$B \to (\rho/\omega)\,\gamma$ **at** *BABAR*

by

Karsten Köneke

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

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author .....
Department of Physics
May 25, 2007

Certified by .....
Gabriella Sciolla
Assistant Professor
Thesis Supervisor

Accepted by .....
Thomas J. Greytak
Associate Department Head for Education

$B \to (\rho/\omega)\,\gamma$ at *BABAR*

by

Karsten Köneke

Submitted to the Department of Physics
on May 25, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

This document describes the measurements of the branching fractions and isospin violations of the radiative electroweak penguin decays $B \to (\rho/\omega)\,\gamma$ at the asymmetric-energy e^+e^- PEP-II collider with the *BABAR* detector. Together with the previously measured branching fractions of the decays $B \to K^*\gamma$ the ratio of CKM-matrix elements $|V_{td}/V_{ts}|$ are extracted and the length of the far side of the unitarity triangle is determined.

Thesis Supervisor: Gabriella Sciolla
Title: Assistant Professor

"[...] Daß ich erkenne, was die Welt
Im Innersten zusammenhält, [...]"[1]

Acknowledgments

To my family

First and foremost, I want to thank my family for their never ending support during all the years leading up to this thesis. I would have never come this far without them.

Furthermore, I want to thank my advisor for her support during the research leading up to the results presented in this document.

And off course I want to thank everyone who helped out academically or morally during the past years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Currently, two high precision experiments designed to study the B meson sector are in operation: *BABAR* and *Belle*. Both experiments are operating at asymmetric energy electron-positron colliders running at a center of momentum (CM) energy of 10.58 GeV. This energy corresponds to the mass of the $\Upsilon(4S)$ resonance, a meson which is a bound state of a b quark and an anti-b quark in the 4S configuration. Both *BABAR* and *Belle* are high luminosity experiments with a current peak luminosity of $1.21 \cdot 10^{34} cm^{-2} s^{-1}$ and $1.65 \cdot 10^{34} cm^{-2} s^{-1}$ respectively.

Also, there is data available from the CLEO collaboration, the previous-generation experiment studying B mesons at an e^+e^- collider. This experiment was a symmetric machine also running at the $\Upsilon(4S)$ resonance. It was acquiring data at a much lower rate, the peak luminosity was about $8.5 \cdot 10^{32} cm^{-2} s^{-1}$.

Furthermore, the two detectors at the Tevatron, CDF and D$\emptyset$, are also studying the physics in the b quark sector. These two experiments are not as clean as the other three mentioned above since the Tevatron is a hadron collider and thus the initial state is not known like it is in the e^+e^- machines. Also, there are a lot more final state particles in a hadron collider than in a e^+e^- collider. Thus it is much harder to isolate a photon from the rest of the event which will be needed for this analysis. But the Tevatron can measure another quantity, the oscillation frequency of B_s mixing, where the B_s meson is a bound state of a b anti-quark and an s quark. As described later, that measurement and the measurement of the branching fraction

of the decay $B \to (\rho/\omega)\,\gamma$ are tied to the same quantities according to our current understanding of elementary particle physics, the Standard Model. Thus, these two measurements should come to the same conclusions, if the Standard Model is correct.

1.1 Theoretical Motivation

1.1.1 The Standard Model

Today's understanding of the physics of elementary particles is described in terms of the Standard Model of Elementary Particle Physics (SM). This model is dealing with two types of spin-$\frac{1}{2}$ particles (quarks and leptons) and three types of forces (strong, electromagnetic and weak force). The forces are mediated by three different types of force carriers: 8 gluons, the photon, and the $W^{\pm}$ and Z^0 vector (=spin-1) bosons, respectively. The SM is a gauge theory. The invariance of the SM lagrangian under local transformation belonging to the $U(1)_Y \otimes SU(2)_L$ group yields the electroweak part of the SM, while the local transformation belonging to the $SU(3)_C$ group leads to the part of the SM dealing with strong interactions, as described by Quantum Chromo Dynamics (QCD). To be more precise, the required invariance of the lagrangian under these local gauge transformations leads to the three types of force carriers, the vector bosons. Therefore, these force carriers are also known under the name "gauge bosons".

Up to now, the theoretical predictions of the SM agree to a high precision with experimental measurements. Deviations are expected at some (unknown) level since this model is intrinsically incomplete. It becomes inconsistent at high energies of about 1 TeVand higher.

1.1.2 The CKM Matrix

In the SM, the quark mass (or flavor) matrices are in general not diagonal. In order to transform the quark flavor states into the weak eigenstates needed in the lagrangian, one performs a rotation in the quark flavor space. This rotation of the flavor eigenstates leads to the diagonal weak eigenstates which are used in the lagrangian. This

leads to a non-vanishing rotation matrix in the weak charged current term of the lagrangian, the Cabibbo-Kobayashi-Maskawa (CKM) matrix. This unitary matrix is usually interpreted as a redefinition of the down-type eigenstates of the weak interaction (primed) as a linear combination of the down-type flavor eigenstates (unprimed).

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

This CKM matrix is a unitary complex matrix and it has for the case of three quark generations 18 parameters. But since the matrix is unitary, only nine of these 18 parameters are independent. Furthermore, one can always redefine the individual phases of the six quark fields and thus remove five more independent parameters. This leaves us with four independent parameters, three real parameters (angles) and one irreducible phase factor. This phase factor is the source of CP violation in the SM.

With this remaining four independent parameters, one can expand the CKM matrix in the so-called Wolfenstein parameterization of the CKM matrix [2][3]:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.2}$$

The unitarity conditions of the CKM matrix are

$$\sum_{i=1}^{3} V_{ji} V_{ki}^* = 0 = \sum_{i=1}^{3} V_{ij} V_{ik}^* \quad \text{for } j \neq k. \tag{1.3}$$

These can be visualized with six triangles. But for only two of those the length of all three sides are of the same order of λ ($\lambda \approx |V_{us}| \approx |V_{cd}| \approx 0.22$). One of these two triangles is usually referred to as "The Unitary Triangle". This triangle originates from the CKM unitarity condition for the first and third column of the CKM matrix,

i.e., the ones relevant in the B meson system

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.4}$$

The triangle is furthermore normalized by dividing the above condition by $V_{cd}V_{cb}^*$. This final standard unitarity triangle is shown in Figure 1-1 in the $(\bar{\rho}, \bar{\eta})$ plane where $\bar{\rho}$ and $\bar{\eta}$ are related to the usual Wolfenstein parameters as

$$\begin{aligned} \bar{\rho} &= \rho - \tfrac{1}{2}\rho\lambda^2 + \mathcal{O}(\lambda^4) \\ \bar{\eta} &= \eta - \tfrac{1}{2}\eta\lambda^2 + \mathcal{O}(\lambda^4). \end{aligned} \tag{1.5}$$

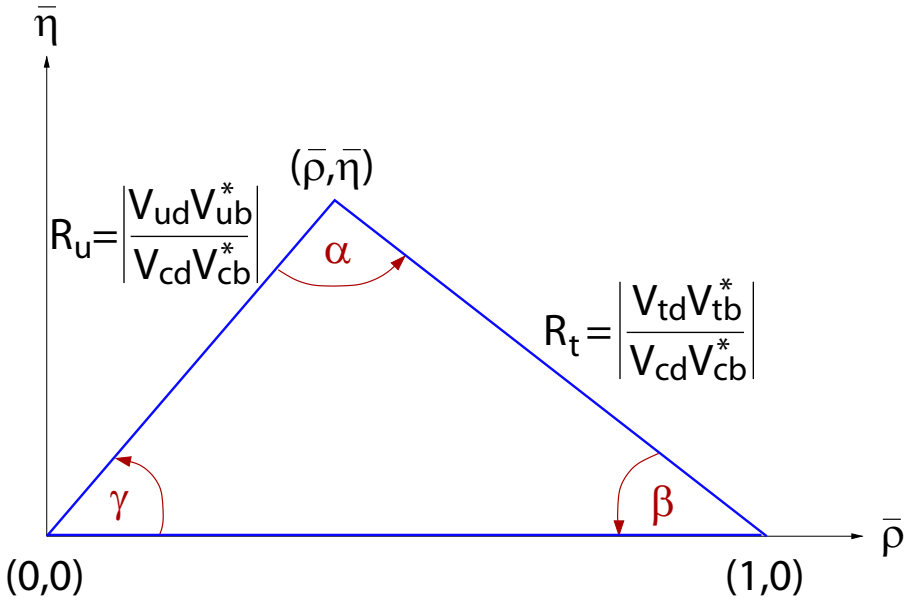


Figure 1-1: The CKM unitarity triangle. Modified from [4].

1.1.3 The $B \to (\rho/\omega)\,\gamma$ Radiative Penguin Decay

A penguin decay is a flavor changing neutral current (FCNC) process. FCNC processes in the SM only take place at higher order, *i.e.,* with at least one loop in the Feynman diagram. The name "penguin decay" was first used as a result of a bet in [5]. In the case of $B \to (\rho/\omega)\,\gamma$, a W boson and an up-type quark are in the loop (see

Figure 1-2).

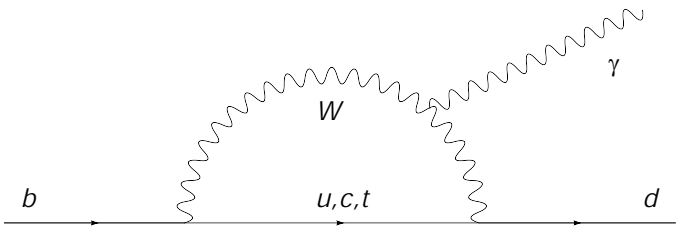


Figure 1-2: Feynman diagram for the $b \to d\gamma$ transition. This is the leading order "penguin" diagram for this transition.

There are several motivations to study $B \to (\rho/\omega)\,\gamma$ (and also $B \to K^*\gamma$) decays. These decays proceed at leading order via an electroweak penguin loop shown in Figure 1-2. Since the top-quark penguin diagrams are the largest contributions to these processes, the measurement of observables in these decays are probes of the top-quark couplings.

Theory Framework

In order to compute observables of the decay $B \to (\rho/\omega)\,\gamma$, an effective theory is used. In this particular case, the heavy quark effective theory is used. This effective theory integrates out the heavy fields from the full theory. The effective interaction Hamiltonian is [6]

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{td}^* V_{tb} \sum_{i=1}^{10} C_i(\mu)\mathcal{O}_i(\mu)\ , \tag{1.6}$$

where G_F is the Fermi constant, V_{tj} are CKM matrix elements, $C_i(\mu)$ are Wilson coefficient (dependent on the renormalization scale μ) and $\mathcal{O}_i(\mu)$ are the operators. Unitarity of the CKM matrix has been assumed and everything beyond leading order in α_{em}, m_b/m_W, m_d/m_b and V_{ub}/V_{cb} has been neglected. The short distance QCD effects due to hard gluon exchanges between the quark lines of the leading order one loop Feynman diagram is contained in the Wilson coefficients which can be calculated perturbativly. The Hamiltonian relates the final ρ/ω meson state to the initial B meson state.

After integrating out the heavy degrees of freedom (*i.e.,* the top quark and the

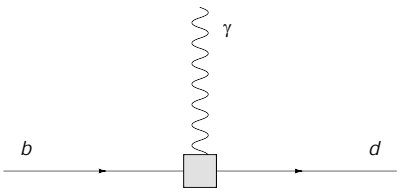


Figure 1-3: Effective Feynman diagram for the $b \to d\gamma$ transition. After integrating out the heavy degrees of freedom, this effective diagram remains.

W boson) from the dominant Feynman diagram shown in Figure 1-2, the effective Feynman diagram shown in Figure 1-3 remains. This diagram is described in the Hamiltonian (see Equation 1.6) by

$$\mathcal{O}_7(\mu) = \frac{e\ m_b\,(\mu)}{8\pi^2}\left(\bar{d}_\alpha \sigma^{\mu\nu}\,(1+\gamma_5)\,b_\alpha\right) F_{\mu\nu}, \tag{1.7}$$

where b_α is the initial b quark state, $\bar{d}_\alpha$ is the final d quark state, $F_{\mu\nu}$ is the photon field strength tensor, e is the electromagnetic coupling constant, $m_b\,(\mu)$ is the scale-dependent b quark mass and $\sigma^{\mu\nu}$ is the only possible antisymmetric tensor needed to contract $F_{\mu\nu}$ with. In fact, this is the simplest possible operator that describes the transition of a b quark into a d quark and a photon.

Ratio of Branching Fractions and $|V_{td}/V_{ts}|$

The ratio of the CKM matrix elements V_{td}/V_{ts} can be calculated from the ratio of branching fractions of the two exclusive radiative penguin decays $B \to (\rho/\omega)\,\gamma$ and $B \to K^*\gamma$ [7]:

$$\frac{\mathcal{B}(B\to\rho\gamma)}{\mathcal{B}(B\to K^*\gamma)} = S_\rho \left|\frac{V_{td}}{V_{ts}}\right|^2 \frac{\left(1-m_\rho^2/M^2\right)^3}{\left(1-m_{K^*}^2/M^2\right)^3}\zeta^2\left[1+\Delta R(\rho/K^*)\right] \tag{1.8}$$

where $\zeta = \xi_\perp^\rho(0)/\xi_\perp^{K^*}(0)$ is the ratio of form factors computed in Heavy Quark Effective Theory (HQET), $S_\rho = 1(1/2)$ are isospin weights for the charged (neutral) ρ meson and $\Delta R(\rho/K^*)$ is a dynamical function calculated for example in [7] which accounts for different dynamics in the three decays like vertex, hard-spectator, weak annihilation or weak exchange contributions. The weak annihilation contribution to

the $B^+ \to \rho^+\gamma$ decay and the W exchange contribution to the neutral B meson decay are shown in Figure 1-4.

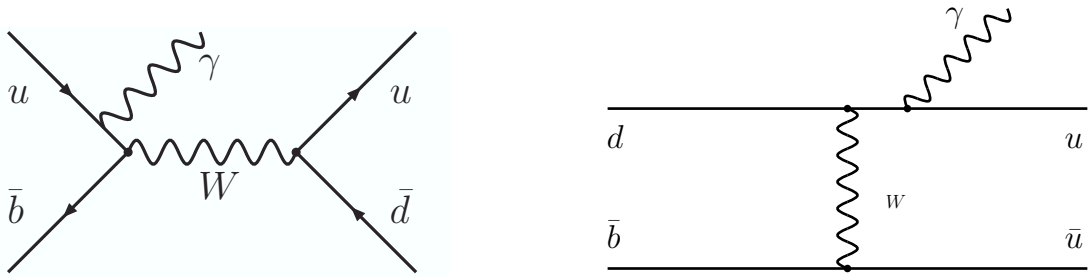


Figure 1-4: Second order contributions to the $B \to (\rho/\omega)\,\gamma$ decay. The Feynman diagram on the left shows the weak annihilation contribution to the $B^+ \to \rho^+\gamma$ decay mode and the Feynman diagram on the left shows the W exchange contribution to the neutral B meson decay.

The latest theory calculation [8] (based on Light-Cone Sum Rules) of the ratio of form factors gives

$$\xi_\rho = \zeta_\rho^{-1} = \frac{T_1^{B\to K^*(0)}}{T_1^{B\to\rho(0)}} = 1.17 \pm 0.09 \tag{1.9}$$

and

$$\xi_\omega = \zeta_\omega^{-1} = \frac{T_1^{B\to K^*(0)}}{T_1^{B\to\omega(0)}} = 1.30 \pm 0.10. \tag{1.10}$$

Reference [9] argues that $\Delta R(\rho/K^*)$ is actually depending on both QCD and CKM parameters. In particular, for the determination of this quantity, a set of input parameters R_b and γ is needed where

$$R_b = \left(1 - \frac{\lambda^2}{2}\right)\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right| \tag{1.11}$$

is the short side of the unitarity triangle and γ is the angle between that side and the base of the unitarity triangle. This puts $\Delta R(\rho/K^*)$ in the neighborhood of 0.1 ± 0.1, but for the actual determination of the branching fraction ratio, the full description with appropriate errors has been used.

This same ratio of CKM matrix elements is determined from the ratio of $\Delta m_d/\Delta m_s$ where $\Delta m_{(d/s)}$ are matter-antimatter oscillation frequencies of the neutral $B_{(d/s)}$ meson system. These $B\bar{B}$ oscillations proceed in the SM at leading order via the Feyn-

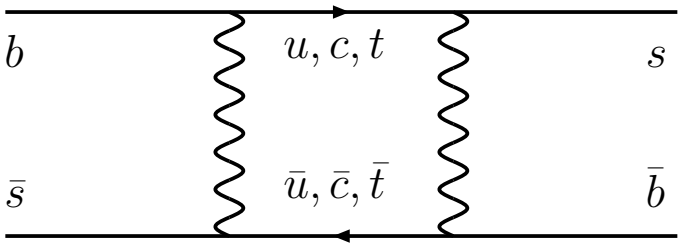


Figure 1-5: Feynman diagram for $B\bar{B}$ oscillations.

man diagram shown in Figure 1-5. The oscillation frequency of the B_d system has been precisely measured by the B-Factories [10, 11, 12, 13, 14, 15, 16, 17] while the oscillation frequency of the B_s system has been recently measured by the CDF collaboration [18]. Even though the underlying dynamics of $B\bar{B}$ oscillations and radiative penguin decays are very different, in the framework of the SM, both ways of measuring $|V_{td}/V_{ts}|$ yield the same result. It would be a clear indication of physics beyond the SM if a difference is found between these two independent results.

Branching Fractions

One can easily imagine a new (e.g. supersymmetric) particle in the penguin loop. This non-SM physics contribution could change the branching fraction of this decay measurably, but it does not necessarily change the ratio of branching fractions between $B \to (\rho/\omega)\,\gamma$ and $B \to K^*\gamma$ decays if the new particle couples the with the same strength to the s and d quark. If this would be the case, the new physics contributions would be better visible in the $B \to K^*\gamma$ case due to the larger branching fraction. If the couplings are different, *e.g.,* only slightly effecting the $B \to K^*\gamma$ branching fraction, but significantly effecting the $B \to (\rho/\omega)\,\gamma$ branching fraction, then this would be visible in the determination of the ratio of branching fractions discussed above. Thus the actual measured value of the branching fractions is, taken on its own, less interesting due to the large theory uncertainties. These large uncertainties of theoretical calculations (see Table 1.1) are due to difficulties in calculating the hadronization of the final state mesons. Also QCD corrections to the penguin loop

itself are hard to compute.

Calculation	$\mathcal{B}(B^+ \to \rho^+\gamma)$	$\mathcal{B}(B^0 \to \rho^0\gamma)$	$\mathcal{B}(B^0 \to \omega\gamma)$
Ali & Parkhomenko [19]	$1.37 \pm 0.26(\text{th.}) \pm 0.09(\text{exp.})$	$0.65 \pm 0.12(\text{th.}) \pm 0.03(\text{exp.})$	$0.53 \pm 0.12(\text{th.}) \pm 0.02(\text{exp.})$
Bosch & Buchalla [20]	$1.58^{+0.53}_{-0.46}$	$0.76^{+0.26}_{-0.23}$	
Ball, Jones & Zwicky [9]	1.16 ± 0.26	0.55 ± 0.13	0.44 ± 0.10

Table 1.1: Next to leading order predictions for the $B \to (\rho/\omega)\,\gamma$ decay modes. All branching fractions are in units of 10^{-6}.

