}} = 0.06693 \pm 0.00147\,. \tag{4.53}$$

Thus, the additional weights correcting the a_1 lineshape increase the signal reconstruction efficiency by a factor of $0.06693/0.06614 = 1.012$. Therefore, a relative uncertainty of $+1.2\,\%$ has been assigned on $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ due to the residual 3π lineshape deviations between data and MC after application of the *Dalitz*-plot reweighting factors.

4.7.9.6 Branching Fraction Reweighting

Since the branching fraction weighting factors of the B_{tag} reconstruction modes have been applied in the calculation of $\varepsilon_{\mathrm{tag}}^{\mathrm{corr}}$, $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$, as well as $N_{\mathrm{bkg}}^{\mathrm{corr}}$, uncertainties of the branching fractions affect all three quantities.

In order to study the impact on these quantities, 500 different branching fraction sets of the decay channels contributing to the B_{tag} reconstruction have been generated, where every single branching fraction has been allowed to vary according to a gaussian distribution with mean equal to the central value and width equal the error as they have been listed in Tab. 4.9. This results in 500 different branching fraction reweighting factors for every MC event. These 500 weights have then been used to compute the quantities of interest in order to investigate their variations. Fig. 4.89 illustrates the resulting distribution of the expected number of background events, where for every single branching fraction set the correction factor C_{bkg} has been recalculated. The dashed line illustrates the result computed with the central values of the branching fractions.

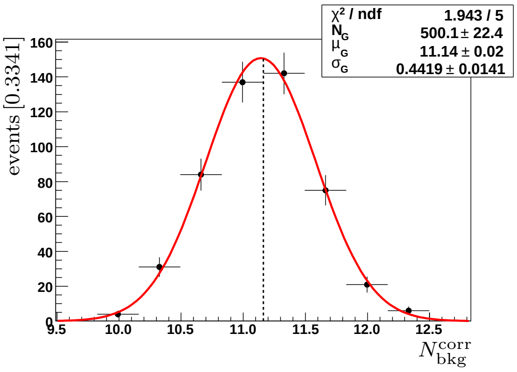


Figure 4.89: Variation of $N_{\mathrm{bkg}}^{\mathrm{corr}}$ due to Branching Fraction Reweighting

This distribution has been fitted to a gaussian with normalization N_G, mean μ_G, and width σ_G represented by the solid line. The fitted mean value μ_G has been found to be consistent with the central value of $N_{\mathrm{bkg}}^{\mathrm{corr}}$ confirmed quantitatively by Tab. 4.23. This table summarizes the central value, the fitted mean value, the fitted width, as well as the ratio σ_G/μ_G. It further contains the corresponding results of the fits to the distributions of $\varepsilon_{\mathrm{tag}}^{\mathrm{corr}}$ and $\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}$ illustrated in Fig. 4.90. Again, the dashed lines indicate the central values and the fitted mean values of the gaussians have been found to be consistent with them in both cases. One should remark that the small uncertainty of $\varepsilon_{\mathrm{tag}}^{\mathrm{corr}}$ is expected since the corrections extracted from the double-tagged sample have also been recalculated for every of the 500 branching fraction sets.

The ratios of the widths σ_G and the mean values μ_G of the fitted gaussians listed in Tab. 4.23 have been assigned as systematic uncertainties due to the branching fraction reweighting.

Quantity	**Central Value**	**Mean Value μ_G**	**Width σ_G**	**σ_G/μ_G**
$\varepsilon_{\text{tag}}^{\text{corr}}$	$0.6958\,\%$	$(0.6960 \pm 0.0002)\,\%$	$(0.0055 \pm 0.0002)\,\%$	$\pm 0.8\,\%$
$\varepsilon_{\text{sig}}^{\text{corr}}$	$6.614\,\%$	$(6.612 \pm 0.003)\,\%$	$(0.061 \pm 0.002)\,\%$	$\pm 0.9\,\%$
$N_{\text{bkg}}^{\text{corr}}$	11.16	11.14 ± 0.02	0.44 ± 0.01	$\pm 3.9\,\%$

Table 4.23: Variations of reconstruction efficiencies and expected number of background events obtained using 500 different branching fraction sets

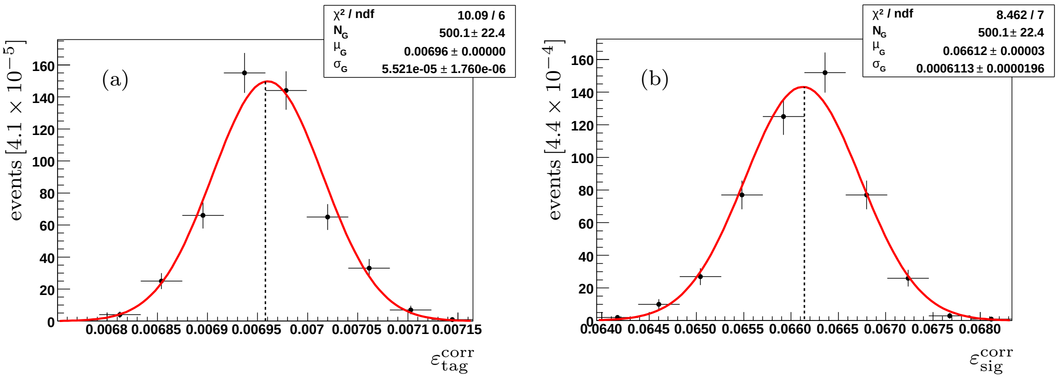


Figure 4.90: Variation of (a) $\varepsilon_{\text{tag}}^{\text{corr}}$ and (b) $\varepsilon_{\text{sig}}^{\text{corr}}$ due to Branching Fraction Reweighting

4.7.9.7 B Counting

Since the number of $\Upsilon(4S) \to B\overline{B}$ events included in the On Peak data sample is needed to calculate the number of signal events corresponding to a given branching fraction $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$, one has to take into account uncertainties of $N_{B\overline{B}}$ extracted by the B counting procedure introduced in Sec. 3.1. The systematic uncertainty of this number is quoted in Ref. [70] to be ±1.1 %, which directly translates into a relative systematic uncertainty of the number of signal events. The statistical error of $N_{B\overline{B}}$ has been neglected.

4.7.9.8 Branching Fraction $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$

Besides $N_{B\overline{B}}$ the branching fraction of the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ has to be accounted for the calculation of the number of signal events. As already noted in Sec. 2.5.1, the current world average of this branching fraction is given as [13]

$$\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) = (9.33 \pm 0.08)\,\% \tag{4.54}$$

corresponding to a relative uncertainty of ±0.9 %. This value has been assigned as a systematic uncertainty.

4.7.9.9 Cross-feed from Other τ Decays

The term "cross-feed" denotes $B^- \to \tau^-\overline{\nu}_\tau$ events with other τ final states X than $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ surviving the final selection. In principle, such events increase the signal efficiency corresponding to the $\tau \to X$ branching fractions. The relative shift of $\varepsilon_{\text{sig}}^{\text{corr}}$ can be calculated by

$$1 + \frac{\Delta\varepsilon_{\text{sig}}^{\text{corr}}}{\varepsilon_{\text{sig}}^{\text{corr}}} = \frac{\sum_{i=1}^{n} \mathcal{B}(\tau \to X_i) \cdot \varepsilon_i^{\text{corr}}}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) \cdot \varepsilon_{\text{sig}}^{\text{corr}}}\,, \tag{4.55}$$

where the index i denotes the τ decay channels (Tab. 2.6) with their branching fractions $\mathcal{B}(\tau \to X_i)$. In order to determine the reconstruction efficiencies $\varepsilon_i^{\text{corr}}$ of cross-feed events, the entire selection has been performed on a set of 641000 MC events of type $B^+ \to X$ vs. $B^- \to \tau^-\overline{\nu}_\tau$. Assuming a branching fraction of $\mathcal{B}_{\text{exp}}(B^- \to \tau^-\overline{\nu}_\tau) = 10^{-4}$ this sample corresponds to an integrated luminosity of about 28 times the total Run14 On Peak data sample.

No cross-feed has been found for $\tau^- \to e^-\overline{\nu}_e\nu_\tau$, $\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau$, $\tau^- \to \pi^-\nu_\tau$, $\tau^- \to \pi^-\pi^0\nu_\tau$, and $\tau^- \to \pi^-\pi^0\pi^0\nu_\tau$. Moreover, other τ decays not explicitly listed in Tab. 2.6, e.g. $\tau^- \to K^{*-}\nu_\tau$, are completely rejected by the B_{sig} selection. In general, one would expect a portion of $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ events to pass the entire selection since such events exhibit a similar signature as $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decays. However, extra neutral energy induced by the π^0 and the different 3π kinematics of this channel leads to strong suppression. In fact, one single $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ event of the given MC sample has been found in the $E_{\text{rem}}^{\text{CN}} - \mathcal{XNN}^{3\pi}$ signal region. This single event translates into a reconstruction efficiency of $\varepsilon_{3\pi\pi^0}^{\text{corr}} = (0.315 \pm 0.315)\,\%$, where the corrected B_{tag} reconstruction efficiency $\varepsilon_{\text{tag}}^{\text{corr}}$ determined from the signal MC sample has been used and corrections (branching fraction reweighting and track reconstruction efficiency correction) have been taken into account.

Using the given value of $\varepsilon_{3\pi\pi^0}^{\text{corr}}$ the relative shift of $\varepsilon_{\text{sig}}^{\text{corr}}$ can be determined from Eq. 4.55 with $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ as the only additional channel contributing to the signal region

$$1 + \frac{\Delta\varepsilon_{\text{sig}}^{\text{corr}}}{\varepsilon_{\text{sig}}^{\text{corr}}} = 1 + \frac{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau) \cdot \varepsilon_{3\pi\pi^0}^{\text{corr}}}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) \cdot \varepsilon_{\text{sig}}^{\text{corr}}}\,, \tag{4.56}$$

$$\frac{\Delta\varepsilon_{\text{sig}}^{\text{corr}}}{\varepsilon_{\text{sig}}^{\text{corr}}} = \frac{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau) \cdot \varepsilon_{3\pi\pi^0}^{\text{corr}}}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) \cdot \varepsilon_{\text{sig}}^{\text{corr}}}\,, \tag{4.57}$$

$$\frac{\Delta\varepsilon_{\text{sig}}^{\text{corr}}}{\varepsilon_{\text{sig}}^{\text{corr}}} = 2.3\,\% \tag{4.58}$$

with $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau) = 4.59\,\%$, $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) = 9.33\,\%$ (Tab. 2.6), and $\varepsilon_{\text{sig}}^{\text{corr}} = 6.614\,\%$. In principle, this relative shift has to be taken into account as an additional correction of $\varepsilon_{\text{sig}}^{\text{corr}}$, but given the high luminosity the aforementioned MC sample corresponds to, the single $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ event has been neglected and $\Delta\varepsilon_{\text{sig}}^{\text{corr}}/\varepsilon_{\text{sig}}^{\text{corr}}$ has been assigned as a systematic uncertainty of $+2.3\,\%$.

4.7.9.10 Summary of Systematic Uncertainties

The systematic uncertainties of $N_{\text{bkg}}^{\text{corr}}$, $\varepsilon_{\text{sig}}^{\text{corr}}$, and $\varepsilon_{\text{tag}}^{\text{corr}}$ are summarized in Tab. 4.24. The uncertainties due to the B counting (±1.1 %) and the branching fraction of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ (±0.9 %) are not listed.

Source	$\boldsymbol{\Delta N_{\text{bkg}}^{\text{corr}}/N_{\text{bkg}}^{\text{corr}}}$	$\boldsymbol{\Delta\varepsilon_{\text{sig}}^{\text{corr}}/\varepsilon_{\text{sig}}^{\text{corr}}}$	$\boldsymbol{\Delta\varepsilon_{\text{tag}}^{\text{corr}}/\varepsilon_{\text{tag}}^{\text{corr}}}$
Limited MC Statistics	±17.7 %	±2.1 %	±0.6 %
N_{bkg} Correction	±22.3 %	–	–
ε_{sig} Correction	–	±4.3 %	–
ε_{tag} Correction	–	–	±1.5 %
N^{CN} Correction	$^{+0}_{-6.8\,\%}$	$^{+1.5\,\%}_{-2.1\,\%}$	–
Dalitz-Plot Reweighting	–	±3.4 %	–
3π Lineshape	–	$^{+1.2\,\%}_{-0}$	–
Cross-feed	–	$^{+2.3\,\%}_{-0}$	–
Branch. Frac. Reweighting	±3.9 %	±0.9 %	±0.8 %
Total	$^{+28.7\,\%}_{-29.5\,\%}$	$^{+6.5\,\%}_{-6.3\,\%}$	±1.8 %

Table 4.24: Summary of systematic uncertainties: The contributions have been added in quadrature to obtain the total systematic uncertainties.

The uncertainties of the corrected tagging efficiency have been found to be of order 2 %, where the main contribution arises from the statistical uncertainty of the double-tagged sample used to correct for the data-MC normalization difference at the stage of the B_{tag} selection. No uncertainties related to event preselection or B_{tag} selection requirements have to be accounted for since such contributions are absorbed by the ε_{tag} correction procedure.

The most prominent uncertainties of the corrected signal efficiency are represented by uncertainties of the data-MC deviations of the track reconstruction efficiency and the statistical errors of the *Dalitz*-plot weights extracted from the τ control sample. The relative efficiency of the recoil selection is assumed to be well-reproduced by the N^{CN} corrected MC. $E_{\text{rem}}^{\text{CN}}$ has been validated from the double-tagged events. Furthermore, all selection requirements of the a_1 preselection as well as the multivariate a_1 selection are based on kinematical properties, which have been checked in the τ control sample and found to be in reasonable agreement after the *Dalitz*-plot reweighting. The residual difference seen in the 3π lineshape has been included as one systematic uncertainty. The total uncertainty has been calculated to be of order 6.5 %.

The largest uncertainties have been found for $N_{\text{bkg}}^{\text{corr}}$ and in fact the uncertainties of $\varepsilon_{\text{tag}}^{\text{corr}}$ and $\varepsilon_{\text{sig}}^{\text{corr}}$ play a secondary role in terms of the upper limit extraction. The total systematic uncertainty of $N_{\text{bkg}}^{\text{corr}}$ is quoted as about 30 % mainly driven by the statistical error of the N_{bkg} correction factor and the small number of MC background events surviving the final selection cuts.

Although the total systematic uncertainties are slightly asymmetric, they have been symmetrized in order to simplify the incorporation into the upper limit calculation. Thus, in the following the uncertainties have been used as

$$\frac{\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}}{N_{\mathrm{bkg}}^{\mathrm{corr}}} = \pm 29.5\,\%\,, \quad \frac{\Delta \varepsilon_{\mathrm{sig}}^{\mathrm{corr}}}{\varepsilon_{\mathrm{sig}}^{\mathrm{corr}}} = \pm 6.5\,\%\,, \quad \frac{\Delta \varepsilon_{\mathrm{tag}}^{\mathrm{corr}}}{\varepsilon_{\mathrm{tag}}^{\mathrm{corr}}} = \pm 1.8\,\%\,. \tag{4.59}$$

The next section illustrates the effect of these systematic uncertainties on $\mathrm{UL}_{\mathrm{exp}}^{14}$ and compares it to the corresponding results of the inclusive and semileptonic reconstruction techniques.

4.8 Comparison of Reconstruction Techniques

The systematic uncertainties have been determined and one can incorporate them into the expected upper limit calculation.

As already mentioned, the $N_{\mathrm{bkg}}^{\mathrm{corr}}$ uncertainties dominate other uncertainties reflected in the increase of $\mathrm{UL}_{\mathrm{exp}}^{14}$ after incorporation of $\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}$ relative to the corresponding increase after inclusion of all systematic uncertainties. $\Delta N_{\mathrm{bkg}}^{\mathrm{corr}}$ leads to

$$\mathrm{UL}_{\mathrm{exp}}^{14} = 8.28 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,, \tag{4.60}$$

while the incorporation of all systematic uncertainties yields

$$\mathrm{UL}_{\mathrm{exp}}^{14} = 8.30 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{4.61}$$

This clarifies the importance to estimate the expected number of background events with high accuracy. It should be noted that although the background expectations extracted within the inclusive and the semileptonic techniques have been corrected for overall normalization deviations between data and MC, no extensive systematic studies have been performed. In general, one expects additional contributions to the systematic uncertainty of N_{bkg} than it has been assumed in the expected upper limit calculations within both techniques (Secs. 4.5.4 and 4.6.3).

However, the qualities of the three selection techniques have been compared with respect to $\mathrm{UL}_{\mathrm{exp}}^{14}$ including systematic uncertainties. Tab. 4.25 summarizes the reconstruction efficiencies, the expected number of signal and background events (N_{sig}, N_{bkg}), the significances S_{exp}^{14}, the relative uncertainties N_{bkg}, and the expected upper limits $\mathrm{UL}_{\mathrm{exp}}^{14}$ with and without the N_{bkg} uncertainties. As mentioned above, for the inclusive and semileptonic techniques the uncertainties of N_{bkg} only include the statistical errors and the uncertainties of the systematic corrections of the continuum and $\Upsilon(4S)$ expectations. Hence, the N_{bkg} uncertainties given in Tab. 4.25 for the inclusive and semileptonic reconstruction do not reflect the accuracy of the background extraction. However, the limits including these N_{bkg} uncertainties are not competitive to the resulting $\mathrm{UL}_{\mathrm{exp}}^{14}$ extracted from the $B \to D^{(*)}\ell\nu_{\ell}$-tagged sample, which contains the entire systematic uncertainties of N_{bkg} from various sources as has been described in the previous section.

Quantity	**Inclusive**	**Semileptonic**	**Recoil of $B \to D^{(*)}\ell\nu_\ell$**
$\varepsilon_{\rm tag}/\%$	–	–	0.6958 ± 0.0125
$\varepsilon_{\rm sig}/\%$	–	–	6.614 ± 0.430
$\varepsilon_{\rm tot}/\%$	1.9361 ± 0.0096	0.2903 ± 0.0037	0.0460 ± 0.0031
$N_{\rm sig}$	41.87 ± 0.62	6.279 ± 0.119	0.995 ± 0.069
$N_{\rm bkg}$	21323.9 ± 265.4	1451.77 ± 29.89	11.16 ± 3.29
$S^{14}_{\rm exp}$	0.081	0.027	0.081
$\mathrm{UL}^{14}_{\rm exp}$	5.6×10^{-4}	8.8×10^{-4}	6.7×10^{-4}
$\Delta N_{\rm bkg}/N_{\rm bkg}$	$\pm 1.2\,\%$	$\pm 2.1\,\%$	$\pm 29.5\,\%$
$\mathrm{UL}^{14}_{\rm exp}$ (syst.)	10.3×10^{-4}	11.2×10^{-4}	8.3×10^{-4}

Table 4.25: Comparison of reconstruction efficiencies, expected numbers of events, significances, and upper limits extracted by the three different reconstruction techniques: For the inclusive and semileptonic reconstruction $N_{\rm sig}$ and $N_{\rm bkg}$ have been determined from the Run12 data set and have then been extrapolated to Run14 using the ratio of the integrated luminosities $\mathcal{L}^{14}_{\rm on}/\mathcal{L}^{12}_{\rm on} = 2.64$. The limits correspond to 90 % C.L.

In summary, three different searches for the decay $B^- \to \tau^-\overline{\nu}_\tau$ have been performed and the qualities of these methods have been compared with respect to the expected upper limit including uncertainties from the background estimation. The best result is provided by the recoil technique due to the very low background level and equal significance compared to the inclusive reconstruction. The vertex separation has been found to improve the multivariate a_1 selection within the inclusive and semileptonic reconstruction. On the other hand, the vertexing no longer provides separation power against $\Upsilon(4S) \to B\overline{B}$ after a $B_{\rm tag}$ selection performed within the recoil technique. However, a multivariate combination of kinematical properties of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay has been used to filter out signal events from the enormous background. The entire selection has been developed on the MC samples and systematic corrections of selection efficiencies and background estimations have been applied. The number of observed On Peak data events in the signal region and the calculation of the observed upper limit UL^{14} are given in the results chapter (Chap. 5).

Chapter 5

Physics Results

This chapter presents the number of events observed in the On Peak data sample after the final selection of the recoil technique and the translation into an upper limit for $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$. Moreover, the consequences of the result for B physics and prospects for future $B^- \to \tau^- \overline{\nu}_\tau$ searches are discussed.

5.1 Number of Observed Events

The two-dimensional $E_{\text{rem}}^{\text{CN}}$ vs. $\mathcal{XNN}^{3\pi}$ distribution extracted from the Run14 On Peak data sample after the $\mathcal{NN}^{3\pi} > 0$ selection is shown in Fig. 5.1. In contrast

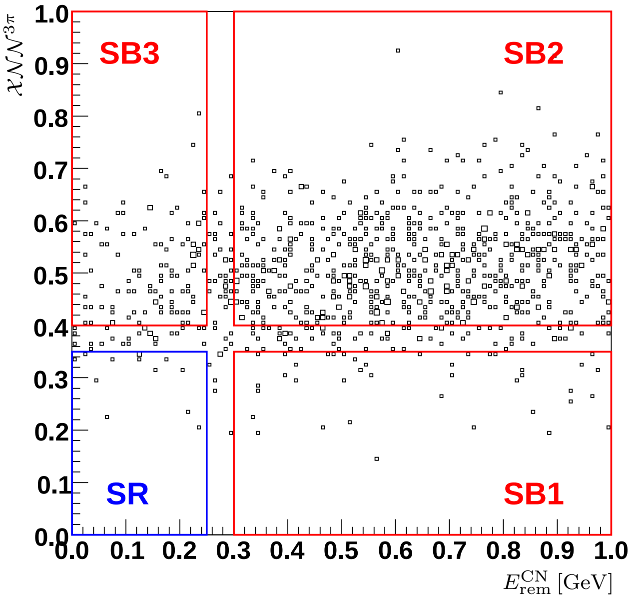


Figure 5.1: $E_{\text{rem}}^{\text{CN}}$ vs. $\mathcal{XNN}^{3\pi}$ from On Peak data after $\mathcal{NN}^{3\pi} > 0$

to Fig. 4.86b, the signal region is uncovered and the observed number of On Peak events after the final selection cuts has been found to be

$$N_{\mathrm{obs}} = 13\,, \tag{5.1}$$

which is consistent with the corrected number of background events determined from the MC samples (Eq. 4.51). Since no significant signal contribution is seen in data, N_{obs} has been translated into an upper limit.

5.2 Upper Limit Extraction

The technique to extract upper limits used in this thesis is based on the ***Frequentist*** approach in the sense that confidence intervals have been constructed, which include the true value of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ with a probability greater or equal to a specified level. The procedure described below has been originally introduced by *Neyman* [94] and the discussion follows the argumentation given in the statistics review section of Ref. [13].

5.2.1 *Neyman* Construction for Confidence Intervals

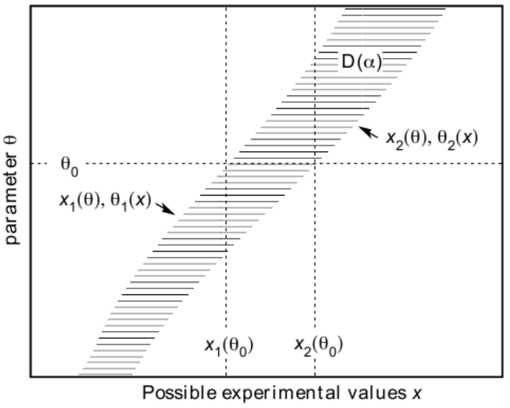


Figure 5.2: Construction of Confidence Intervals [13]

Given a function $f(x;\theta)$ describing the probability to measure a quantity x representing an estimator for an unknown parameter θ, a set of values $x_1(\theta,\alpha)$ and $x_2(\theta,\alpha)$ can be found with

$$\mathcal{P}(x_1 < x < x_2) = 1-\alpha = \int_{x_1}^{x_2} f(x;\theta)\mathrm{d}x \tag{5.2}$$

and a pre-specified probability $1-\alpha$. The intervals corresponding to a given α $[x_1(\theta,\alpha), x_2(\theta,\alpha)]$ form a confidence belt $D(\alpha)$ as visualized in Fig. 5.2 in the $\theta - x$ plane. The obtained value x_0 determined by a measurement of x corresponds to a vertical line in Fig. 5.2 and the confidence interval of θ to the given α is defined as the set of values of θ for which this line intersects the corresponding line segments $[x_1(\theta,\alpha), x_2(\theta,\alpha)]$. The resulting interval $[\theta_1(x,\alpha), \theta_2(x,\alpha)]$ is said to have a confidence level of $1-\alpha$.

An assumed true value θ_0 within $[\theta_1(x), \theta_2(x)]$ requires a measured x_0 between $x_1(\theta_0)$ and $x_2(\theta_0)$. Furthermore, the two events are of same probability

$$1-\alpha = \mathcal{P}(x_1(\theta) < x < x_2(\theta)) = \mathcal{P}(\theta_1(x) < \theta < \theta_2(x))\,, \tag{5.3}$$

where the subscript "0" has been dropped since the above argumentation is true for all values of θ_0. While θ stands for an unknown constant, the endpoints θ_1 and θ_2 in

Eq. 5.3 are random numbers. If one repeats the x measurement n times, $[\theta_1(x), \theta_2(x)]$ varies and for large n this interval covers the true value θ in $n \cdot (1-\alpha)$ experiments.

The coverage condition (Eq. 5.2) is not sufficient to extract x_1 and x_2 and an additional criterion is needed, e.g. the criterion of central intervals, which requires the probabilities below x_1 and above x_2 to be $\alpha/2$. If one wants to determine upper limits, the probability excluded below x_1 is set to zero and vice versa for lower limits.

The next section gives a detailed explanation, how upper limits have been calculated using the aforementioned *Frequentist* approach.

5.2.2 Upper Limit Extraction Procedure

Instead of constructing the entire confidence belt $D(\alpha)$ for a fixed confidence level C.L. $= 1-\alpha$, a scan of the quantity of interest $\theta = \mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ has been performed and upper limits have been extracted from hypothesis tests using the observed number of events N_{obs}.

For a given branching fraction $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ the number of signal events in a given data set can be calculated as

$$N_{\text{sig}}(\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)) = \left[2 \cdot \frac{N_{B\overline{B}}}{2} \cdot \varepsilon_{\text{tot}} \cdot \mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)\right] \cdot \mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)\,. \tag{5.4}$$

Here, the factor $2 \cdot \frac{N_{B\overline{B}}}{2}$ again accounts for the fact that both B mesons of the $\Upsilon(4S)$ event can decay via $B^- \to \tau^-\overline{\nu}_\tau$ and the half of all $\Upsilon(4S)$ mesons decay into a charged B pair (Sec. 4.1).

In order to perform a hypothesis test of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$, the branching fraction has been varied between zero and 20×10^{-4} in steps of 10^{-5}. For every scan point $n_{\text{toy}} = 10000$ toy experiments have been generated for $N_{\text{sig}}(\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau))$ and N_{bkg}, respectively, where *Poisson* distributed event numbers have been assumed

$$f(k; N_{\text{sig}}) = \frac{e^{-N_{\text{sig}}} N_{\text{sig}}^k}{k!}\,, \qquad f(l; N_{\text{bkg}}) = \frac{e^{-N_{\text{bkg}}} N_{\text{bkg}}^l}{l!}\,. \tag{5.5}$$

For every argument $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ this results in n_{toy} different expected numbers of observed events $k+l$. In order to illustrate the development of $f(k; N_{\text{sig}})$ as well as the $k+l$ distribution $f(k+l; N_{\text{sig}}, N_{\text{bkg}})$ with $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$, Fig. 5.3 shows the resulting distributions of k, l, and $k+l$ for six different branching fractions using the results of the recoil analysis. Naturally, the mean value of the background distribution $f(l; N_{\text{bkg}})$ does not depend on $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ and stays at $N_{\text{bkg}} \approx 11.16$. Since systematic uncertainties are not incorporated yet, the width of $f(l; N_{\text{bkg}})$ is given by the statistical error of $\sqrt{N_{\text{bkg}}} = 3.34$. In contrast, the mean value of $f(k; N_{\text{sig}})$ is proportional to $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ (Eq. 5.4) and therefore the width $\sqrt{N_{\text{sig}}}$ increases with increasing branching fraction.