Besides the already mentioned transitions of the type $b \to s\gamma$, there are other decays related to $B \to (\rho/\omega)\,\gamma$. Two of these decays are $B^0 \to \phi\gamma$ and $B^0 \to J/\psi\,\gamma$. The difference is that these decays proceed at leading order via the annihilation type Feynman diagrams. Thus their branching fractions are expected to be unobservable at the current B factories. Based on the SM, they are expected to be $\mathcal{B}\,(B^0 \to \phi\gamma) = \left(2.7^{+0.3+1.2}_{-0.6-0.6}\right) \times 10^{-11}$ and $\mathcal{B}\,(B^0 \to J/\psi\,\gamma) = \left(4.5^{+0.6+0.7}_{-0.5-0.6}\right) \times 10^{-7}$ [21]. Even though the later branching fraction is sizable, the branching fraction of $\mathcal{B}\,(J/\psi \to l^+l^-) = (5.94 \pm 0.06)\,\%$ [22] (where $l^+ = e^+, \mu^+$) is too small to allow for a detection of this decay at the current B factories.

The CKM Angle α

The CKM angle α is the angle at the apex of the unitarity triangle. It is defined as

$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \tag{1.12}$$

where V_{ij} are CKM matrix elements. Since the weak annihilation diagram (see Figure 1-4) contributing to the $B^+ \to \rho^+\gamma$ decay is proportional to $V_{ud}V_{ub}^*$ and the top quark dominated penguin diagram (see Figure 1-2) is proportional to $V_{td}V_{tb}^*$, the ratio of these two diagrams is proportional to

$$\frac{V_{ub}V_{ud}^*}{V_{tb}V_{td}^*} = -\left|\frac{V_{ub}V_{ud}^*}{V_{tb}V_{td}^*}\right| e^{i\alpha}. \tag{1.13}$$

Thus, the ratio of branching fraction of $B^+ \to \rho^+\gamma$ to $B^0 \to \rho^0\gamma$ is dependent on α. More precisely, the quantity measuring isospin violation between the two decays $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ defined as

$$\Delta \equiv \frac{\Gamma(B^+ \to \rho^+\gamma)}{2\Gamma(B^0 \to \rho^0\gamma)} - 1 \tag{1.14}$$

is sensitive to $\cos\alpha$ [19].

1.2 Previous Measurements

Prior to this analysis, only the *Belle* collaboration found a positive result in one channel, namely an approximately five sigma result in the neutral $B^0 \to \rho^0\gamma$ decay mode. No other experiment was able to observe this decay mode nor has any of the other two decay modes ($B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$) been observed. This situation is summarized in Table 1.2.

Experiment	$\mathcal{B}(B^+ \to \rho^+\gamma)$	$\mathcal{B}(B^0 \to \rho^0\gamma)$	$\mathcal{B}(B^0 \to \omega\gamma)$	Combined result
CLEO[23]	< 13	< 17	< 9.2	
BABAR[24]	< 1.76	< 0.36	< 0.97	< 1.16
BELLE[25]	$0.55^{+0.42\ +0.09}_{-0.36\ -0.08}$	$1.25^{+0.37\ +0.07}_{-0.33\ -0.06}$	$0.56^{+0.34\ +0.05}_{-0.27\ -0.10}$	$1.32^{+0.34\ +0.10}_{-0.31\ -0.09}$

Table 1.2: Current experimental results for $\mathcal{B}(B \to \rho\gamma)$ and $\mathcal{B}(B \to \omega\gamma)$. Limits are shown at 90% confidence level. The Belle paper [25] does not provide upper limits for the ρ^+ and ω modes. Where appropriate, the first error is statistical, the second systematic. All numbers are in units of 10^{-6}.

For this analysis the charged decay mode $B^+ \to \rho^+\gamma$ is assumed to have a branching fraction of $1.0 \cdot 10^{-6}$ and the branching fractions of both neutral decay modes $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$ are assumed to be $0.5 \cdot 10^{-6}$ as summarized in Tabel 1.3. This values are used for the purpose of optimizing the analysis, *i.e.,* for selecting cuts, choosing the fit strategy, etc.

Mode	Assumed branching fractions
$B^+ \to \rho^+\gamma$	$1.0 \cdot 10^{-6}$
$B^0 \to \rho^0\gamma$	$0.5 \cdot 10^{-6}$
$B^0 \to \omega\gamma$	$0.5 \cdot 10^{-6}$

Table 1.3: Assumed branching fractions for this analysis.

Chapter 2

The *BABAR* Experiment

The *BABAR* experiment is based on an asymmetric energy e^+e^- collider at the Stanford Linear Accelerator Center (SLAC) operating at a center of momentum (CM) energy of 10.58 GeV, coinciding with the $\Upsilon(4S)$ resonance. *BABAR* started taking data in 1999 and will continue running until fall of 2008.

2.1 The Stanford Linear Accelerator Center

The Stanford Linear Accelerator Center (SLAC) was established in 1962 and is located in Menlo Park, California. In 1966, the main linear accelerator (Linac)went into operation. This is a three kilometer long electron and positron accelerator now used to inject the electron and positron beams into the PEP-II storage rings. The linac is still the world's largest and most powerful linear accelerator and it can provide beam energies of up to about 50 GeV. A schematic overview of the experimental site at SLAC can be seen in Figure 2-1.

2.2 The PEP-II Collider

The PEP-II collider is a two storage ring machine. One ring is an upgrade of the previously existing PEP collider which now stores a 9.0 GeV electron beam. The second storage ring is a new ring storing a 3.1 GeV positron beam. The PEP-II

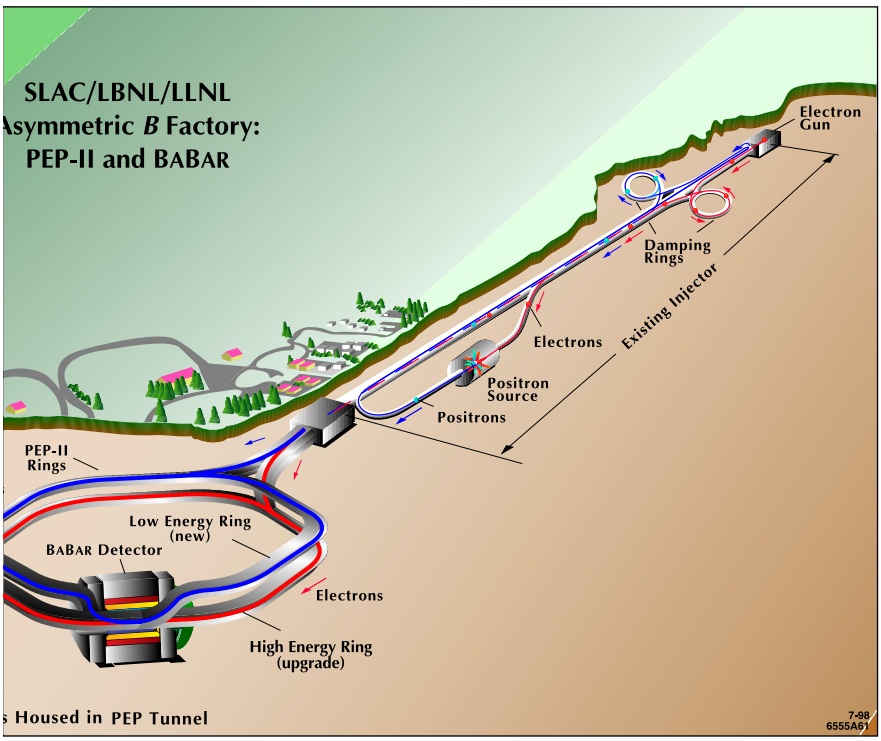


Figure 2-1: A schematic overview of the experimental site at SLAC. Courtesy of Stanford Linear Accelerator Center.

collider was completed in July 1998.

The different energies of the electron and positron beams result into a CM frame which is moving in the laboratory frame with a Lorentz boost $\beta\gamma = 0.56$. This boost is crucial to study the B-meson system. Since the $\Upsilon(4S)$ is only about 22 MeV heavier than the two resulting B-mesons, the B-mesons are produced almost at rest in the $\Upsilon(4S)$ rest frame. But with the above mentioned Lorentz boost of the CM frame (which is the $\Upsilon(4S)$ rest frame), it is possible to measure the difference in the decay length of the two B-mesons which is due to a difference in their decay times. The Lorentz boost results in a spacial separation along the z-axis of the two B meson decay vertices of $\Delta z \approx 250\,\mu$m.

The design peak luminosity of the PEP-II collider is $3 \times 10^{33}\,cm^{-2}s^{-1}$. As of today, the PEP-II record is $12.069 \times 10^{33}\,cm^{-2}s^{-1}$, achieved August 16th 2006, which already exceeds the design goal by a factor of four. The production cross section of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is 1.05 nb. Thus, the high luminosity could lead to about 200

Figure 2-2: The PEP-II storage ring. The upper ring is the low energy positron ring and the lower ring is the high energy electron ring. Courtesy of Stanford Linear Accelerator Center.

million $\Upsilon(4S)$ produced by PEP-II each year.

The two beams in the PEP-II machine collide head on, *i.e.,* there is no crossing angle at the interaction point (IP). And the time between two bunch crossings is 4.2 ns. A picture of the PEP-II storage ring is shown in Figure 2-2.

2.3 The *BABAR* Detector

The *BABAR* detector is a modern high energy particle detector designed specifically for an asymmetric electron positron collider running at the $\Upsilon(4S)$ resonance. Because of the boost along the e^- direction, the detector is asymmetric with respect to the collision point. The forward-backward asymmetry of the detector is designed in such a way that it covers approximately the same solid angle in the forward and backward direction in the CM frame.

I am going to describe the different subsystems of the *BABAR* detector following

a particle produced at the interaction point in the center of the detector to the outermost part of the detector. A schematic overview of the *BABAR* detector with its subsystems can be seen in Figure 2-3. The information in these sections is coming mostly from "The *BABAR* Physics Book" [26] and "The *BABAR* Detector" [27].

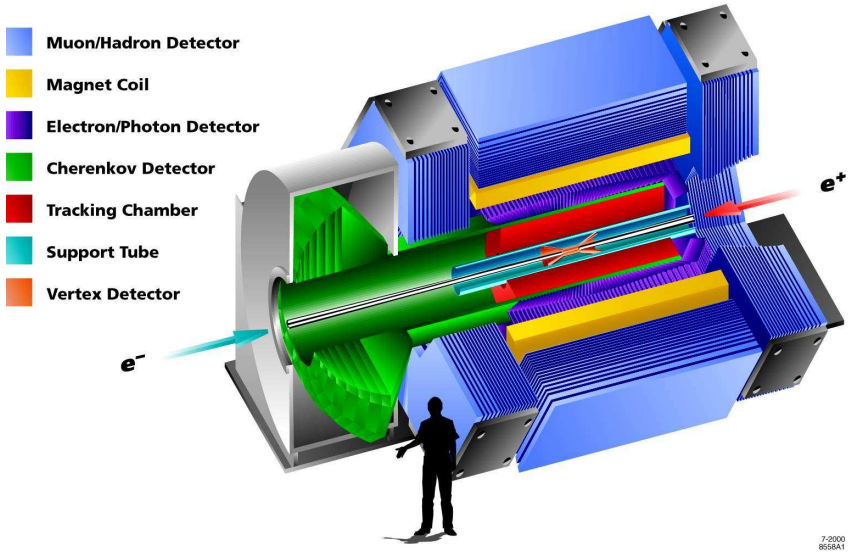


Figure 2-3: The *BABAR* detector. Courtesy of Stanford Linear Accelerator Center.

2.3.1 The Silicon Vertex Tracker

The silicon vertex tracker (SVT) is the innermost part of the *BABAR* detector. It consists of five layers of planar silicon detectors ordered in a cylindrical geometry around the beam axis at the interaction point. The innermost layer is in radial direction only 3.3 cm away from the interaction point and the outermost layer is 14.6 cm away. The SVT provides a spatial resolution in the z-direction of less than 70 μm [27] which is absolutely crucial in order to be able to resolve the separated vertices for the two B-mesons and thus being able to study CP-violation in the B-meson system. The resolution in the x-y plane is better than 100 μm.

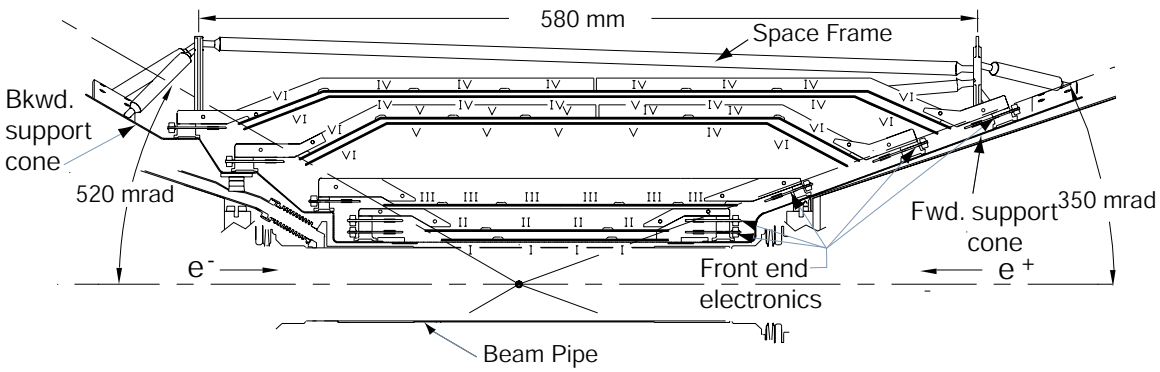


Figure 2-4: The *BABAR* Silicon Vertex Tracker [27] (side view).

2.3.2 The Drift Chamber

The *BABAR* drift chamber (DCH) is a 280 cm long cylinder with an inner radius of 23.6 cm and an outer radius of 80.9 cm (see Figure 2-5). The drift chamber is, like other parts of the detector, asymmetric in the forward-backward design in order to account for the asymmetric beam energies. The wires in the tracking volume are arranged in 10 super-layers of 4 layers of wires each, summing up to a total of 40 layers of wires. The super-layers have a slightly different orientation with respect to each other in order to be able to achieve a three-dimensional track reconstruction. The first super-layer is oriented exactly along the beam axis, the second and third one are tilted with an opposite angle with respect to the beam axis. The fourth layer is again oriented along the beam axis and this pattern is continued until the last super-layer is again oriented along the beam axis. The volume around the wires is filled with a Helium-based drift gas. A schematic overview of the DCH is shown in Figure 2-5.

With this configuration, the drift chamber performance for spatial resolution is better than 140 μm. Charged particles produced at the interaction point need at least a transverse momentum of about 100 MeV in order to reach the DCH. The tracking device (SVT and DCH) is located inside of an axial 1.5 T magnetic field.

From the curvature of a reconstructed track in the tracking systems of the *BABAR* detector (SVT and DCH), one can deduce the momentum of the associated particle.

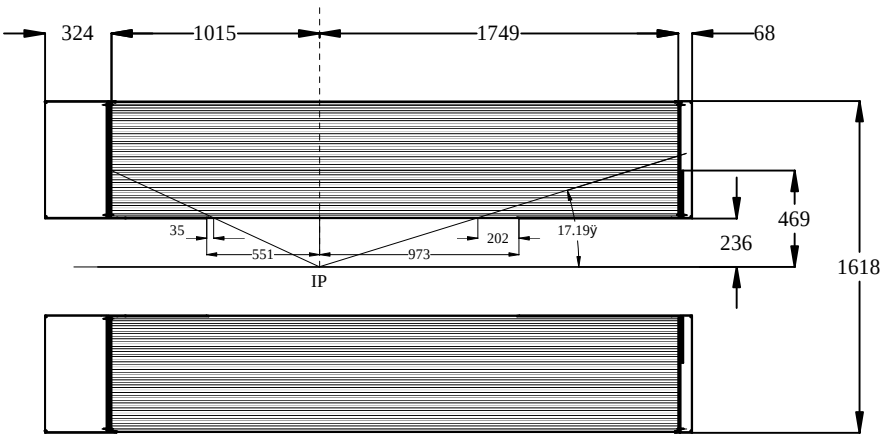


Figure 2-5: The *BABAR* drift chamber (side view). All measurements are in units of millimeter.

The resolution in the transverse momentum is found to be [27]

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

The Lorentz force acting on a particle with charge q and velocity $\vec{v}$ propagating in a magnetic field $\vec{B}$ is (without the presence of an electric field)

$$\vec{F}_{Lorentz} = q\left(\vec{v} \times \vec{B}\right) \ . \tag{2.2}$$

The centripetal force on a particle with mass m moving with momentum $\vec{p}$ on a curvature with radius r can be deduced with Newton's second law

$$\vec{F} = \frac{d\vec{p}}{dt} \tag{2.3}$$

and the infinitesimal change in momentum direction $d\vec{p}$ by an angle $d\vec{\theta}$

$$d\vec{p} = \vec{p} \times d\vec{\theta} \ , \tag{2.4}$$

as can be seen in Fig. 2-6. It follows that the centripetal force is

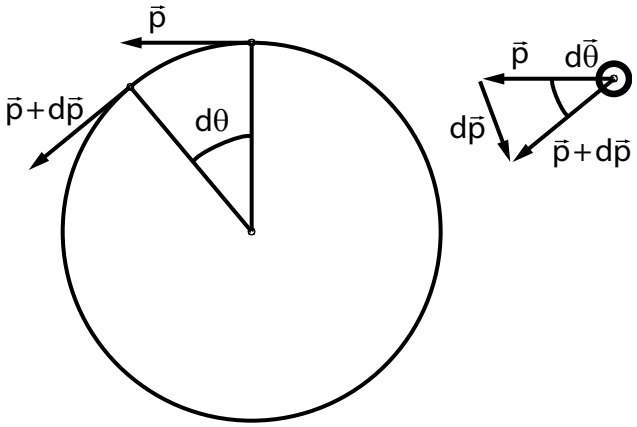


Figure 2-6: A charged particle moving perpendicular to a magnetic field.

$$\vec{F}_{centripetal} = \frac{d\vec{p}}{dt} = \vec{p} \times \frac{d\vec{\theta}}{dt} = -\vec{p}\frac{|\vec{v}|}{r} \ . \tag{2.5}$$

Since both forces have equal magnitude and opposite directions, one can combine these two equations in order to get a relation between the charge q and the momentum $\vec{p}$ of the particle, the magnetic field, the angle between magnetic field and particle momentum $\alpha_{\vec{B}-\vec{p}}$ and the radius of the track curvature

$$|\vec{p}| = rq \left|\vec{B}\right| \sin \alpha_{\vec{B}-\vec{p}} \ . \tag{2.6}$$

2.3.3 The Detector of Internally Reflected Cherenkov Light

The detector of internally reflected Cherenkov light (DIRC) is designed to separate charged kaons from charged pions. If a charged particle travels in a medium faster than the speed of light in that medium, it emits Cherenkov light. The relation between the momentum of the charged particle and the angle between the particle's flight direction and the Cherenkov light cone emission angle is

$$\cos \theta_{Cherenkov} = \frac{1}{\beta n} \ , \tag{2.7}$$

where $n = 1.473$ is the refraction index of the medium and β is the velocity of the charged particle normalized to the vacuum speed of light.

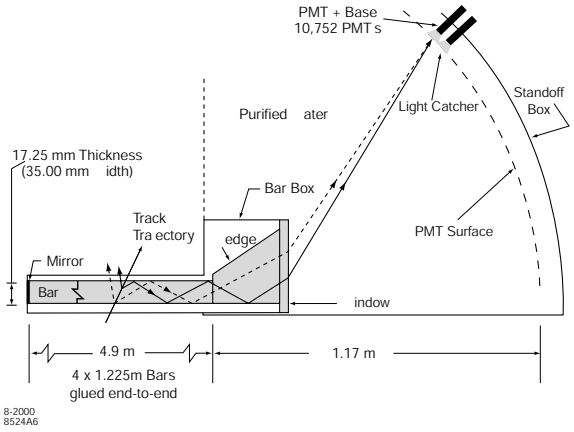


Figure 2-7: The *BABAR* Detector of Internally Reflected Cherenkov Light (side view) [27].

The DIRC is an array of 4.9 m long, 3.5 cm wide and 1.7 cm high bars made of synthetic fused silica, a translucent material in which Cherenkov light can be produced. There are a total of 144 bars arranged in 12 groups of 12 bars each. The Cherenkov light travels inside these bars towards a toroidal water tank at the backward end of the detector. The emission angle is preserved in these internat reflections due to the high accuracy of the parallel quartz-bar surfaces. Finally, an array of photomultiplier tubes detects the Cherenkov light at the backward end of this water tank. From the position and time of the detection in the photomultiplier tubes, an image of the initially produced Cherenkov light cone is inferred. This process is illustrated in Figure 2-7.

The separation between charged pions and kaons relies on their different masses. Particles with the same momentum but with different masses have different velocities and thus produce Cherenkov light at different angles. The separation only works if the charged particles are faster than the speed of light in the bars. A charged pion starts producing Cherenkov light if its momentum is larger than 129 MeV/c and a charged kaon has to have a momentum larger than 457 MeV/c in order to produce Cherenkov light. The number of Cherenkov photons produced increases with the momentum of the charged particle and a sufficient number of photons is needed in order to separate them from the background. Thus, charged pions can be separated

from charged kaons reliably only, if the momentum of these particles is larger than about 600 MeV.

The achieved Cherenkov angle resolution per track is 2.5 mr. The K-π separation at a momentum of 3 GeV/c is 4.2σ and increasing with lower momenta.

2.3.4 The Electromagnetic Calorimeter

The electromagnetic calorimeter (EMC) is an array of 6580 Thallium-doped Cesium-Iodide (CsI(Tl)) crystals. This part of the detector covers a polar angle of $-0.775 \leq \cos(\theta) \leq 0.962$ in the laboratory frame, corresponding to $-0.916 \leq \cos(\theta) \leq 0.895$ in the center of momentum frame. If a particle hits the very forward or backward end of the EMC, a part of the energy of this particle will not be deposited in a crystal of the EMC. Some of this particle energy will escape on these edges of the EMC. Thus, for analysis purposes, the solid angle coverage is slightly smaller. Two photodiodes are mounted at the rear end of each crystal. These photodiodes convert scintillation light produced by an electromagnetic shower inside the crystals into a measurable electric pulse. In the radial direction, the calorimeter is placed between the DIRC and the magnet cryostat. A schematic overview of the EMC design is shown in Figure 2-8.

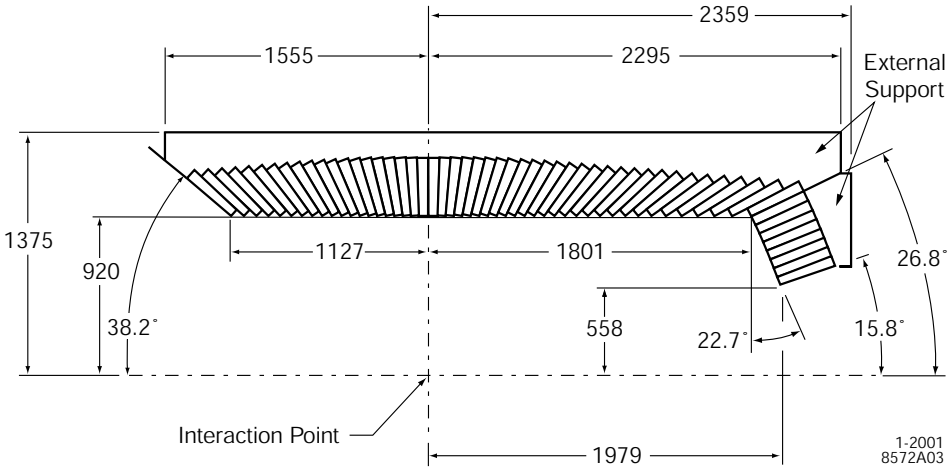


Figure 2-8: The *BABAR* Electromagnetic Calorimeter (side view) [27].

The EMC is designed to be capable of detecting photons (coming from π^0 and η decays) very efficiently in the energy range of 20 MeV up to 9 GeV. The design energy

resolution is of the order of 1-2%. Also a high angular resolution was a design goal. The achieved position resolution of a few mm translates into an angular resolution of a few mrad. The energy-dependent resolutions are for the energy resolution ($\oplus$ means sum in quadrature)

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus b \tag{2.8}$$

and for the angular resolution

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} + d \tag{2.9}$$

with $a = (2.32 \pm 0.30)\,\%$, $b = (1.85 \pm 0.12)\,\%$, $c = (3.87 \pm 0.07)\,mrad$ and $d = (0.00 \pm 0.04)\,mrad$ [27].

The crystals are arranged in two sections. The first section is a cylindrical arrangement of 48 rings with 120 crystals each. This barrel covers in polar angle in the laboratory frame $-0.775 \leq \cos(\theta) \leq 0.892$. A conical end cap consisting of 8 rings with a total of 820 crystals is mounted in addition in the forward direction. This end cap covers in the forward direction in the laboratory frame $0.893 \leq \cos(\theta) \leq 0.962$ in polar angle. The gap between barrel and end cap is of the order of 2 mm.

Most of the support structure and all of the electronics is mounted at the radial outer end of the crystals in order to minimize the material in front of the crystals. This results in less than $0.3 - 0.6\,X_0$(radiation length) in front of the crystals.

2.3.5 The Instrumented Flux Return

The outer part of the *BABAR* detector has three main purposes:

- Magnetic flux return in the iron yoke,
- Muon detection,
- Neutral hadron detection.

The Instrumented Flux Return (IFR) consists of a barrel and two end caps. The iron for the magnetic yoke is separated into 18 plates in the radial direction. In the

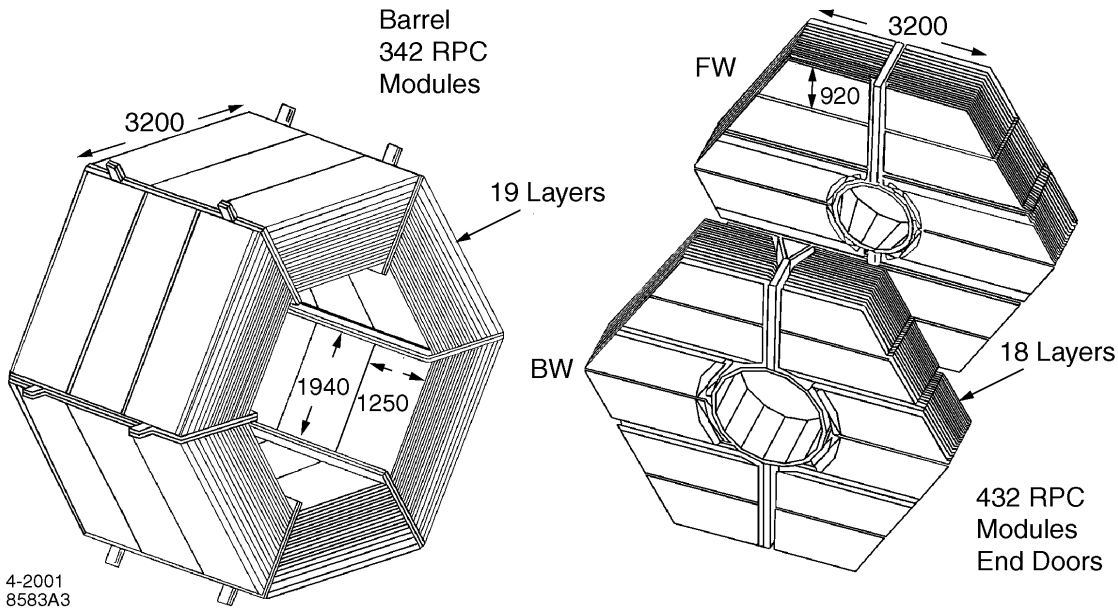


Figure 2-9: The *BABAR* Instrumented Flux Return [27].

barrel, the nine innermost plates are 2 cm thick and 3.5 cm apart, the next four plates are 3 cm thick and 3.2 cm apart, followed by three 5 cm thick plates and two 10 cm thick plates, all 3.2 cm apart. The end caps have a similar layout, the two differences are that all plates are 3.2 cm apart and that the outer two plates are 5 cm and 10 cm thick. The barrel is furthermore segmented in azimuthal direction into six sectors forming a uniform hexagon. The two end caps as well as the barrel layout can be seen if Figure 2-9.

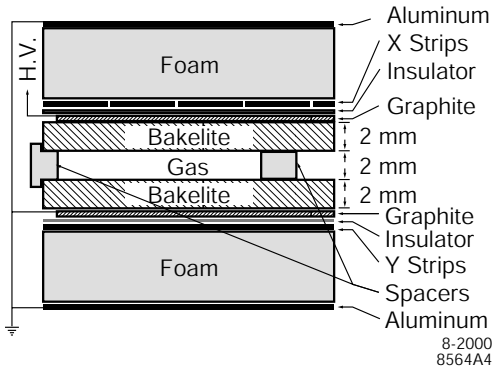


Figure 2-10: The *BABAR* Resistive Plate Counters. These particle detectors are mounted in the gaps between the iron plates of the IFR [27].

The IFR was originally instrumented with resistive plate counter (PRCs), but its muon detection performance was found to be degrading faster than expected. Thus, the decision was made to replace some of the RPCs with limited streamer tubes (LST).

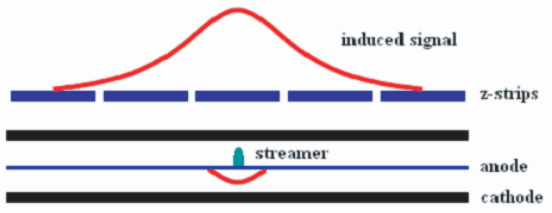
Resistive Plate Counters

The gaps between the iron plates are instrumented with resistive plate counters (RPCs) [28] in order to detect muons and neutral hadrons. The basic layout of these RPCs is the following. Two graphite plates are separated by a 6 mm thin gap. One of the plates is electrically grounded and the other is at an 8 kV electric potential. In between these two graphite plates is first a 2 mm thin layer of a bakelite with a high bulk resistivity ($10^{10} - 10^{11} \Omega\,\mathrm{cm}$), followed by a 2 mm thin gap filled with a gas with high absorption coefficient for ultraviolet light, followed by another 2 mm thin layer of the same bakelite. Aluminum strips are glued on the outside of the graphite plates, separated by an insulator. The signal is read out capacitively from these strips. The design of the RPCs can be seen in Figure 2-10.

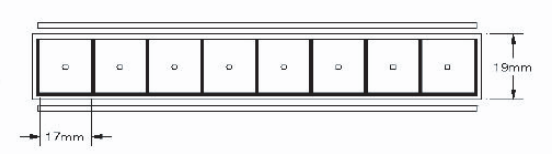
A signal is induced when a charged particle traverses the RPC. The charged particle can either be a muon coming from the inside of the detector or a charged particle coming from a hadronic shower due to a hadron interacting in the material of the IFR (or upstream). A discharge is produced at the point where the charged particle traversed the gas, due to the high electric field. When the discharge occurs, the electrons travel to the electrode and ionize more gas molecules on their way. When they arrive at the bakelite, they remain there for a sufficient time due to the high resistivity of the bakelite and locally, there is now only an electric field between the surface of the bakelite facing the gas gap (where the electrons accumulated) and the graphite electrode on the other side of the bakelite. But there is no electric field in the gas any more, thus the discharge is stopped. On the other hand, the high absorption coefficient for ultraviolet light of the gas prevents photons of this discharge to travel in the gas and produce a secondary discharge away from the primary one. Since the spacing between the two charges (the graphite electrode on one side and

the accumulated electrons on the other side) differs from the larger distance between the two graphite plates, the capacitance of this system changes and this signal is read out by the strips on the outside.

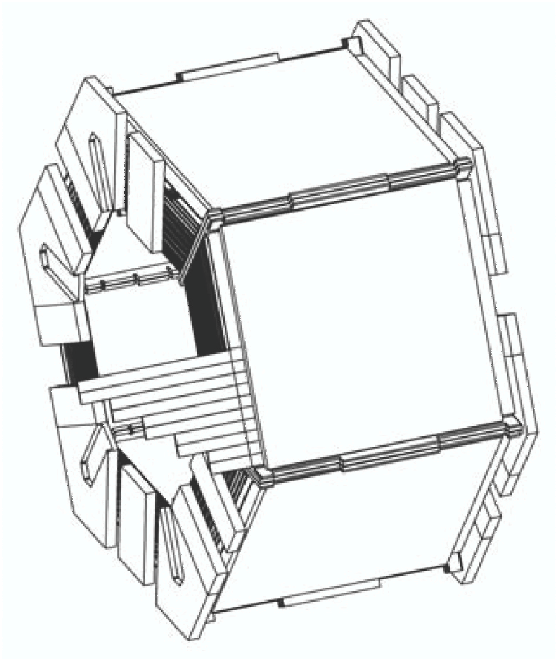
Limited Streamer Tubes



(a) Side view on a LST.



(b) View along the LST axis.



(c) Location of the LSTs in the barrel.

Figure 2-11: The *BABAR* Limited Streamer Tubes. These particle detectors are mounted in the gaps between the iron plates of the barrel IFR. The top left figure shows a side view of a single tube with a schematic working of the signal. The bottom left figure shows a 8 cell layout, looking along the direction of the tube. The figure on the right shows the location of the LSTs in the IFR barrel of the *BABAR* detector.

After some period of data taking, it was realized that the efficiency of the RPCs was slowly degrading. After some unsuccessful attempts were made to recover the full efficiency of the RPCs, the decision was made to replace the barrel RPCs with a different and well established technology: limited streamer tubes (LST) [29, 30, 31]. In the summer of 2004, the first two sextants (top and bottom) of the IFR PRCs were replaced with LSTs. The remaining four sextants followed in the summer and fall of 2006. The location of the LSTs is illustrated in Figure 2-11 (c).

A single LST is a rectangular cell 15×17 mm made of graphite-painted PVC (cathode) and a silver plated wire at the center of the cell (anode) (see Figure 2-11 (b)). Each cell is about 4 m long. A voltage of 5500 V is applied and the LSTs are filled with a gas mixture consisting of CO_2 (89%, electronegative quench gas to capture excess electrons to prevent spurious avalanches), Isobutane (8%, the quench gas to capture UV photons to prevent distant secondary avalanches) and Argon (3%, high gain).

When a charged particle traverses the gas inside an LST, it creates electron-ion pairs. Due to the large electric field, the electron is accelerated to an energy such that it will produce secondary electron-ion pairs. An electromagnetic avalanche is formed. The secondary electron-ion pairs quickly produce an electric field comparable to the applied external field and thus cancel it. Thus, this avalanche saturates. Now, the electric field is between the tip of the avalanche and the anode wire. In this space, new avalanches form from photo-ionized electrons and electric field is now only present between the tips of the new avalanches and the anode wire. The electron-ion pairs from the previous avalanche recombine. This cascade of avalanches, called streamer, propagate to the anode wire and produce a relatively large and detectable signal.

The z position is determined by an array of copper strips (≈35 mm wide) aligned perpendicular to the LSTs ontop of them. The electromagnetic streamer in the gas of the LSTs will induce a signal in these z planes and thus, the z position of the streamer can be determined. This method is illustrated in Figure 2-11 (a).

2.4 Monte Carlo Simulation

It needs to be understood what kind of experimental signatures a theoretical model like the SM or also PBSM would produce. This is the main reason to perform simulations of the theoretical physics processes, to see what experimental signals the detector will produce. These simulations are generically called “Monte Carlo” (MC). The probabilities of production and decay properties of many physics processes are stored in the software and what type of event is actually simulated is chosen randomly

by the computer, base upon these stored probabilities. According to this property of the simulation, the name "Monte Carlo" is used which is the name of a famous gambling and casino city in the Mediterranean principality of Monaco.

The production of particles in the e^+e^- collision at the IP and the subsequent decay of these particles is done with an ensemble of several software packages, most notably `EvtGen` [32] and `JETSET 7.4` [33]. `EvtGen` handles the decays of B-mesons based on decay amplitudes and not decay probabilities. This has the advantage of correctly simulating angular distributions of and correlations between the decay products. It also has the ability to correctly simulate CPV. At *BABAR*, `JETSET 7.4` simulates generically physics processes in the high-energy e^+e^- collisions, fragmentations of hadrons and decays of particles which are not specifically simulated with `EvtGen`.

The propagation of simulated particles through the *BABAR* detector, the interaction of these particles with the detector material and the electronic response of the *BABAR* detector are simulated with `GEANT4`. The same *BABAR* reconstruction software is used for MC as well as for real data.

The final output of the simulation are MC files that resemble the real data as close as possible. Over time, adjustments of the MC are performed in order to match the MC better with what is observed in real data. An important property of MC samples is the knowledge of which physics process was simulated for each event. Also, a simulated particle candidate found by the reconstruction software can be matched to a generated particle and thus, ideally, the true identity and the true four-momentum of each reconstructed particle is known. It is said that the simulated particle is "truth–matched".

Chapter 3

Analysis Overview

3.1 Introduction

This analysis aims to exclusively reconstruct three decays modes:

- $B^+ \to \rho^+\gamma$ with $\rho^+ \to \pi^+\pi^0$ and $\pi^0 \to \gamma\gamma$;
- $B^0 \to \rho^0\gamma$ with $\rho^0 \to \pi^+\pi^-$;
- $B^0 \to \omega\gamma$ with $\omega \to \pi^+\pi^-\pi^0$ and $\pi^0 \to \gamma\gamma$.

The branching fractions for these decays are $\mathcal{B}(\rho^0 \to \pi^+\pi^-) \approx 100\%$, $\mathcal{B}(\rho^+ \to \pi^+\pi^0) \approx 100\%$, $\mathcal{B}(\omega \to \pi^+\pi^-\pi^0) = (89.1 \pm 0.7)\%$, and $\mathcal{B}(\pi^0 \to \gamma\gamma) = (98.798 \pm 0.032)\%$ [22].

The general strategy for this analysis is to first reconstruct the three signal decay modes individually and form a set of B-meson candidates in each event. A first set of loose cuts is applied to suppress completely uninteresting events in order to have a smaller subset of the data that is easier to handle. This part of the analysis is described in Chapter 4.

Then, a set of specific techniques is developed to suppress background events. These techniques include likelihood based methods, neural networks and multidimensional cut-optimizations. They are documented in Chapter 5.

To extract the observables from data, maximum likelihood fits are performed on the final small data set. This stage of the analysis is described in Chapter 7.

The systematic uncertainties of the analysis are evaluated and presented in Chapter 8. Chapter 9 presents the results and Chapter 10 concludes.

3.2 Event Signatures

The most striking signature of a $B \to (\rho/\omega)\,\gamma$ decay is the presence of a very high-energy photon in the event. In the system of the B meson decaying into the signal mode, the photon energy is about half the B meson mass. This still holds for the center of momentum (CM) frame of the whole event since the B mesons have only a momentum of 341 MeV/c in the CM frame. In each event, the highest energy photon is selected and required to have an energy between 1.5 GeV and 3.5 GeV in the CM frame. In the reconstruction, a photon is identified as a cluster in the electromagnetic calorimeter with no associated charged track.