The hypothesis of one given $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ to be the true branching fraction value for the decay $B^- \to \tau^-\overline{\nu}_\tau$ can be tested using the observed number of events.

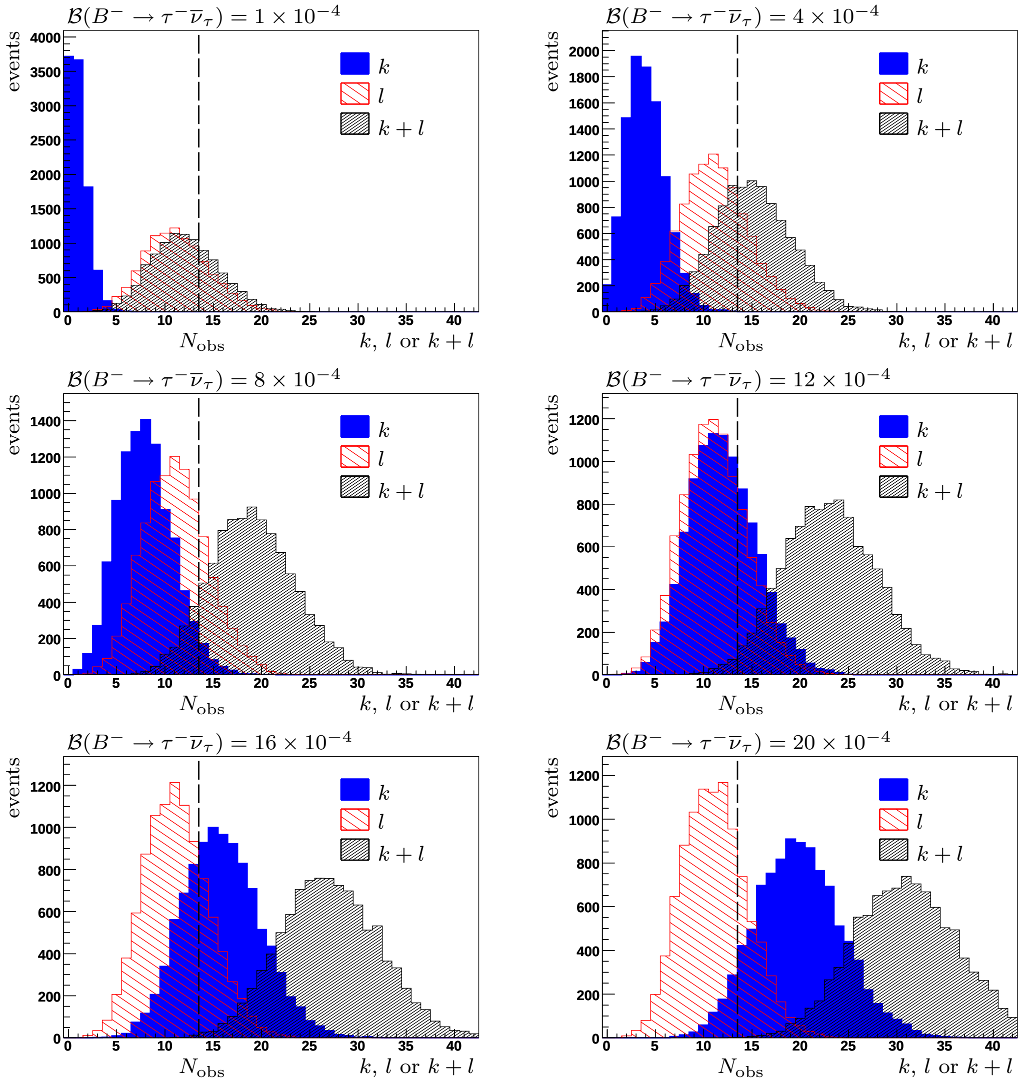


Figure 5.3: Random event numbers for upper limit extraction: The dashed line illustrates the observed number of events as extracted from On Peak data. The branching fractions used to calculate the mean value N_{sig} of $f(k; N_{\mathrm{sig}})$ are given above the figures. The distribution of background events $f(l; N_{\mathrm{bkg}})$ does not depend on $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$.

The corresponding confidence level can be determined from

$$\mathrm{C.L.} = 1 - \alpha = \mathcal{P}(k + l > N_{\mathrm{obs}}) = \frac{\int_{N_{\mathrm{obs}}}^{+\infty} f(k+l; N_{\mathrm{sig}}, N_{\mathrm{bkg}}) \mathrm{d}(k+l)}{\int_{-\infty}^{+\infty} f(k+l; N_{\mathrm{sig}}, N_{\mathrm{bkg}}) \mathrm{d}(k+l)} \tag{5.6}$$

with the probability $\mathcal{P}(k + l > N_{\mathrm{obs}})$ to find a total number of events $k + l$ above N_{obs}. Since the total number of observed events is dominated by the background expectation, an upper limit instead of a central interval has been determined, i.e. $\mathcal{P}(k + l \leq N_{\mathrm{obs}}) = \alpha$ and therefore $\mathcal{P}(k + l > N_{\mathrm{obs}}) = 1 - \alpha$.

The C.L. of Eq. 5.6 is a function of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ since the mean value and width of $f(k+l; N_{\mathrm{sig}}, N_{\mathrm{bkg}})$ increase with increasing branching fraction as already illustrated in Fig. 5.3. Technically, the C.L. has been calculated from the counted number $n_{\mathrm{toy}}(k+l > N_{\mathrm{obs}})$ of toy experiments with $k+l > N_{\mathrm{obs}}$. The C.L. is then equal to the fraction

$$\mathrm{C.L.} = \frac{n_{\mathrm{toy}}(k+l > N_{\mathrm{obs}})}{n_{\mathrm{toy}}} \tag{5.7}$$

at a given $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ scan point.

The dependence of $1-\mathrm{C.L.}$ on the branching fraction is illustrated in Fig. 5.4 using the background expectation and the reconstruction efficiencies determined within the recoil analysis. That value of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ with $1-\mathrm{C.L.} = 0.1$ is called the upper limit at 90 % C.L. Since the $1-\mathrm{C.L.}$ curve has been constructed from a $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ scan, it is not a continuous but a discrete function. Thus, in order to find that branching fraction value with exactly $1-\mathrm{C.L.} = 0.1$, it has been linearly interpolated between two scan points. However, as shown by the lines in Fig. 5.4 for the results of the recoil analysis an upper limit of

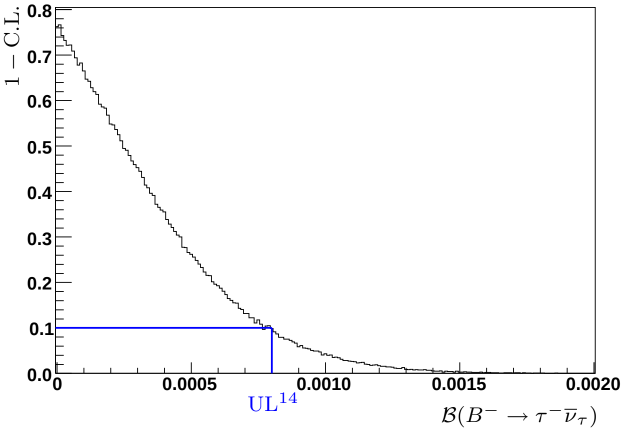


Figure 5.4: $1-\mathrm{C.L.}$ as a function of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$: The lines illustrate the upper limit at 90 % C.L.

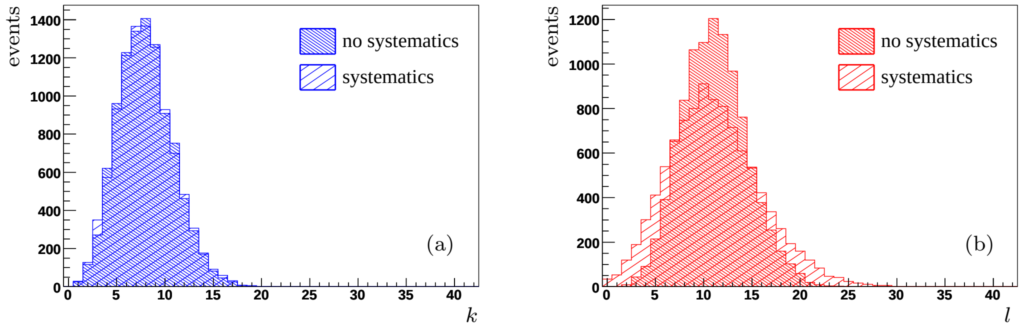
$$\mathrm{UL}^{14} = 8.0 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.}) \tag{5.8}$$

has been calculated from $N_{\mathrm{obs}} = 13$ representing the entire Run14 On Peak data sample using the above procedure.

It should be noted here that all expected upper limits mentioned in the analysis section have been calculated using this algorithm, where N_{obs} has been set to the sum of the expected numbers of background and signal events. The signal expectation has always been calculated assuming a branching fraction value of $\mathcal{B}_{\mathrm{exp}}(B^- \to \tau^- \overline{\nu}_\tau) = 10^{-4}$. On the other hand, it has been claimed several times that systematic uncertainties affect the upper limit considerably. Sec. 5.2.3 explains how such uncertainties have been included in the upper limit calculations.

5.2.3 Incorporation of Systematic Uncertainties

The upper limit at 90 % C.L. obtained from the observed number of On Peak data events $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) < 8.0 \times 10^{-4}$ only corresponds to statistical fluctuations of $N_{\rm bkg}$ as well as $N_{\rm sig}$ introduced by the *Poisson* distributed random experiments (Eq. 5.5). But the reconstruction efficiencies, the number of $B\overline{B}$ events included in the data set, the branching fraction of $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, and the expected number of background events are not perfectly known (Sec. 4.7.9.10) leading to uncertainties of the mean values of the *Poisson* distributions $f(k; N_{\rm sig})$ and $f(l; N_{\rm bkg})$.



In order to incorporate the uncertainties, these quantities are randomly smeared, i.e. for every toy experiment to a given branching fraction new values of $N_{\rm bkg}^{\rm corr}$, $\varepsilon_{\rm tag}^{\rm corr}$, $\varepsilon_{\rm sig}^{\rm corr}$, $N_{B\overline{B}}$, as well as $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$ have been generated assuming the systematic uncertainties to be gaussian distributed. In general, systematic uncertainties are not forced to satisfy gaussian distributions and in fact the total uncertainties of $N_{\rm bkg}^{\rm corr}$ and $\varepsilon_{\rm sig}^{\rm corr}$ have been found to be slightly asymmetric (Tab. 4.24). On the other hand, this assumption is justifiable since the dominant contributions induced by the limited MC statistics and the systematic corrections are related to the statistical errors of the MC and control samples and are therefore can be assumed to be gaussian-like.

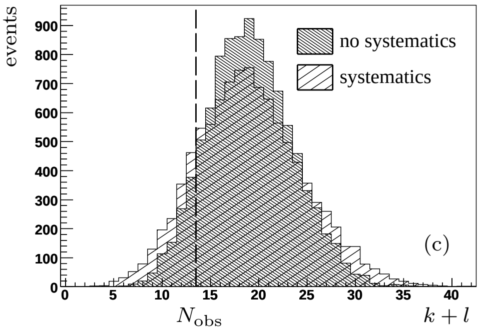


Figure 5.5: Effect of the incorporation of systematic uncertainties on (a) $f(k; N_{\rm sig})$, (b) $f(l; N_{\rm bkg})$, and (c) $f(k+l; N_{\rm sig}, N_{\rm bkg})$ for the scan point $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) = 8 \times 10^{-4}$. The dashed line in (c) illustrates $N_{\rm obs}$.

Exemplarily, the effect of this gaussian smearing on $f(k; N_{\rm sig})$ and $f(l; N_{\rm bkg})$ for $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) = 8 \times 10^{-4}$ is shown in Figs. 5.5a,b. As already mentioned in Sec. 4.8, the influence of uncertainties of the quantities affecting $N_{\rm sig}$, such as $\varepsilon_{\rm tag}^{\rm corr}$, $\varepsilon_{\rm sig}^{\rm corr}$, $N_{B\overline{B}}$, and $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$, are small compared to the $N_{\rm bkg}$ uncertainty. The resulting broadening of $f(k+l; N_{\rm sig}, N_{\rm bkg})$ illustrated in Fig. 5.5c is mainly driven by the uncertainty of $N_{\rm bkg}$. Moreover, Fig. 5.5c clarifies that due to the increase of the fraction of toy experiments with $k + l \leq N_{\mathrm{obs}}$ the broadening necessarily results in a higher upper limit.

The effect of the systematic uncertainties on $1 - \mathrm{C.L.}$ is given in Fig. 5.6 and it confirms the increase of UL^{14}. The final result of the 90 % C.L. upper limit determined from the search in the recoil of $B \to D^{(*)}\ell\nu_\ell$ is

$$\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau) < 9.4 \times 10^{-4} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{5.9}$$

In the next section the quality of this result has been judged with respect to the results of former analyses and the usability of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel for future analyses.

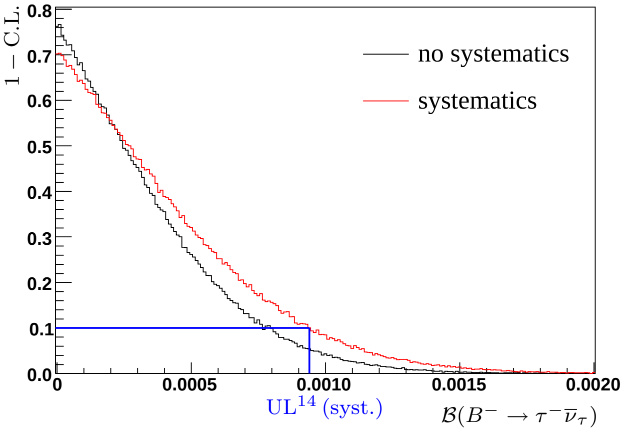


Figure 5.6: Comparison of the $1 - \mathrm{C.L.}$ curves before and after incorporation of systematic uncertainties: The lines illustrate the upper limit at 90 % C.L.

5.3 Comparison with Other Analyses

A comparison of the final upper limit given in the previous section with the recent results from the *BABAR* and *Belle* collaborations (Tab. 2.4) clarifies that the search for the decay $B^- \to \tau^-\overline{\nu}_\tau$ presented in this thesis is not competitive to these analyses. On the other hand, one has to note here that the presented upper limit of $\mathrm{UL}^{14} = 9.4 \times 10^{-4}$ has been obtained from only one τ decay channel with a branching fraction of roughly 10 %. The recent measurements performed at the B factories used combinations of the most dominant τ decay channels, such as $\tau^- \to e^-\overline{\nu}_e\nu_\tau$, $\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau$, $\tau^- \to \pi^-\nu_\tau$, $\tau^- \to \pi^-\pi^0\nu_\tau$, and $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. These five channels account for about 80 % of the total τ decay rate (Tab. 2.6).

Tab. 2.4 contains the result of a former *BABAR* analysis published in spring 2006 [32], which has been performed on the same data sample (Run14). This analysis is based on a semileptonic reconstruction of one B meson in $B \to D^*\ell\nu_\ell$ and searches for $B^- \to \tau^-\overline{\nu}_\tau$ decays in the recoil in a quite similar way as it has been described here.

This measurement used all five aforementioned τ decay channels, which automatically increases the signal sample considerably. However, in order to judge the quality of the selection technique presented in this thesis, one has to compare the results obtained from $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ only.

Quantity	**This Analysis**	**Former *BABAR* Analysis (2006)**
$\varepsilon_{\rm tag}/\%$	0.6958 ± 0.0125	0.1750 ± 0.0090
$\varepsilon_{\rm sig}/\%$	6.614 ± 0.430	14.469 ± 2.069
$\varepsilon_{\rm tot}/\%$	0.0460 ± 0.0031	0.0253 ± 0.0038
$N_{\rm sig}$	0.995 ± 0.069	0.548 ± 0.083
$N_{\rm bkg}$	11.16 ± 3.29	21.62 ± 3.01
$S^{14}_{\rm exp}$	0.081	0.014
$\mathrm{UL}^{14}_{\rm exp}$	6.7×10^{-4}	14.4×10^{-4}
$\mathrm{UL}^{14}_{\rm exp}$(syst.)	8.3×10^{-4}	16.4×10^{-4}
$N_{\rm obs}$	13	26
UL^{14}	8.0×10^{-4}	22.3×10^{-4}
UL^{14}(syst.)	9.4×10^{-4}	25.1×10^{-4}

Table 5.1: Comparison with former *BABAR* analysis [32]: The upper limits corresponding to the results of the former *BABAR* analysis have been calculated by the procedure explained in Sec. 5.2.2 since no τ decay channel dependent limits have been quoted in Ref. [32]. The signal reconstruction efficiency of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel is given in Ref. [32] as $\varepsilon_{\rm sig} = (1.4 \pm 0.07 \pm 0.05)\,\%$. For comparison only this value has been divided by $\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$ since the analysis presented in this thesis determined $\varepsilon_{\rm sig}$ from a signal MC sample, where the τ always decays via $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. All upper limits correspond to 90 % C.L.

Tab. 5.1 summarizes the results of the presented selection compared to the corresponding values obtained from the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ sample of the former *BABAR* analysis. The lower tagging efficiency is caused by the fact that the former analysis explicitly restricts the $B_{\rm tag}$ to be reconstructed in $B \to D^*\ell\nu_\ell$. The $B \to D\ell\nu_\ell$ channel has not been reconstructed. Furthermore, a cut-based selection has been applied resulting in a less efficient $B_{\rm tag}$ reconstruction. Even though the cut-based 3π candidate selection in the recoil of the $D^*\ell$ candidate yields higher signal side reconstruction efficiency, $S^{14}_{\rm exp}$ of the former *BABAR* selection is not competitive to the expected significance achieved by the selection technique presented in this thesis. This is not a consequence of the additional $B_{\rm tag}$ reconstruction mode $B \to D\ell\nu_\ell$. It is rather accomplished by the exploitation of kinematical properties and their combination using multivariate algorithms.

In comparison, while the expected number of signal events assuming a branching fraction of 10^{-4} has been doubled, the background expectation dropped by a factor of two. The observed upper limit UL^{14} at 90 % C.L. on $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ from $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ including systematic uncertainties decreased by a factor of three. Hence, the reconstruction techniques presented in this thesis result in a considerable improvement of the $B^- \to \tau^-\overline{\nu}_\tau$ selection for the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel and multivariate algorithms have been established as important tools for sufficient background suppression.

Due to the very low significance of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel found by the *BABAR* analysis mentioned before, this channel has not been reconstructed by the most recent *BABAR* measurements, which have been presented on ICHEP06 [33] and FPCP07 [41]. As one important conclusion it can be stated at this point that the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay has been reopened for usage within future searches for the decay $B^- \to \tau^-\overline{\nu}_\tau$, even if an improvement of the significance of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ can probably only be achieved by combinations of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel with the other main τ channels.

5.4 Implications for *B* Physics

As discussed in Sec. 2.4, a precise measurement of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ has consequences for the understanding of the physics on the B sector since it implies information about the B decay constant f_B, the CKM matrix element V_{ub}, and further possibly allows access to a non-SM *Higgs* sector. Although no significant signal has been found, the observed upper limit on $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ can be used to set a limit on f_B (Sec. 5.4.1), to map out allowed regions in the $\bar{\rho}-\bar{\eta}$ plane (Sec. 5.4.2), and to exclude charged *Higgs* masses depending on $\tan\beta$ (Sec. 5.4.3).

5.4.1 Consequences for f_B

Assuming absence of non-SM contributions to the decay $B^- \to \tau^-\overline{\nu}_\tau$ the branching fraction directly measures $f_B|V_{ub}|$ (Sec. 2.4.1). Using the observed 90 % confidence limit $\mathrm{UL}^{14} = 9.4 \times 10^{-4}$ an upper limit for $f_B|V_{ub}|$ can be calculated from Eq. 2.33

$$\frac{f_B}{191\,\mathrm{MeV}}\frac{|V_{ub}|}{0.00410} < \sqrt{\frac{\mathrm{UL}^{14}}{1.08 \times 10^{-4}}} \tag{5.10}$$

$$\frac{f_B}{191\,\mathrm{MeV}}\frac{|V_{ub}|}{0.00410} < 2.95 \tag{5.11}$$

and one obtains

$$f_B|V_{ub}| < 2.31\,\mathrm{MeV} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{5.12}$$

Thus, a limit of f_B particularly depends on $|V_{ub}|$. The average $|V_{ub}|$ value of the inclusive and exclusive measurements (Eq. 2.32) has been used to extract an upper limit of f_B. Since the uncertainty of $|V_{ub}|$ is dominated by the theoretical error, a value of $|V_{ub}| = 3.71 \times 10^{-3}$ has been assumed, where the theoretical uncertainty of 0.39×10^{-3} has been subtracted from the central value. We obtain (Eq. 5.12)

$$f_B < 622\,\mathrm{MeV} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{5.13}$$

In order to demonstrate the potential of the decay $B^- \to \tau^-\overline{\nu}_\tau$, f_B has also been calculated using the central value of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ extracted from the combined likelihoods of the summer 2006 BABAR and *Belle* results (Eq. 2.34) to

$$f_B|V_{ub}| = (0.87 \pm 0.16_{\text{ex}})\,\text{MeV}\,. \tag{5.14}$$

The error of 0.16 MeV is induced by the experimental uncertainties of the branching fraction from the combined likelihoods[1]. Using the $|V_{ub}|$ average from Eq. 2.32 f_B takes a value of

$$f_B = (213 \pm 38_{\text{ex}} \pm 20_{\text{th}})\,\text{MeV}\,, \tag{5.15}$$

where the relative experimental uncertainties of $f_B|V_{ub}|$ (Eq. 5.14) and $|V_{ub}|$ (Eq. 2.32) have been added in quadrature. The precision of f_B is still limited by the statistical error of the branching fraction measurements. Hence, more data or more sophisticated analyses are required to improve the experimental knowledge of f_B, e.g. multivariate procedures as described in this thesis.

5.4.2 Constraint from $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ and Δm_d

It has been claimed in Sec. 2.4.2 that the measured $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ in combination with the $B^0 - \overline{B}^0$ oscillation frequency Δm_d is usable to exclude areas in the $\bar{\rho} - \bar{\eta}$ plane with low theoretical uncertainties.

Three SM fits with different sets of constraints have been performed on the $\bar{\rho} - \bar{\eta}$ plane [10, 11, 95]. For all of them the CKM matrix elements $|V_{ud}|$, $|V_{us}|$, and $|V_{cb}|$ have been constrained to the current experimental values [12] and the B decay constant f_B is taken from an unquenched LQCD calculation [22] as given in Tab. 2.3. The three fits (F1, F2, and F3) have been performed with different additional constraints from

- F1: the $1 -$ C.L. curve including the systematic uncertainties[2] from Fig. 5.6
- F2: Δm_d as extracted from the $B^0 - \overline{B}^0$ oscillation measurements (Eq. 2.41)
- F3: a combination of both aforementioned quantities

Fig. 5.7a illustrates $1 -$ C.L. scans over the $\bar{\rho} - \bar{\eta}$ plane for every of the three SM fits, where the corresponding constraints from $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ represented by the $1 -$ C.L. curve and Δm_d are indicated by different colours. As already mentioned in Fig. 2.2a, $|V_{ub}|$ is related to the length of the side of the unitarity triangle joining $(0, 0)$ and the apex. Since for a given f_B an upper limit on $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ is equivalent to an upper limit of $|V_{ub}|$, the constraint from the $1 -$ C.L. curve results in an allowed region indicated by a disk around $(0, 0)$. In contrast, Δm_d measures the magnitude of the CKM matrix element V_{td} and therefore the experimental value constrains the apex on a ring around $(1, 0)$ as illustrated by the yellow region in Fig. 5.7a. The blue disk as well as the yellow ring both include the theoretical uncertainties of f_B, while the Δm_d constraint additionally contains uncertainties of the bag parameter B_{B^0} as

---

[1] The relative error of $f_B|V_{ub}|$ is equal to the relative error of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ divided by two since $f_B|V_{ub}|$ is proportional to $\sqrt{\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)}$.

[2] For the fits $1 -$ C.L. has been translated into a χ^2 function.

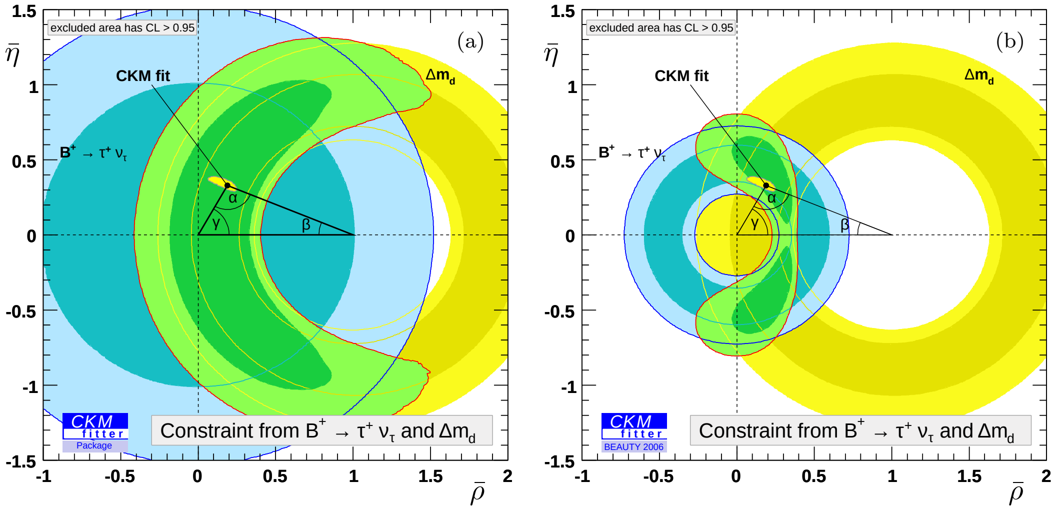


Figure 5.7: Constraints from $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ and Δm_d [10, 11, 95]: (a) shows the constraints from $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ (blue disk), from Δm_d (yellow ring), and from the combination of both quantities (green banana) [95]. (b) illustrates the same constraints, where $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ is represented by the combined likelihoods of the summer 2006 *BABAR* and *Belle* measurements [10]. Excluded areas are at C.L > 95 %. For comparison, the result of the global SM fit is overlaid (Fig. 2.2b). The dark shaded areas illustrate the 1σ contours.

introduced in Sec. 2.4.2. If one combines the constraints from $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ and Δm_d, the allowed region is given by the overlap of both constraints resulting in the green "banana" visualized in Fig. 5.7a. Due to the combination of both constraints the dependence on f_B drops out as claimed in Sec. 2.4.2 and therefore this region does not rely on any theoretical uncertainties of f_B, even though it still includes the uncertainties induced by B_{B^0}.

The abovementioned procedure has been repeated using the combined likelihoods of the summer 2006 *BABAR* and *Belle* measurements instead of the 1 − C.L. curve determined in this analysis. Since from the combined likelihood function $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) = 0$ is excluded by about 2.8 standard deviations, the blue disk consequently shrinks to a ring around $(0, 0)$ as can be seen in Fig. 5.7b, which results in a combined constraint indicated by the green coil. A more precise measurement of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ results in a more narrow blue ring around $(0, 0)$ and therefore in a more narrow allowed region for the apex $(\bar\rho, \bar\eta)$ of the unitarity triangle. Thus, the decay $B^- \to \tau^- \overline{\nu}_\tau$ considerably contributes to the understanding of B physics within the framework of the SM. However, as already mentioned in Sec. 2.4.3.2 this decay also possibly allows access to non-SM physics.

5.4.3 Limits on Charged *Higgs* Boson Effects

Within the 2HDM model (Sec. 2.4.3.2) charged *Higgs* boson contributions to the decay $B^- \to \tau^- \overline{\nu}_\tau$ could result in either enhancement or suppression, i.e. such effects cause deviations of the measured branching fraction of $B^- \to \tau^-\overline{\nu}_\tau$ and the corresponding SM prediction as given in Eq. 2.33. It has been shown that within this model the ratio of the measured and predicted branching fraction r_H depends on $\tan\beta/m_{H^-}$ (Eq. 2.50) with the ratio of the vacuum expectation values of the two *Higgs* fields $\tan\beta$ and the charged *Higgs* mass m_{H^-}. The observed upper limit UL^{14} can be used to set an upper limit on r_H, which consequently limits $\tan\beta/m_{H^-}$.

In order to find a limit of $\tan\beta/m_{H^-}$, the r_H limit have been calculated from UL^{14} and $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)$ from Eq. 2.33. Using the same argument as given for $|V_{ub}|$ for the f_B limit calculation in Sec. 5.4.1, a value of $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau) = 0.74 \times 10^{-4}$ has been used, where the theoretical uncertainty of 0.34×10^{-4} has been subtracted from the central value of $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau) = 1.08 \times 10^{-4}$. An upper limit of

$$r_H < \frac{\mathrm{UL}^{14}}{\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)} = 12.7 \quad (90\,\%\,\mathrm{C.L.}) \tag{5.16}$$

has been obtained. With the "+" solution of Eq. 2.51 and a charged B meson mass of $m_{B^-} = 5.279\,\mathrm{GeV}/c^2$ this limit on r_H translates into

$$\frac{\tan\beta}{m_{H^-}} < 0.404\,(\mathrm{GeV}/c^2)^{-1} \quad (90\,\%\,\mathrm{C.L.})\,. \tag{5.17}$$

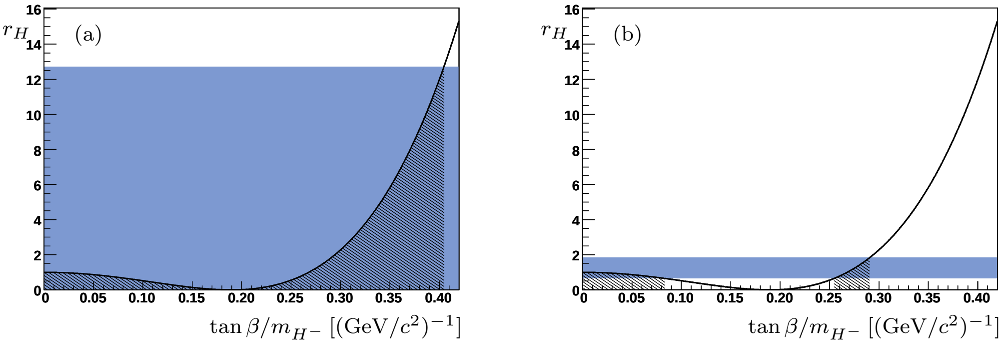


Figure 5.8: Limits on charged *Higgs* contributions in $B^- \to \tau^-\overline{\nu}_\tau$: Both plots show r_H as a function of $\tan\beta/m_{H^-}$ (black line). The blue bands indicate (a) the upper limit on r_H calculated from the observed upper limit UL^{14} from this analysis and (b) $r_H \pm 1\sigma$ calculated from $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ extracted from the combined summer 2006 *BABAR* and *Belle* likelihoods (Eq. 2.34). The bands contain the uncertainties of the SM expectation induced by f_B and $|V_{ub}|$. The hatched areas illustrate the allowed regions of $\tan\beta/m_{H^-}$.

This result is illustrated in Fig. 5.8a. It shows the function $r_H(\tan\beta/m_{H^-})$ given by Eq. 2.50 with the upper limit on r_H overlaid. The r_H shape clarifies that enhancement ($r_H > 1$) requires $\tan\beta/m_{H^-} > 0.27\,(\mathrm{GeV}/c^2)^{-1}$ and suppression implies two solutions for $\tan\beta/m_{H^-}$. However, all values below the given limit are allowed.

In contrast, Fig. 5.8b displays the constraint on $\tan\beta/m_{H^-}$ from r_H calculated from the combined *BABAR* and *Belle* results at the state of ICHEP06. Here, r_H

has been extracted from the HFAG value (Eq. 2.34) and the aforementioned SM expectation to be

$$r_H = \frac{\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)}{\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)} = 1.24\,^{+0.45}_{-0.44}\,^{+0.31}_{-0.39}\,. \tag{5.18}$$

The first errors are experimental dominated by the errors of the *BABAR* and *Belle* branching fraction measurements. The second ones are the theoretical uncertainties of $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)$ induced by the uncertainties of f_B and $|V_{ub}|$. These errors have been added in quadrature resulting in the $\pm 1\sigma$ band illustrated in Fig. 5.8b. Since the measured central value of 1.34×10^{-4} exceeds the SM prediction of 1.08×10^{-4}, one would expect only one solution for $\tan\beta/m_{H^-}$ as has been claimed in Sec. 2.4.3.2. On the other hand, due to the high uncertainties of about 50 % $r_H < 1$ is not excluded resulting in an allowed area in the low $\tan\beta/m_{H^-}$ region.

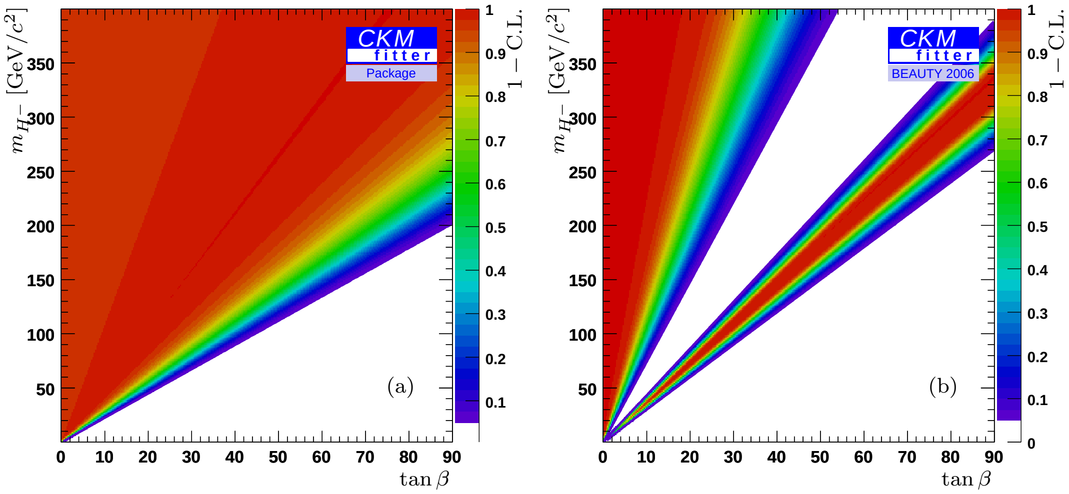


Figure 5.9: Constraint from $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ in the $\tan\beta - m_{H^-}$ plane [10, 11, 95]: (a) illustrates $1 - $C.L. from the fit using the result of this analysis. (b) shows the fit result, where the $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ constraint is represented by the combination of the *BABAR* and *Belle* likelihoods. The excluded (white) areas are at C.L. > 95 %.

In order to visualize the constraints from the measured upper limit UL[14]in the $\tan\beta - m_{H^-}$ plane, a fit has been performed [10, 11, 95], while in contrast to the SM fit mentioned in the previous section the fit model explicitly includes charged *Higgs* boson effects in the decay $B^- \to \tau^-\overline{\nu}_\tau$ corresponding to the 2HDM model. Again, the CKM parameters $|V_{ud}|$, $|V_{us}|$, and $|V_{cb}|$ have been constrained to the experimental values. Furthermore, the average value from the inclusive and exclusive results of $|V_{ub}|$ (Eq. 2.32) and the unquenched LQCD calculation of f_B [22] have been used. Similarly to the SM fit discussed in Sec. 5.4.2, the branching fraction measurement is represented by the $1 - $C.L. curve extracted from the analysis presented in this thesis. After the fit the $\tan\beta - m_{H^-}$ plane has been scanned for the confidence level. Fig. 5.9a illustrates the constraint from the presented measurement in the $\tan\beta - m_{H^-}$ plane.

This figure illustrates the behaviour of $\tan\beta/m_{H^-}$ already mentioned in Fig. 5.8a since the upper limit of $\tan\beta/m_{H^-}$ (Eq. 5.17) translates into a lower limit of

$$\frac{m_{H^-}}{\tan\beta} > \frac{1}{0.404\,(\mathrm{GeV}/c^2)^{-1}} = 2.48\,\mathrm{GeV}/c^2 \quad (90\,\%\,\mathrm{C.L.}) \tag{5.19}$$

confirmed by Fig. 5.9a. Moreover, since the $1-\mathrm{C.L.}$ curve including systematic uncertainties given in Fig. 5.6 is maximal at zero, the maximum of $1-\mathrm{C.L.}$ in Fig. 5.9a corresponds to $r_H = 0$, which implies $m_{H^-}/\tan\beta = m_{B^-}$ (Eq. 2.50).

The aforementioned fit has been repeated with the combined *BABAR* and *Belle* likelihoods representing the $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ measurements. As already claimed in Fig. 5.8b, there are two ambiguities visible in the $1-\mathrm{C.L.}$ scan (Fig. 5.9b); the low ("$-$") and the high ("$+$") $\tan\beta/m_{H^-}$ solution of Eq. 2.51. An improved knowledge of r_H results in more narrow allowed regions in the $\tan\beta - m_{H^-}$ plane. If $r_H < 1$ could be completely excluded, the low $\tan\beta/m_{H^-}$ solution vanishes and only one narrow band remains. However, as given in Eq. 5.18, currently the experimental and theoretical uncertainties are of compareable size. Thus, an improvement of r_H not only requires a more precise direct measurement of $\mathcal{B}(B^- \to \tau^-\overline{\nu}_\tau)$ but also demands an improved knowledge of f_B and $|V_{ub}|$. Currently, these parameters limit the precision of the SM expectation value $\mathcal{B}_{\mathrm{SM}}(B^- \to \tau^-\overline{\nu}_\tau)$.

5.5 Prospects of $B^- \to \tau^-\overline{\nu}_\tau$ using $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$

Although the search for the process $B^- \to \tau^-\overline{\nu}_\tau$ presented in this thesis has been found to be not competitive to former analyses due to the limitation on the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel, this section briefly discusses chances to improve the significance using this τ decay mode.

5.5.1 Maximum Likelihood Fit

The search in the recoil of $B \to D^{(*)}\ell\nu_\ell$ has been performed as a typical Cut-and-Count analysis. An increase of the significance achieved by this technique could possible be provided by a maximum likelihood fit in the $E_{\mathrm{rem}}^{\mathrm{CN}} - \mathcal{XNN}^{3\pi}$ plane. As given in Sec. 4.7.7.2, analytical functions have been fitted to the corresponding distributions used later on for the optimization of the final cut values. Such functions can be fitted to the corresponding distributions of signal MC as well as all relevant background components. After normalization on the fit ranges the resulting ***Probability Density Functions*** **(PDF's)** can be combined to a likelihood function using the product ansatz. Afterwards, the contributions from signal and background components to the On Peak data sample can be determined from an extended maximum likelihood fit.

Although such a fit increases the statistical significance, additional systematic uncertainties are expected to enter. $E_{\mathrm{rem}}^{\mathrm{CN}}$ and $\mathcal{XNN}^{3\pi}$ are not uncorrelated, which is assumed by using a product ansatz for the combination of the PDF's. Moreover, especially the $E_{\mathrm{rem}}^{\mathrm{CN}}$ shape of the background expectation has been found to be imperfectly modeled by the MC simulation (Fig. 4.87) and the limited MC statistics further induces uncertainties of the PDF's describing the background shapes. It remains unclear, if the expected increase of the significance compensates such additional systematic effects related to a maximum likelihood fit resulting in an improved upper limit compared to the Cut-and-Count analysis presented in this thesis.

5.5.2 Partial $B \to D^* \ell \nu_\ell$ Reconstruction

It has been shown that the significance of a $B^- \to \tau^- \overline{\nu}_\tau$ search on the $\Upsilon(4S)$ resonance strongly depends on the treatment of the second B meson. The three techniques presented in this thesis have been compared and the best expected upper limit has been achieved by the recoil technique. However, an additional B_{tag} reconstruction technique should be mentioned, which possibly could provide an improvement of the significance.

The technique of partially reconstructed $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$ decays is based on the kinematical properties of the two-body $\overline{D}^{*0} \to \overline{D}^0 \pi^0$ decay. In the $\overline{D}^{*0}$ rest frame the magnitude of the π^0 momentum is well-defined by a combination of the nominal values of the $\overline{D}^{*0}$, the $\overline{D}^0$, and the π^0 masses. The *Lorentz*-Transformation of the π^0 momentum vector from the $\overline{D}^{*0}$ rest into the c.m. frame results in an allowed region of the magnitude of the $\overline{D}^{*0}$ three-momentum $\vec{p}_{D^*}$. The borders $f^{\pm}(|\vec{p}_\pi|)$ of this region depend on the magnitude of the transformed π^0 three-momentum $\vec{p}_\pi$ in the c.m. frame, i.e. $|\vec{p}_{D^*}|$ is limited to a region defined by $|\vec{p}_\pi|$ as measured in the c.m. frame. This fact is illustrated in Fig. 5.10a. It shows $|\vec{p}_{D^*}|$ vs. $|\vec{p}_\pi|$ in the c.m. frame determined from a sample of true MC $\overline{D}^{*0} \to \overline{D}^0 \pi^0$ decays, where the $\overline{D}^{*0}$ are originated in the process $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$. The aforementioned borders, which correspond to parallel $(f^+(|\vec{p}_\pi|))$ or antiparallel $(f^-(|\vec{p}_\pi|))$ emission of the π^0 in the $\overline{D}^*$ rest frame relative to the $\overline{D}^{*0}$ flight direction in the c.m. frame, are nonlinear functions of $|\vec{p}_\pi|$ and can be used to calculate an estimator for $|\vec{p}_{D^*}|$ as the average $[f^+(|\vec{p}_\pi|) + f^-(|\vec{p}_\pi|)]/2$, which linearly depends on $|\vec{p}_\pi|$. The $\overline{D}^{*0}$ momentum

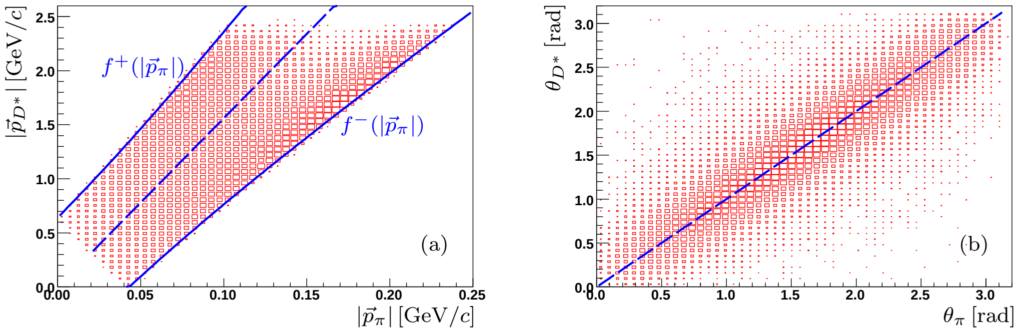


Figure 5.10: Prospect of Partial B_{tag} Reconstruction: (a) illustrates the dependence of the magnitude of the D^* momentum on the magnitude of the soft π^0 momentum in the c.m. frame. (b) shows the relation of the polar angles of the D^* and the π^0 in the c.m. frame. The $|\vec{p}_{D^*}|$ and θ_{D^*} estimators are illustrates by the dashed lines. The solid lines in (a) indicate the calculated borders of $|\vec{p}_{D^*}|$ as a function of $|\vec{p}_\pi|$.

resolution achieved by this approximation has been found to be about 800 MeV/c. Due to the low $D^* - D$ mass difference, the π^0 is emitted with low momentum discussed in Sec. 4.7.3.2. Consequently, the flight direction of the soft π^0 nearly corresponds to the direction of the $\overline{D}^{*0}$ three-momentum vector as illustrated for the polar angle θ in Fig. 5.10b. Therefore, the $\overline{D}^{*0}$ four-momentum vector p_{D^*} can be approximated using the estimated $|\vec{p}_{D^*}|$ and the flight direction of the soft π^0. The angular resolutions in θ and the azimuthal angle ϕ are of compareable size and have been found to be about 17° (300 mrad).

To be more explicit, using the abovementioned approximation for $\overline{D}^{*0} \to \overline{D}^0\pi^0$ decays p_{D^*} can be calculated from the detected $\pi^0 \to \gamma\gamma$ candidate without any reconstruction of the $\overline{D}^0$. Once p_{D^*} has been extracted, it can be combined with the four-momentum of a lepton candidate detected in the tracking devices and all $D^*\ell$ quantities explained within the recoil selection (Sec. 4.7.3.4) can be exploited, e.q. in multivariate methods. Furthermore, a $B_{\rm tag}$ decay vertex can be reconstructed from all tracks not belonging to a given $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ candidate and the vertex separations in xy and z may be useful for background rejection. Moreover, the charges of the 3π and the lepton candidate are anti-correlated for correctly reconstructed $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$ vs. $B^- \to \tau^-\overline{\nu}_\tau$ events.

The advantage of this technique lies in the fact that the $\overline{D}^0$ is not explicitly reconstructed resulting in higher $B_{\rm tag}$ reconstruction efficiencies since all $\overline{D}^0$ decay channels enter the selection. On the other hand, due to the ignorance of the $\overline{D}^0$ a higher background level compared to the search in the recoil of $B \to D^{(*)}\ell\nu_\ell$ is expected since the track and neutral candidates originated in the $\overline{D}^0$ decay can not be removed from the entire event to perform a recoil analysis. However, this technique in combination with a multivariate $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ selection could possibly increase the significance of a search for the decay $B^- \to \tau^-\overline{\nu}_\tau$.

Chapter 6

Summary and Conclusion

This thesis presented three different methods to perform searches for the purely leptonic decay $B^- \to \tau^- \overline{\nu}_\tau$. All of them used the 3-prong τ lepton decay channel $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, which has been found to provide a high potential for these searches. Kinematical properties of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay as well as the separation of the decay vertices of the τ decay and the decay of the companion B meson (Δ^{xy},Δ^z) in the $\Upsilon(4S) \to B^+B^-$ event have been found to be exploitable for the reconstruction of $B^- \to \tau^- \overline{\nu}_\tau$ events. The difference of the three techniques manifest in the treatment of the companion B meson (B_{tag}).

Within the inclusive reconstruction the B_{tag} has not been restricted to particular decay modes and its flavour has been determined from an inclusive tagging algorithm. Kinematical features, e.g. the correlation of the invariant masses of the two neutral $\pi\pi$ combinations of a given $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ candidate (*Dalitz*-plot), have been combined with the aforementioned vertex separation and the information returned by the tagging algorithm using multivariate methods (ANN). These methods have been established as important tools since they allow more efficient signal event selection and background suppression.

Furthermore, within the semileptonic technique events have been required to include high-energetic leptons aiming at the selection of events of type $B^+ \to X^0\ell^+\nu_\ell$ vs $B^- \to \tau^- \overline{\nu}_\tau$. Again, the kinematical properties of the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel and the vertex separation have been exploited using multivariate algorithms, but this method has been found to be not competitive with the aforementioned inclusive selection.

The inclusive as well as the semileptonic reconstructions both suffer from high background levels and the separation of the τ and B_{tag} decay vertices have been found to be not sufficient for background rejection, even though the multivariate selections have been improved by including Δ^{xy} and Δ^z. Thus, within a third selection technique $B^- \to \tau^- \overline{\nu}_\tau$ events have been searched for in the recoil of explicitly reconstructed B_{tag} candidates ($B \to D^{(*)}\ell\nu_\ell$). Although such a recoil selection restricts the companion B meson to a set of exclusive reconstruction modes resulting in low selection efficiencies, this technique has been found to provide compareable significance as achieved by the inclusive reconstruction. Simultaneously, the background level has been decreased considerably. Therefore, the recoil technique has been chosen to search for $B^- \to \tau^- \overline{\nu}_\tau$ events in the data sample.

All selection steps have been developed using the MC simulation and extensive systematic studies have been performed in order to determine uncertainties of the expected number of background events N_{bkg} and the reconstruction efficiencies (ε_{tag}, ε_{sig}). A relative systematic uncertainty of $\pm 29.5\,\%$ has been extracted for $N_{\text{bkg}} = 11.16$. After the entire recoil selection we find $N_{\text{obs}} = 13$ events in the data sample, which is consistent with the aforementioned background expectation, i.e. no significant signal has been found.

Using the *Frequentist* approach a method has been developed to calculate an upper limit from N_{bkg}, ε_{tag}, ε_{sig}, and N_{obs} for a given data sample and one obtains

$$\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) < 8.0 \times 10^{-4} \quad (90\,\%\,\text{C.L.})\,. \tag{6.1}$$

This upper limit only includes statistical fluctuations of the number of signal and background events. After incorporation of the systematic uncertainties the limit increases to

$$\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) < 9.4 \times 10^{-4} \quad (90\,\%\,\text{C.L.}) \tag{6.2}$$

mainly driven by the uncertainties of N_{bkg}.

This result is not competitive with former measurements performed at the B factories (*BABAR* and *Belle*). On the other hand, this upper limit has been obtained using one τ decay channel only, while the most recent measurements reconstructed at least four of the most prominent τ decay channels. The most recent *BABAR* result has been achieved using the decay channels $\tau^- \to e^- \overline{\nu}_e \nu_\tau$, $\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$, $\tau^- \to \pi^- \nu_\tau$ and $\tau^- \to \pi^- \pi^0 \nu_\tau$. Since a former *BABAR* analysis published in spring 2006 claimed a negligible contribution of the decay $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ to the significance of the measurement, this channel has been explicitly excluded from the most recent *BABAR* measurement. In comparison with the aforementioned former *BABAR* analysis, the significance of signal events reconstructed in $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ and the resulting upper limit of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ have been considerably improved by the selection techniques presented in this thesis.

One can conclude that the usage of multivariate algorithms exploiting the kinematical properties of the decay $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ reopened this τ decay channel to perform future searches for the decay $B^- \to \tau^- \overline{\nu}_\tau$. On the other hand, an improvement of the current precision of $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ can probably only be achieved by combinations of the $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ channel with other prominent τ decay channels. However, using new techniques, e.g. partial $B \to D^* \ell \nu_\ell$ reconstruction, in combination with multivariate procedures a precise measurement of the branching fraction of the decay $B^- \to \tau^- \overline{\nu}_\tau$ could possibly succeed. This thesis explained the importance of such a measurement and it is worth to exploit all alternatives to measure $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ in order to extend our knowledge of physics within the framework of the SM as well as beyond.

Bibliography

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

[2] Glashow, S. L., *Partial Symmetries of Weak Interactions*, **Nucl. Phys. 22**(1961):579–588.

[3] Weinberg, S., *A Model of Leptons*, **Phys. Rev. Lett. 19**(1967):1264–1266.

[4] Goldstone, J., et al., *Broken Symmetries*, **Phys. Rev. 127**(1962):965–970.

[5] **Super-Kamiokande** Collaboration, Ashie, Y. et al., *A measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I*, **Phys. Rev. D71**(2005):112005, hep-ex/0501064.

[6] **SNO** Collaboration, Ahmed, S. N. et al., *Measurement of the total active B-8 solar neutrino flux at the Sudbury Neutrino Observatory with enhanced neutral current sensitivity*, **Phys. Rev. Lett. 92**(2004):181301, nucl-ex/0309004.

[7] Higgs, P. W., *Spontanous symmetry breakdown without massless bosons*, **Phys. Rev. 145**(1966):1156–1163.

[8] Schmuser, P., *Feynman diagrams and gauge theories for experimentists* (Springer, Stuttgart, Leipzig, 1988).

[9] Wolfenstein, L., *Parametrization of the Kobayashi-Maskawa Matrix*, **Phys. Rev. Lett. 51**(1983):1945.

[10] **CKMfitter Group** Collaboration Available from: `http://ckmfitter.in2p3.fr`.

[11] **CKMfitter Group** Collaboration, Charles, J. et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, **Eur. Phys. J. C41**(2005):1–131, hep-ph/0406184.

[12] **CKMfitter Group** Collaboration, *Updated Results on the CKM Matrix and the Unitarity Triangle* (2006). Available from: `http://www.slac.stanford.edu/xorg/ckmfitter/plots_beauty06/ckmEval_results_beauty06.pdf`.

[13] Yao, W.-M. et al., *Review of Particle Physics*, **J. Phys. G33**(2006):1+. Available from: `http://pdg.lbl.gov`.

[14] (HFAG), T. H. F. A. G., *Updates of semileptonic results for Summer 2006* Available from: `http://www.slac.stanford.edu/xorg/hfag/semi/summer06/home.shtml`.

[15] Lange, B. O., et al., *Theory of charmless inclusive B decays and the extraction of* V_{ub}, **Phys. Rev. D72**(2005):073006, hep-ph/0504071.

[16] Buchmuller, O. et al., *Fits to moment measurements from* $B \to X_c \ell \nu$ *and* $B \to X_s \gamma$ *decays using heavy quark expansions in the kinetic scheme*, **Phys. Rev. D73**(2006):073008, hep-ph/0507253.

[17] **HPQCD** Collaboration, Gulez, E., et al., *B meson semileptonic form factors from unquenched lattice QCD*, **Phys. Rev. D73**(2006)(7):074502. Available from: `http://link.aps.org/abstract/PRD/v73/e074502`.

[18] Okamoto, M. et al., *Semileptonic* $D \to \pi/K$ *and* $B \to \pi/D$ *decays in* $2+1$ *flavor lattice QCD*, **Nucl. Phys. Proc. Suppl. 140**(2005):461–463, hep-lat/0409116.

[19] Abada, A. et al., *Heavy* → *light semileptonic decays of pseudoscalar mesons from lattice QCD*, **Nucl. Phys. B619**(2001):565–587, hep-lat/0011065.

[20] Ball, P. et al., *New results on* $B \to \pi K \eta$ *decay form factors from light-cone sum rules*, **Phys. Rev. D71**(2005)(1):014015. Available from: `http://link.aps.org/abstract/PRD/v71/e014015`.

[21] **HPQCD** Collaboration, Gray, A. et al., *The B meson decay constant from unquenched lattice QCD*, **Phys. Rev. Lett. 95**(2005):212001, hep-lat/0507015.

[22] **JLQCD** Collaboration, Aoki, S. et al., $B^0 \overline{B}^0$ *mixing in unquenched lattice QCD*, **Phys. Rev. Lett. 91**(2003):212001, hep-ph/0307039.

[23] **CP-PACS** Collaboration, Khan, A. A. et al., *B meson decay constant from two-flavor lattice QCD with non-relativistic heavy quarks*, **Phys. Rev. D64**(2001):054504, hep-lat/0103020.

[24] Penin, A. A. et al., *Heavy-light meson decay constant from QCD sum rules in three-loop approximation*, **Phys. Rev. D65**(2002):054006, hep-ph/0108110.

[25] Davies, C., *Precision lattice QCD calculations and predictions of fundamental physics in heavy quark systems* (2006), hep-lat/0608018.

[26] Miller, D. E., *Lattice QCD calculation for the physical equation of state* (2006), hep-ph/0608234.

[27] **L3** Collaboration, Acciarri, M. et al., *Measurement of* $D_s^- \to \tau^- \overline{\nu}_\tau$ *and a new limit for* $B^- \to \tau^- \overline{\nu}_\tau$, **Phys. Lett. B396**(1997):327–337.