Also, the event has to contain at least one well-reconstructed charged track which is identified as a pion. The excellent charged particle identification capabilities of the *BABAR* detector is very useful for this purpose. The most important information for pion identification is provided by the DIRC and supplemented by the measurement of dE/dx from the DCH and SVT.

The last type of particle needed to be able to reconstruct the signal B meson is a π^0, except for the decay mode $B^0 \to \rho^0\gamma$. π^0 candidates are identified by computing the invariant mass of all pairs of photons in the event (which have to fulfill some quality requirements) and requiring that the computed invariant mass of this photon pair be close to the nominal π^0 mass of $(134.9766 \pm 0.0006)\,\mathrm{MeV}/c^2$ [22].

3.3 Major Backgrounds

The major background for this analysis originates from $e^+e^- \to q\bar{q}$ events (where q = u,d,s,c). The most outstanding signature of a signal event, the high-energy photon,

can be mimicked in the following ways:

- The process $e^+e^- \to q\bar{q}\gamma$ can occur due to initial state radiation (ISR) or final state radiation. In particular, a photon due to ISR can have very high energy and thus fake a photon from a signal event.

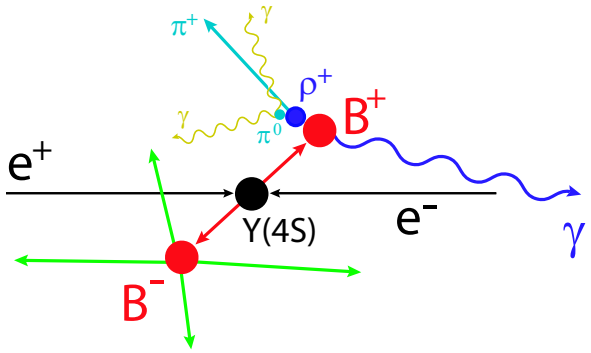
- The high-energy photon can be a decay product of a high-energy $\pi^0(\eta)$. This high-energy $\pi^0(\eta)$ can decay into a photon pair with very asymmetric energies in the lab frame and the low-energy photon can be lost, e.g. by going along the beam pipe. Or the two photons from the high-energy $\pi^0(\eta)$ can be in the same cluster in the calorimeter and thus be identified as only one photon.

Due to the jetlike structure of these so-called continuum events, the high-energy photon is highly correlated with the energy and momentum flow of the rest of the event. This is not the case for an isotropic signal event since the B mesons have only a momentum of about 341 MeV/c in the CM frame.

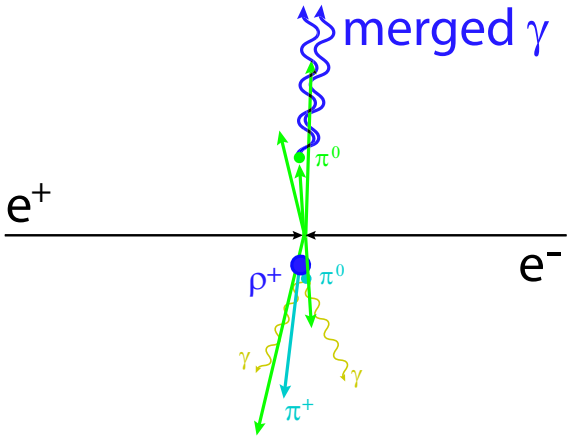
Potentially dangerous are so-called "peaking backgrounds". These are real B meson events that decay into a mode very similar to the true signal, *e.g.,* $B \to K^*\gamma$, $K^* \to K\pi$ where the kaon was mis-identified as a pion.

3.4 Expected Yields

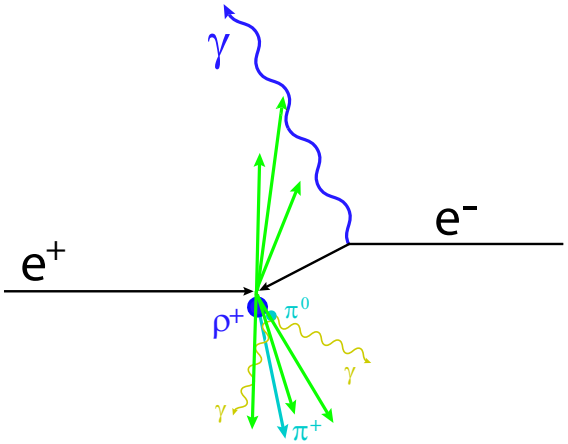
In the 316 fb^{-1} of data used for this analysis, one expects about 350, 175, and 175 signal events produced in the channels $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$, respectively. The expected number of potentially reconstructed signal events is further reduced due to the branching fraction of the $\omega \to \pi^+\pi^-\pi^0$ decay of $(89.1 \pm 0.7)\,\%$ [22], the branching fraction of the $\pi^0 \to \gamma\gamma$ decay of $(98.798 \pm 0.032)\,\%$ [22] and the detector hermeticity of about 90%. Ignoring angular correlations and momentum distributions, one expects about 70% of the initial $B^+ \to \rho^+\gamma$ candidates to be found in the detector, 75% of the initial $B^0 \to \rho^0\gamma$ candidates and about 60% of the $B^0 \to \omega\gamma$ candidates.



(a) A signal event.



(b) A continuum background event with a high energy π^0.



(c) A continuum background event with ISR.

Figure 3-1: Event signatures. All plots are shown in the CM frame.

Also, *w.r.t.* the previous *BABAR* analysis on $B \to (\rho/\omega)\,\gamma$ [24], the continuum background suppression is expected to be improved due to an intended increase in the high-energy photon purity, a use of more signal-background separating variables in a higher-dimension neural network and a set of newly optimized cuts.

3.5 A Blind Analysis

This analysis is done "blind", which means that all the analysis optimizations and considerations are based on so called MC simulations The real data is looked at only after the whole analysis procedure is fixed. The continuum MC can be verified with off–resonance data. This data has been collected in a mode where the collision energy $\sqrt{s}$ is reduced by 40 MeV. It is sufficiently away from the $\Upsilon(4S)$ resonance so that no B mesons can be produced.

In MC events the type of a generated particle is known. This MC truth information is used to optimize the analysis with respect to the efficiency of the true signal events, referred to as truth matched MC hereafter.

3.6 Data Samples

3.6.1 Monte Carlo

In this analysis, Monte Carlo simulations have been used to optimize the different selection criteria of the analysis. They are referred to as SP8 Monte Carlo or simply MC in the remainder of this thesis. There are several categories of MC:

- First, there is so-called signal MC which simulates $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ where one B meson is constrained to decay into only a specific decay mode under consideration, e.g. $B^0 \to \rho^0\gamma$, but the other B meson in the event is allowed to decay generically into all final states. This type of MC is available in rather large amount, in terms of equivalent integrated data luminosity, due to the usually small branching fraction of the considered specific B decay.

Decay	Branching Fraction (in 10^{-6})	Generated Events	Corresponding Luminosity (fb^{-1})
$B^+ \to \rho^{\pm}\gamma$	1	280000	266667
$B^0 \to \rho^0\gamma$	0.5	328000	312381
$B^0 \to \omega\gamma$	0.5	328000	312381
$B^0 \to K^{*0}\gamma$	39.2 ± 3.1	2155000	52357
$B^+ \to K^{*+}\gamma$	38.7 ± 3.8	2019000	49682
$B^0 \to X^0_{sd}\gamma$	352	434000	1233
$B^+ \to X^+_{su}\gamma$	352	434000	1233

Table 3.1: Signal Monte Carlo modes used in this analysis. The assumed cross section for the process $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is 1.05nb [26].

Decay	Assumed cross section (nb) [26]	Generated Events	Corresponding Luminosity (fb^{-1})
$e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$	0.525	428558000	816.301
$e^+e^- \to \Upsilon(4S) \to B^+B^-$	0.525	416022000	792.423
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	2.09	535274000	256.112
$e^+e^- \to c\overline{c}$	1.30	497006000	382.311
$e^+e^- \to \tau^+\tau^-$	0.94	272228000	289.606

Table 3.2: Generic Monte Carlo modes used in this analysis.

- Second, there is so-called generic B MC. Here, the generic process $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is simulated and both B mesons in the event are allowed to decay into all possible final states. This type of MC is available in amount of about a few times the equivalent integrated data luminosity.

- Last, but not least, there is the so-called continuum MC. This is the simulation of non-resonant physics processes under the $\Upsilon(4S)$ peak, i.e. $e^+e^- \to f\overline{f}$, where f is a either a quark lighter than the b quark (u, d, s, c) or a charged lepton. Due to the large cross-section of these processes the available MC is usually only about one times the equivalent integrated data luminosity.

The MC samples used in this analysis are listed in Table 3.1 and Table 3.2 for signal MC and the generic B and continuum MC, respectively.

Time taken	Refered to as	on–resonance		off–resonance
		B-counting	Luminosity	Luminosity
01/2000 – 10/2000	Run1		21fb^{-1}	2.4fb^{-1}
02/2001 – 06/2002	Run2		61fb^{-1}	6.8fb^{-1}
11/2002 – 06/2003	Run3		32fb^{-1}	2.4fb^{-1}
09/2003 – 07/2004	Run4		101fb^{-1}	10.1fb^{-1}
04/2005 – 06/2006	Run5		101fb^{-1}	5.5fb^{-1}
Total		347×10^6	$316\,\text{fb}^{-1}$	27.2fb^{-1}

Table 3.3: Data used in this analysis.

3.6.2 Data

The data used for this analysis is the combined Run1 to Run4 and most of Run5 data set. Within that sample, there are 347×10^6 $\Upsilon(4S) \rightarrow B\bar{B}$ events, which corresponds to an integrated luminosity of $316\,\text{fb}^{-1}$ (see Table 3.3). An equal B^+/B^0 production at the $\Upsilon(4S)$ resonance has been assumed .

Chapter 4

Event Reconstruction

4.1 Skim Requirements

In order to reduce the relevant (MC and real) data to a computationally manageable size, a crude first selection called "skim" is made. The skim requirements are selections which keep most of the signal events but remove events from the continuum background like Bhabhas, radiative Bhabhas, dimuons and many hadronic events with no high-energy photon. The skim also requires the reconstruction of a signal B candidate imposing some very loose kinematic cuts. The selection requirements used in this skim are:

- The event passes either the `BGFMutliHadron` or the `BGFNeutralHadron` tag filter (see Section 4.2.1).
- There are at least two or more tracks in the `GoodTracksLoose` list (see Section 4.2.2).
- The ratio of the 2^{nd} to 0^{th} Fox-Wolfram moment of the event is $R_2^{All} < 0.9$ (see Section 4.2.3).
- The primary photon energy measured in the CM frame is $1.5 < E_\gamma^* < 3.5\,\mathrm{GeV}$ and it is selected from the `GoodPhotonLoose` list (see Section 4.4).

- The vector meson candidate mass is in the range $0.5 < m_{\pi^+\pi^0} < 1.3\,\mathrm{GeV}/c^2$ and $0.5 < m_{\pi^+\pi^-} < 1.2\,\mathrm{GeV}/c^2$ for the ρ^+ and ρ^0, respectively. For the ω meson, the mass $m_{\pi^+\pi^-\pi^0}$ must be within $50\,\mathrm{MeV}/c^2$ around the nominal ω mass. Also, the kinematics of the ρ^0 daughters must be compatible with a common vertex (see Section 4.7).

- The reconstructed B candidates pass the loose kinematic cuts $5.1 < m_{ES}^0 < 5.5\,\mathrm{GeV}/c^2$ and $-0.6 < \Delta E < 0.6\,\mathrm{GeV}$ (see Section 4.8).

On top of the skim requirements, additional selections are made to further reduce background events. Some of these selections only tighten a cut already applied in the skim whereas others utilize variables not used in the skim. In addition to simply reducing background, some cuts are also intended to assure the quality of a candidate. All selections are described in the following sections, starting with the selections that are based on the whole event (Section 4.2), over the individual neutral and charged particles (Sections 4.4 and 4.6) and the light mesons (Sections 4.5 and 4.7) to the final B-meson candidates (Section 4.8).

4.2 Event Level Requirements

This section describes selection criteria which are based on a whole event, independent of the signal B candidate. They are intended to select only events which are likely to have the underlying physics process $e^+e^- \to \Upsilon(4S) \to B\overline{B}$.

4.2.1 Tag Filter

Before a recorded event is fully reconstructed in the *BABAR* software environment, some basic charged track finding is done using the information from the drift chamber, and some basic cluster finding is performed using information from the electromagnetic calorimeter. Together with information from the Level 1 and Level 3 triggers, this information is used in the *BABAR* offline reconstruction to decide whether the event will be fully reconstructed or not. Several independent algorithms are used and

the output of each of these is a simple "yes" or "no". A single algorithm returning a positive answer is sufficient for the event to be fully reconstructed. These individual "yes" or "no" flags are also stored in the data and are called "tag filter bits".

For this analysis, only those events are selected which passed the `BGFMutliHadron` or the `BGFNeutralHadron` tag filter. Both of these tag filters make use of a list called `RecoGoodTracksLoose`, defined as a list of reconstructed charged particles that fulfill the following requirements:

- The radial distance in the x-y plane of the point of closest approach to the z-axis is less than 1.5 cm.
- The distance in the z-direction of that point to the origin is less than 10 cm.
- The transverse momentum of the track is greater than 100 MeV/c.

Using this definition, the `BGFMutliHadron` filter requires:

- At least three charged particles reconstructed in the `RecoGoodTracksLoose` list.
- The ratio of the 2^{nd} to 0^{th} Fox-Wolfram moment R_2^{ch} must be less than 0.98. For a definition of R_2, see Section 4.2.3. For the calculation of R_2^{ch}, only entries in the `RecoGoodTracksLoose` list have been used.

This filter is clearly intended to select events with a large multiplicity of charged particles, i.e. every event where a $B^0 \to \omega\gamma$ candidate is reconstructed should pass this filter.

The `BGFNeutralHadron` tag filter is designed as an orthogonal filter w.r.t. the `BGFMultiHadron` filter by selecting events with less than three charged particles reconstructed in the `RecoGoodTracksLoose` list. In addition, those energy deposits in the electromagnetic calorimeter that could not be matched to a reconstructed charged particle are considered. When a singly-connected energy deposition exceeds 100 MeV, it is called a "cluster". When it exceeds 500 MeV, it is called a "photon". Events with two entries in the `RecoGoodTracksLoose` list are accepted if also at least two photons and three clusters are found. If only one entry is found in the `RecoGoodTracksLoose`

list, at least two photons and four clusters are needed in order for the event to be accepted. If no entry is in the `RecoGoodTracksLoose` list, the event passes this filter if at least three photons and six clusters are found. In addition R_2 is calculated using all entries in the `RecoGoodTracksLoose` list and all clusters and photons (for a definition of R_2, see Section 4.2.3). The requirement on this variables is $R_2 < 0.95$. This filter is important for the reconstruction of the $B^+ \to \rho^+\gamma$ mode in events with one or no other reconstructed charged particle.

4.2.2 Number of Reconstructed Tracks

For an event to be used in this analysis, it must contain at least two charged tracks which fulfill the following criteria (referred to as "GoodTracksLoose" (GTL) criteria in *BABAR* lingo):

- The number of hits in the DCH must be more than 12.
- The transverse momentum must be larger than 100 MeV/c.
- The track must have come closer than 1.5 cm to the beam axis.
- The track must have come closer than 10 cm to the nominal beam spot, measured along the z (beam) direction.
- The momentum of the track must be less than 10 GeV/c.

The first two requirements ensure that the particle traverses a sufficiently long path through the tracking device (SVT and DCH) to be reliably reconstructed. The last three requirements ensure that the track is not due to a cosmic ray passing through the detector or a charged particle originating from an interaction of the PEP-II beam with the residual gas in the beam pipe or the beam pipe itself.

4.2.3 R_2

Also, for each event, the ratio R_2 of the 2^{nd} to the 0^{th} Fox-Wolfram moment is computed in the CM frame as

$$R_2 = \frac{H_2}{H_0}, \tag{4.1}$$

with H_l being the Fox-Wolfram moment defined in [34] as

$$H_l = \sum_{i,\ j} \frac{|\vec{p}_i|\,|\vec{p}_j|}{s} P_l(\cos\phi_{ij}), \tag{4.2}$$

where $\vec{p}_{i/j}$ are the momenta of two particles in the CM frame, ϕ_{ij} is the angle between these two momenta, P_l are Legendre polynomials, and s is the total CM energy squared. This variable is a measure of an event being jet-like. Continuum background events ($e^+e^- \to f\overline{f}$ where f = u, d, s, c or a charged lepton) are in the CM frame in general two rather collimated bunches of particles back to back. This is due to the large excess of kinetic energy and the conservation of energy and momentum. On the other hand, $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events are almost isotropic since the $\Upsilon(4S)$ resonance is only about $20\,\mathrm{MeV}/c^2$ heavier than the sum of the two B meson masses and thus there is almost no excess of kinetic energy available which could go into the momenta of the B mesons. R_2 is required to be less than 0.9 in the skim selection. This cut is actually tightened to 0.7 after the skim to removes some significant portion of continuum background events without loosing almost any signal events.

4.3 Photon Selection

The high energy photon is detected as an energy deposition in the EMC.A photon produces an electromagnetic shower via a cascade of pair productions, compton scatterings and the photoelectric effect in the crystals of the EMC. In general, this electromagnetic shower spreads out over several EMC crystals. A number of adjoining (connected at at least one corner) crystals with energy deposited in them is called a cluster. A local maximum (bump) within this cluster is identified as a photon, if

no charged-particle track points directly to it.

4.3.1 The `GoodPhotonLoose` List

Photons for this analysis are selected as clusters which contain only a single bump with no charged-particle track associated with it. These photons are listed for each event in the `GoodPhotonLoose` list which has additional requirements imposed on the photon candidate:

- The calibrated, but uncorrected, raw energy deposited in the EMC be more than 30 MeV.
- The lateral shower moment LAT be less than 0.8. This variable is defined in the following section.

All photons used in this analysis are taken from this `GoodPhotonLoose` list, both for the high-energy photon candidate originating directly from the decay of the signal B meson and also the photons used for reconstruction π^0s.

The Lateral Shower Shape Variable

Electromagnetic showers have a different lateral energy distribution than hadronic showers. In general, a hadronic shower has a wider lateral distribution. This difference can be quantified in the LAT variable defined as [35]

$$LAT = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=3}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \tag{4.3}$$

where

- N is the number of crystals hit by the shower,
- E_i is the energy deposited in the i-th crystal, ordered in decreasing energy starting with 1 being the highest energy.
- r_i is the lateral radius between the centroid of the shower and the i-th crystal,

- $r_0 = 5\,\mathrm{cm}$, which is roughly the width of a crystal and thus the distance between two crystal centers.

Since the Molière radius for the CsI(Tl) crystals of the EMC is 3.8 cm, most of the energy of an electromagnetic shower is deposited within only 2-4 crystals. This is not the case for a hadronic shower. A hadronic shower spreads out over more crystals in the EMC and thus the LAT variable is on average larger than for an electromagnetic shower. Only photons where this variable is less than 0.8 are selected.

4.4 High Energy Photon Selection

A photon originating directly from a two-body B meson decay has a rather unique signature. To exploit that signature and to assure a high quality of the reconstructed photon candidate, cuts on several variables are applied. These cuts are described in the following sections.

4.4.1 Energy of the High-Energy Photon

In a signal event, the decay of the B meson is a two-body decay into a ρ or ω meson and a high energy photon. Thus, the energy of this photon is in the rest frame of the B meson approximately half of the B meson mass. Furthermore, since the B meson has a very low momentum in the CM frame due to the small mass difference between the $\Upsilon(4S)$ and the sum of two B meson masses, the CM energy of this photon E^*_γ is approximately half of the B meson mass. Thus, the CM energy distribution of the high energy photon peaks at approximately half of the B mass, with a long tail on the low-energy side due to the imperfect reconstruction in the calorimeter and thus sometimes, not the full photon energy is captured inside the crystals.

A minimum requirement of $E^*_\gamma > 1.5\,\mathrm{GeV}$ is applied in order to suppress events without a high-energy photon, *e.g.,* high multiplicity continuum events. On the high-energy side, the cutoff is chosen at 3.5 GeV. This reduces events such as $e^+e^- \to \gamma\gamma$ or also $e^+e^- \to \mu^+\mu^-\gamma$.

This cut is rather loose. Since the photon CM momentum is used to compute m_{ES} and ΔE (see Section 4.8), a tight cut on the CM energy would introduce correlations between m_{ES} and ΔE and distort these distributions. Since m_{ES} and ΔE are variables in the final fit, a correlation between them and a cut has to be avoided.

4.4.2 Reconstruction Quality of the High-Energy Photon

Not all of the 6580 EMC crystals work properly. There are some crystals where both photodiodes are broken, or where the electronics associated with that crystal is too noisy to perform reliable measurements. In order to avoid these effects, the following requirements are imposed:

- The cluster in the EMC caused by the photon is required not to have any noisy or dead channels.
- If a noisy crystal is overlooked by the EMC monitoring, it can fake an energy deposition and thus a photon. Because of this, it is required that the photon has deposited energy in more than four crystals with a circular lateral energy distribution.

4.4.3 Geometric Acceptance

The EMC covers a polar angle of $-0.775 \leq \cos(\theta) \leq 0.962$ in the laboratory frame. An electromagnetic shower spreads out over several crystals in the EMC. In order to ensure that the shower is fully contained inside, the photon is required to be away from the edges of the EMC. Furthermore, the photon needs to be within the acceptance of the tracking devices (SVT and DCH) in order to ensure that the energy deposited in the EMC is not due to a charged track. Also, beam backgrounds are possible sources of energy deposition in the EMC close to the beam pipe. Thus, the photon is required to be within $-0.74 \leq \cos(\theta) \leq 0.93$ in the laboratory frame.

4.4.4 Distance to Next Energy Deposition

The closest charged or neutral bump is required to be at least 25 cm away from the centroid of the bump caused by the photon in consideration. This requirement reduces contamination due to hadronic split-off as well as high energy π^0s and ηs decaying into a close pair of photons. A hadronic split-off is due to a hadron producing a hadronic shower inside the calorimeter, and one of the hadrons in this shower traverses a few crystals before inducing a secondary hadronic shower. The crystals detecting the secondary hadronic shower are not necessarily connected to the crystals detecting the primary shower.

4.4.5 The 2^{nd} Moment

A single photon causes a shower and thus the distribution of energy deposited in the crystals is circular around the centroid of the bump. The energy deposition caused by two very close photons (e.g. a π^0 or η decay with very collimated decay product photons due to Lorentz boost) is more ellipsoidal. This is measured by the second moment of the energies E_i deposited in the crystals centered around the maximum energy of the bump. This variable is defined as

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

where θ_C and ϕ_C are the angular coordinates of the centroid of the bump and θ_i and ϕ_i are the angular coordinates of the ith crystal of the bump. A bump caused by a single photon has a smaller second moment than a bump caused by two photons. Thus, the second moment of the photon originating directly from a B meson decay tends to be smaller than a photon from a high-momentum π^0 or η decay. The cut on this variable will be optimized for each of the three decay modes individually.

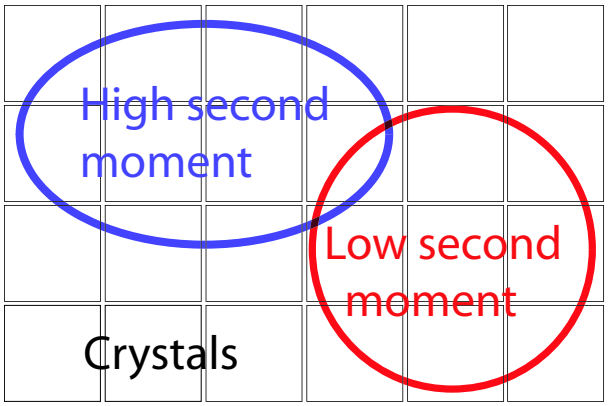


Figure 4-1: Illustration of the second moment variable. The low second moment is caused by a single photon and thus circular. If two photons hit the EMC very close, an elliptical distribution and thus a higher second moment is the result.

4.4.6 Ratio of Energies Deposited in Crystals

In order to get another handle at the shape of the electromagnetic shower deposited in the crystals of the EMC, the ratio of energy deposited in a 3×3 array of crystals centered around the shower maximum to the energy deposited in a 5×5 array of crystals centered around the same shower maximum is computed. This ratio is called $\frac{s9}{s25}$. If the shower is very spread out in the lateral direction, *e.g.,* as in a diffuse hadronic shower, this ratio will be small; if instead the shower is very confined, the ratio will be close to unity.

4.5 π^0 Selection

A π^0 decays into two photons $(98.798 \pm 0.032)\%$ of the time after a mean life τ of only $(8.4 \pm 0.6) \times 10^{-17}$ s [22]. Thus, in order to reconstruct a π^0 the four-momenta of two photons are combined. Both of these photons are taken from the `GoodPhotonLoose` list (Section 4.3.1), with an increased cut on the minimum photon-energy of 50 MeV. Only those two-photon pairs are kept whose invariant mass falls in the window $100\,\mathrm{MeV}/c^2 < m_{\gamma\gamma} < 170\,\mathrm{MeV}/c^2$. Finally, the nominal π^0 mass of $134.9766\,\mathrm{MeV}/c^2$ [22] is assigned to the resulting π^0 and a minimum π^0 energy requirement of 250 MeV is applied.

An additional variable of interest is the cosine of the opening angle of the two photons in the laboratory frame. A cut on this variable as well as a more refined cut on the two-photon invariant mass will be optimized at a later stage of this analysis.

4.6 π^+ Selection

A reconstructed track of a charged particle is used as a charged pion candidate for this analysis if it is listed in the `GoodTracksLoose` list (see Section 4.2.2). The identification of this particle is described in the following section.

4.6.1 Particle Identification

One of the outstanding capabilities of the *BABAR* detector is the excellent charged particle identification system, and especially the capability of separating charged pions from charged kaons and protons. This excellent separation power is in a large part due to the DIRC.

The particle identification chosen for this analysis is based on a likelihood. Separate likelihoods for different particle species and also for the three different detector sub-systems SVT, DCH and DIRC are constructed and in the end multiplied to obtain the final likelihood

$$L_i = L_i^{SVT} \cdot L_i^{DCH} \cdot L_i^{DIRC}. \tag{4.5}$$

Here, i are the different particle hypothesis (K^+, π^+, p^+, e^- and μ^-) and L_i are likelihood functions. The charged-particle identification in the SVT and DCH is based on the energy-loss dE/dx, whereas the likelihood function for the DIRC is based on the emitted Cherenkov light. For more detailed information about the the SVT and DCH likelihood selectors, see Appendix D.

The DIRC likelihood is constructed from a lookup table that is binned in three dimensions as

- laboratory momentum of the considered particle – 100 MeV/c bins,
- Cherenkov angle of the photon produced in the DIRC – 3 bins,

- number of detected Cherenkov photons and quality of the charged particle trajectory.

The three Cherenkov angle bins correspond to the pion, kaon and proton bands. The third bins of the above list are evaluated based on the last layer of the DCH that detected the charged particle, the Poisson probability for the number of observed Cherenkov photons, and the energy deposited in the EMC. This binned likelihood is For charged particles with momenta greater than $1.5\,\mathrm{GeV}/c$, this binned likelihood is multiplied by a Gaussian likelihood formed from the Cherenkov angle.

The total likelihood is computed according to Equation 4.5. For the final selection, the likelihood ratios $L_K/\left(L_K + L_\pi\right)$ and $L_p/\left(L_p + L_\pi\right)$ are computed. The π^+ selection for the $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ mode of this analysis requires

- $\frac{L_K}{L_K+L_\pi} < 0.2$,
- $\frac{L_p}{L_p+L_\pi} < 0.5$,
- for particles with a transverse (to the beam) momentum of $p_T > 0.6\,\mathrm{GeV}/c$, a Poisson photon consistency of the number of detected Cherenkov photons w.r.t the expected number of photons is calculated and required to be > 0.0001,
- and the particle not be identified as an electron or muon.

For the $B^0 \to \omega\gamma$ decay mode, the two charged pions only have to fulfill

- $\frac{L_K}{L_K+L_\pi} < 0.5$,
- $\frac{L_p}{L_p+L_\pi} < 0.98$,
- and the particle not be identified as an electron.

The above described algorithm ensures a charged pion efficiency between 80% and 86% in the momentum range relevant for this analysis. In the same momentum range, the charged kaon misidentification rate is below 1.5%.

4.7 Vector Meson Selection

The vector mesons ρ^+, ρ^0 and ω are reconstructed by combining the four-momenta of the appropriate charged and neutral pions. The π^0 and π^+ selections are described in Section 4.5 and Section 4.6, respectively.

The ρ^+ is reconstructed as a combination of a π^+ and a π^0. The decay point of the ρ^+ is constrained to be the primary vertex of the event and the invariant mass of π^+ π^0 pair has to be $0.5\,\mathrm{GeV}/c^2 < m_{\pi^+\pi^0} < 1.3\,\mathrm{GeV}/c^2$. This invariant-mass cut is actually applied in the skim selection, but this cut-range will be optimized later on.

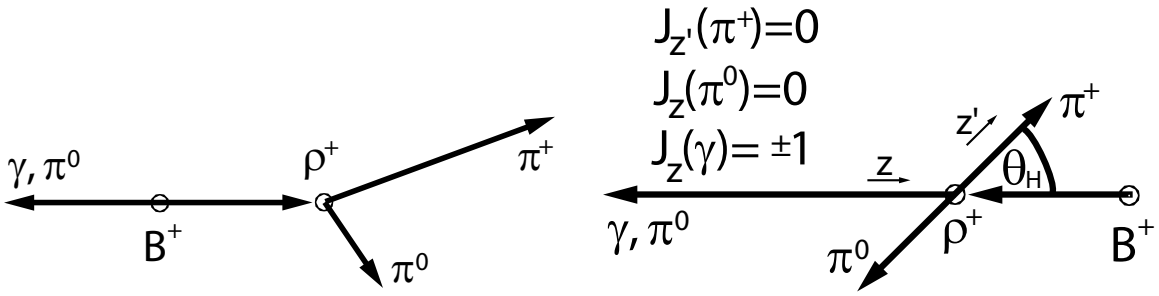
A combination of a π^+ and a π^- that originate from a common vertex is used to reconstruct the ρ^0 meson. A geometric fit is performed to determine the vertex $\chi^2_{\rho^0}$ probability of the two particles actually originating from a common point. Also, the skim requires that the invariant mass of the π^+ π^- pair be $0.5\,\mathrm{GeV}/c^2 < m_{\pi^+\pi^-} < 1.2\,\mathrm{GeV}/c^2$. Both the cut on the vertex probability and the invariant mass cut will be optimized at a later stage.

The ω meson is reconstructed in the $\pi^+\pi^-\pi^0$ decay channel which accounts for $(89.1 \pm 0.7)\%$ of all ω decays [22]. The π^+ π^- combination is required to originate from a common vertex. Both the ρ^+ and the ρ^0 meson have a rather large width of $\sim$150 MeV/c^2. The ω meson on the contrary is very narrow with a width of only $\sim$8.5 MeV/c^2 [22]. The detector resolution is comparable to the ω width. Thus, a cut on the invariant mass of the $\pi^+\pi^-\pi^0$ combination will be very powerful in rejecting non-ω background candidates. In the skim selection, the allowed mass range is $\left|m_{\pi^+\pi^-\pi^0} - m_\omega^{PDG}\right| < 50\,\mathrm{MeV}/c^2$ where m_ω^{PDG} is the nominal mass of the ω meson according to [22]. Another computed quantity is the flight distance of the ω meson away from the primary vertex. This quantity is powerful in rejecting background candidates from long-lived $K^0_S \to \pi^+\pi^-$ decays. Later on, the cut on the χ^2_ω probability, the mass window around m_ω^{PDG} and the ω flight distance will be optimized.

4.7.1 Helicity Angle

For all three vector mesons, a helicity angle is calculated. For both the ρ^+ and the ρ^0 meson, the angle is defined as the angle between the π^+ flight direction and the negative B flight direction, both in the rest-frame of the ρ^+ (ρ^0) meson.

For the ω meson, the helicity angle is defined as the angle between the normal to the π^+ π^- plane and the inverse B meson flight direction, both in the ω rest frame.



(a) The B meson rest frame. (b) The ρ^+ meson rest frame.

Figure 4-2: Graphical representation of the helicity angle.

The reaction $B \to \gamma\pi^+\pi^0$ proceeds via the intermediate resonance of the spin 1 vector meson ρ^+ for the two pion system. Due to the conservation of total angular momentum, there must be a specific relationship between the direction of the B meson and the direction of the π^+ in the rest frame of the ρ^+. For this decay, there is a total angular momentum along the axis of the photon and B meson z of $J_z(\gamma B) = \pm 1$ in the rest frame of the ρ^+, as can be seen in Fig. 4-2. This has to be connected to a total angular momentum of $J_{z'}(\pi^+\pi^0) = 0$ along the z' axis, again in the rest frame of the ρ^+ meson.

The connection is done via a rotation around the y axis (perpendicular to the reaction plane):

$$d^j_{m,m'} = \left\langle jm' \left| e^{-i\theta J_y} \right| jm \right\rangle, \tag{4.6}$$

where where j is the spin of the intermediate resonance, is this case $j = 1$ for the ρ^+ meson, and m and m' are the 3-components of the total angular momentum along the z and z' axis, respectively. θ is the rotation angle, which is called the helicity angle in this analysis, and J_y is the angular momentum operator along the y axis.

The coefficients $d^j_{m,m'}$ are tabulated in [22] and for the case of the $B \to \gamma\pi^+\pi^0$ decay, $j = 1$, $m = \pm 1$ and $m' = 0$. The decay rate is proportional to the square of the transition matrix $\mathcal{M}$ and thus to the square of the $d^j_{m,m'}$ coefficients. The square of the transition matrixes for two classes of decays relevant to this analysis are:

$$\begin{aligned} |\mathcal{M}(B \to (\rho/\omega)\,\gamma)|^2 &\propto \left(d^1_{1,0}\right)^2 = \sin^2(\Theta_H) \\ |\mathcal{M}(B \to (\rho/\omega)\,(\pi^0/\eta))|^2 &\propto \left(d^1_{0,0}\right)^2 = \cos^2(\Theta_H). \end{aligned} \tag{4.7}$$

Thus, the helicity angle in the $B \to (\rho/\omega)\,\gamma$ decays follows a $\sin^2(\Theta_H)$ distribution whereas potentially dangerous backgrounds like $B \to (\rho/\omega)\,(\pi^0/\eta)$ are expected to follow a $\cos^2(\Theta_H)$ distribution.

Also, for background events with a fake vector meson, the $\cos\Theta_H$ distribution is expected to be flat. Thus, the absolute value of the cosine of the helicity angle is for all three modes required to be $|\cos\Theta_H| < 0.75$.

4.7.2 Dalitz Angle

Since the $\omega \to \pi^+\pi^-\pi^0$ decay is a three-body decay, there are two independent angles. The second angle is the Dalitz angle Θ_D. It is defined as the angle between the π^+ and the π^0 computed in the $\pi^+\pi^-$ rest frame.

When boosting into the $\pi^+\pi^-$ rest frame, the flight direction of the π^0 and the ω mesons are collinear along the z axis and generally rotated *w.r.t.* the $\pi^+\pi^-$ z' axis. The z component of the total angular momentum along the z axis is $J_z(\omega\pi^0) = 0$ and the total angular momentum along the z' axis is $J_{z'}(\pi^+\pi^-) = 0$. As in the case of the Helicity Angle (Section 4.7.1), connecting the two reference frames by a rotation around the y axis leads to a $\sin^2(\Theta_D)$ distribution for true ω decays, whereas for fake ω decays, the $\cos\Theta_D$ distribution is roughly flat.

4.8 B Meson Selection

The B meson candidate is reconstructed by combining the four momenta of the high-energy photon with that of the ρ/ω candidate. Several variables are computed using the B meson candidate, two of which are also used in the final fit performed to extract the signal yield. These two variables are ΔE (Section 4.8.1) and m_{ES} (Section 4.8.2). Also the z separation between the decay point of the signal B meson and the assumed other B meson decay point, Δz, is calculated (Section 4.8.3).

4.8.1 ΔE of the B Meson

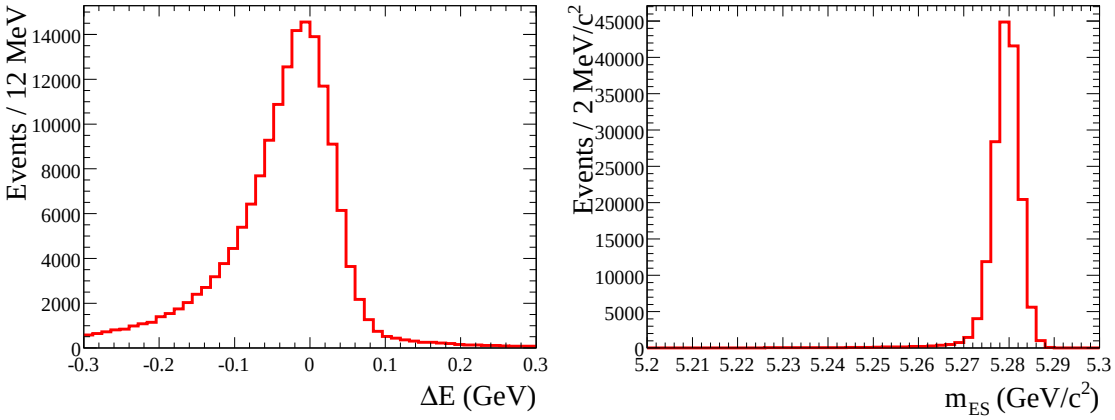
ΔE is defined as the difference between the reconstructed energy of the B meson candidate and the expected energy of the B meson, both computed in the CM frame. The expected energy of the B meson is simply half of the total CM energy $c\sqrt{s}/2$, which is the energy of one of the incoming beams in the CM frame. Thus, the definition of ΔE is

$$\Delta E = E^*_\gamma + E^*_{\rho/\omega} - c\frac{\sqrt{s}}{2} \ . \tag{4.8}$$

Here, the "$*$" refers to a quantity in the CM frame. With this definition, the correctly reconstructed signal events should form a peak centered at zero in ΔE. The energies of the charged particles in the B meson decay tree are computed from the precisely measured momenta and the assigned particle mass hypothesis via $E^* = \sqrt{m^2c^4 + |\vec{p}^*c|^2}$. Thus, the resolution of ΔE for signal events is dominated by the reconstruction of the photons and therefore by the energy resolution of the EMC. The peak formed by signal events in ΔE has a width of about 30-50 MeV and a significant tail on the negative side of ΔE. This tail is due to leakage of the electromagnetic shower in the EMC. The signal distribution of ΔE is shown in Figure 4-3(a).