[28] **ALEPH** Collaboration, Barate, R. et al., *Measurements of* $\mathcal{B}(b \to \tau^- \overline{\nu}_\tau X)$ *and* $\mathcal{B}(b \to \tau^- \overline{\nu}_\tau D^{*+-} X)$ *and Upper Limits on* $\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau)$ *and* $\mathcal{B}(b \to s\nu\overline{\nu})$, **Eur. Phys. J. C19**(2001):213–227, hep-ex/0010022.

[29] **CLEO** Collaboration, Browder, T. E. et al., *A search for* $B \to \tau\nu$, **Phys. Rev. Lett. 86**(2001):2950–2954, hep-ex/0007057.

[30] ***BABAR*** Collaboration, Aubert, B. et al., *A search for* $B^- \to \tau^- \overline{\nu}_\tau$ *recoiling against a fully reconstructed* B (2003), hep-ex/0304030.

[31] ***BABAR*** Collaboration, Aubert, B. et al., *Search for the rare leptonic decay* $B^- \to \tau^- \overline{\nu}_\tau$, **Phys. Rev. Lett. 95**(2005):041804, hep-ex/0407038.

[32] ***BABAR*** Collaboration, Aubert, B. et al., *A search for the decay* $B^+ \to \tau^+ \nu_\tau$, **Phys. Rev. D73**(2006):057101, hep-ex/0507069.

[33] ***BABAR*** Collaboration, Aubert, B., *A search for* $B^+ \to \tau^+ \nu_\tau$ *recoiling against* $B^- \to D^0 \ell^- \overline{\nu}_\ell X$ (2006), hep-ex/0608019.

[34] ***BABAR*** Collaboration, Aubert, B., *A search for* $B^+ \to \tau^+ \nu_\tau$ (2007), arXiv:0705.1820 [hep-ex]. Available from: `http://www.arxiv.org/abs/0705.1820`.

[35] **Belle** Collaboration, Abe, K. et al., *Search for* $B^- \to \tau^- \overline{\nu}_\tau$ *at Belle* (2004), hep-ex/0408144.

[36] **Belle** Collaboration, Abe, K. et al., *Search for* $B \to \tau\nu$ *and* $B \to K\nu\overline{\nu}$ *decays with a fully reconstructed* B *at Belle* (2005), hep-ex/0507034.

[37] **Belle** Collaboration, Ikado, K., *Hot topics from Belle experiment*, **ECONF C060409**(2006):003, hep-ex/0605068.

[38] **Belle** Collaboration, Ikado, K. et al., *Evidence of the purely leptonic decay* $B^- \to \tau^- \overline{\nu}_\tau$ (2006), hep-ex/0604018.

[39] **Belle** Collaboration, Browder, T., *Rare B decays with missing energy at Belle* Prepared for XXXIII INTERNATIONAL CONFERENCE ON HIGH ENERGY PHYSICS, Moscow, Russia, 26. July - 2. August 2006. Available from: `http://ichep06.jinr.ru/reports/287_8s5_15p18_Browder(1).ppt`.

[40] (HFAG), T. H. F. A. G., *Updates on electroweak radiative B decays, August 2006* Available from: `http://www.slac.stanford.edu/xorg/hfag/rare/ichep06/radll/OUTPUT/TABLES/radll.pdf`.

[41] ***BABAR*** Collaboration, Gritsan, A., *Hot topics from BABAR* Prepared for 5th Flavour Physics and CP Violation Conference, Bled, Slovenia, 12. - 16. May 2007. Available from: `http://indico.cern.ch/materialDisplay.py?contribId=2&sessionId=29&materialId=slides&confId=5905`.

[42] ***BaBar*** Collaboration, Aubert, B. et al., *Measurement of the B^0 lifetime and the $B^0\overline{B}^0$ oscillation frequency using partially reconstructed $\overline{B}^0 \rightarrow D^{*+}\ell^-\overline{\nu}_\ell$ decays*, **Phys. Rev. D73**(2006):012004, hep-ex/0507054.

[43] Inami, T. et al., *Effects of superheavy quarks and leptons in low-energy weak processes $K_L \rightarrow \mu\overline{\nu}_\mu$, $K^+ \rightarrow \pi^+\nu\overline{\nu}$ and $K^0 \longleftrightarrow \overline{K}^0$*, **Prog. Theor. Phys. 65**(1981):297.

[44] Buras, A. J. et al., *Heavy Flavours II* (World Scientific, Physics department of the Technical University Munich, 1998).

[45] Buras, A. J., et al., *$B^0\overline{B}^0$ Mixing, CP Violation and the B meson decay*, **Nucl. Phys. B245**(1984):369.

[46] Buras, A. J., et al., *Leading and next-to-leading QCD corrections to epsilon parameter and $B^0\overline{B}^0$ mixing in the presence of a heavy top quark*, **Nucl. Phys. B347**(1990):491–536.

[47] Shrock, R. E. et al., *$K^0 \longleftrightarrow \overline{K}^0$ transition amplitude in the MIT bag model*, **Phys. Rev. D19**(1979):2148–2157.

[48] (HFAG), T. H. F. A. G., *B Lifetime and Oscillation Parameters: Results after the summer and fall 2006 conferences* Available from: `http://www.slac.stanford.edu/xorg/hfag/osc/end_2006/`.

[49] Gunion, J. F., et al., *The Higgs Hunter's Guide* SCIPP-89/13.

[50] Marin, C. A., *Higgs phenomenology in the two Higgs doublet model of type II* (2004), hep-ph/0409207.

[51] Chen, C.-H. et al., *Charged Higgs on $B^- \rightarrow \tau\overline{\nu}_\tau$ and $\overline{B} \rightarrow P(V)\ell\overline{\nu}_\ell$* (2006), hep-ph/0608166.

[52] Hou, W.-S., *Enhanced charged Higgs boson effects in $B^- \rightarrow \tau^-\overline{\nu}_\tau$, $\mu^-\overline{\nu}_\mu$ and $b \rightarrow \tau\overline{\nu}_\tau + X$*, **Phys. Rev. D48**(1993):2342–2344.

[53] Perl, M. L. et al., *Evidence for anomalous lepton production in e^+e^- annihilation*, **Phys. Rev. Lett. 35**(1975):1489–1492.

[54] Schmidtler, M., *Review of the three pion structure in τ decay*, **Nucl. Phys. Proc. Suppl. 76**(1999):271–282.

[55] **OPAL** Collaboration, Akers, R. et al., *Measurement of the hadronic decay current in $\tau^- \rightarrow \pi^-\pi^-\pi^+\nu_\tau$*, **Z. Phys. C67**(1995):45–56.

[56] **OPAL** Collaboration, Ackerstaff, K. et al., *A measurement of the hadronic decay current and the ν_τ helicity in $\tau^- \rightarrow \pi^-\pi^-\pi^+\nu_\tau$*, **Z. Phys. C75**(1997):593–605.

[57] Davier, M., et al., *The physics of hadronic τ decays* (2005), hep-ph/0507078.

[58] Kuhn, J. H. et al., *τ decays to pions*, **Z. Phys. C48**(1990):445–452.

[59] Isgur, N., et al., *The a_1 in τ decay*, **Phys. Rev. D39**(1989):1357.

[60] Feindt, M., *Measuring Hadronic Currents and Weak Coupling constants in $\tau \to \nu 3\pi$*, **Z. Phys. C48**(1990):681–688.

[61] Li, B. A., *Theory of τ mesonic decays*, **Phys. Rev. D55**(1997):1436–1452, hep-ph/9606402.

[62] Jadach, S., et al., *The τ decay library TAUOLA: Version 2.4*, **Comput. Phys. Commun. 76**(1993):361–380.

[63] **CLEO** Collaboration, Coan, T. et al., *Measurement of α_s from τ decays*, **Phys. Lett. B356**(1995):580–588.

[64] **DELPHI** Collaboration, Abreu, P. et al., *A Study of the hadronic resonance structure in the decay $\tau \to 3\pi + \nu_\tau$*, **Phys. Lett. B426**(1998):411–427.

[65] **MAC** Collaboration, Band, H. et al., *Measurement of τ decays to three pions*, **Phys. Lett. B198**(1987):297–301.

[66] **ARGUS** Collaboration, Albrecht, H. et al., *Analysis of the decay $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$ and determination of the $a_1(1260)$ resonance parameters*, **Z. Phys. C58**(1993):61–70.

[67] **ARGUS** Collaboration, Albrecht, H. et al., *Determination of the Michel parameters ξ and δ in leptonic τ decays*, **Phys. Lett. B349**(1995):576–584.

[68] ***BABAR*** Collaboration, Aubert, B. et al., *The BABAR detector*, **Nucl. Instrum. Meth. A479**(2002):1–116, hep-ex/0105044.

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

[70] ***BABAR*** Collaboration, Hearty, C., *Measurement of the Number of $\Upsilon(4S)$ Mesons Produced in Run 1 (B Counting)*, ***BABAR* Analysis Document 134** (2001).

[71] Brun, R., et al., *ROOT, an object oriented data analysis framework* Prepared for CERN School of Computing (CSC 2000), Marathon, Greece, 17-30 Sep 2000.

[72] ***BABAR*** Collaboration, De Nardo, G. et al., *User defined data in the new analysis model of the BABAR experiment* To appear in the proceedings of IEEE 2003 Nuclear Science Symposium (NSS) and Medical Imaging Conference (MIC) including Sessions on Nuclear Power Systems, Portland, Oregon, 19-24 Oct 2003.

[73] ***BABAR*** Collaboration, Aubert, B. et al., *Measurement of branching fractions and charge asymmetries in* B^+ *decays to* $\eta\pi^+$, ηK^+, $\eta\rho^+$ *and* $\eta'\pi^+$, *and search for* B^0 *decays to* ηK^0 *and* $\eta\omega$, **Phys. Rev. Lett. 95**(2005):131803, hep-ex/0503035.

[74] **Belle** Collaboration, Chang, P. et al., *Measurements of branching fractions and* CP *asymmetries in* $B \to \eta h$ *decays*, **Phys. Rev. D71**(2005):091106, hep-ex/0412043.

[75] Sjostrand, T., *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, **Comput. Phys. Commun. 82**(1994):74–90.

[76] Andersson, B., et al., *Parton fragmentation and string dynamics*, **Phys. Rept. 97**(1983):31.

[77] Barberio, E. et al., *PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0*, **Comput. Phys. Commun. 79**(1994):291–308.

[78] Lange, D. J., *The EvtGen particle decay simulation package*, **Nucl. Instrum. Meth. A462**(2001):152–155.

[79] **GEANT4** Collaboration, Agostinelli, S. et al., *GEANT4: A simulation toolkit*, **Nucl. Instrum. Meth. A506**(2003):250–303.

[80] ***BABAR*** Collaboration, Banerjee, S., *CM2 skims for* $e^+e^- \to \mu^+\mu^-$ *and* $e^+e^- \to \tau^+\tau^-$, ***BABAR* Analysis Document 760** (2003).

[81] Fox, G. C. et al., *Event shapes in* e^+e^- *annihilation*, **Nucl. Phys. B149**(1979):413.

[82] Fisher, R. A., *The use of multiple measurements in taxonomic problems*, **Annals Eugen. 7**(1936):179–188.

[83] Gay, P., et al., *Tagging Higgs bosons in hadronic LEP2 events with neural networks* Prepared for 4th International Workshop on Software Engineering and Artificial Intelligence for High-energy and Nuclear Physics (AIHENP 95), Pisa, Italy, 3-8 April 1995.

[84] ***BABAR*** Collaboration, Martinez-Vidal, F., *The BABAR Vertexing*, ***BABAR* Analysis Document 102** (2001).

[85] ***BABAR*** Collaboration, Aubert, B. et al., *A study of time dependent CP-violating asymmetries and flavor oscillations in neutral B decays at the* $\Upsilon(4S)$, **Phys. Rev. D66**(2002):032003, hep-ex/0201020.

[86] ***BABAR*** Collaboration, Aubert, B. et al., *Improved measurement of CP asymmetries in* $B^0 \to (c\bar{c})K^{(*)0}$ *decays*, **Phys. Rev. Lett. 94**(2005):161803, hep-ex/0408127.

[87] ***BABAR*** Collaboration, Bruinsma, M., *A B Tagging Algorithm for CP Violation Measurements with the BABAR Experiment*, ***BABAR*** **Analysis Document 1025** (2005).

[88] ***BABAR*** Collaboration, Lange, D., *Studies towards an Improved Tagging Algorithm*, ***BABAR*** **Analysis Document 730** (2004).

[89] **CLEO** Collaboration, Adam, N. E. et al., *Determination of the* $\overline{B}^0 \rightarrow D^*\ell\overline{\nu}_\ell$ *decay width and* $|V_{cb}|$, **Phys. Rev. D67**(2003):032001, hep-ex/0210040.

[90] ***BABAR*** Collaboration, Jackson, P. D., *Tagging using* $B \rightarrow D\ell\nu(X)$ *decays*, ***BABAR*** **Analysis Document 537** (2003).

[91] ***BABAR*** Collaboration, Allmendinger, T., *Tracking Efficiency Studies in Release 12 and 14*, ***BABAR*** **Analysis Document 867** (2004).

[92] ***BABAR*** Collaboration, *Particle Id* Available from: `http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/pid.html`.

[93] ***BABAR*** Collaboration, Allen, M., *A Measurement of* π^0 *Efficiency using* $\tau \rightarrow \rho\nu$ *and* $\tau \rightarrow \pi\nu$ *Decays*, ***BABAR*** **Analysis Document 870** (2004).

[94] Neyman, J., **Phil. Trans, Royal Soc. London, Series A 236**(1937).

[95] Jantsch, A., *Private Communication*, special thanks to Andreas Jantsch from the CKMfitter group for performing the CKM fits and producing the corresponding plots.

[96] ***BABAR*** Collaboration, Brandt, T., *Likelihood Based Electron Identification*, ***BABAR*** **Analysis Document 396** (2002).

[97] ***BABAR*** Collaboration, Mohapatra, A., *Studies of a Neural Net Based Muon Selector for the BABAR Experiment*, ***BABAR*** **Analysis Document 474** (2004).

[98] ***BABAR*** Collaboration, Mancinelli, G., *Kaon Selection at the BABAR Experiment*, ***BABAR*** **Analysis Document 116** (2001).

[99] *The Particle Detector Brief Book* Available from: `http://rkb.home.cern.ch/rkb/PH14pp/node91.html#90`.

Acknowledgements

Ich möchte es nicht versäumen, mich an dieser Stelle bei all jenen zu bedanken, die bei der Entstehung dieser Arbeit mitgeholfen haben.

Prof. Dr. Bernhard Spaan gebürt besonderer Dank für die Bereitstellung des interessanten Themas sowie für seine Hilfe bei der Klärung fachlicher Fragen. Desweiteren möchte ich mich bei PD Dr. Reiner Klingenberg für die Bereitschaft bedanken, als Gutachter dieser Arbeit zu fungieren. Dr. Heiko Lacker ist ebenfalls zu nennen, der in zahlreichen Diskussionen insbesondere zur Interpretation des Ergebnisses beigetragen hat. Weiterhin möchte ich Andreas Jantsch von der CKMfitter-Gruppe für die Durchführung der globalen Fits unter Benutzung des Ergebnisses dieser Arbeit mit dem CKMfitter-Paket und die Bereitstellung der entsprechenden Abbildungen danken. Für ihre zahlreichen Anmerkungen zu Inhalt, Ausdruck und Form dieser Arbeit sollen hier auch die Doktoranden des Instituts für Kern- und Teilchenphysik der TU Dresden erwähnt werden. Im speziellen möchte ich mich auf diesem Weg bei Andreas Petzold, Verena Klose und Markus Warsinsky bedanken.

Neben der oben erwähnten fachlichen Unterstützung darf aber auch der moralische Zuspruch seitens meiner gesamten Familie, im besonderen meiner Eltern, nicht unerwähnt bleiben. Ganz besonderer Dank gebührt Astrid Nobis für ihr unerschütterliches Vetrauen und ihre Geduld. Vielen Dank.

Appendix A

Definitions of Candidates

A.1 Charged Candidates

Charged track candidates satisfying different criteria are collected in different quality lists. In this section the definitions of candidates entering these quality lists are summarized:

1. ***Charged Tracks* (CT)**: This list contains all charged track objects reconstructed in the tracking system. For the *Kalman* track fits $\pi^{\pm}$ mass hypothesis has been assigned.

2. ***GoodTracksVeryLoose* (GTVL)**: Here, all CT candidates satisfying

   - $|\vec{p}| < 10\,\mathrm{GeV}/c$,
   - $\sqrt{\mathrm{POCA}_x^2 + \mathrm{POCA}_y^2} < 1.5\,\mathrm{cm}$, and
   - $-10 < \mathrm{POCA}_z < 10\,\mathrm{cm}$

   are collected, where $\vec{p} = (p_x, p_y, p_z)$ denotes the fitted track three-momentum vector and $\mathrm{POCA} = (\mathrm{POCA}_x, \mathrm{POCA}_y, \mathrm{POCA}_z)$ stands for the point of closest approach of the fitted track to the origin of the *BABAR* coordinate system.

3. ***GoodTracksLoose* (GTL)**: Candidates are GTVL candidates with

   - $\sqrt{p_x^2 + p_y^2} > 0.1\,\mathrm{GeV}/c$ and
   - $N_{\mathrm{DCH}} \geq 12$.

   Here, N_{DCH} denotes the number of DCH hits used to perform the *Kalman* track fit.

A.2 Neutral Candidates

Similarly to charged candidates, neutral objects are organized in quality lists:

1. ***CalorNeutral* (CN)**: This list includes all energy deposits found in the EMC. In the case of clusters with only one maximum the bump energy is equal to the full cluster energy. If more than one local energy minimum has been found within a cluster, it is splitted into bumps. The CN list contains all bumps regardless of which type of cluster they belong to. Photon mass hypothesis is assigned to all CN candidates.

2. ***GoodPhotonLoose* (GPL)**: Candidates are CN bumps with

    - $E > 0.03\,\mathrm{GeV}$ and
    - $\mathrm{LAT} < 0.8$.

    The bump energy E is equal to the sum of the n energy deposits E_i found to belong to the bump

    $$E = \sum_{i=1}^{n} E_i\,. \tag{A.1}$$

    The shower shape variable LAT is given as

    $$\mathrm{LAT} = \frac{\sum_{i \neq j,k} E_i r_i^2}{\sum_i E_i r_i^2}\,, \tag{A.2}$$

    where j, k denote the crystal indices with the two highest energy deposits within the bump. The energy entries are weighted by squares of the distances r_i of the center of the i-th crystal to the energetic centroid of the bump. The energy requirement rejects background photons radiated from the beams, while the LAT selection suppresses bumps induced by interactions of hadrons, e.g. π or K_L^0, with the EMC material. Typically, such hadronic showers are of different shape compared to electromagnetic showers induced by photons or electrons. The different production mechanism of hadronic showers results in higher LAT values.

A.3 Particle Identification

Here, the criteria of $e^\pm$, $\mu^\pm$, and $K^\pm$ identification are summarized:

$e^\pm$: The ***PidLHElectrons*** selector has been used to identify electrons out of the pool of CT candidates combining information from the DCH, the EMC, and the DIRC. Based on a set of pure samples of particles of types $\xi \in \{e;\pi;K;p\}$ likelihoods $L(\xi)$ have been constructed using five variables; the track energy loss per path $(\mathrm{d}E/\mathrm{d}x)_{\mathrm{DCH}}$ measured in the DCH, the ratio of the deposited energy of the associated EMC bump and the track momentum measured in the DCH $E_{\mathrm{EMC}}/|\vec{p}_{\mathrm{DCH}}|$, the lateral shape of the EMC bump LAT (Eq. A.2), the longitudinal shape of the EMC bump $\Delta\Phi$, and the *Cherenkov* angle Θ^C_{DIRC} measured by the DIRC. A detailed description of the selection of the pure samples and the procedure to construct $L(\xi)$ is given in Ref. [96]. For a given track one can calculate the likelihood ratio

$$f_L = \frac{p_e L(e)}{\sum_\xi p_\xi L(\xi)} \quad \text{with} \quad \xi \in \{e; \pi; K; p\}\,, \tag{A.3}$$

where p_ξ stands for a priori probabilities ($p_e : p_\pi : p_K : p_p = 1 : 5 : 1 : 0.1$). CT candidates are identified as $e^\pm$ if $f_L > 0.95$.

$\boldsymbol{\mu^\pm}$: A neural network with eight input variables has been developed to select muons. The variables are: the energy of the associated EMC bump, the number of interaction lengths traversed by the track in the *BABAR* detector λ_{meas}, $\Delta\lambda = \lambda_{\mathrm{exp}} - \lambda_{\mathrm{meas}}$ with the expected number of interaction lengths λ_{exp} using muon hypothesis, the χ^2/n_{dof} of the IFR hit strips w.r.t. a third order polynomial fit to the IFR cluster, the χ^2/n_{dof} of the IFR hit strips w.r.t. the track extrapolation, the track continuity in the IFR, the average multiplicity of hit strips per layer $\overline{m}$, and the standard deviation $\sigma_{\overline{m}}$ of $\overline{m}$. Details of the algorithm as well as exact definitions of the input variables are given in Ref. [97]. The criterion chosen to identify muons is "Tight". Identified muons are collected in the corresponding PID list named ***muNNTight***.

$\boldsymbol{K^\pm}$: Similarly to electrons, charged kaon identification is based on a likelihood calculation, where the likelihood consists of SVT, DCH, and DIRC contributions. The SVT and DCH part is constructed from $\mathrm{d}E/\mathrm{d}x$ measured in both subsystems, respectively. The PDF describing the DIRC properties consists of two parts. One part includes the measured *Cherenkov* angle Θ^C_{DIRC}, its error $\sigma_{\Theta^C_{\mathrm{DIRC}}}$, and the expected *Cherenkov* angle for a given particle type (Eq. 3.1). The second part includes the number of *Cherenkov* photons emitted by the track passing the radiator material. Ref. [98] gives a complete overview of the procedure. In analogy to the $e^\pm$-ID algorithm, likelihoods are computed for different particle hypotheses $\xi \in \{e; \mu; \pi; K; p\}$. In this analysis CT candidates are identified as kaons if

$$L(K) > p_\pi(\vec{p}_{\mathrm{lab}})L(\pi) \quad \vee \quad L(p) > p_\pi(\vec{p}_{\mathrm{lab}})L(\pi)\,, \tag{A.4}$$

with a momentum dependent a priori probability for pions $p_\pi(\vec{p}_{\mathrm{lab}})$. The requirement of Eq. A.4 corresponds to the "NotAPion" criterion (***KLHNotPion***) and is optimized with respect to kaon efficiency by rejecting pions.

In Fig. A.1 the PID efficiencies and fake rates of the aforemetioned PID selectors are summarized. Electrons are selected with an average efficiency of more than 90 % and very low misidentification rates of hadrons. The μ-ID efficiency is about 70 %, while in particular the $\pi \rightarrow \mu$ mis-ID rate is worse compared to electrons. Kaons are identified with excellent efficiency with $\pi \rightarrow K$ fake rates of order 10 %. Due to the definition of the "NotAPion" criterion (Eq. A.4) almost all protons pass this PID selector.

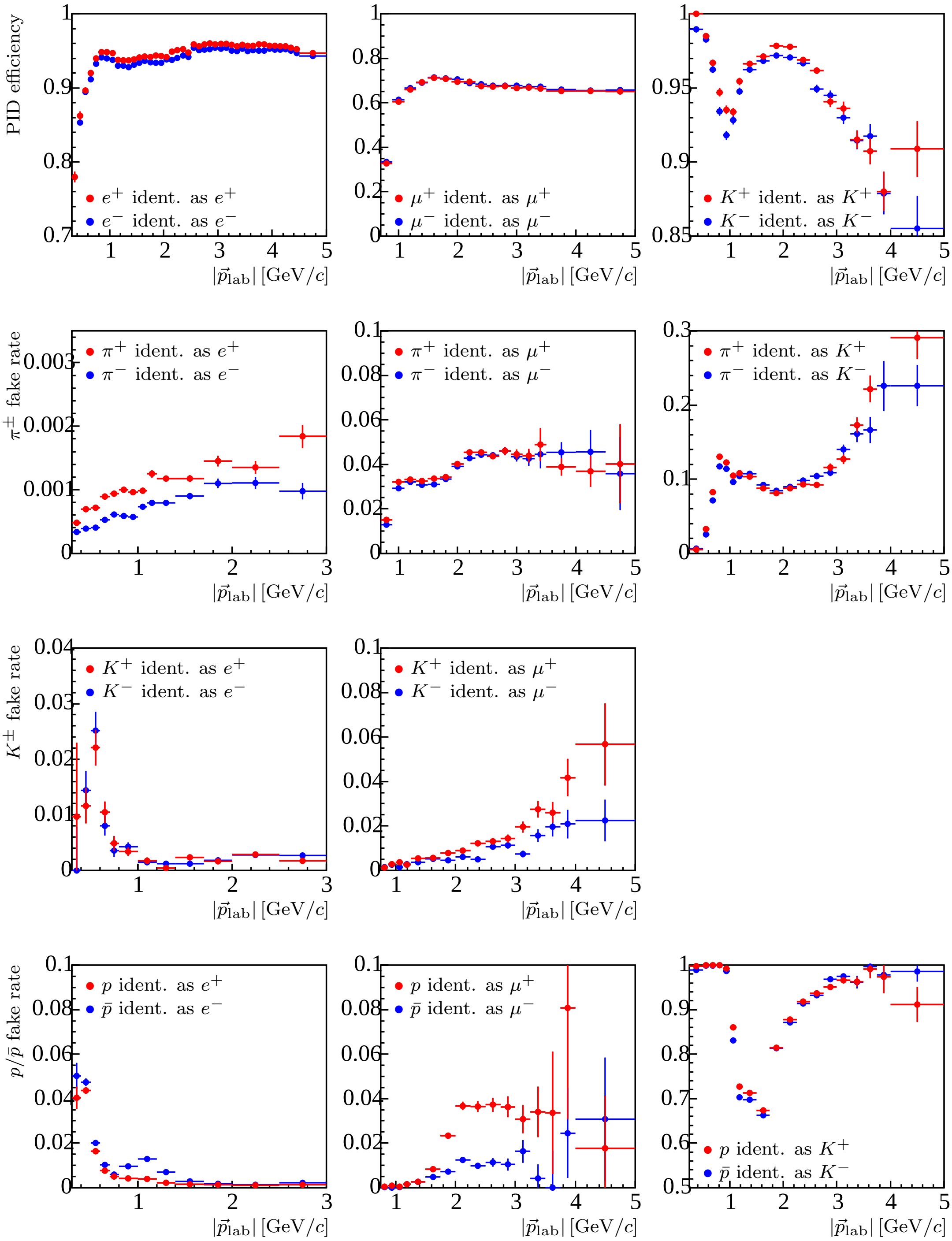


Figure A.1: PID efficiencies and fake rates for positively and negatively charged particles as a function of the magnitude of the track three-momentum $\vec{p}_{\text{lab}}$ measured in the laboratory frame [92]: The upper row illustrates the relative rates of true $e^\pm$, $\mu^\pm$, and $K^\pm$ correctly identified by the PID selectors. The lower three rows show the relative rates of true $\pi^\pm$, $K^\pm$, and protons or antiprotons $(p, \bar{p})$ entering the given PID lists.

A.4 π^0 Candidates

The selection of neutral pions is needed at different stages of the B_{tag} reconstruction within the $B \to D^{(*)}\ell\nu_\ell$ recoil analysis. This section defines two different types of π^0 candidates:

1. ***pi0AllDefaultMass* (π^0_{all})**: This list contains combinations of two photon candidates satisfying the GPL requirements (App. A.2) with an invariant mass of $0.115 < m_{\gamma\gamma} < 0.150\,\text{GeV}/c^2$. After this selection the $\gamma\gamma$ combinations have been fitted, where $m_{\gamma\gamma}$ is constrained to the nominal π^0 mass of $m_{\pi^0} = 0.13498\,\text{GeV}/c^2$ [13]. π^0_{all} candidates are used to reconstruct neutral D candidates in the $\overline{D}^0 \to K^+\pi^-\pi^0$ channel.