4.8.2 m_{ES} of the B Meson

The variable m^0_{ES} is referred to as "beam energy substituted mass" (or also "beam constrained mass"). In order to compute it, one invokes the relativistic mass-energy-



(a) ΔE (b) m_{ES}

Figure 4-3: ΔE and m_{ES} for $B^0 \to \rho^0\gamma$ signal MC. The left plot shows truth matched signal MC in ΔE and the right plot shows truth matched signal MC in m_{ES}.

momentum relation, the usual way of calculating the invariant mass from its energy and momentum. But with one important trick: instead of using the measured energy of the B meson, the knowledge of the initial state is used, namely the B meson energy is substituted with half of the total available CM energy which is equal to the energy of one of the incoming beams in the CM frame. The momentum of the B meson candidate is very small, thus an error on the momentum measurement does not have so much of an impact as an error on the energy measurement will have. And the beam energies are known to a precision of the order of 1-2 MeV, much better than any energy measurement of a final state particle in the detector. Using this knowledge, the definition of m_{ES}^0 becomes

$$m_{ES}^0 = \frac{1}{c}\sqrt{\frac{s}{4} - \left(\vec{p}^*_{\rho/\omega} + \vec{p}^*_{\gamma}\right)^2} \ . \tag{4.9}$$

An additional improvement of this variable can be made by assuming that the m_{ES}^0 resolution is dominated by the energy resolution of the high-energy photon. The impact of the limited resolution can be reduced by rescaling the four momentum of the photon in such a way as to satisfy $\Delta E = 0$ for the B meson candidate

$$p^{*\prime}_{\gamma} = \frac{\frac{c\sqrt{s}}{2} - E^*_{\rho/\omega}}{E^*_{\gamma}} \cdot p^*_{\gamma} \ , \tag{4.10}$$

and by using this momentum for the computation of a rescaled energy-substituted mass

$$m_{ES} = \frac{1}{c}\sqrt{\frac{s}{4} - \left(\vec{p}^{\,*}_{\rho/\omega} + \vec{p}^{\,*\prime}_{\gamma}\right)^2} \,. \tag{4.11}$$

The improvement in the energy resolution can be seen by comparing the truth matched signal Monte Carlo peak in m^0_{ES} with the peak in m_{ES} (see Figure 4-4). And the signal distribution of m_{ES} is shown in Figure 4-3(b).

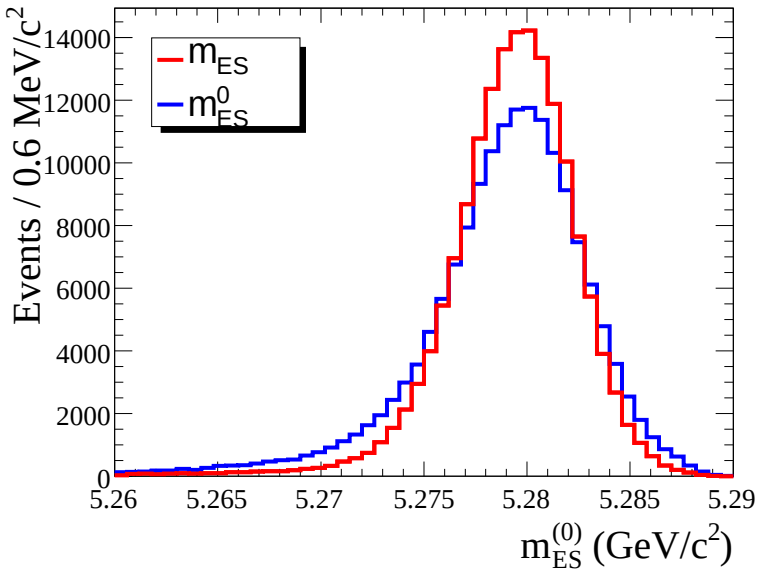


Figure 4-4: Comparison of the resolution of m^0_{ES} with m_{ES} for $B^0 \to \rho^0\gamma$ signal MC. The blue histogram shows m^0_{ES} and the red histogram shows m_{ES}. Only truth matched signal MC candidates have been selected.

4.8.3 The Δz Separation Between the two B Decay Vertices

Since the leading order decays of B mesons proceed via the electroweak force and not the strong force, they are rather long-lived, $c\tau = 491.1\,\mu$m ($458.7\,\mu$m) for the B^+ (B^0) meson. However, the two B mesons in an $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ event have only a very small momentum in the CM frame and thus do not travel far before they decay. But since the CM frame is moving in the laboratory due to the Lorentz boost of $\beta\gamma = 0.56$, the two B mesons are traveling on average a distance of about $250\,\mu$m

before they decay. Thus, the separation of the two B meson decay vertices along the z axis can be measured.

In order to compute this separation, the decay vertex of the hypothetical other B meson in the event must be reconstructed. This other B meson is called the *"tag B"*. This is done by considering all charged tracks that are not used to reconstruct the signal B meson under consideration. A Kalman fit is performed on these trajectories, that is after an initial seed, charged-particle trajectories are added one at a time and a lest χ^2 fit is performed. Only tracks that do not contribute too much to the overall χ^2 of the vertex are kept in the fit.

To achieve an unbiased fit result for Δz, the final fitter needs to have the four-momentum of the *"tag B"* meson unconstrained. This is achieved by adding a hypothetical missing particle to the tracks used for the vertex fit of the *"tag B"* meson. Together with the signal-side B meson candidate, an $\Upsilon(4S)$ candidate is formed. The $\Upsilon(4S)$ is constrained to be originating from the beam spot. With this constraint, a final Kalman fit is performed to the whole decay tree of the $\Upsilon(4S)$ candidate, including the decay tree of the signal side B meson. All particle lifetimes are properly considered in this fit, in fact the sum of the two fitted B meson lifetimes t_B is constrained to be twice the nominal B meson lifetime τ_B

$$2 \cdot t_B = 2 \cdot \tau_B. \tag{4.12}$$

This fit can also determine the vertex separation in the $B^+ \to \rho^+\gamma$ mode even though the signal B meson candidate contains only one charged particle trajectory. For this case, the constraint on the sum of the two B meson lifetimes is crucial as is the information to the fit that the B meson will not travel far away from the z axis due to the low CM momentum of the B mesons. This later information basically favors the point of closest approach of the single charged-particle trajectory to the z axis as the signal B meson decay vertex.

One result of this fit is the separation of the two B meson vertices along the z axis Δz, including its error $\sigma(\Delta z)$. In order to reduce unphysical results and failed

fits, two cuts are applied:

- $|\Delta z| < 0.4\,\mathrm{cm}$
- $\sigma\left(\Delta z\right) < 0.04\,\mathrm{cm}$

4.8.4 Analysis Regions

This analysis makes use of an optimization region, a fit region and a signal region. They are defined as:

- Optimization region:

  $(-0.3\,\mathrm{GeV} < \Delta E < 0.3\,\mathrm{GeV}) \otimes (5.20\,\mathrm{GeV}/c^2 < m_{ES} < 5.30\,\mathrm{GeV}/c^2)$

- Fit region:

  $(-0.3\,\mathrm{GeV} < \Delta E < 0.3\,\mathrm{GeV}) \otimes (5.22\,\mathrm{GeV}/c^2 < m_{ES} < 5.30\,\mathrm{GeV}/c^2)$

- Signal region:

  $(-0.2\,\mathrm{GeV} < \Delta E < 0.1\,\mathrm{GeV}) \otimes (5.27\,\mathrm{GeV}/c^2 < m_{ES} < 5.29\,\mathrm{GeV}/c^2)$

Chapter 5

Background Suppression

The major challenge for this analysis is the suppression of the continuum background. To maximize the separation between this type of background and signal, three methods are used:

- Vetoing candidates where the high energy photon candidate actually originated from a π^0 or η decay with a two-dimensional likelihood method.
- Rejecting continuum background with a high-dimensional neural network (NN) utilizing mostly information from the reconstructed particles not associated with the signal B candidate.
- Performing a simultaneous multidimensional cut optimization on selected variables.

5.1 π^0 and η Veto for the High-Energy Photon

In a large fraction of background events, the high-energy photon candidate does not originate directly from a B decay, but rather from the decay of a high-energy π^0 or η to a pair of photons. The relative sizes of these two sources of backgrounds to the high-energy photon *w.r.t.* everything else can be seen in Fig. 5-1.

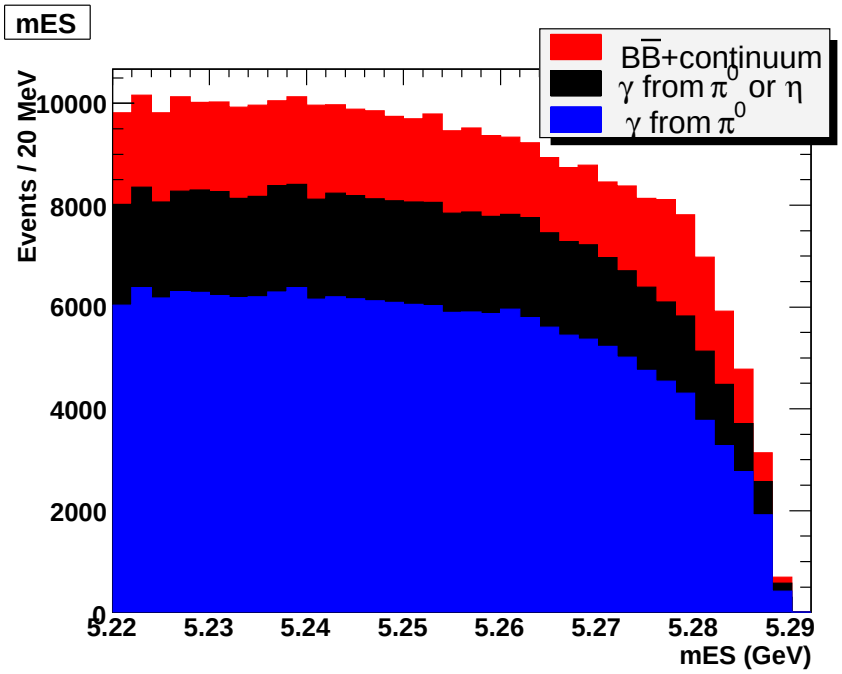


Figure 5-1: Background composition in terms of origins of the high energy photon candidates.

5.1.1 Likelihood Based Method

To suppress these background sources, a two-dimensional likelihood based method has been developed. The high-energy photon candidate (γ_B) is paired with any other photon in the event from the `CalorNeutral` list (γ_2), that is with any other energy deposition in the EMC not associated with a charged track.

Two variables are considered. The invariant mass of the two-photon pair

$$m_{\gamma_B\gamma_2} = |p_{\gamma_B} + p_{\gamma_2}|\,, \tag{5.1}$$

where p_{γ_B} and p_{γ_2} are the four-momenta of the high-energy photon candidate and the other photon, respectively. The second variable is the energy of the second photon γ_2 in the laboratory frame. The different shapes of these variables between a true signal high-energy photon and a high-energy photon candidate originating from a π^0 or η can be seen in Fig. 5-2.

A two-dimensional fit to these two variables is performed for each of the three

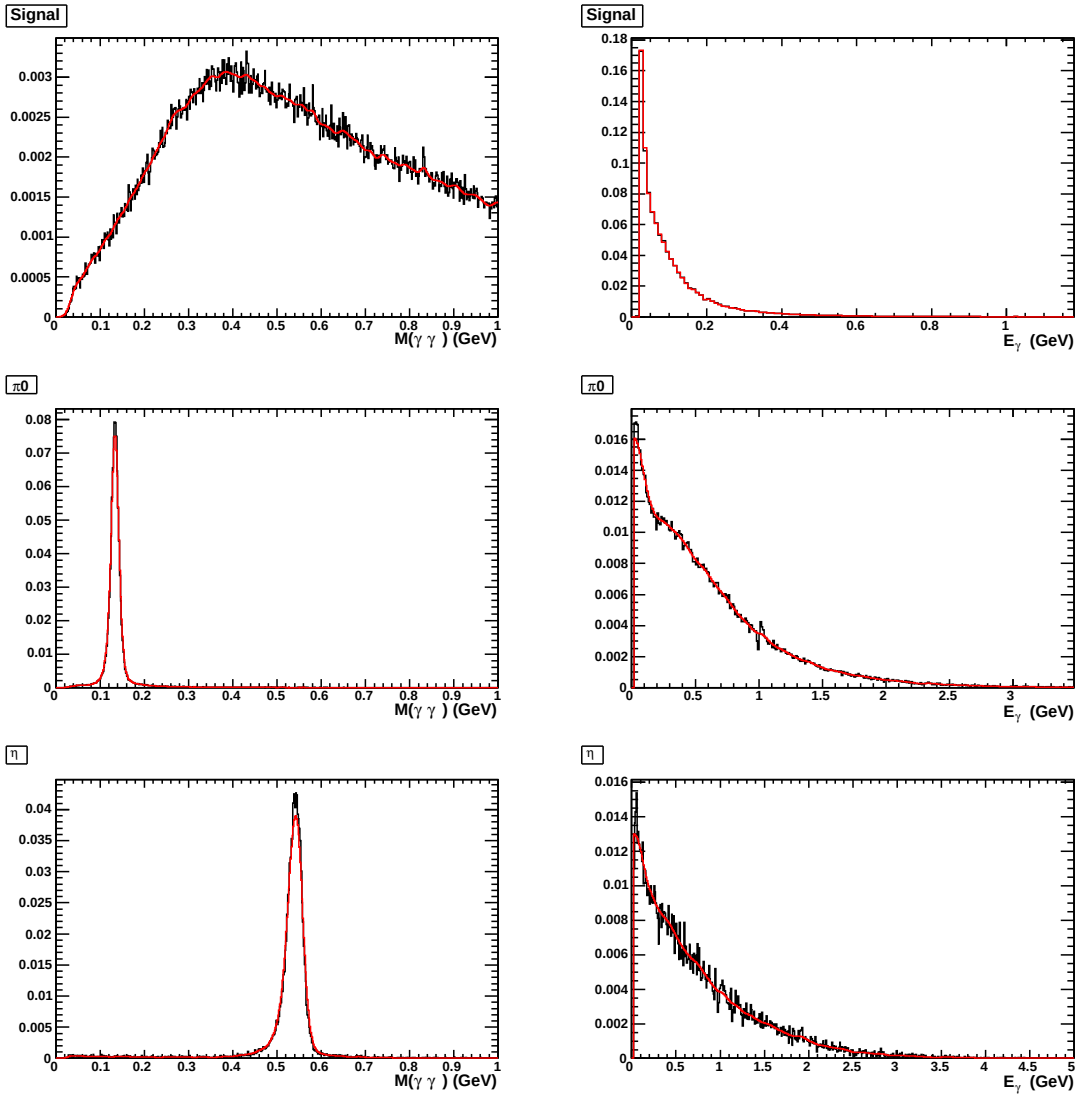


Figure 5-2: Distributions of $m_{\gamma_B\gamma_2}$ (left) and E_{γ_2} (right) for truth matched signal events, and high-energy photons originating from π^0 and η decays. The curves are projections of the two dimensional PDFs obtained from these MC events.

categories (true signal from signal MC; high-energy photon originating from a π^0 or η decay where both photons are truth matched to the same mother π^0 or η using continuum MC) individually using a two-dimensional non-parametric 2D-Keys probability density function (PDF). The 1-dimensional projections of these 2-dimensional PDFs are also shown in Fig. 5-2. In order to technically handle these PDFs, they are binned with a bin width of $2\,\mathrm{MeV}/c^2$ in $m_{\gamma_B\gamma_2}$ and a bin width of $10\,\mathrm{MeV}$ in E_{γ_2}.

From these 2-dimensional binned PDFs, the two likelihood ratios ($\mathcal{LR}$) are built as the ratio of the likelihood of a γ_B originating from a true π^0 (η) decay to the sum of likelihood of that same γ_B being from a true $B \to \rho\gamma$ decay and the likelihood of being from a true π^0 (η)

$$\mathcal{LR}(i) = \frac{PDF(M(\gamma_B\gamma_2), E_{\gamma_2}|i)}{PDF(M(\gamma_B\gamma_2), E_{\gamma_2}|\rho\gamma) + PDF(M(\gamma_B\gamma_2), E_{\gamma_2}|i)}, \tag{5.2}$$

where i is either π^0 or η. The larger this $\mathcal{LR}$ is (the closer to 1), the more likely a high-energy photon candidate originates from a π^0 (η). For each high-energy photon candidate, only one combination with another photon in the event is retained for each of the two $\mathcal{LR}(\pi^0)$ and $\mathcal{LR}(\eta)$, the one with the highest value of the corresponding $\mathcal{LR}$. The two resulting $\mathcal{LR}$ for true signal and true background can be seen in Fig. 5-3. Also, the resulting background $\mathcal{LR}$ are compared to off–resonance data.

5.1.2 Converted Photons

Another source of photons that can form a true π^0 or η candidate when combined with the high-energy photon candidate, are converted photons. Photons can in the presence of matter convert into an e^+e^- pair. The high-energy photon candidate is combined with any e^+e^- pair consistent with a converted photon. An e^+e^- pair is considered to be a converted photon if

- its invariant mass $m_{e^+e^-} < 50\,\mathrm{MeV}/c^2$,
- the distance of the e^+e^- vertex to the primary vertex be $< 0.5\,\mathrm{cm}$ in the xy-plane,

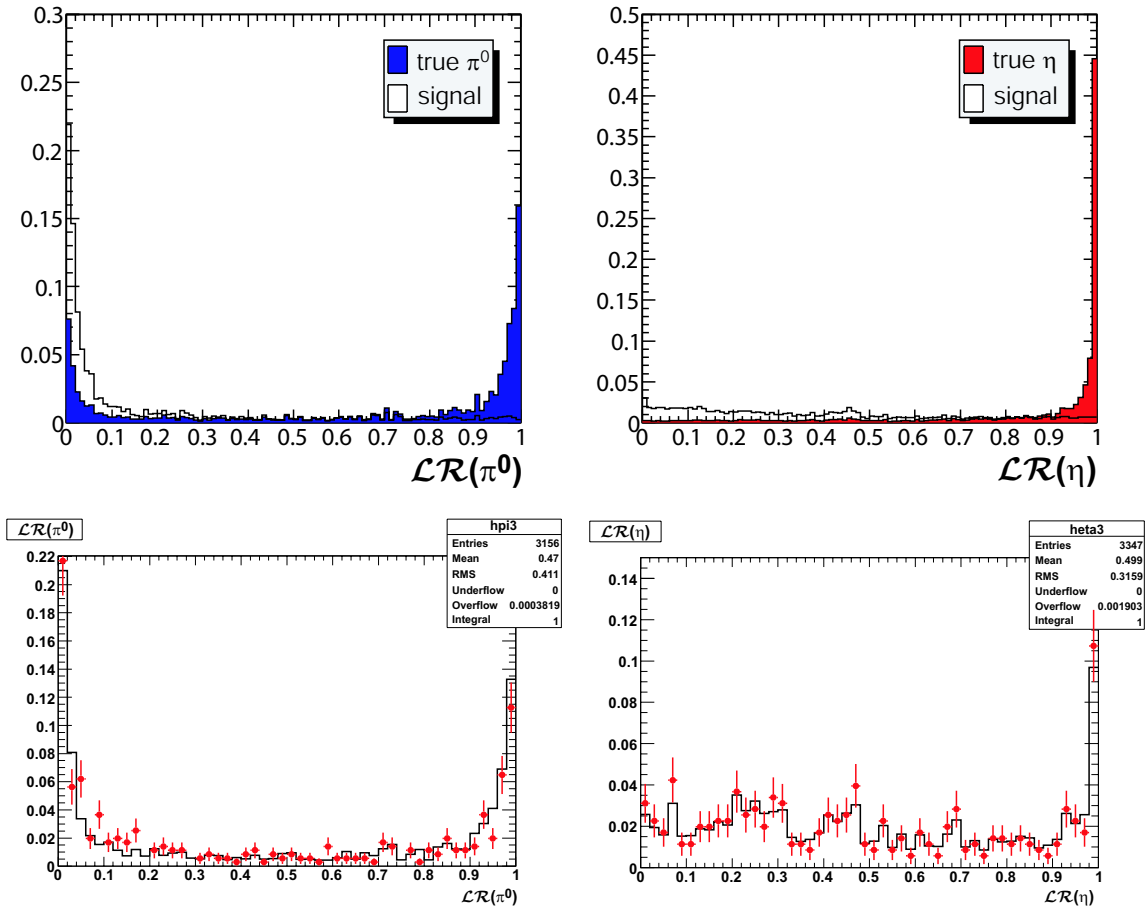


Figure 5-3: π^0 and η likelihood ratios for MC and off-resonance data for the ρ^0 mode. In the top two plots, signal MC is depicted as an open histogram and the π^0 and η background MC is depicted in a filled histogram. The bottom two plots show the likelihood ratio distributions for off-resonance data (points with error bars) and continuum MC (histogram) for the ρ^0 mode. On the left hand side is the π^0 likelihood ratio and on the right is the η likelihood ratio.