2. ***pi0SoftDefaultMass* (π^0_{soft})**: The $\gamma\gamma$ candidates included in this list satisfy the same requirements as given for the π^0_{all} list, but one additional selection criterion is applied. The magnitude of the three-momentum vector of the $\gamma\gamma$ combination $\vec{p}_{\gamma\gamma}$ is restricted to $|\vec{p}_{\gamma\gamma}| < 0.450\,\text{GeV}/c$ before the photon candidates are refitted. Such low-energetic π^0 candidates are needed for reconstruction of D^* mesons in the $\overline{D}^{*0} \to \overline{D}^0\pi^0$ and $D^{*-} \to D^-\pi^0$ decay modes.

A.5 K^0_S Candidates

The reconstruction of D candidates in the exclusive decay channels $\overline{D}^0 \to K^0_S\pi^-\pi^+$ and $D^- \to K^0_S\pi^-$ demands the selection of K^0_S. In our case they are defined as combinations of two oppositely charged CT candidates (App. A.1) with an invariant mass of $0.3 < m_{\pi^+\pi^-} < 0.7\,\text{GeV}/c^2$. After this selection the two tracks are fitted to a common vertex using a geometric constraint. The refitted combination is required to own an invariant mass in a mass window of $0.025\,\text{GeV}/c^2$ around the nominal K^0 mass of $m_{K^0} = 0.4977\,\text{GeV}/c^2$ [13].

Appendix B

Artificial Neural Network (Inclusive Reconstruction)

This section summarizes the input variables of the event (App. B.1) as well as the a_1 candidate based ANN's (Apps. B.2, B.3, and B.4) used within the inclusive reconstruction. All distributions correspond to the entire Run12 MC and data samples and are scaled to the Run12 On Peak luminosity.

B.1 Input Variables of Event based ANN

The input variables of the event based ANN are:

1. $\boldsymbol{R_2}$ calculated from all CT and CN objects (Figs. B.1a-c)
2. event thrust $\boldsymbol{T}$ calculated from all CT and CN objects (Figs. B.1d-f)
3. cosine of polar angle of thrust axis $\boldsymbol{\cos\theta_T}$ (Figs. B.1g-i)
4. missing energy $\boldsymbol{E_{\mathrm{miss}}}$ (Figs. B.1j-l)
5. transverse component of missing three-momentum vector $\boldsymbol{p^{xy}_{\mathrm{miss}}}$ (Figs. B.2a-c)
6. z component of missing three-momentum vector $\boldsymbol{p^{z}_{\mathrm{miss}}}$ (Figs. B.2d-f)
7. cosine of angle between event thrust axis and missing three-momentum vector $\boldsymbol{\cos\Theta_{T,\mathrm{miss}}}$ (Figs. B.2g-i)
8. GTL multiplicity $\boldsymbol{N^{\mathrm{GTL}}}$ (Figs. B.2j-l)
9. GPL multiplicity $\boldsymbol{N^{\mathrm{GPL}}}$ (Figs. B.2m-o)

The dimensionless event thrust is defined as [99]

$$T = 2 \cdot \max \frac{\sum_{i,\vec{p_i}\cdot\vec{n}>0} \vec{p}^{\,L}_i}{\sum_i |\vec{p_i}|} \,, \tag{B.1}$$

calculated from a set of candidates i. $\vec{p}^{\,L}_i$ denotes the longitudinal component of $\vec{p_i}$ along a unit vector $\vec{n}$, which defines an axis in the three-dimensional space. That $\vec{n}$, which maximizes $\sum_{i,\vec{p_i}\cdot\vec{n}>0} \vec{p}^{\,L}_i$ is called thrust axis and has been used to calculate $\cos\theta_T$ and $\cos\Theta_{T,\mathrm{miss}}$.

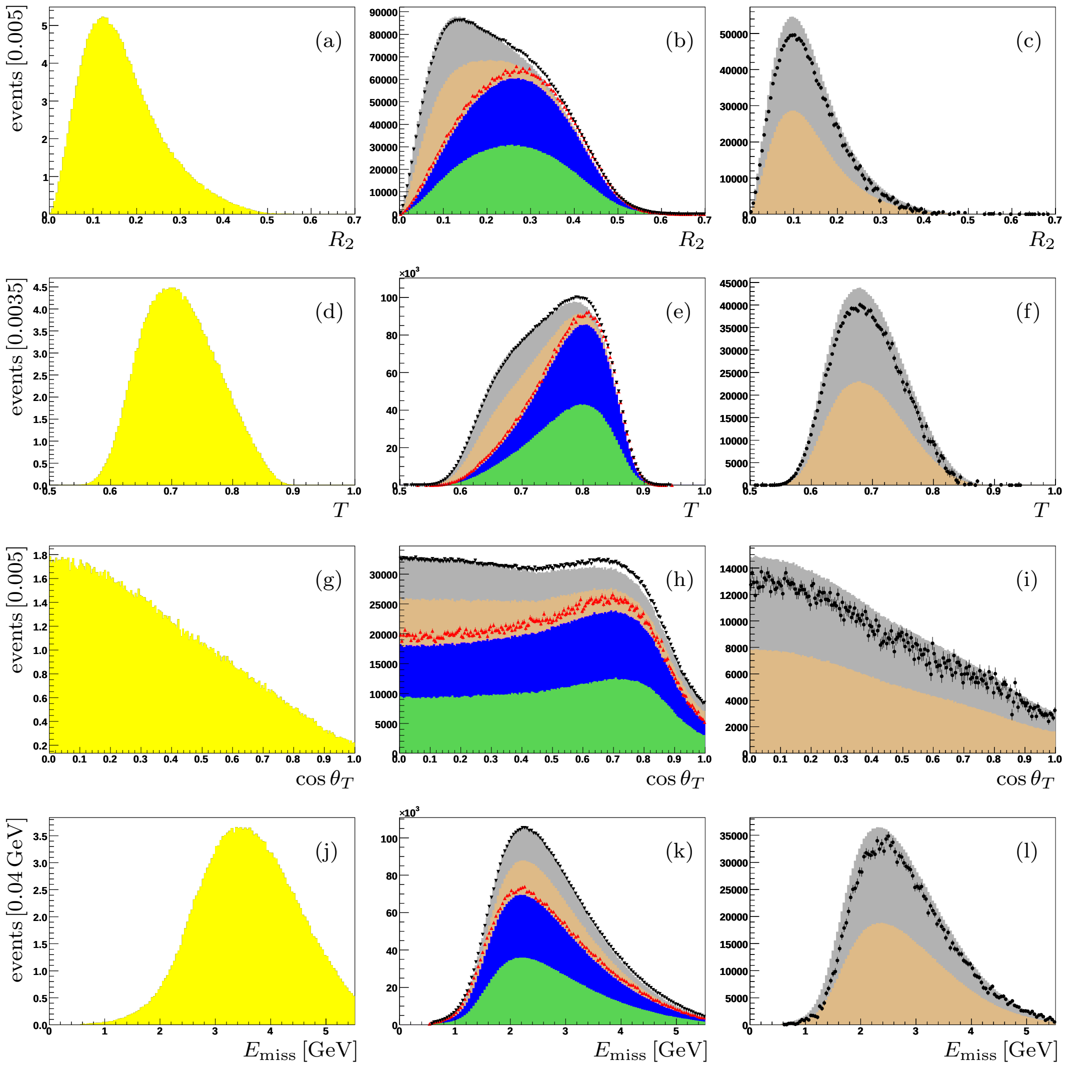


Figure B.1: Input variables of event based ANN: Distributions of (a-c) R_2, (d-f) T, (g-i) $\cos\theta_T$, and (j-l) E_{miss} after event preselection. (signal [🟨], $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

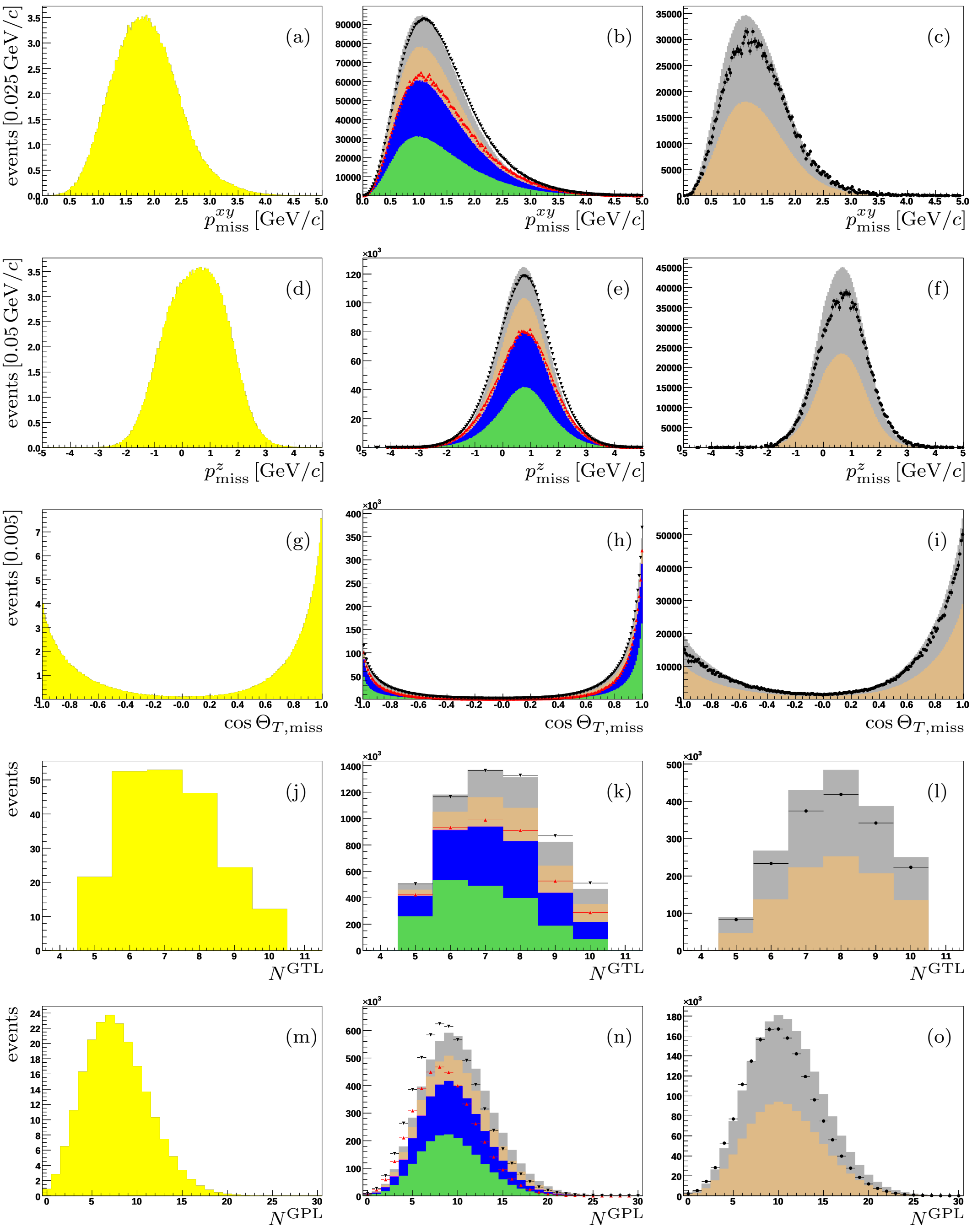


Figure B.2: Input variables of event based ANN: Distributions of (a-c) p^{xy}_{miss}, (d-f) p^{z}_{miss}, (g-i) $\cos\Theta_{T,\mathrm{miss}}$, (j-l) N^{GTL}, and (m-o) N^{GPL} after event preselection. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

B.2 Kinematical Input Variables of a_1 based ANN

This is the list of the kinematical input variables of the a_1 candidate based ANN:

1. invariant a_1 mass $\boldsymbol{m_{3\pi}}$ (Figs. B.3a-c)
2. magnitude of a_1 three-momentum $|\boldsymbol{\vec{p}_{3\pi}}|$ (Figs. B.3d-f)
3. cosine of polar angle of a_1 three-momentum $\boldsymbol{\cos\theta_{3\pi}}$ (Figs. B.4a-c)
4. cosine of angle between a_1 and missing momentum $\boldsymbol{\cos\Theta_{3\pi,\mathrm{miss}}}$ (Figs. B.4d-f)
5. difference of a_1 and missing momentum $|\boldsymbol{\vec{p}_{\mathrm{miss}}}| - |\boldsymbol{\vec{p}_{3\pi}}|$ (Figs. B.4g-i)
6. invariant mass of first neutral $\pi\pi$ combination $\boldsymbol{\sqrt{s_1}}$ (Figs. B.4j-l)
7. invariant mass of second neutral $\pi\pi$ combination $\boldsymbol{\sqrt{s_2}}$ (Figs. B.4m-o)
8. zeroth angular moment $\boldsymbol{M_0}$ (Figs. B.5a-c)
9. second angular moment $\boldsymbol{M_2}$ (Figs. B.5d-f)
10. product of magnitudes of π three-momenta $|\boldsymbol{\vec{p}_{\pi_1}}| \cdot |\boldsymbol{\vec{p}_{\pi_2}}| \cdot |\boldsymbol{\vec{p}_{\pi_3}}|$ (Figs. B.5g-i)
11. magnitude of sum of three-momenta of all CN and CT not belonging to a_1 candidate $|\boldsymbol{\vec{p}_{\mathrm{tag}}}|$ (Figs. B.5j-l)
12. cosine of angle between thrust axis (Eq. B.1) calculated from a_1 tracks and thrust axis calculated from all CN and CT not belonging to a_1 candidate $\boldsymbol{\cos\Theta_{T_{3\pi},T_{\mathrm{tag}}}}$ (Figs. B.5m-o).

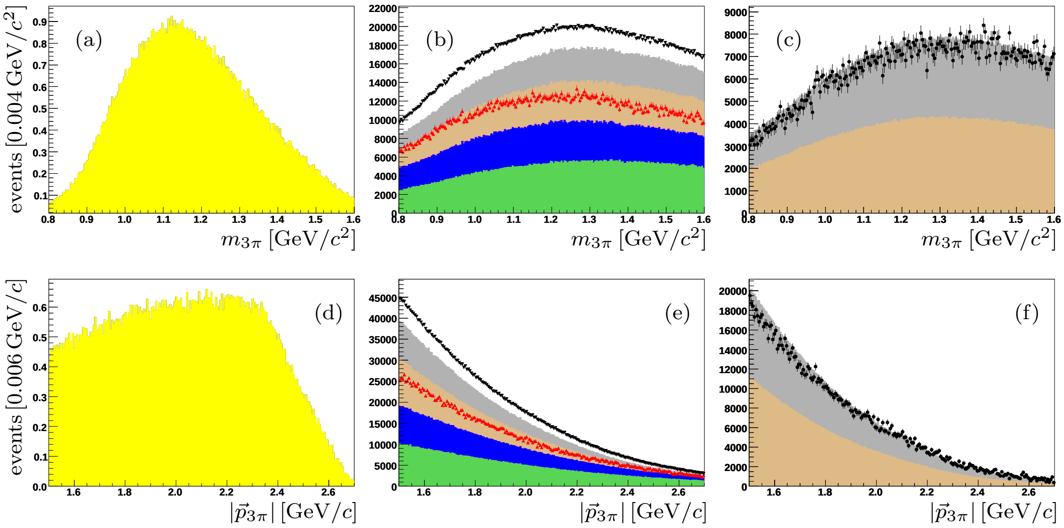


Figure B.3: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $m_{3\pi}$ and (d-f) $|\vec{p}_{3\pi}|$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

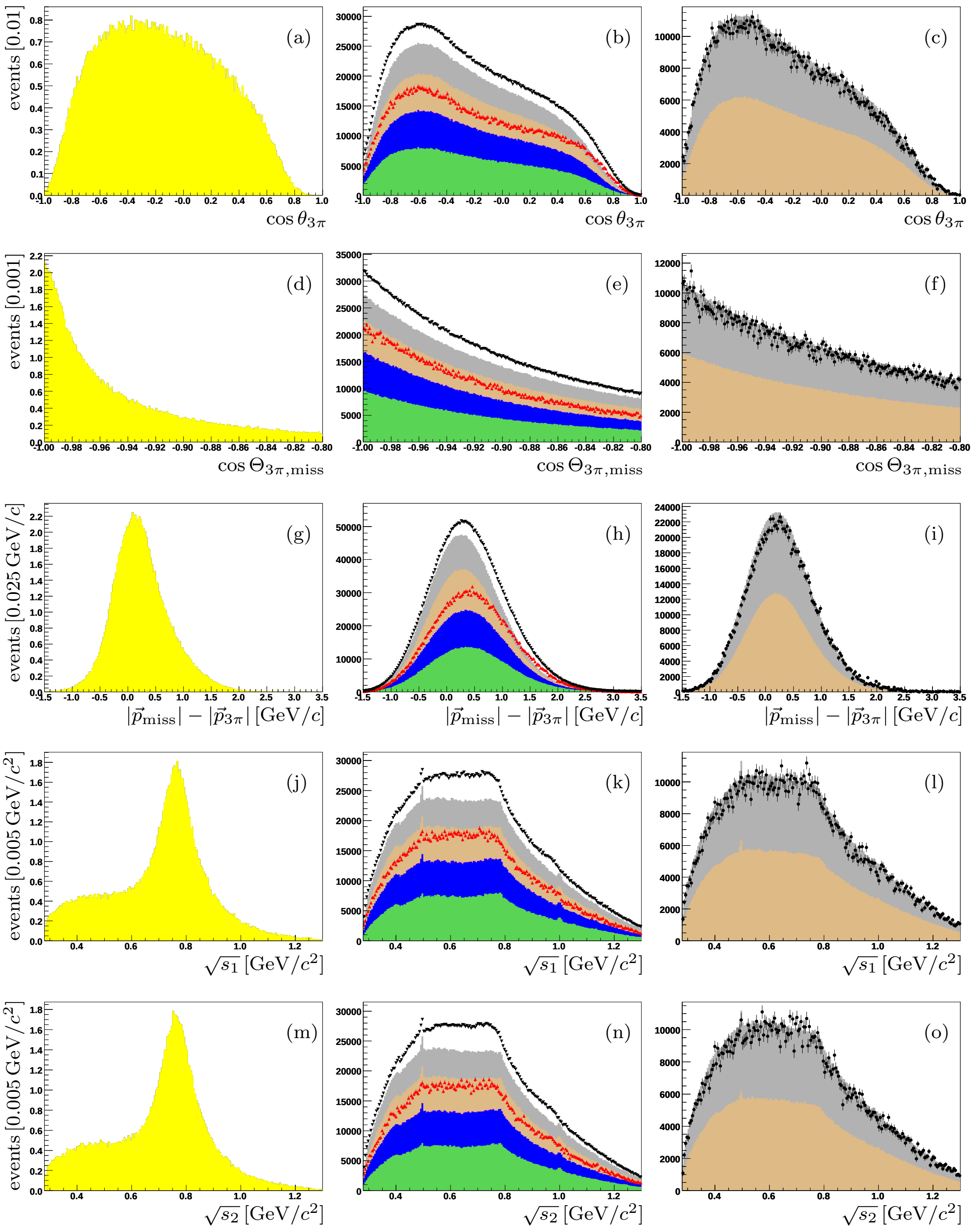


Figure B.4: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $\cos\theta_{3\pi}$, (d-f) $\cos\Theta_{3\pi,\mathrm{miss}}$, (g-i) $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$, (j-l) $\sqrt{s_1}$, and (m-o) $\sqrt{s_2}$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

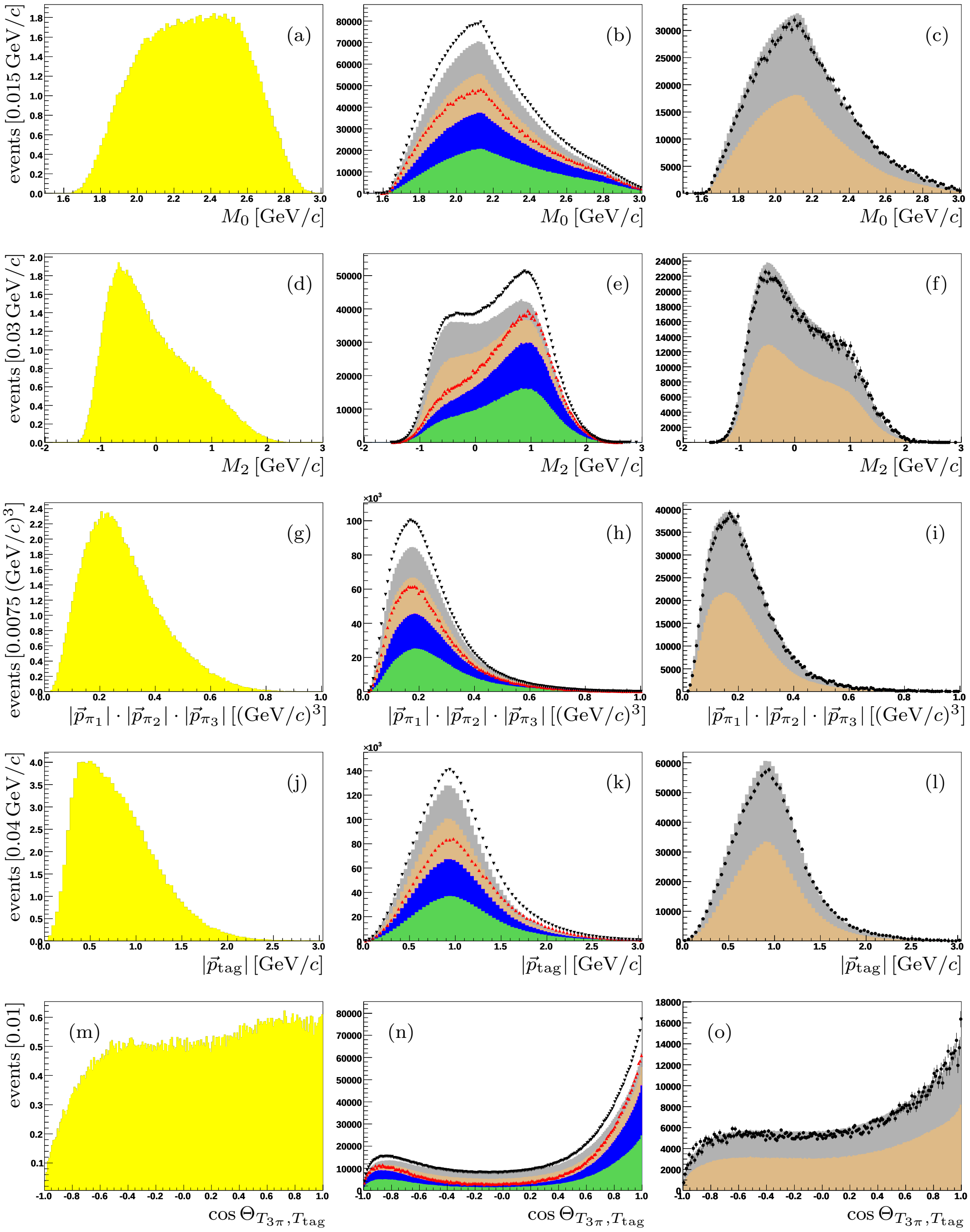


Figure B.5: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) M_0, (d-f) M_2, (g-i) $|\vec{p}_{\pi_1}|\cdot|\vec{p}_{\pi_2}|\cdot|\vec{p}_{\pi_3}|$, (j-l) $|\vec{p}_{\rm tag}|$, and (m-o) $\cos\Theta_{T_{3\pi},T_{\rm tag}}$ after $\mathcal{NN}>0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

B.3 Vertexing Input Variables of a_1 Candidate based ANN

The vertexing input of the a_1 candidate based ANN is represented by:

1. separation of B_{tag} and a_1 decay vertices in transverse plane $\boldsymbol{\Delta^{xy}}$ (Figs. B.6a-c)
2. separation of B_{tag} and a_1 decay vertices along beam axis $\boldsymbol{\Delta^{z}}$ (Figs. B.6d-f)

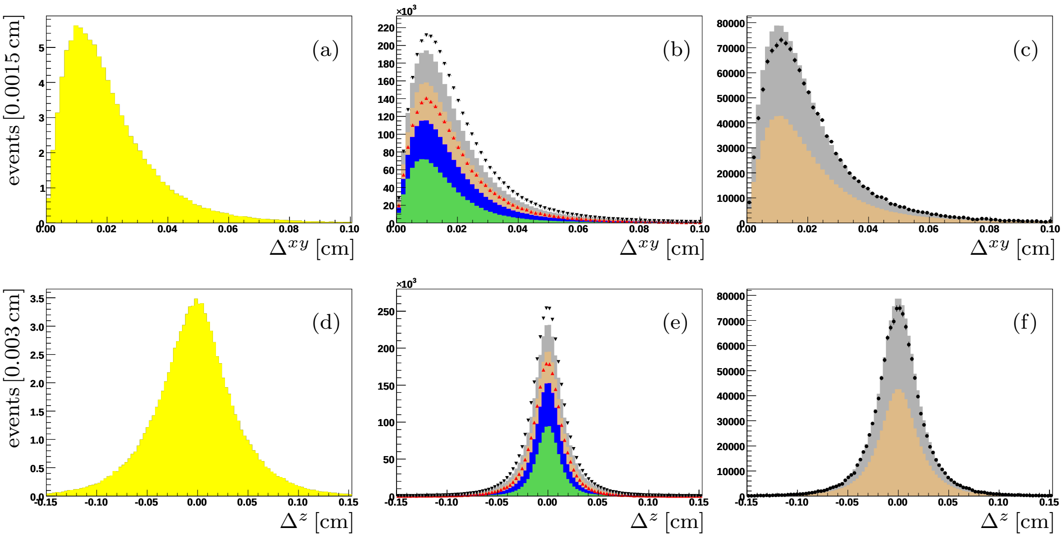


Figure B.6: Vertexing input variables of a_1 candidate based ANN: Distributions of (a-c) Δ^{xy} and (d-f) Δ^{z} after $\mathcal{NN} > 0$. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^{+}\tau^{-}$ [🟪], $B^{0}\overline{B}^{0}$ [🟫], $B^{+}B^{-}$ [■], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

B.4 Flavour Tagging Input Variables of a_1 Candidate based ANN

The B_{tag} flavour tagging quantities of the a_1 candidate based ANN are:

1. B_{tag} tagging category $\boldsymbol{\mathcal{C}}$ returned by standard algorithm (Figs. B.7a-c)
2. B_{tag} tagging probability $\boldsymbol{\mathcal{P}}$ returned by standard algorithm (Figs. B.7d-i)
3. reconstructed a_1 charge $\boldsymbol{c_{3\pi}}$

The reconstructed 3π charge is included in order to exploit the correlation of the tagging probability and $c_{3\pi}$. $\mathcal{P}^+$ denotes the tagging probability corresponding to a_1^+, while $\mathcal{P}^-$ corresponds to a_1^- candidates. As one expects, the distribution of $\mathcal{P}^+$ can be transformed into $\mathcal{P}^-$ by mirroring at $\mathcal{P} = 0$ apart from statistical fluctuations.

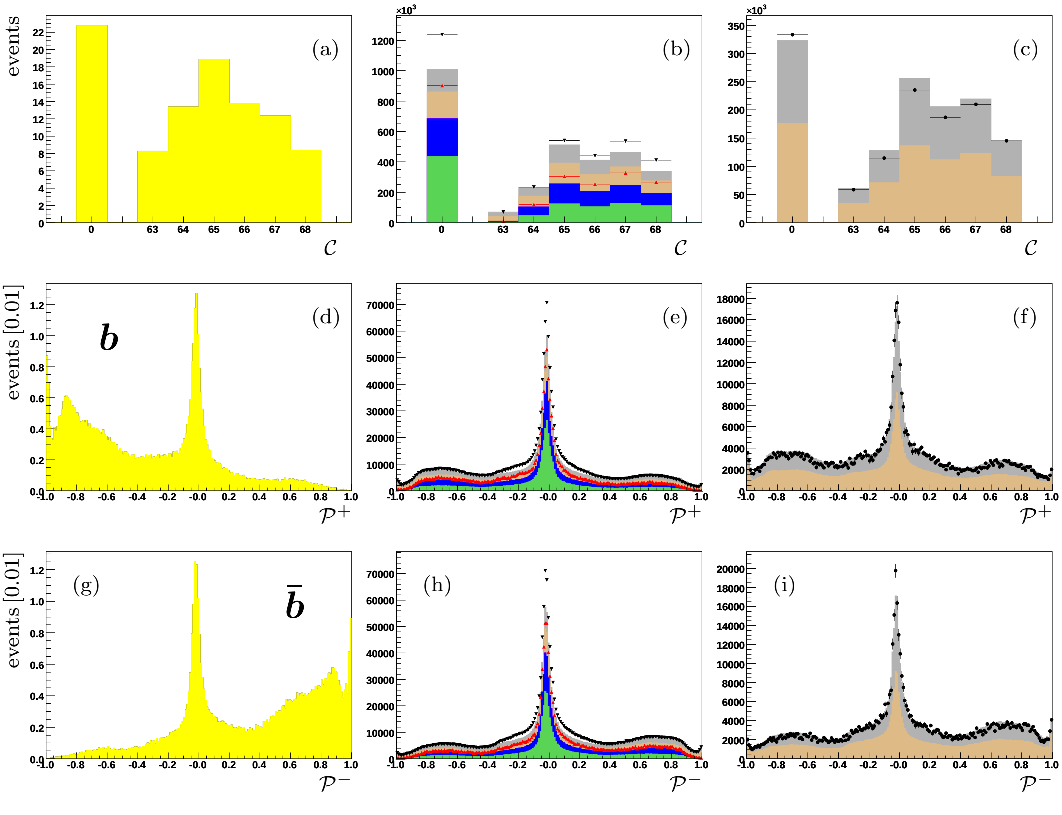


Figure B.7: Flavour Tagging input variables of a_1 candidate based ANN: Distributions of (a-c) $\mathcal{C}$, (d-f) $\mathcal{P}^+$, and (g-i) $\mathcal{P}^-$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Appendix C

Artificial Neural Network (Semileptonic Reconstruction)

This section summarizes the input variables of the event (App. C.1) as well as the a_1 candidate based ANN's (Apps. C.2 and C.3) used within the semileptonic reconstruction. Since the B_{tag} flavour is supposed to be already determined by the charge of the high-energetic lepton candidate required in the event preselection, no flavour tagging quantities are included in the a_1 candidate based ANN. All distributions correspond to the entire Run12 MC and data samples and are scaled to the Run12 On Peak luminosity.

C.1 Input Variables of Event based ANN

The multivariate event selection within the semileptonic recontruction uses the same set of input variables as given for the inclusive reconstruction (App. B.1):

1. $\boldsymbol{R_2}$ calculated from all CT and CN objects (Figs. C.1a-c)
2. event thrust $\boldsymbol{T}$ calculated from all CT and CN objects (Figs. C.1d-f)
3. cosine of polar angle of event thrust axis $\boldsymbol{\cos\theta_T}$ (Figs. C.1g-i)
4. missing energy $\boldsymbol{E_{\mathrm{miss}}}$ (Figs. C.1j-l)
5. transverse component of missing three-momentum vector $\boldsymbol{p^{xy}_{\mathrm{miss}}}$ (Figs. C.2a-c)
6. z component of missing three-momentum vector $\boldsymbol{p^{z}_{\mathrm{miss}}}$ (Figs. C.2d-f)
7. cosine of angle between event thrust axis and missing three-momentum vector $\boldsymbol{\cos\Theta_{T,\mathrm{miss}}}$ (Figs. C.2g-i)
8. GTL multiplicity $\boldsymbol{N^{\mathrm{GTL}}}$ (Figs. C.2j-l)
9. GPL multiplicity $\boldsymbol{N^{\mathrm{GPL}}}$ (Figs. C.2m-o)

The thrust T is calculated using Eq. B.1.

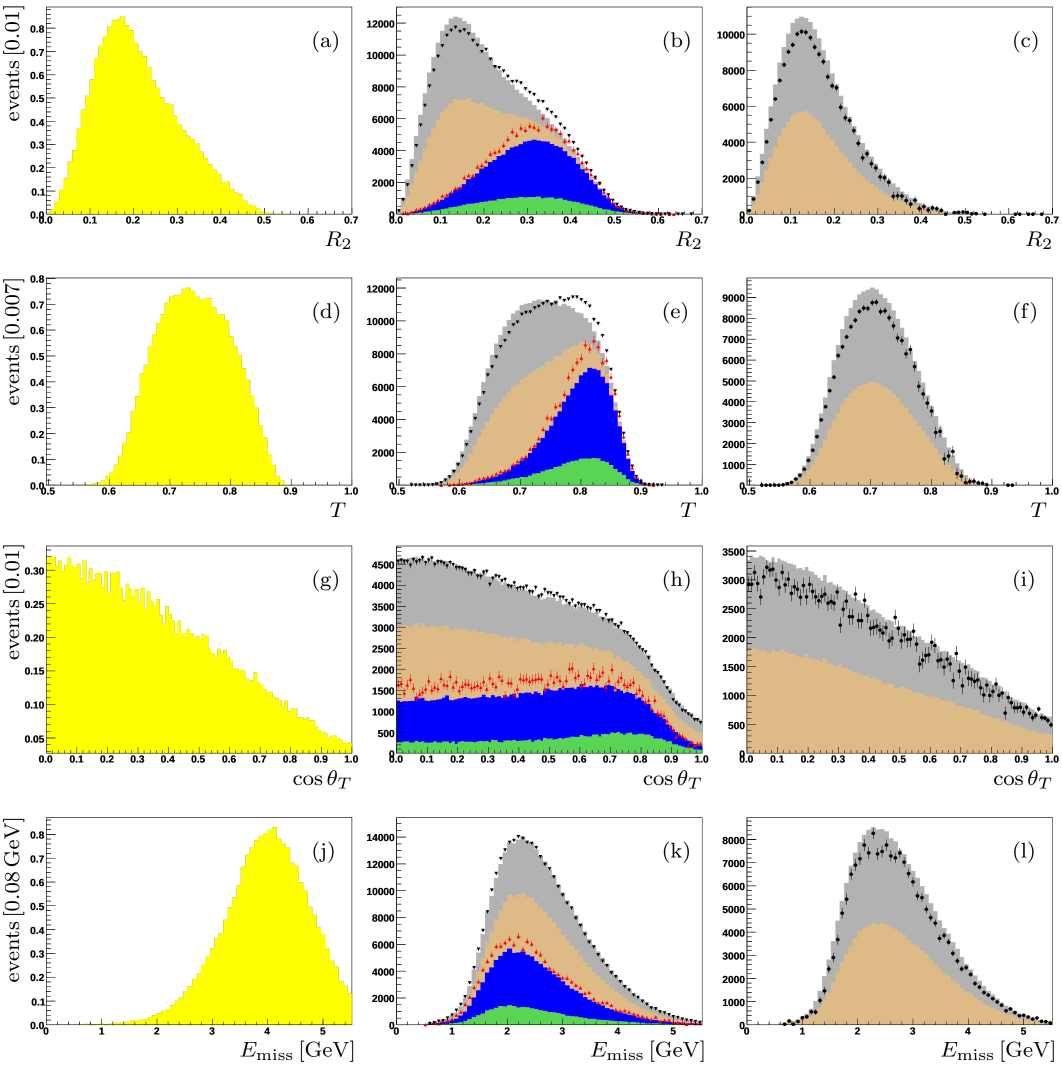


Figure C.1: Input variables of event based ANN: Distributions of (a-c) R_2, (d-f) T, (g-i) $\cos\theta_T$, and (j-l) $E_{\rm miss}$ after event preselection. (signal [🟨], $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

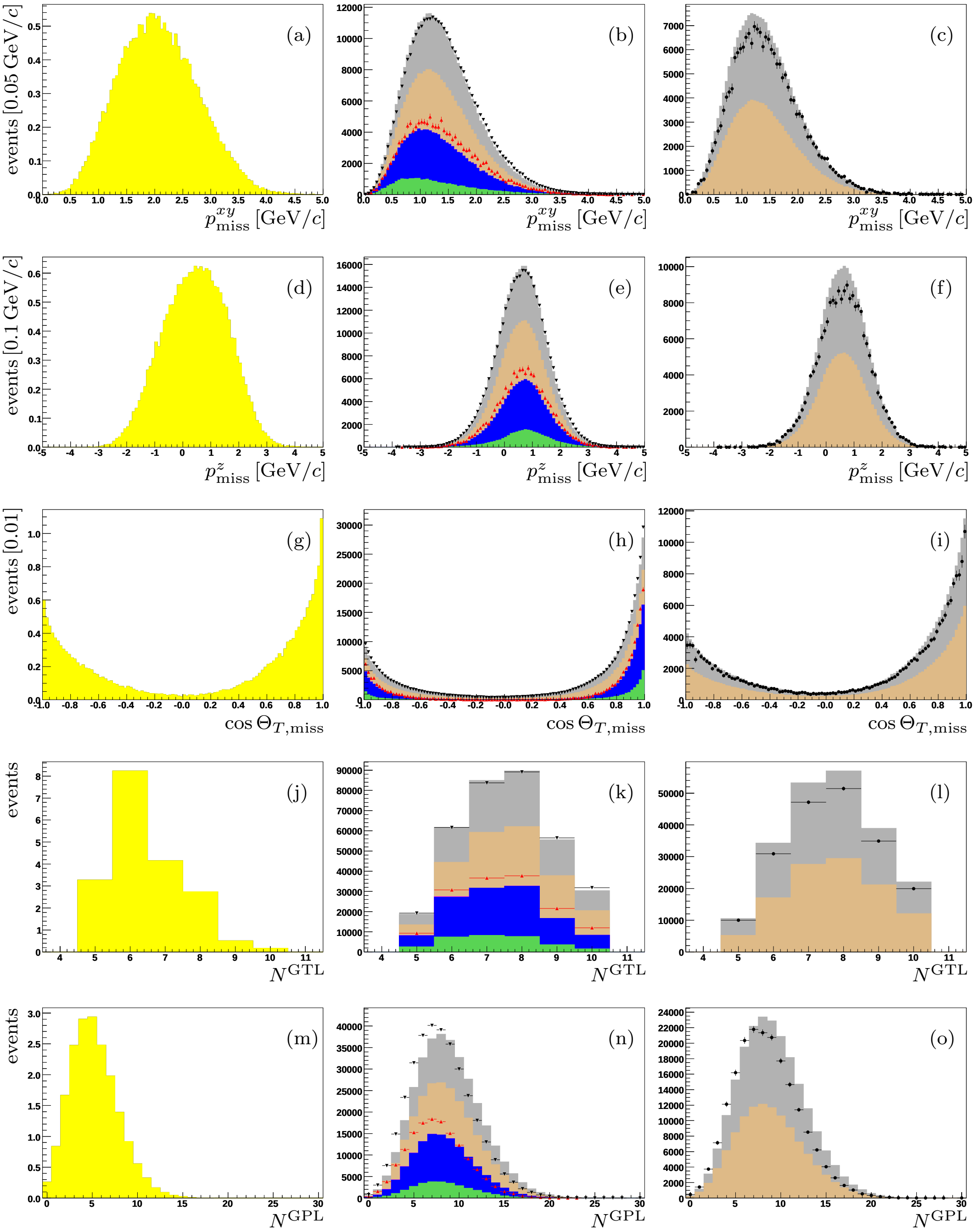


Figure C.2: Input variables of event based ANN: Distributions of (a-c) $p_{\rm miss}^{xy}$, (d-f) $p_{\rm miss}^{z}$, (g-i) $\cos\Theta_{T,\rm miss}$, (j-l) $N^{\rm GTL}$, and (m-o) $N^{\rm GPL}$ after event preselection. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

C.2 Kinematical Input Variables of a_1 based ANN

The kinematical variables included in the a_1 candidate based ANN are the same as already used for the inclusive reconstruction (App. B.2):

1. invariant a_1 mass $\boldsymbol{m_{3\pi}}$ (Figs. C.3a-c)
2. magnitude of a_1 three-momentum $|\boldsymbol{\vec{p}_{3\pi}}|$ (Figs. C.3d-f)
3. cosine of polar angle of a_1 three-momentum $\mathbf{cos}\,\boldsymbol{\theta_{3\pi}}$ (Figs. C.4a-c)
4. cosine of angle between a_1 and missing momentum $\mathbf{cos}\,\boldsymbol{\Theta_{3\pi,\mathrm{miss}}}$ (Figs. C.4d-f)
5. difference of a_1 and missing momentum $|\boldsymbol{\vec{p}_{\mathrm{miss}}}| - |\boldsymbol{\vec{p}_{3\pi}}|$ (Figs. C.4g-i)
6. invariant mass of first neutral $\pi\pi$ combination $\sqrt{\boldsymbol{s_1}}$ (Figs. C.4j-l)
7. invariant mass of second neutral $\pi\pi$ combination $\sqrt{\boldsymbol{s_2}}$ (Figs. C.4m-o)
8. zeroth angular moment $\boldsymbol{M_0}$ (Figs. C.5a-c)
9. second angular moment $\boldsymbol{M_2}$ (Figs. C.5d-f)
10. product of magnitudes of π three-momenta $|\boldsymbol{\vec{p}_{\pi_1}}| \cdot |\boldsymbol{\vec{p}_{\pi_2}}| \cdot |\boldsymbol{\vec{p}_{\pi_3}}|$ (Figs. C.5g-i)
11. magnitude of sum of three-momenta of all CN and CT not belonging to a_1 candidate $|\boldsymbol{\vec{p}_{\mathrm{tag}}}|$ (Figs. C.5j-l)
12. cosine of angle between thrust axis (Eq. B.1) calculated from a_1 tracks and thrust axis calculated from all CN and CT not belonging to a_1 candidate $\mathbf{cos}\,\boldsymbol{\Theta_{T_{3\pi},T_{\mathrm{tag}}}}$ (Figs. C.5m-o).

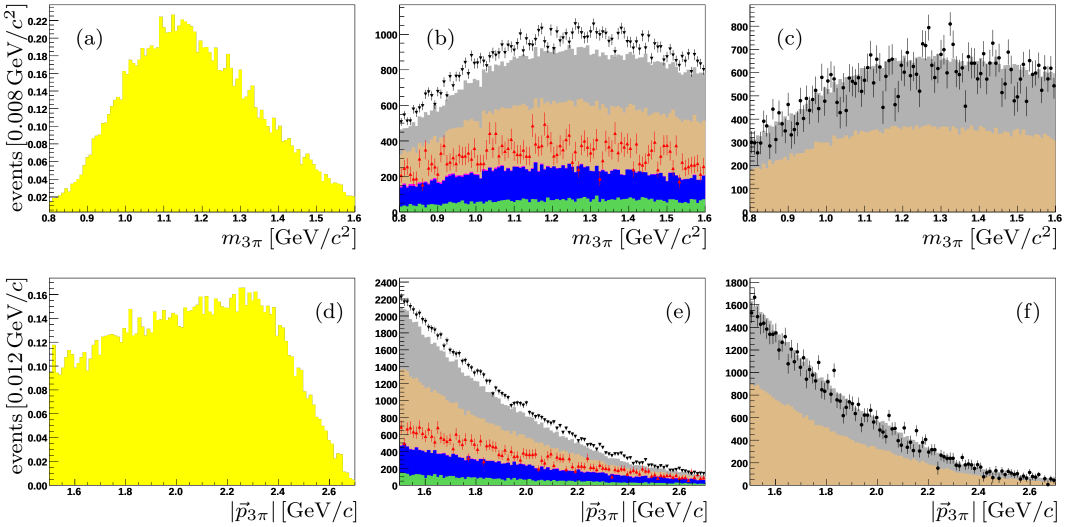


Figure C.3: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $m_{3\pi}$ and (d-f) $|\vec{p}_{3\pi}|$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

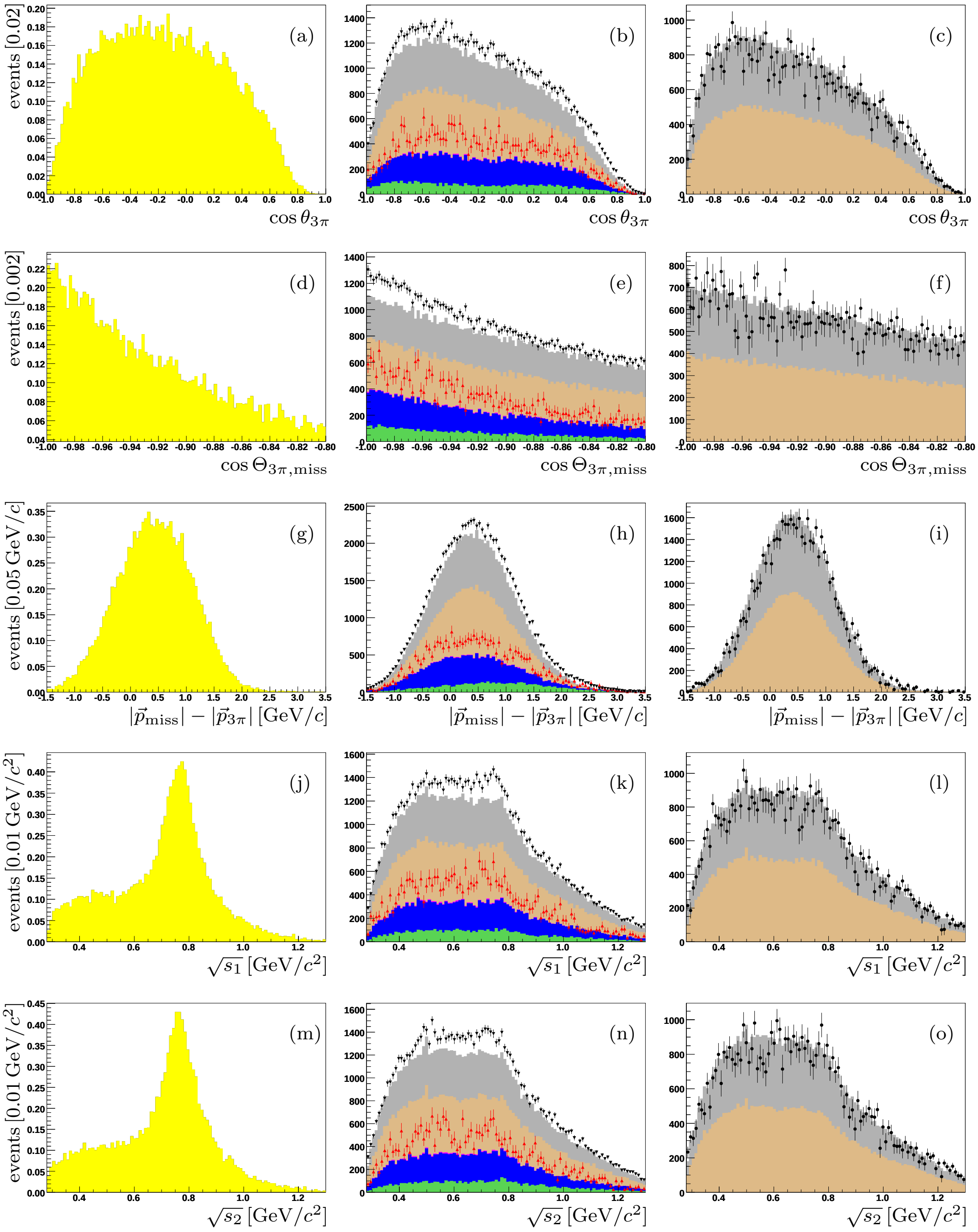


Figure C.4: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $\cos\theta_{3\pi}$, (d-f) $\cos\Theta_{3\pi,\text{miss}}$, (g-i) $|\vec{p}_{\text{miss}}| - |\vec{p}_{3\pi}|$, (j-l) $\sqrt{s_1}$, and (m-o) $\sqrt{s_2}$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

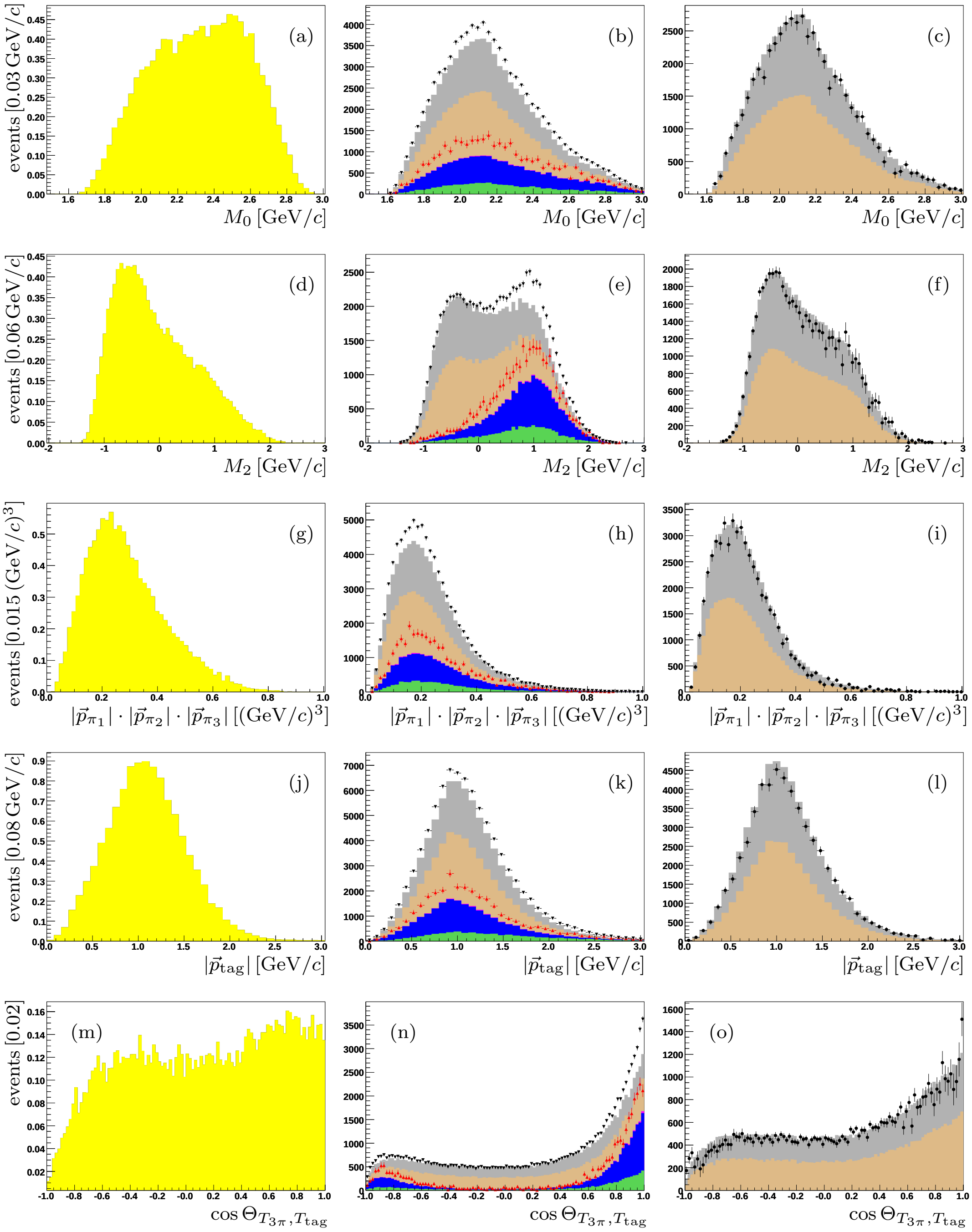


Figure C.5: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) M_0, (d-f) M_2, (g-i) $|\vec{p}_{\pi_1}|\cdot|\vec{p}_{\pi_2}|\cdot|\vec{p}_{\pi_3}|$, (j-l) $|\vec{p}_{\rm tag}|$, and (m-o) $\cos\Theta_{T_{3\pi},T_{\rm tag}}$ after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

C.3 Vertexing Input Variables of a_1 Candidate based ANN

The vertexing input variables are the same as used for the multivariate a_1 selection within the inclusive reconstruction (App. B.3):

1. separation of B_{tag} and a_1 decay vertices in transverse plane $\boldsymbol{\Delta^{xy}}$ (Figs. C.6a-c)
2. separation of B_{tag} and a_1 decay vertices along beam axis $\boldsymbol{\Delta^{z}}$ (Figs. C.6d-f)

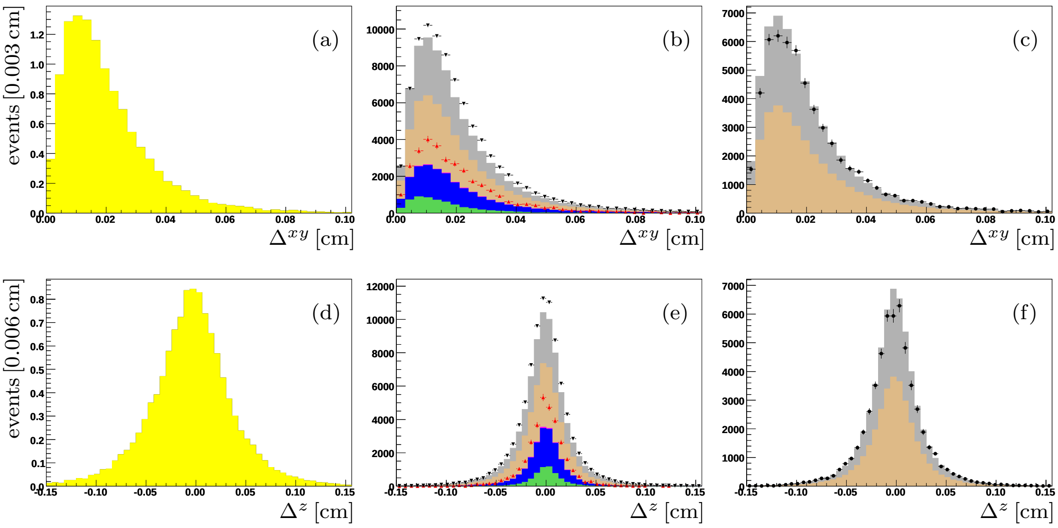


Figure C.6: Vertexing input variables of a_1 candidate based ANN: Distributions of (a-c) Δ^{xy} and (d-f) Δ^{z} after $\mathcal{NN} > 0$. (signal [🟨], $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [◼], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Appendix D

Artificial Neural Network (Recoil of $B \rightarrow D^{(*)}\ell\nu_\ell$)

Within the recoil analysis two ANN's have been used, whereas the multivariate event selection has been taken over by a $D\ell$ candidate based ANN (App. D.1). 3π candidates found in the recoil of the best $D\ell$ candidate are again selected by a multivariate a_1 selection using the already introduced input variables (Apps. D.2 and D.3).

D.1 Input Variables of $D\ell$ Candidate based ANN

The $D\ell$ candidate based ANN has been constructed with the following input variables:

1. invariant mass of reconstructed D candidate $\boldsymbol{m_D}$ (Fig. D.1)
2. magnitude of lepton three-momentum vector $|\vec{\boldsymbol{p}}_{\boldsymbol{\ell}}|$ (Fig. D.2)
3. magnitude of hadron three-momentum vector $|\vec{\boldsymbol{p}}_{\boldsymbol{h}}|$ (Fig. D.3)
4. cosine of angle between three-momentum vectors of hadron and lepton candidates $\boldsymbol{\cos\Theta_{h,\ell}}$ (Fig. D.4)
5. cosine of angle between three-momentum vectors of B meson and $D\ell$ candidates $\boldsymbol{\cos\Theta_{B,D\ell}}$ (Fig. D.5)
6. D reconstruction index $\boldsymbol{i_D}$
7. D^* reconstruction index $\boldsymbol{i_{D^*}}$
8. $D\ell$ reconstruction index $\boldsymbol{i_{D\ell}}$

The inclusion of the reconstruction indices i_D, i_{D^*}, and $i_{D\ell}$ representing the reconstruction modes ensures the mode-by-mode training of the ANN. Therefore, the ANN reacts on B_{tag} decay mode dependencies resulting in an improved separation power of the ANN output.