- and the distance of the e^+e^- vertex to the primary vertex be $< 1.0\,\mathrm{cm}$ along the z-axis.

The invariant mass of the high-energy photon candidate and the e^+e^- pair is calculated. Simple veto cuts are applied in order to reduce the background from π^0 and η decays. The vetoed invariant mass $m_{\gamma_B e^+e^-}$ regions are $[100, 160]\,\mathrm{MeV}/c^2$ around the π^0 and $[500, 590]\,\mathrm{MeV}/c^2$ around the η. These cut criteria are 99.8% efficient for signal MC. Distributions for signal and continuum MC are shown in Fig. 5-4. For the $B^0 \to \rho^0\gamma$ mode about 10.4% of the uds, 11.7% of the $c\bar{c}$, 9.0% for the $\tau^+\tau^-$ and

8.6% of the signal MC events contain a converted photon.

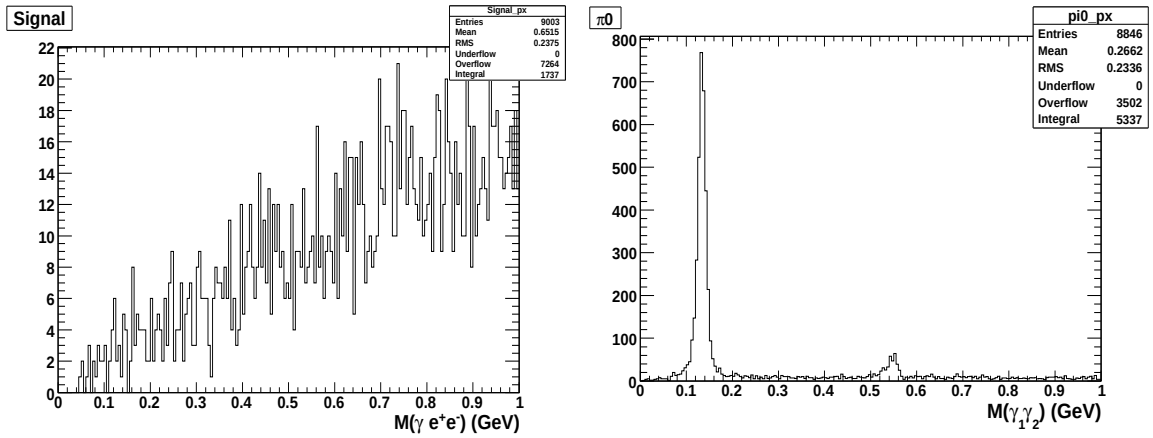


Figure 5-4: $m_{\gamma_B e^+e^-}$ distributions for $\rho^0\gamma$ signal (left) and continuum MC (right).

5.1.3 Combined Veto Performance

In order to gauge the performance of this two-fold π^0/η veto technique ($\mathcal{LR}$ and the use of converted photons), a comparison with simple veto cuts analogous to a previous analysis [24] is performed. The veto regions are $105 < m_{\gamma_B e^+e^-} < 155\,\mathrm{MeV}/c^2$ for energies of the second photon E_{γ_2} greater than 30 MeV and $0.500 < m_{\gamma_B e^+e^-} < 0.590\,\mathrm{MeV}/c^2$ for E_{γ_2} greater than 250 MeV. The efficiency-efficiency curve for the three methods (old cut based veto, $\mathcal{LR}$ based veto alone and $\mathcal{LR}$ based veto plus converted photons veto) can be seen in Fig. 5-5. Clearly, the combination of the $\mathcal{LR}$ and the use of converted photons is superior to the other two in both the π^0 and η case.

Another illustration of the improvement is shown in Table 5.1 where efficiencies of off–resonance data are compared between the old cut-based veto method and the combined $\mathcal{LR}$ and converted photons method. The cut on the $\mathcal{LR}$ is adjusted such that the signal efficiencies (for all three decay modes individually) are the same as for the old cut-based method. It can be clearly seen that the off–resonance data efficiencies are lower for the new $\mathcal{LR}$ based method.

Overall, this new method of utilizing a two-dimensional $\mathcal{LR}$ and in addition using also photons converted to an e^+e^- pair is gaining approximately 10% in background

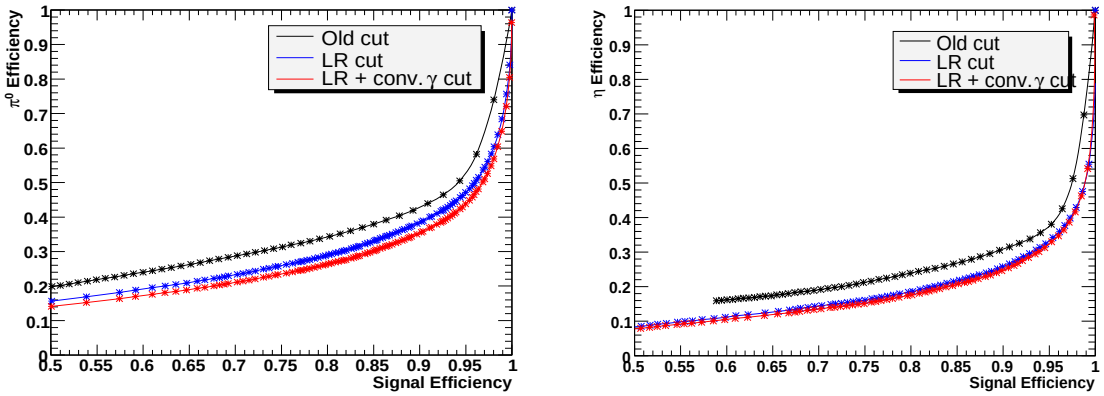


Figure 5-5: π^0/η efficiency vs. signal efficiency for different veto methods. The black curve is for the veto using only $M(\gamma_1\gamma_2)$, the blue curve is for the likelihood method, and the red curve is for the likelihood veto plus converted π^0/η veto.

rejection over a simple cut-based veto method used in the old analysis.

	π^0 Veto		η Veto		Combined	
	$\mathcal{LR}(\pi^0)$	$m_{\gamma_B\gamma_2}$	$\mathcal{LR}(\eta)$	$m_{\gamma_B\gamma_2}$	$\mathcal{LR}(\pi^0/\eta)$	$m_{\gamma_B\gamma_2}$
$B \to \rho^0\gamma$						
Signal Eff%	91.2±0.2	91.1±0.2	95.2±0.2	95.1±0.2	86.8±0.2	86.5±0.2
Off-res Eff%	58.4±0.1	63.4±0.1	73.2±0.2	73.6±0.2	42.8±0.1	46.7±0.1
Cont. MC%	54.3±0.1	59.0±0.1	75.5±0.2	76.0±0.2	41.0±0.1	44.8±0.1
$B \to \rho^+\gamma$						
Signal Eff%	90.9±0.3	90.7±0.3	94.9±0.3	94.9±0.3	86.3±0.3	86.1±0.3
Off-res Eff%	58.6±0.1	63.3±0.1	73.0±0.2	73.5±0.2	42.7±0.1	46.6±0.1
$B \to \omega\gamma$						
Signal Eff%	91.0±0.3	90.8±0.3	94.8±0.3	94.9±0.3	86.3±0.3	86.2±0.3
Off-res Eff%	59.6±0.2	64.2±0.2	72.2±0.3	72.7±0.3	43.0±0.2	46.6±0.2

Table 5.1: Veto efficiencies for off-resonance data and continuum MC. Comparison of the π^0/η veto efficiencies on Run 1-5 off-resonance data and continuum MC. The other photon quality requirements imposed are those listed in Table 5.5, and the requirement second moment > 0.002. The $m_{\gamma_B\gamma_2}$ veto regions are the same as used in the previous *BABAR* analysis [24] [105, 155] MeV/c^2 for the photon energies greater than 30 MeV and [500, 590] MeV/c^2 for the photon energies greater than 250 MeV. For the purpose of comparison, the $\mathcal{LR}$ cut values for this table are adjust to have the same signal efficiency as the $m_{\gamma_B\gamma_2}$ cuts.

5.2 Background Suppression with a Neural Network

A multitude of variables with signal–background separating power are available. But there are correlations amongst these variables. Thus, simple cuts on each variable is not an optimal solution. A multivariate algorithm that is able to take correlations into account is a better solution. The multivariate algorithm chosen for this analysis is a neural network. It can not only take the correlations into account, but also achieve excellent signal–background separation in a single output variable by taking all input variables into account for a given event. This single output variable ranges from zero to one where an output closer to one is more likely to be a signal event and and output closer to zero is more likely to be a background event. A pedagogical introduction to neural networks is given in Appendix A.

5.2.1 Input Variables to the Neural Network

The neural networks used for this analysis rely on the following class of input variables:

- Event shape variables computed on the whole event,
- rest-of-event (ROE) shape variables computed on the whole event excluding the reconstructed signal B meson candidate,
- ROE variables related to the tagging of the flavor of the presumed other B meson in the event,
- variables computed in the frame recoiling against the high-energy photon candidate,
- angular distributions of the B meson and spacial separation between the signal B meson decay vertex and the decay vertex of the presumed other B meson in the event.

All these variables have discrimination power, which means that their shape is different for signal and continuum background events. A brief description of all used variables is given below. Unless otherwise noted, all the variables described below are used as inputs to the three neural networks for the three decay modes. A summary of the used input variables for all three modes can be seen in Table 5.2. Comparison plots between signal MC and continuum background MC on the one hand and continuum background MC and off–resonance data on the other hand are shown for all neural network input variables in Appendix B.

The Second-to-Zeroth Fox-Wolfram Moment of the Event

The second-to-zeroth Fox-Wolfram moment using all charged and neutral candidates, as described in Section 4.2.3, is used as an input to all three neural networks. In order to suppress $e^+e^- \to \tau^+\tau^-$ events without loosing a large fraction of signal events, a loose pre-cut of $R_2^{All}<0.7$ is applied.

The Second-to-Zeroth Fox-Wolfram Moment in the Photon Recoil Frame

The R_2' variable is similar to R_2^{All} and it is computed in almost the same way. The only difference is that it is not computed in the CM frame as R_2^{All} is, but rather, the high-energy photon is removed from the event and the rest frame of the remaining particles is used for computing R_2'. The prime is an indication of the use of the high-energy photon recoil frame for this variable. All charged particle candidates and all neutral particle candidates with energies greater than 30 MeV are used, excluding the candidates used for reconstructing the signal B meson. The idea of "removing" the high-energy photon from the event is to be sensitive to ISR events. By going to the rest frame of the initial e^+e^- system recoiling against the high energy photon coming from ISR, one recovers two back-to-back final state particle jets in the event.

Legendre Moments in the Photon Recoil Frame

Another attempt at separating signal events from ISR events is made by computing Legendre moments in the frame recoiling against the high-energy photon candidate.

Legendre moments are defined as

$$L'_i = \frac{\sum_{j \text{ in ROE}} |\vec{p}'_j| \cdot |\cos(\theta'_j)|^i}{\sum_{j \text{ in ROE}} |\vec{p}'_j|}, \tag{5.3}$$

where θ'_j is the angle between the j^{th} particle in the ROE and the high-energy photon momentum direction. $\vec{p}'_j$ is the three-momentum vector of the j^{th} particle in the ROE boosted to the frame recoiling against the high-energy photon candidate.

The 0^{th} to 3^{rd} Legendre moments are computed and the ratios L'_1/L'_0 and L'_2/L'_0 are used as inputs for the neural networks for all three modes, whereas the ratio L'_3/L'_0 is only used as an input for the $B^+ \to \rho^+\gamma$ neural network.

The Thrust Angle

The thrust angle Θ_T is the angle between the high-energy photon candidate and the thrust axis of the rest of the event computed in the CM frame. The thrust axis of the event $\hat{T}$ is defined as the axis which maximizes the sum of the particle momenta along this axis. The thrust T is related to this axis by

$$T = \frac{\sum_i \left|\hat{T} \cdot \vec{p}_i\right|}{\sum_i |\vec{p}_i|}\ , \tag{5.4}$$

where $0.5 < T < 1.0$. T close to 1 corresponds to a jet like event and T close to 0.5 corresponds to an isotropic B meson event. The daughters of the signal B meson are excluded for the determination of the thrust axis, but all other neutral and charged candidates in the event are used.

Due to the small 3-momentum of the B meson, signal events are very isotropic. Thus, the momentum direction of the high energy photon is uncorrelated with the daughters of the other B meson. The distribution of $\cos\Theta_T$ is thus expected to be flat. This is not the case for continuum background events. As mentioned before, a continuum background event is mostly a pair of two back-to-back jets. The high-energy photon originates usually from a high energy π^0 or η decaying in one of these jets. Thus, the momentum direction of the high energy photon is almost identical

with the direction of one of the jets. For continuum background events, Θ_T is thus expected to be close to 0° or 180°. Thus, $\cos\Theta_T$ is expected to peak around ± 1 for a continuum background event.

Legendre Moments in the CM Frame *w.r.t.* the Thrust Axis of the ROE

In order to separate signal events from continuum background events, Legendre moments are computed on the ROE in the CM frame. These Legendre moments are defined as

$$L_i = \frac{\sum_{j \text{ in ROE}} |\vec{p}_j| \cdot |\cos(\theta_j)|^i}{\sum_{j \text{ in ROE}} |\vec{p}_j|}, \tag{5.5}$$

where θ_j is the angle between the j^{th} particle in the ROE and the thrust axis of the ROE $\hat{T}$. $\vec{p}_j$ is the CM three-momentum vector of the j^{th} particle in the ROE.

The ratios L_1/L_0 and L_2/L_0 are used as inputs for the neural networks for all three modes, whereas the ratio L_3/L_0 is only used as an input for the $B^+ \to \rho^+\gamma$ neural network.

The Polar Angle of the B Meson

The variable $\Theta_B$is the angle between the momentum vector of the B meson candidate and the beam direction computed in the CM frame. This is simply the helicity angle for the $\Upsilon(4S)$ meson. Since the $\Upsilon(4S)$ is a vector meson and the two B mesons are scalars, the expected distribution of Θ_B is a $\sin^2\Theta_B$ distribution. Since there is no true B meson candidate in a continuum event, the distribution of $\cos\Theta_B$ is expected to be flat.

The Vertex Separation Significance between the two B Mesons

A variable that is rather independent from the ones described above is the separation of the two B meson decay vertices along the z axis, Δz, described in Section 4.8.3. In order to also incorporate the uncertainty on Δz, the significance of Δz is computed as

$$\Sigma_{\Delta z} = \frac{\Delta z}{\sigma\left(\Delta z\right)}. \tag{5.6}$$

For true signal events, this quantity is wider than for continuum background events due to the long lifetime of the B meson. Since the sign of this quantity has no physical meaning in separating signal events from continuum background events, the absolute value of $\Sigma_{\Delta z}$ is fed into the neural network.

Leptons in the ROE

B mesons decay via the weak interaction. Approximately 10.6% of these decays are semi-leptonic [22]. This is quite different from a typical continuum background event where leptons are produced almost exclusively in charmed hadron decays. This fact is utilized in this analysis by searching for muons and electrons in the ROE which could originate from a semi-leptonic B meson decay. If multiple electrons or multiple muons are found, the two most energetic ones of each species are further considered.

In addition to the absolute value of the CM momentum (multiplied by the electric charge of the lepton), two more variables are fed into the neural networks: E_{90}^{W} and $\cos\theta_{miss}$. E_{90}^{W} is defined as the CM energy in the hemisphere defined by the direction of the virtual W^+ boson in the assumed semi-leptonic B decay. To calculate this, the four-momentum of the neutrino p_ν^μ is determined from the four-momentum of the lepton p_l^μ, the four-momentum of the assumed other B meson in the event (assumed to be at rest in the CM frame) $p_B^\mu \approx \left(m_B, \vec{0}\right)$ and the sum of the four-momenta of all other charged and neutral candidates (excluding the signal B meson candidate) p_X^μ. With these definitions, E_{90}^{W} is determined as

$$\begin{aligned} p_W^\mu &= p_l^\mu + p_\nu^\mu \\ p_X^\mu &= \sum_{i \neq l} p_i^\mu \\ p_B^\mu &= p_W^\mu + p_X^\mu \approx \left(m_B, \vec{0}\right) \\ E_{90}^{W} &= \sum_{i \in X,\ \vec{p}_i \cdot \vec{p}_W > 0} E_i. \end{aligned} \tag{5.7}$$

The second variable $\cos\theta_{miss}$ is calculated as the cosine between the reconstructed

lepton momentum $\vec{p}_l$ and the missing momentum $\vec{p}_{miss}$, both in the CM frame

$$\vec{p}_{miss} = \vec{p}_B - \vec{p}_X - \vec{p}_l \approx -\left(\vec{p}_X + \vec{p}_l\right). \tag{5.8}$$

These three distributions (absolute value of the CM momentum multiplied by lepton's electric charge, E_{90}^W and $\cos\theta_{miss}$) are shown in Figure **??** for reconstructed electrons and muons.

In some cases, a charged lepton is reconstructed, but not identified as a lepton. In order to account for those events, the CM momenta multiplied by the electric charge of the two charged particle candidates with the highest CM momentum are also used as inputs to the neural networks.

Kaons in the ROE

$(78.9 \pm 2.5)\%$ of all B mesons produce at least one charged kaon in their decay chain, and $(64 \pm 4)\%$ produce at least one neutral kaon in their decay chain [22]. They usually originate from a cascade decay of the form $B \to D \to K$. Thus, their momenta are generally lower than in an $e^+e^- \to s\bar{s}$ or $e^+e^- \to c\bar{c}$ event that indicate high-momentum D mesons decaying into kaons. The CM momenta of the two most energetic charged kaons and the two most energetic K_S^0 are used as inputs to the neural networks. The momenta of the charged kaons are multiplied by the electric charge of the kaon. The mass of K_S^0 with the highest momentum is also used as an input to the neural networks.

The Outputs of the Sub-taggers

Being able to determine the flavor of the *tag* B meson is crucial for time-dependent CP analysis. At *BABAR*, several neural networks have been trained to do this, based on different physics processes. For time-dependent CP analysis, the output of these so-called "sub-taggers" is combined into one final neural network which determines the overall flavor tag of the *tag* B.