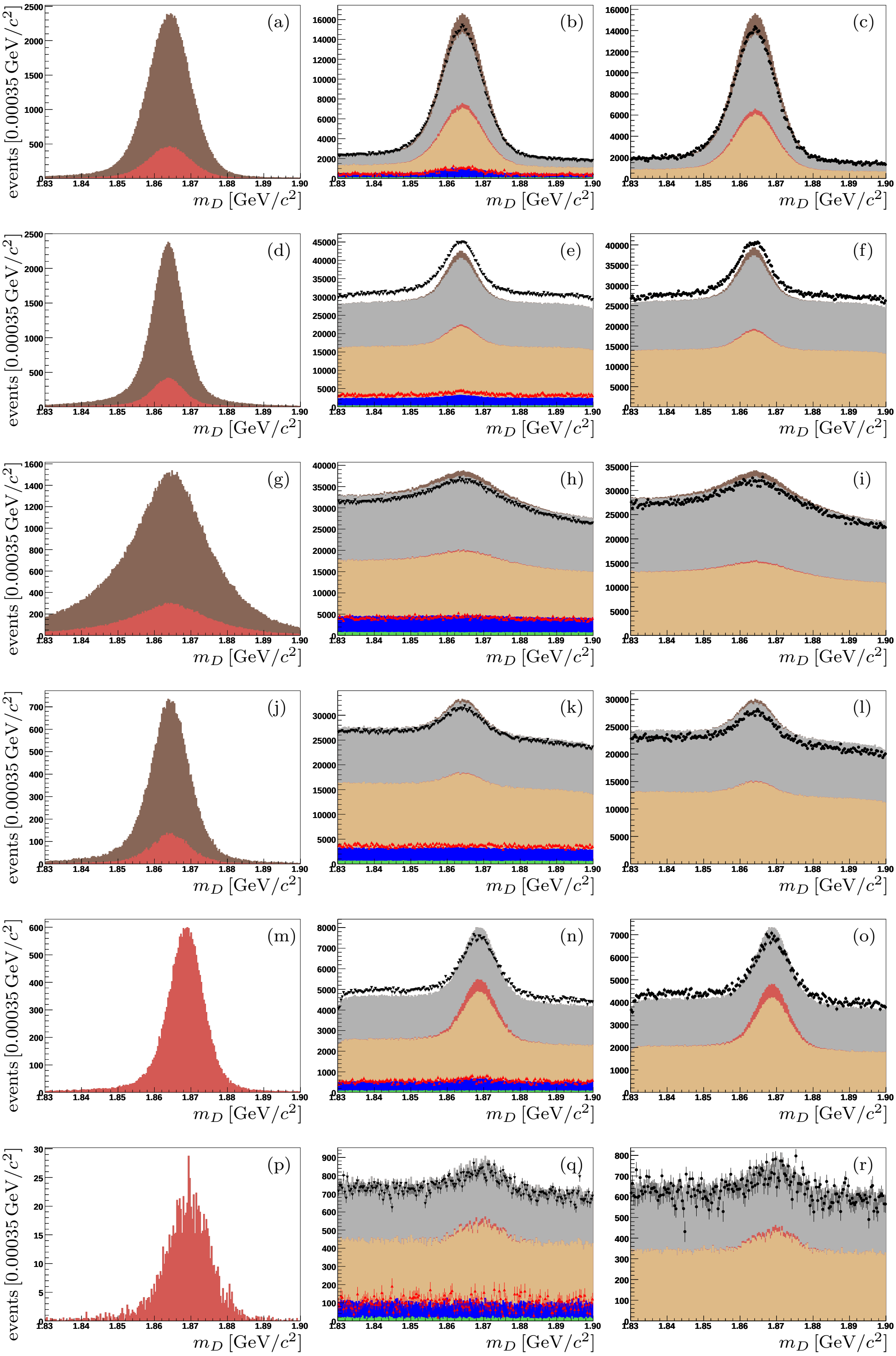


Figure D.1: m_D of correctly reconstructed (left) and all $D\ell$ candidates after $D\ell$ preselection for (a-c) $\overline{D}^0 \to K^+\pi^-$, (d-f) $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, (g-i) $\overline{D}^0 \to K^+\pi^-\pi^0$, (j-l) $\overline{D}^0 \to K^0_S\pi^-\pi^+$, (m-o) $D^- \to K^+\pi^-\pi^-$, and (p-r) $D^- \to K^0_S\pi^-$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

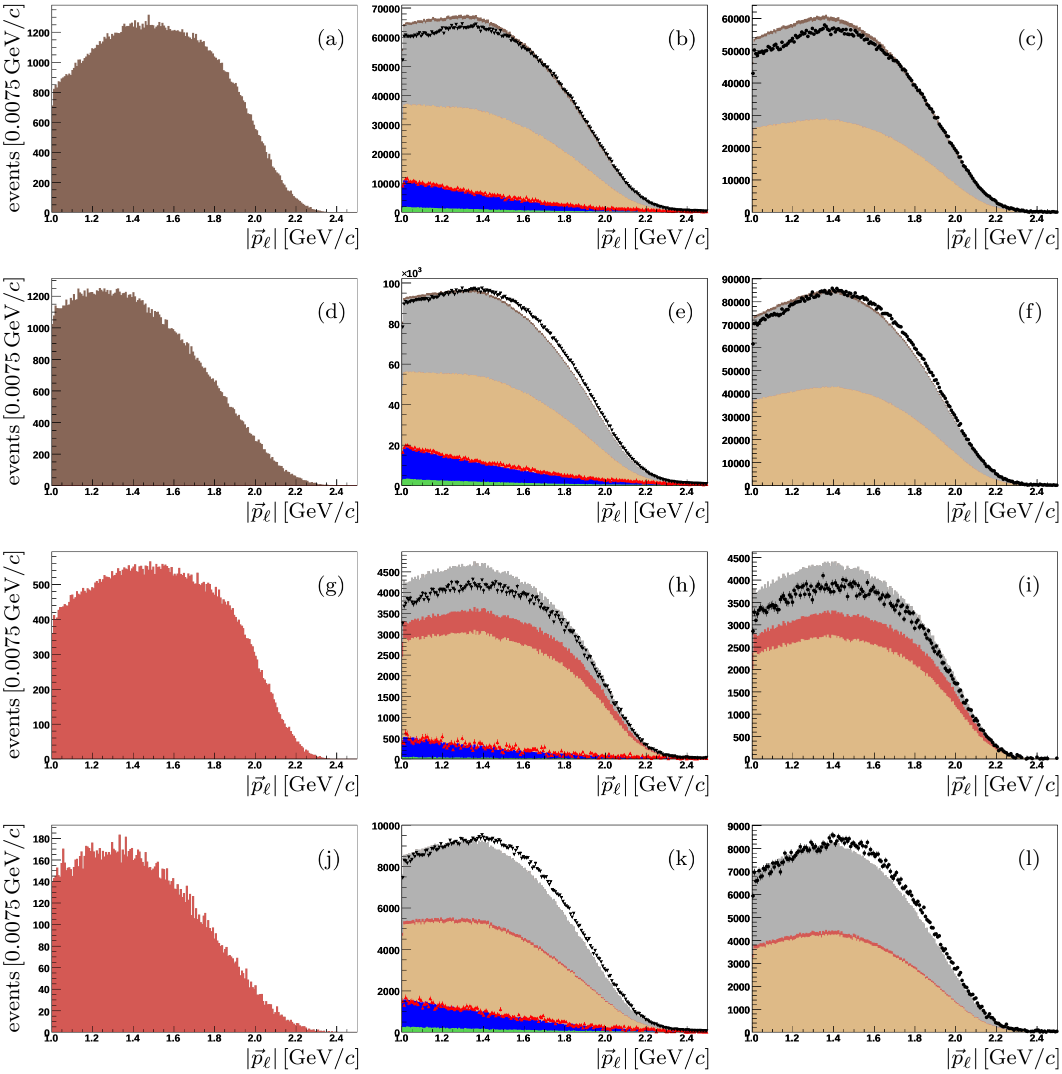


Figure D.2: $|\vec{p}_\ell|$ of correctly reconstructed (left) and all $D\ell$ candidates after $D\ell$ preselection for (a-c) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (d-f) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (g-i) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (j-l) $B^0 \to D^-\ell^+\nu_\ell$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

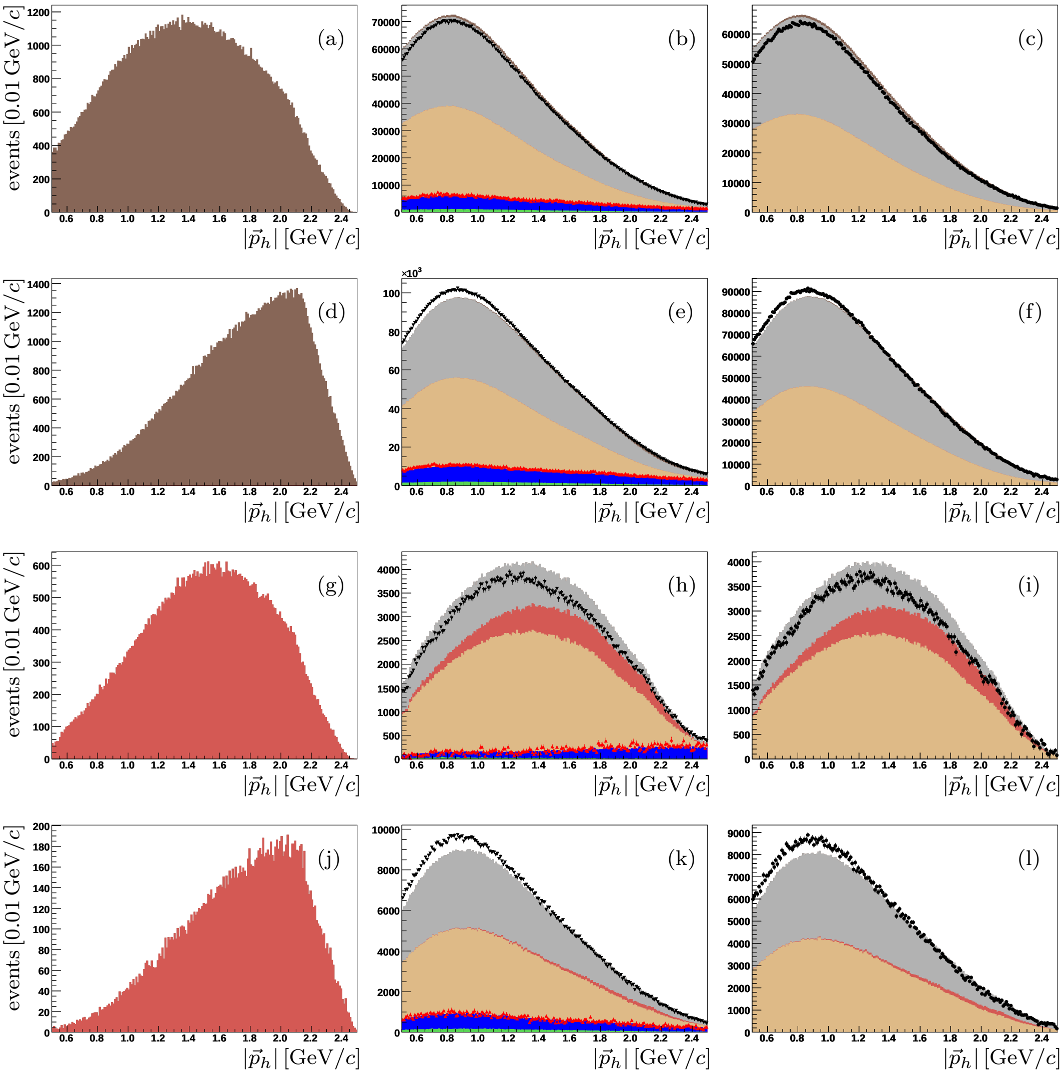


Figure D.3: $|\vec{p}_h|$ of correctly reconstructed (left) and all $D\ell$ candidates after $D\ell$ preselection for (a-c) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (d-f) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (g-i) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (j-l) $B^0 \to D^-\ell^+\nu_\ell$. ($u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

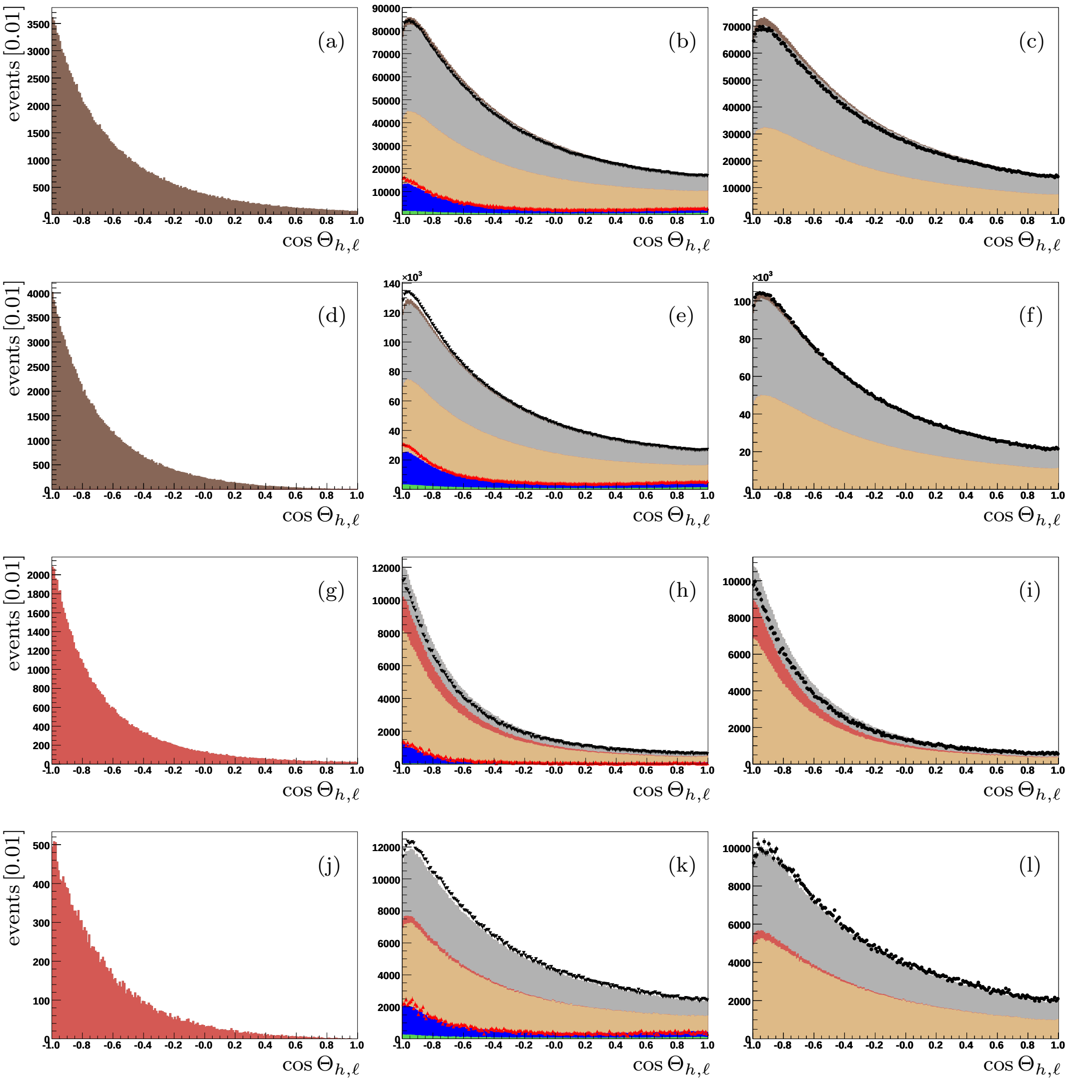


Figure D.4: $\cos\Theta_{h,\ell}$ of correctly reconstructed (left) and all $D\ell$ candidates after $D\ell$ preselection for (a-c) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (d-f) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (g-i) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (j-l) $B^0 \to D^-\ell^+\nu_\ell$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟨], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

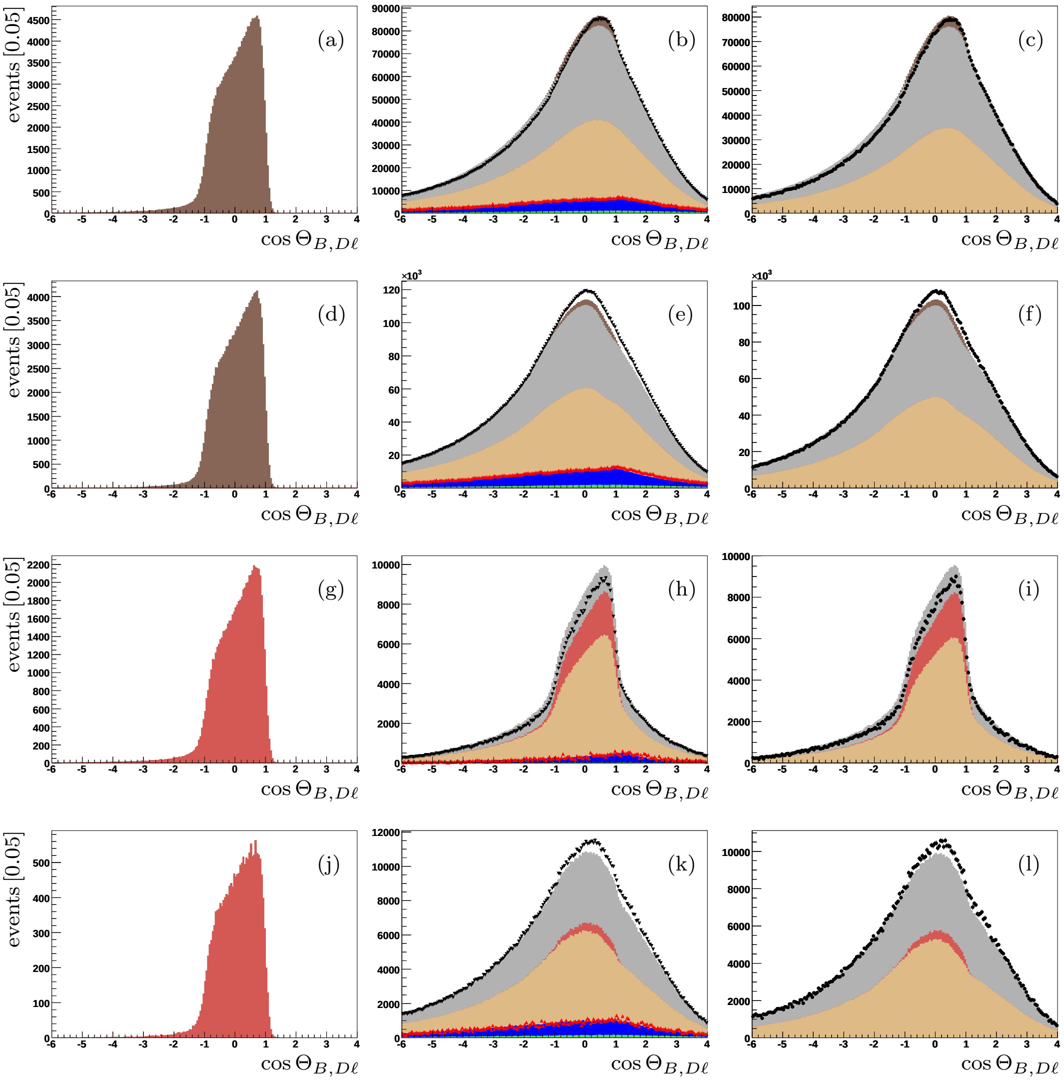


Figure D.5: $\cos\Theta_{B,D\ell}$ of correctly reconstructed (left) and all $D\ell$ candidates after $D\ell$ preselection for (a-c) $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$, (d-f) $B^+ \to \overline{D}^0\ell^+\nu_\ell$, (g-i) $B^0 \to D^{*-}\ell^+\nu_\ell$, and (j-l) $B^0 \to D^-\ell^+\nu_\ell$. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0 \to D^{(*)-}\ell^+\nu_\ell$ [🟥], other $B^0\overline{B}^0$ [🟧], $B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell$ [🟫], other B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

D.2 Kinematical Input Variables of a_1 based ANN

This kinematical input variable list contains a quite similar set of quantities as given in App. B.2, whereas the properties obtained from the inclusive B_{tag} reconstruction $|\vec{p}_{\text{tag}}|$ and $\cos\Theta_{T_{3\pi},T_{\text{tag}}}$ have been excluded. Hence, the a_1 candidate based ANN combines the ten following kinematical variables:

1. invariant a_1 mass $\boldsymbol{m_{3\pi}}$ (Figs. D.6a-c)
2. magnitude of a_1 three-momentum $|\vec{\boldsymbol{p}}_{\boldsymbol{3\pi}}|$ (Figs. D.6d-f)
3. cosine of polar angle of a_1 three-momentum $\boldsymbol{\cos\theta_{3\pi}}$ (Figs. D.7a-c)
4. cosine of angle between a_1 and modified missing momentum $\boldsymbol{\cos\Theta_{3\pi,\text{miss}}}$ (Figs. D.7d-f)
5. difference of a_1 and modified missing momentum $|\vec{\boldsymbol{p}}_{\textbf{miss}}| - |\vec{\boldsymbol{p}}_{\boldsymbol{3\pi}}|$ (Figs. D.7g-i)
6. invariant mass of first neutral $\pi\pi$ combination $\sqrt{\boldsymbol{s_1}}$ (Figs. D.7j-l)
7. invariant mass of second neutral $\pi\pi$ combination $\sqrt{\boldsymbol{s_2}}$ (Figs. D.7m-o)
8. zeroth angular moment $\boldsymbol{M_0}$ (Figs. D.8a-c)
9. second angular moment $\boldsymbol{M_2}$ (Figs. D.8d-f)
10. product of magnitudes of π three-momenta $|\vec{\boldsymbol{p}}_{\boldsymbol{\pi_1}}| \cdot |\vec{\boldsymbol{p}}_{\boldsymbol{\pi_2}}| \cdot |\vec{\boldsymbol{p}}_{\boldsymbol{\pi_3}}|$ (Figs. D.8g-i)

The missing four-momentum vector has been modified by subtracting the neutrino four-momentum vector calculated from the reconstructed best $D\ell$ candidate assuming the B_{tag} to be in rest (Eq. 4.40).

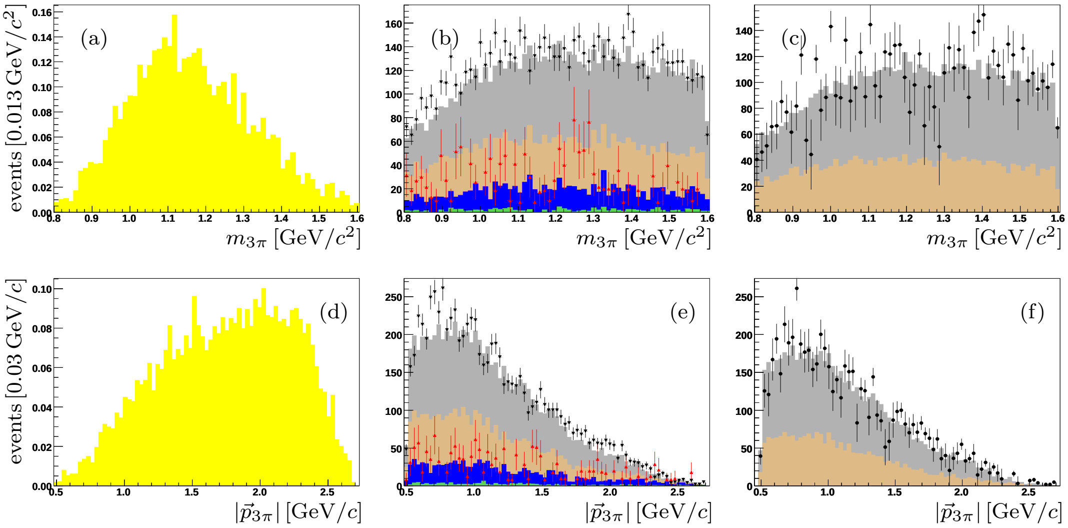


Figure D.6: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $m_{3\pi}$ and (d-f) $|\vec{p}_{3\pi}|$ after a_1 candidate preselection. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [▨], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

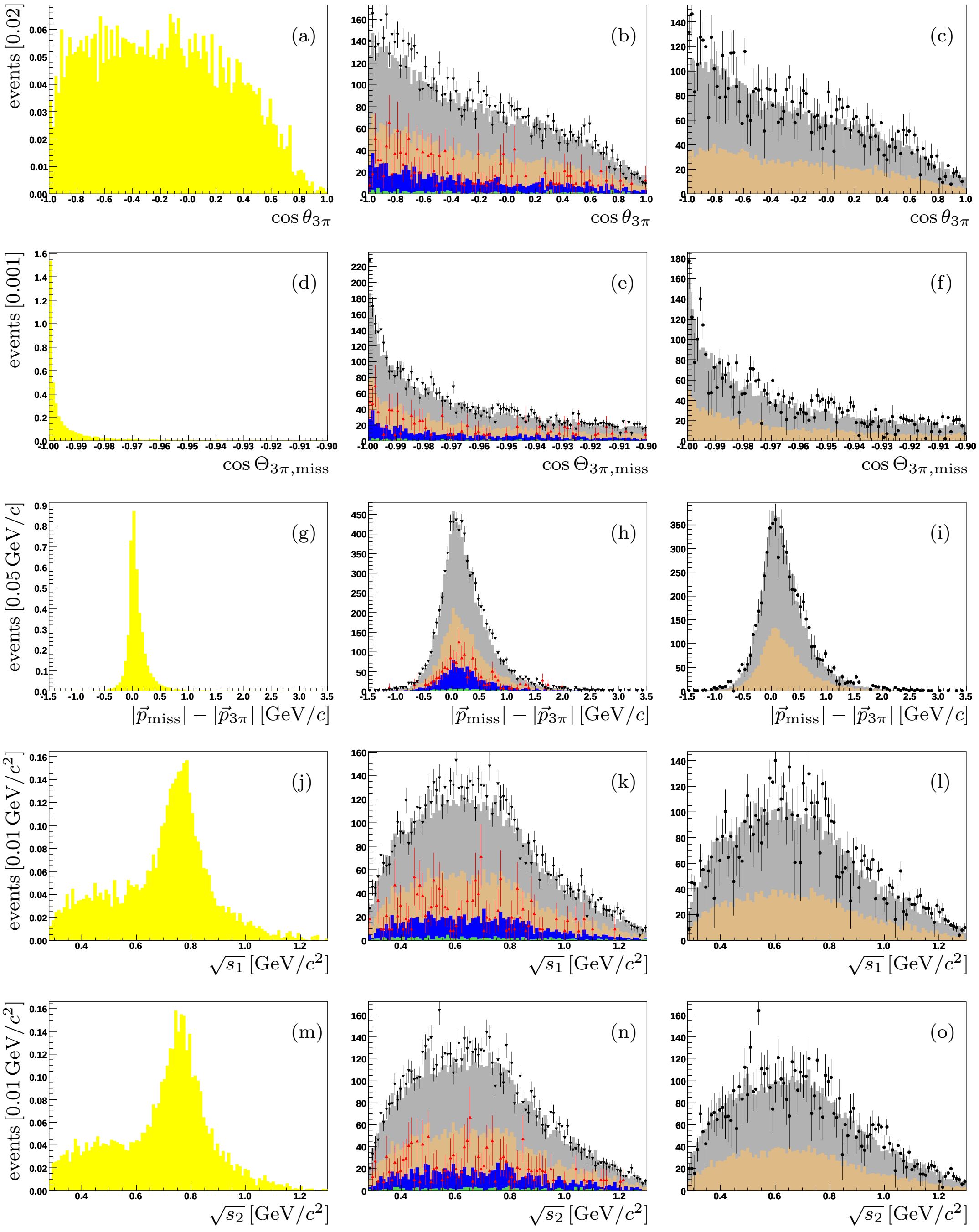


Figure D.7: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) $\cos\theta_{3\pi}$, (d-f) $\cos\Theta_{3\pi,\mathrm{miss}}$, (g-i) $|\vec{p}_{\mathrm{miss}}| - |\vec{p}_{3\pi}|$, (j-l) $\sqrt{s_1}$, and (m-o) $\sqrt{s_2}$ after a_1 candidate preselection. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [▒], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

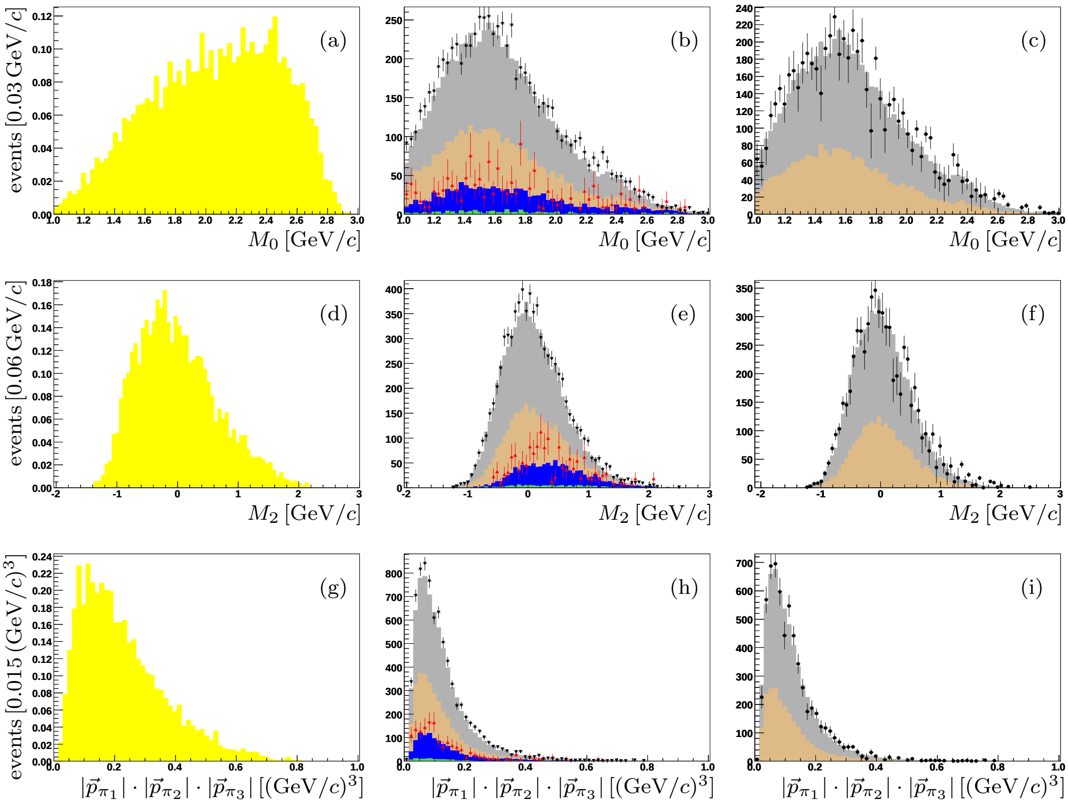


Figure D.8: Kinematical input variables of a_1 candidate based ANN: Distributions of (a-c) M_0, (d-f) M_2, and (g-i) $|\vec{p}_{\pi_1}| \cdot |\vec{p}_{\pi_2}| \cdot |\vec{p}_{\pi_3}|$ after a_1 candidate preselection. (signal [🟨], $u\overline{u}, d\overline{d}, s\overline{s}$ [🟩], $c\overline{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [■], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

D.3 Vertexing Input Variables of a_1 Candidate based ANN

The vertexing input for the a_1 candidate based ANN is represented by:

1. separation of B_{tag} and a_1 decay vertices in transverse plane $\boldsymbol{\Delta^{xy}}$ (Figs. D.9a-c)
2. separation of B_{tag} and a_1 decay vertices along beam axis $\boldsymbol{\Delta^{z}}$ (Figs. D.9d-f)

In contrast to the former B_{tag} vertexing procedure using all remaining tracks not belonging to the 3π candidate, within the recoil analysis the B_{tag} decay vertex has been reconstructed from the tracks and neutrals participating on the B_{tag} reconstruction chain.

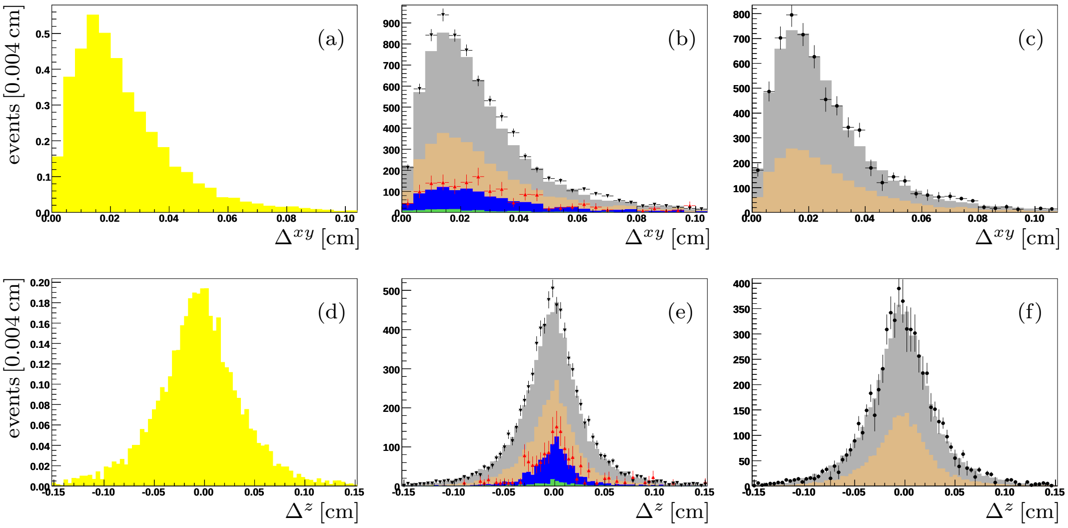


Figure D.9: Vertexing input variables of a_1 candidate based ANN: Distributions of (a-c) Δ^{xy} and (d-f) Δ^{z} after a_1 candidate preselection. (signal [🟨], $u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\overline{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼], On minus Off Peak [▼−▲=●])

Appendix E

Neutral Multiplicity Correction

The N^{CN} correction factors are listed in Tabs. E.1, E.2, and E.3 as derived from the Run13 and Run4 samples at the stage of the B_{tag} selection corresponding to $|c^{D\ell}_{\mathrm{best}}| = 1$. The errors of these factors are determined from the statistical errors of the mean values μ_{Data} and μ_{MC} of the corresponding data and MC $N^{\mathrm{CN}}(E, N^{\mathrm{GTL}})$ distributions. In order to illustrate the development of N^{CN} with energy Fig. E.1 shows $N^{\mathrm{CN}}(E, 8)$ extracted from the Run13 sets.

E/GeV	$N^{\mathrm{GTL}} = 5$		$N^{\mathrm{GTL}} = 6$	
	Run13	**Run4**	**Run13**	**Run4**
$0.02 - 0.03$	0.7240 ± 0.0035	0.7499 ± 0.0035	0.7125 ± 0.0025	0.7304 ± 0.0025
$0.03 - 0.04$	0.8383 ± 0.0044	0.8308 ± 0.0044	0.8251 ± 0.0032	0.8181 ± 0.0032
$0.04 - 0.05$	0.8497 ± 0.0051	0.7954 ± 0.0049	0.8375 ± 0.0036	0.7975 ± 0.0036
$0.05 - 0.06$	0.8597 ± 0.0055	0.8111 ± 0.0054	0.8594 ± 0.0040	0.7908 ± 0.0038
$0.06 - 0.07$	0.8633 ± 0.0058	0.8185 ± 0.0057	0.8601 ± 0.0042	0.7993 ± 0.0041
$0.07 - 0.08$	0.9146 ± 0.0065	0.8603 ± 0.0065	0.9026 ± 0.0047	0.8481 ± 0.0047
$0.08 - 0.09$	0.9223 ± 0.0069	0.8697 ± 0.0068	0.9093 ± 0.0049	0.8646 ± 0.0050
$0.09 - 0.10$	0.9527 ± 0.0074	0.8987 ± 0.0074	0.9439 ± 0.0053	0.8949 ± 0.0055
$0.10 - 0.15$	0.9826 ± 0.0039	0.9452 ± 0.0040	0.9784 ± 0.0028	0.9456 ± 0.0030
$0.15 - 0.20$	1.0216 ± 0.0048	0.9900 ± 0.0050	1.0249 ± 0.0035	1.0024 ± 0.0038
$0.20 - 0.30$	0.9967 ± 0.0042	1.0009 ± 0.0046	0.9996 ± 0.0031	1.0060 ± 0.0034
$0.30 - 0.50$	1.0272 ± 0.0044	1.0155 ± 0.0047	1.0304 ± 0.0033	1.0152 ± 0.0036
$0.50 - 1.00$	1.0094 ± 0.0050	1.0219 ± 0.0055	1.0227 ± 0.0039	1.0297 ± 0.0043
$1.00 - 2.50$	0.9646 ± 0.0093	0.9707 ± 0.0101	0.9877 ± 0.0076	0.9666 ± 0.0083

Table E.1: N^{CN} correction factors $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$ for $N^{\mathrm{GTL}} = 5$ and $N^{\mathrm{GTL}} = 6$

E/GeV	**$N^{\mathrm{GTL}} = 7$**		**$N^{\mathrm{GTL}} = 8$**	
	Run13	**Run4**	**Run13**	**Run4**
$0.02 - 0.03$	0.7026 ± 0.0022	0.7218 ± 0.0023	0.6879 ± 0.0021	0.7103 ± 0.0023
$0.03 - 0.04$	0.8147 ± 0.0028	0.7988 ± 0.0029	0.7962 ± 0.0028	0.7862 ± 0.0029
$0.04 - 0.05$	0.8262 ± 0.0032	0.7788 ± 0.0032	0.8222 ± 0.0032	0.7652 ± 0.0032
$0.05 - 0.06$	0.8437 ± 0.0035	0.7746 ± 0.0035	0.8360 ± 0.0035	0.7613 ± 0.0035
$0.06 - 0.07$	0.8414 ± 0.0037	0.7772 ± 0.0037	0.8306 ± 0.0037	0.7638 ± 0.0037
$0.07 - 0.08$	0.8922 ± 0.0042	0.8288 ± 0.0043	0.8840 ± 0.0042	0.8174 ± 0.0043
$0.08 - 0.09$	0.8969 ± 0.0044	0.8486 ± 0.0046	0.9001 ± 0.0044	0.8325 ± 0.0046
$0.09 - 0.10$	0.9325 ± 0.0048	0.8828 ± 0.0050	0.9227 ± 0.0048	0.8760 ± 0.0051
$0.10 - 0.15$	0.9690 ± 0.0025	0.9332 ± 0.0027	0.9572 ± 0.0025	0.9176 ± 0.0028
$0.15 - 0.20$	1.0160 ± 0.0032	0.9872 ± 0.0035	1.0090 ± 0.0033	0.9805 ± 0.0036
$0.20 - 0.30$	0.9961 ± 0.0029	1.0036 ± 0.0033	0.9876 ± 0.0029	0.9946 ± 0.0034
$0.30 - 0.50$	1.0264 ± 0.0031	1.0215 ± 0.0035	1.0237 ± 0.0032	1.0095 ± 0.0036
$0.50 - 1.00$	1.0328 ± 0.0038	1.0403 ± 0.0044	1.0422 ± 0.0041	1.0448 ± 0.0048
$1.00 - 2.50$	0.9991 ± 0.0081	0.9908 ± 0.0092	1.0336 ± 0.0095	1.0122 ± 0.0109

Table E.2: N^{CN} correction factors $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$ for $N^{\mathrm{GTL}} = 7$ and $N^{\mathrm{GTL}} = 8$

E/GeV	**$N^{\mathrm{GTL}} = 9$**		**$N^{\mathrm{GTL}} = 10$**	
	Run13	**Run4**	**Run13**	**Run4**
$0.02 - 0.03$	0.6759 ± 0.0025	0.7048 ± 0.0028	0.6640 ± 0.0032	0.6883 ± 0.0036
$0.03 - 0.04$	0.7890 ± 0.0033	0.7728 ± 0.0035	0.7839 ± 0.0043	0.7527 ± 0.0046
$0.04 - 0.05$	0.8080 ± 0.0038	0.7511 ± 0.0039	0.7890 ± 0.0049	0.7384 ± 0.0052
$0.05 - 0.06$	0.8212 ± 0.0041	0.7479 ± 0.0042	0.8128 ± 0.0054	0.7330 ± 0.0056
$0.06 - 0.07$	0.8201 ± 0.0044	0.7414 ± 0.0045	0.8040 ± 0.0057	0.7301 ± 0.0059
$0.07 - 0.08$	0.8720 ± 0.0050	0.8001 ± 0.0052	0.8603 ± 0.0065	0.7928 ± 0.0070
$0.08 - 0.09$	0.8807 ± 0.0052	0.8228 ± 0.0056	0.8671 ± 0.0069	0.8124 ± 0.0076
$0.09 - 0.10$	0.9082 ± 0.0057	0.8562 ± 0.0061	0.8898 ± 0.0075	0.8387 ± 0.0082
$0.10 - 0.15$	0.9420 ± 0.0030	0.9048 ± 0.0034	0.9353 ± 0.0040	0.8883 ± 0.0045
$0.15 - 0.20$	0.9977 ± 0.0040	0.9641 ± 0.0045	0.9843 ± 0.0053	0.9617 ± 0.0061
$0.20 - 0.30$	0.9773 ± 0.0035	0.9870 ± 0.0042	0.9709 ± 0.0048	0.9692 ± 0.0058
$0.30 - 0.50$	1.0166 ± 0.0040	1.0049 ± 0.0046	1.0052 ± 0.0054	0.9988 ± 0.0065
$0.50 - 1.00$	1.0438 ± 0.0054	1.0555 ± 0.0065	1.0512 ± 0.0078	1.0506 ± 0.0095
$1.00 - 2.50$	1.0436 ± 0.0139	1.0519 ± 0.0166	1.0768 ± 0.0223	1.0972 ± 0.0275

Table E.3: N^{CN} correction factors $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}}$ for $N^{\mathrm{GTL}} = 9$ and $N^{\mathrm{GTL}} = 10$

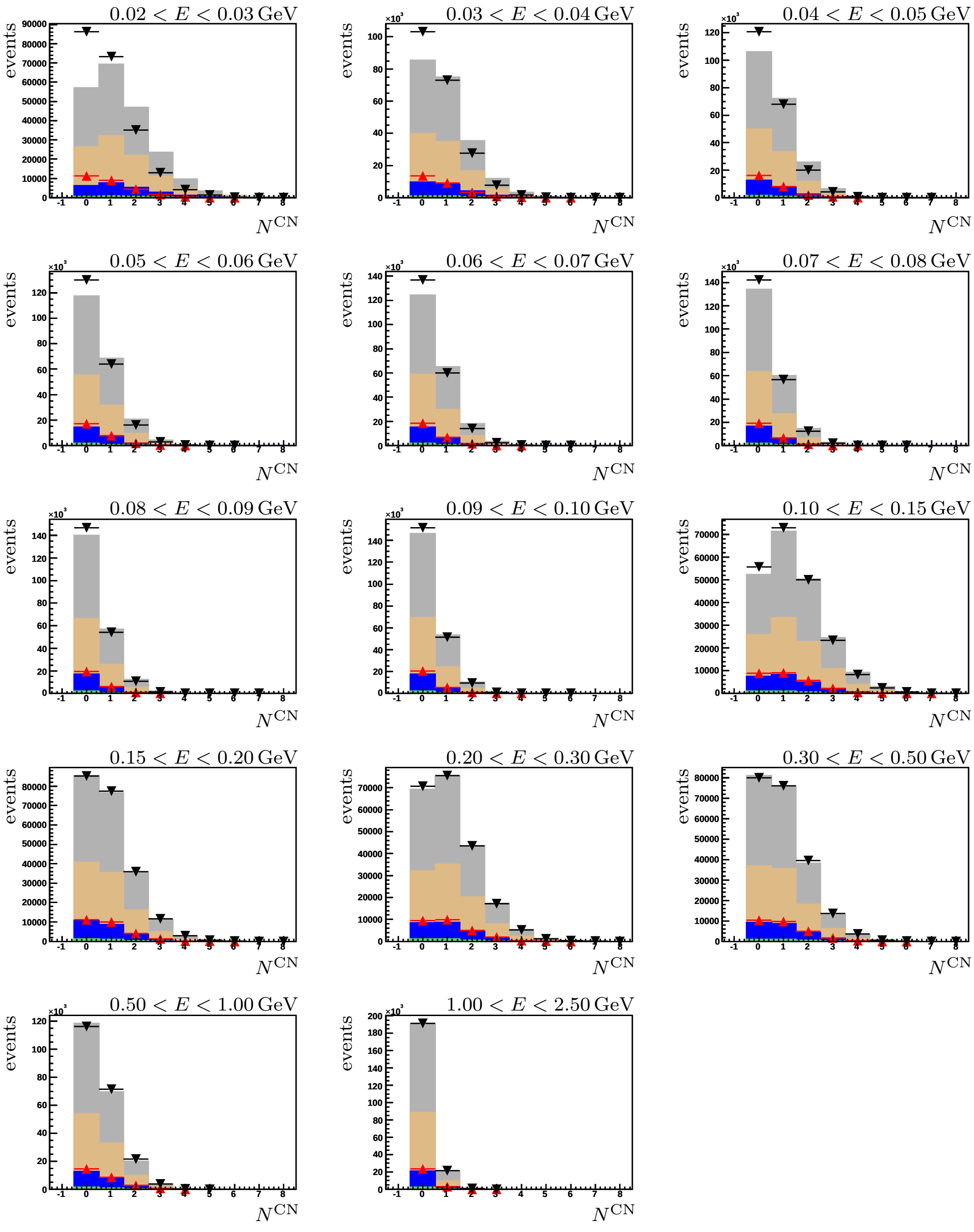


Figure E.1: Neutral Multplicities after $B_{\rm tag}$ selection ($|c^{D\ell}_{\rm best}| = 1$) for events with $N^{\rm GTL} = 8$ extracted from the Run13 sample in bins of the CN energy: The energy ranges are given above the figures. ($u\bar{u}, d\bar{d}, s\bar{s}$ [🟩], $c\bar{c}$ [🟦], $\tau^+\tau^-$ [🟪], $B^0\bar{B}^0$ [🟫], B^+B^- [⬜], Off Peak [▲], On Peak [▼])

Appendix F

Branching Fraction Calculation for $B^+ \to \overline{D}^{*0}\ell^+\nu_\ell$

The branching fraction for a given process $B \to D^*\ell\nu_\ell$ is defined as the ratio of the partial decay width $\Gamma(B \to D^*\ell\nu_\ell)$ and the total B decay rate $\Gamma_{\text{tot}} = 1/\tau_B$

$$\mathcal{B}(B \to D^*\ell\nu_\ell) = \frac{\Gamma(B \to D^*\ell\nu_\ell)}{\Gamma_{\text{tot}}} = \Gamma(B \to D^*\ell\nu_\ell) \cdot \tau_B \tag{F.1}$$

with the mean B meson lifetime τ_B. Eq. F.1 can be formulated separately for charged and neutral B decays

$$\begin{aligned} \mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) &= \Gamma(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) \cdot \tau_{B^+} \,, \\ \mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) &= \Gamma(B^0 \to D^{*-}\ell^+\nu_\ell) \cdot \tau_{B^0} \,. \end{aligned} \tag{F.2}$$

Isospin symmetry implies equal decay rates for the charged and neutral decay since on quark-level both processes are identical apart from the light spectator quark, i.e.

$$\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell) = \Gamma(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) \,. \tag{F.3}$$

Exploiting Eqs. F.2 this equation translates into

$$\frac{\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)}{\tau_{B^0}} = \frac{\mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell)}{\tau_{B^+}} \,. \tag{F.4}$$

From Eq. F.4 follows

$$\frac{\mathcal{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell)}{\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)} = \frac{\tau_{B^+}}{\tau_{B^0}} \,, \tag{F.5}$$

which relates the ratio of the branching fractions with the ratio of the mean lifetimes of charged and neutral B mesons.

Appendix G

$\cos\Theta_{B,D\ell}$ Calculation

For the three-body decay $B \to D^{(*)}\ell\nu_\ell$ the four-momentum of the B meson is given by

$$p_B = (p_D + p_\ell) + p_\nu = p_{D\ell} + p_\nu \tag{G.1}$$

with the four-momentum vector of the combined $D\ell$ system $p_{D\ell}$ and the four-momentum vector of the neutrino p_ν. This leads to

$$p_\nu = p_B - p_{D\ell}\,. \tag{G.2}$$

Eq. G.2 can be squared and with the *Einstein* convention it follows

$$\begin{aligned} p_\nu^2 &= p_B^2 - 2p_B p_{D\ell} + p_{D\ell}^2\,, \\ m_\nu^2 &= m_B^2 - 2E_B E_{D\ell} + 2\vec{p}_B\vec{p}_{D\ell} + m_{D\ell}^2\,. \end{aligned} \tag{G.3}$$

Assuming $m_\nu = 0$ one further calculates

$$0 = m_B^2 + m_{D\ell}^2 - 2E_B E_{D\ell} + 2|\vec{p}_B||\vec{p}_{D\ell}|\cos\Theta_{B,D\ell} \tag{G.4}$$

and one obtains

$$\cos\Theta_{B,D\ell} = \frac{2E_B E_{D\ell} - m_B^2 - m_{D\ell}^2}{2|\vec{p}_B||\vec{p}_{D\ell}|}. \tag{G.5}$$

Appendix H

Parametrization of $E_{\mathrm{rem}}^{\mathrm{CN}}$ and $\mathcal{XNN}^{3\pi}$

This section gives the analytical functions used for the optimization of the final selection within the $B \to D^{(*)}\ell\nu_\ell$ recoil analysis.

H.1 $E_{\mathrm{rem}}^{\mathrm{CN}}$ Parametrization

The function $f(E_{\mathrm{rem}}^{\mathrm{CN}})$ describing the $E_{\mathrm{rem}}^{\mathrm{CN}}$ distribution consists of two parts; a Logarithmic Normal distribution $L(x)$ and an exponential part $E(x)$ given as

$$L(x) = \frac{1}{\sqrt{2\pi}\sigma_L} \exp\left[-\frac{\ln^2\left(1 + \frac{\sinh(\tau_L\sqrt{\ln 4})}{\sqrt{\ln 4}} \frac{x-\mu_L}{\sigma_L}\right)}{2\tau_L^2} - \frac{\tau_L^2}{2} \right] , \tag{H.1}$$

$$E(x) = \frac{b_E}{e^{b_E \cdot x_{\max}} - e^{b_E \cdot x_{\min}}} \cdot e^{b_E \cdot x} . \tag{H.2}$$

It should be noted that $L(x)$ is normalized in the range $[-\infty, +\infty]$, while $E(x)$ is normalized on the fit range $[x_{\min}, x_{\max}]$. The sum of these two components has been used to describe the $E_{\mathrm{rem}}^{\mathrm{CN}}$ shapes in signal as well as combined background MC

$$f(E_{\mathrm{rem}}^{\mathrm{CN}}) = N_L \cdot L(E_{\mathrm{rem}}^{\mathrm{CN}}) + N_E \cdot E(E_{\mathrm{rem}}^{\mathrm{CN}}) . \tag{H.3}$$

Here, the exponential part has been explicitly introduced to approximate the increase towards $E_{\mathrm{rem}}^{\mathrm{CN}} = 0$.

The parameter values after the fits to the signal and background $E_{\mathrm{rem}}^{\mathrm{CN}}$ distributions are summarized in Tab. H.1. The fits yield reasonable χ^2/n_{dof}. The parameters describing $E(E_{\mathrm{rem}}^{\mathrm{CN}})$ in the fit to the background distribution are very uncertain since the exponential part is determined by a few bins in the low $E_{\mathrm{rem}}^{\mathrm{CN}}$ region. The χ^2 probablities $\mathcal{P}_{\chi^2}$ are of order 10 % and the fitted function is assumed to reproduce the fractional integrals used for the cut optimization procedure.

Quantity	**Signal MC**	**Combined Background MC**
N_L	2.010 ± 0.080	39844.478 ± 10421.306
μ_L	0.206 ± 0.001	1.334 ± 0.169
σ_L	0.141 ± 0.001	3.644 ± 0.471
τ_L	0.693 ± 0.009	1.643 ± 0.097
N_E	2.142 ± 0.071	80.941 ± 87.288
b_E	-5.938 ± 0.239	-3.190 ± 4.648
χ^2/n_{dof}	53.6/44	55.1/44
$\mathcal{P}_{\chi^2}$	15.2 %	12.2 %

Table H.1: Fit parameters of the function describing $E_{\mathrm{rem}}^{\mathrm{CN}}$ after fit

H.2 $\boldsymbol{\mathcal{XNN}^{3\pi}}$ Parametrization

The $\mathcal{XNN}^{3\pi}$ distributions have been approximated by a combination of two gaussians $G_1(\mathcal{XNN}^{3\pi})$ and $G_2(\mathcal{XNN}^{3\pi})$

$$f(\mathcal{XNN}^{3\pi}) = N \cdot (1 - Q_{G_2}) \cdot G_1(\mathcal{XNN}^{3\pi}) + N \cdot Q_{G_2} \cdot G_2(\mathcal{XNN}^{3\pi}) \tag{H.4}$$

with

$$G_i(x) = \frac{1}{\sqrt{2\pi}\sigma_{G_i}} \exp\left[-\frac{(x-\mu_{G_i})^2}{2\sigma_{G_i}^2}\right] . \tag{H.5}$$

In order to take into account possible asymmetries of the $\mathcal{XNN}^{3\pi}$ distributions, the means μ_{G_i} are allowed to take different values. Given this parametrization, the function $f(\mathcal{XNN}^{3\pi})$ is normalized in the range $[-\infty, +\infty]$. The parameter N is equal to the integral of this function, while Q_{G_2} describes the contribution of $G_2(\mathcal{XNN}^{3\pi})$ to the entire integral of $f(\mathcal{XNN}^{3\pi})$. The parameter values after the fits are listed in Tab. H.2. The fit to the signal MC yields a low $\mathcal{P}_{\chi^2}$ of about 4 %. However, $f(E_{\mathrm{rem}}^{\mathrm{CN}})$ and $f(\mathcal{XNN}^{3\pi})$ have only been used to find the optimal cut combination. The final numbers of signal and background events have been extracted by counting events passing the optimal cuts on $E_{\mathrm{rem}}^{\mathrm{CN}}$ and $\mathcal{XNN}^{3\pi}$.

Quantity	**Signal MC**	**Combined Background MC**
N_{G_1}	3.870 ± 0.008	852.999 ± 29.445
μ_{G_1}	0.376 ± 0.000	0.499 ± 0.006
σ_{G_1}	0.073 ± 0.001	0.096 ± 0.004
Q_{G_2}	0.581 ± 0.019	0.028 ± 0.031
μ_{G_2}	0.389 ± 0.001	0.722 ± 0.067
σ_{G_2}	0.131 ± 0.001	0.061 ± 0.032
χ^2/n_{dof}	54.1/38	39.6/33
$\mathcal{P}_{\chi^2}$	4.4 %	19.9 %

Table H.2: Fit parameters of the function describing $\mathcal{XNN}^{3\pi}$ after fit