Some of the outputs of the standard *BABAR* sub-taggers are also used as inputs to

the neural networks used in this analysis. Besides the *electron* and *muon* sub-taggers, the *slow pion*, and *kaon slow-pion* correlation sub-taggers are also used. The later two are aimed at finding B meson that decayed to a D^* meson. The D^* usually decays into a D meson and a pion. This pion has a rather low momentum in the rest frame of the D^* meson due to the small mass difference between the D^* and the D meson. And the D meson decays usually into a kaon.

The *slow pion* sub-tagger uses the CM momentum of charged particle candidates, the likelihood of this candidate to be a kaon, and the cosine of the angle between the momentum of the candidate and the thrust axis of the ROE. This sub-tagger was trained to identify true slow pions originating from a $D^* \to D\pi$ decays.

The *kaon slow-pion* correlation sub-tagger is aimed at finding a slow pion and a charged kaon that both originate from the same D^* decay. The inputs to this sub-tagger are particle identification information for the kaon, the charge of the kaon, the charge of the pion, the output of the *slow pion* sub-tagger and the angle between the kaon and slow-pion, computed in the CM frame. The *kaon slow-pion* sub-tagger output is only used as an input to the neural network for the $B^+ \to \rho^+\gamma$ decay mode.

The *electron* and *muon* sub-taggers are used in this analysis in order to find semi-leptonic B decays, but they were originally developed to distinguish a B^0 from a $\overline{B}^0$, as were the *slow pion* and *kaon slow-pion* sub-taggers. The inputs to the *electron* and *muon* sub-taggers are the CM momentum, the electric charge, E_{90}^W and $\cos\theta_{miss}$ for a charged particle candidate identified as an electron or muon.

5.2.2 Neural Network Algorithm and Training Implementation

The "Stuttgart Neural Network Simulator" (SNNS) [36] is the neural network implementation used for this analysis. Three individual neural networks have been trained, one for each one of the three decay modes $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$. The neural networks have been trained on truth-matched signal MC and continuum background MC. The MC samples are split randomly into half. The first half is used for

N	Variable Description	Used in NN for mode		
		$\rho^+\gamma$	$\rho^0\gamma$	$\omega\gamma$
1	$\cos\Theta_T$ angle between ROE thrust axis and high-E γ	Yes	Yes	Yes
2	R_2^{All} of all charged and neutrals in ROE CM frame	Yes	Yes	Yes
3	$\cos\Theta_B$ cosine of the polar angle of the B meson	Yes	Yes	Yes
4	L_2'/L_0' of ROE in high-energy photon recoil frame	Yes	Yes	Yes
5	L_2/L_0 in CM frame *w.r.t.* thrust axis of ROE	Yes	Yes	Yes
6	L_1/L_0 in CM frame *w.r.t.* thrust axis of ROE	Yes	Yes	Yes
7	L_3/L_0 in CM frame *w.r.t.* thrust axis of ROE	Yes	No	No
8	L_1'/L_0' of ROE in high-energy photon recoil frame	Yes	Yes	Yes
9	L_3'/L_0' of ROE in high-energy photon recoil frame	Yes	No	No
10	R_2' of ROE in high-energy photon recoil frame	Yes	Yes	Yes
11	$\Sigma_{\Delta z}$ separation significance between B decay vertices	Yes	Yes	Yes
12	$\cos\theta_{miss}\,(e_1)$ of highest momentum electron	Yes	Yes	Yes
13	$E_{90}^W\,(e_1)$ of highest momentum electron	Yes	Yes	Yes
14	$q\cdot\lvert\vec{p}^*\rvert\,(e_1)$ of highest momentum electron	Yes	Yes	Yes
15	$\cos\theta_{miss}\,(e_1)$ of highest momentum electron	Yes	No	No
16	$E_{90}^W\,(e_2)$ of 2^{nd} highest momentum electron	Yes	Yes	Yes
17	$q\cdot\lvert\vec{p}^*\rvert\,(e_2)$ of 2^{nd} highest momentum electron	Yes	Yes	Yes
18	$\cos\theta_{miss}\,(\mu_1)$ of highest momentum muon	Yes	Yes	Yes
19	$E_{90}^W\,(\mu_1)$ of highest momentum muon	Yes	No	No
20	$q\cdot\lvert\vec{p}^*\rvert\,(\mu_1)$ of highest momentum muon	Yes	Yes	Yes
21	$E_{90}^W\,(\mu_2)$ of 2^{nd} highest momentum muon	Yes	Yes	Yes
22	$q\cdot\lvert\vec{p}^*\rvert\,(\mu_2)$ of 2^{nd} highest momentum muon	Yes	Yes	Yes
23	$q\cdot\lvert\vec{p}^*\rvert\,(K_1)$ of highest momentum charged kaon	Yes	No	No
24	$q\cdot\lvert\vec{p}^*\rvert\,(K_2)$ of 2^{nd} highest momentum charged kaon	Yes	Yes	Yes
25	$\lvert\vec{p}^*\rvert\,\left(K_S^0\ 1\right)$ of highest momentum K_S^0	Yes	Yes	Yes
26	$\lvert\vec{p}^*\rvert\,\left(K_S^0\ 2\right)$ of 2^{nd} highest momentum K_S^0	Yes	Yes	Yes
27	$q\cdot\lvert\vec{p}^*\rvert\,(\max_1)$ of highest momentum charged particle	Yes	Yes	Yes
28	$q\cdot\lvert\vec{p}^*\rvert\,(\max_2)$ of 2^{nd} highest momentum charged particle	Yes	Yes	Yes
29	$m_{\pi^+\pi^-}$ of highest energy K_S^0 in ROE	Yes	Yes	Yes
30	Output of the *electron* sub-tagger	Yes	Yes	Yes
31	Output of the *muon* sub-tagger	Yes	Yes	Yes
32	Output of the *slow pion* sub-tagger	Yes	Yes	Yes
33	Output of the *kaon-slow pion* sub-tagger	Yes	No	No

Table 5.2: NN input variables. Listed in the first column is the input node number N. The second column shows the variable where all variables are described in detail in the preceding sections. The last three columns state whether the variable is used in the mode specific neural network.

training and the second half is used for validating the trained neural network. The amount of continuum background MC corresponds to approximately 250fb^{-1} of data where the relative fractions of $e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$, $e^+e^- \to c\overline{c}$ and $e^+e^- \to \tau^+\tau^-$ MC are taken according to their cross sections. The goal of the training is the separation of signal from continuum background MC.

A two-step iterative process is performed. First, a set of cuts are taken from the previous *BABAR* analysis of $B \to (\rho/\omega)\,\gamma$ [24] to obtain the first set of neural networks. Then, this neural network is used in the cut-optimization stage (see Section 5.3. The second and last round of neural network training uses the new cuts that are determined in this cut optimization. These cuts are listed in Table 5.3.

5.2.3 Neural Network Performance

The output of the neural networks for the three channels is shown in Figure 5-6. The performance of the three neural networks is assessed by cutting at different values of the neural network output (NN) and calculating the signal and background efficiencies for these cut values. From this, plots are made that depict on the ordinate the efficiency of rejecting background and on the abscissa the signal efficiency for the three decay modes. These plotsare shown in Figure 5-6.As can be seen on these plots, for a cut that retains 50% of the signal MC, the background rejection efficiency is greater than 98%.

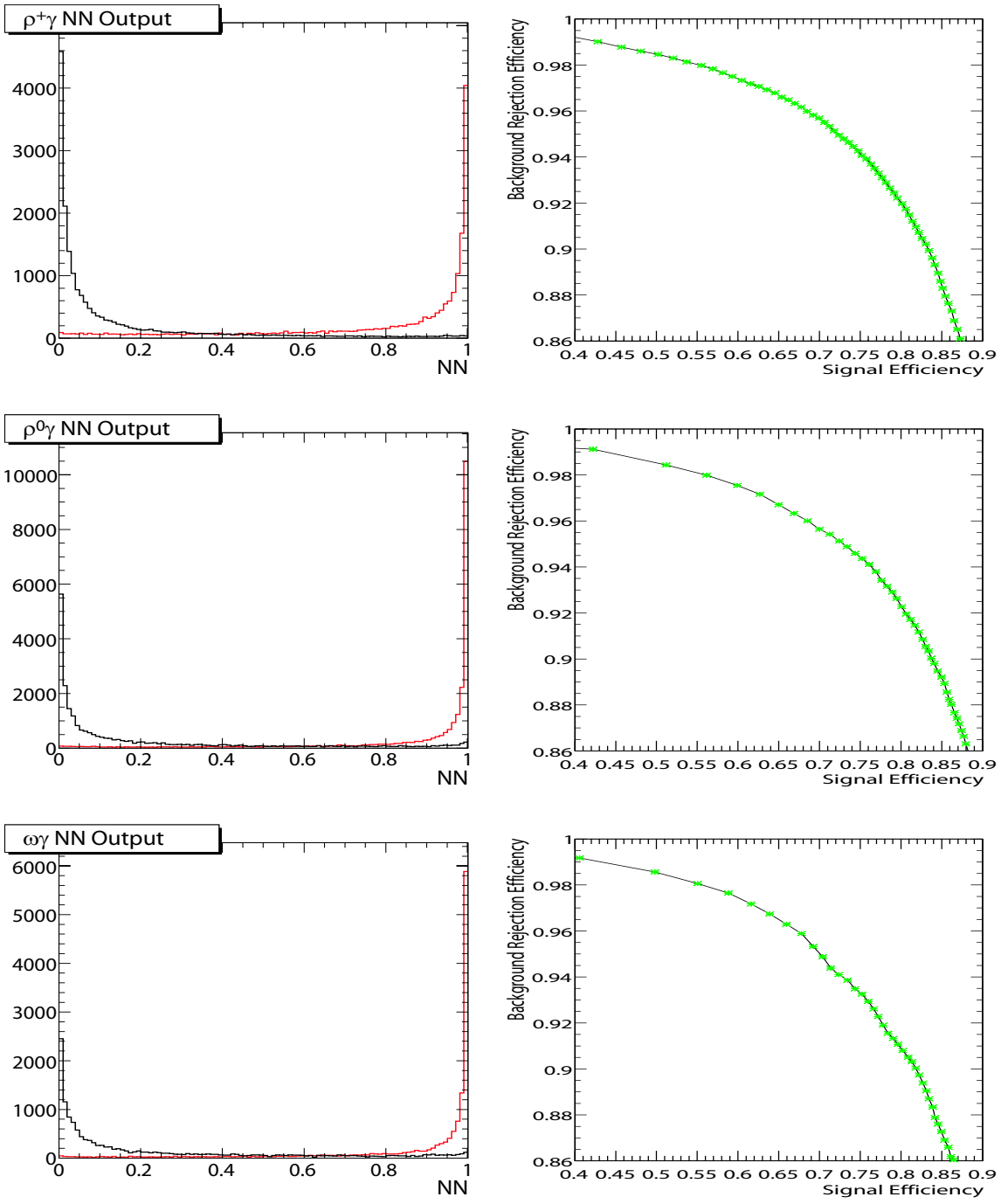


Figure 5-6: Resulting neural network performance. The left column shows the output of the trained neural networks for signal MC (red) and continuum MC (black). The right column shows the performance of the neural networks by showing the background rejection efficiency on the ordinate as a function of the signal efficiency on the abscissa. The first row shows the $B^+ \to \rho^+\gamma$ mode, the second row the $B^0 \to \rho^0\gamma$ mode and the last row the $B^0 \to \omega\gamma$ mode. For all plots, only events not used for training the neural networks have been used.

Description	$B^+ \to \rho^+\gamma$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$
Event	$R_2^{All} < 0.7$		
High-energy photon	$N_{\text{crystals}} > 4$ No problem crystal $-0.74 < \cos(\theta) < 0.93$ $> 25\,\text{cm}$ isolation		
	2^{nd} moment < 0.002 $0.944 < \frac{s9}{s25} < 0.996$	2^{nd} moment < 0.0017 $0.938 < \frac{s9}{s25} < 1.0$	2^{nd} moment < 0.0019 $0.945 < \frac{s9}{s25} < 0.990$
π^0 veto η veto	$\mathcal{LR}(\pi^0) < 0.944$ $\mathcal{LR}(\eta) < 0.932$	$\mathcal{LR}(\pi^0) < 0.737$ $\mathcal{LR}(\eta) < 0.918$	$\mathcal{LR}(\pi^0) < 0.767$ $\mathcal{LR}(\eta) < 0.939$
Converted γ veto	$0.1 < m_{\gamma_B e^+e^-} < 0.16$ and $0.5 < m_{\gamma_B e^+e^-} < 0.59$ (GeV/c^2)		
π^0 selection	in `pi0DefaultMass` list $0.789 < \cos\Theta_{\gamma\gamma} < 0.997$ $117 < m_{\gamma\gamma} < 148\,\text{MeV}/c^2$	n/a	in `pi0DefaultMass` list $0.413 < \cos\Theta_{\gamma\gamma} < 1.0$ $112 < m_{\gamma\gamma} < 1000\,\text{MeV}/c^2$
Tracking	GTL requirement for all charged tracks		
PID	veryTight π ID + DIRC photon consistency		tight π ID
Vector meson	$636 < m_{\pi^+\pi^0} < 932\,\text{MeV}/c^2$	$633 < m_{\pi^+\pi^-} < 957\,\text{MeV}/c^2$ $\text{Prob}\left(\chi^2_{\rho^0}\right) < 0.008$	$764 < m_{\pi^+\pi^-\pi^0} < 795\,\text{MeV}/c^2$ $\text{Prob}(\chi^2_\omega) < 0.000$ Flight distance $< 0.189\,\text{cm}$
Vertexing	$\lvert\Delta z\rvert < 0.4\,\text{cm}$ $\sigma(\Delta z) < 0.04\,\text{cm}$		
Optimization region	$\lvert\Delta E\rvert < 0.3\,\text{GeV}$ $5.2\,\text{GeV}/c^2 < m_{ES} < 5.3\,\text{GeV}/c^2$		

Table 5.3: Analysis cuts applied prior to the training of the final neural networks.

	$B^+ \to \rho^+\gamma$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$
Number of input variables	33	27	27
Number of hidden nodes	66	55	55
Number of hidden layers	1		
Training algorithm	Standard Backpropagation		
Number of training cycles	1400		

Table 5.4: Neural network architecture for the three decay modes.

5.3 Cut Optimization

5.3.1 Introduction

In order to maximize the signal significance for each mode individually, a number of variables that provide separation between signal and background are selected. Simple cuts are applied on these variables. These cuts are optimized *w.r.t.* the signal significance simultaneously in the multidimensional variable space. Here, the signal significance is defined as $\Sigma \equiv S/\sqrt{S+B}$ where S is the number of signal events and B is the number of background events.

5.3.2 The PRIM algorithm

The algorithm chosen to optimize the cuts in the multidimensional variable space is called PRIM [37]. This is a bump-hunting algorithm that searches for a rectangular region with an optimized figure of merit (here: signal significance) in a multi-dimensional variable space. This algorithm is implemented in the StatPattern-Recognition software package [38]. The algorithm works in two steps:

- Shrinkage of the multidimensional signal box of events intended to be kept.
- Expansion of the multidimensional signal box.

In the shrinkage stage, the bump-hunter algorithm tries to improve the figure of merit by looking at all possible cut tightenings in all input variables, one at a time. In a single step, the new cut that improves the figure of merit the most is chosen. The procedure is iteratively continued on all variables until the figure of merit cannot be further improved. The maximum amount by which a cut can be tightened is controlled by a peel parameter which sets the fraction of events that can be cut away from the signal box at a single step.

During the expansion stage, the bump hunter algorithm tries to improve the figure of merit by expanding the signal box, looking at one variable at a time for each step and keeping the relaxed cut which improves the figure of merit by the largest amount.

Description	Section	$B^+ \to \rho^+\gamma$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$
Event	4.2.3	$R_2^{All} < 0.7$		
High-energy photon	4.4.2 4.4.3 4.4.4	$N_{\text{crystals}} > 4$ No problem crystal $-0.74 < \cos(\theta) < 0.93$ $> 25\,\text{cm}$ isolation		
Converted γ veto	5.1.2	$0.1 < m_{\gamma_B e^+e^-} < 0.16$ and $0.5 < m_{\gamma_B e^+e^-} < 0.59$ (GeV/c^2)		
π^0 selection	4.5	in `pi0DefaultMass` list	n/a	in `pi0DefaultMass` list
Tracking	4.2.2	GTL requirement for all charged tracks		
PID	4.6.1	veryTight π ID + DIRC photon consistency		tight π ID
Vertexing	4.8.3	$\lvert\Delta z\rvert < 0.4\,\text{cm}$ $\sigma(\Delta z) < 0.04\,\text{cm}$		
Optimization region	4.8.4	$\lvert\Delta E\rvert < 0.3\,\text{GeV}$ $5.2\,\text{GeV}/c^2 < m_{ES} < 5.3\,\text{GeV}/c^2$		

Table 5.5: Fixed cuts applied before optimization is carried out for the remaining criteria.

Again, this is performed iteratively and stopped when no further improvement can be obtained.

5.3.3 Method and Used Variables

Before the considered variables are optimized, a pre-selection is applied. The applied cuts already introduced in Chapter 4 are listed in Table 5.5 and are mostly meant to assure the quality of the selected signal candidates. The optimization is performed on MC samples. As for the neural network training, all samples are randomly split in half, resulting in a training and a validation sample. Each sample is weighted to a target integrated luminosity of 300fb^{-1} and the used signal MC is truth matched.

Also, in order to have sufficient statistics available for the optimization, all candidates that fall into the optimization region (Section 4.8.4) are considered. But one really wants to optimize the signal significance in the signal region. To simulate this, the signal and background MC samples are weighted not only according to luminosity, but also by a weight factor determined from the ratio of events inside and outside the

signal region. This weight is determined independently for each MC sample.

ΔE and $m_{ES}$are not used in this cut optimization. The other variables used in the fit, NN, $\cos\Theta_H$ and $\cos\Theta_D$, are included in the optimization procedure. This is done so that the optimization algorithm finds the correct signal region for these variables as well. After the cut optimization is complete, the cuts on these variables are relaxed for the fit. This way, the separation these variables will give in the fit is simulated in the cut optimization.

The above mentioned peel parameter is chosen to be 0.3 for the $B^0 \to \rho^0\gamma$ mode and 0.4 for the $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ mode. With this choice, the PRIM algorithm finds the best cuts.

5.3.4 Resulting Cuts

The cuts resulting from the optimization are listed in Tables 5.6 to 5.8. The overall performance is estimated on the signal region of the training and validation MC samples, assuming a cut and count analysis without a final fit. As an indicator for the quality of the cuts, the signal significance is chosen as the figure of merit (FOM). The performance in the signal region after all cuts are applied is summarized in Tabel 5.9.

Variable category	Section	Lower limit	Variable	Upper limit
High-energy photon selection	4.4.5	0	2^{nd} moment	0.002
	4.4.6	0.944	$\frac{s9}{s25}$	0.996
	5.1.1	0.944	$\mathcal{LR}(\pi^0)$	1
		0.932	$\mathcal{LR}(\eta)$	1
π^0 selection	4.5	0.789	$\cos\Theta_{\gamma\gamma}$	0.997
		$0.117\,\mathrm{GeV}/c^2$	$m_{\gamma\gamma}$	$0.148\,\mathrm{GeV}/c^2$
Vector meson selection	4.7	$0.636\,\mathrm{GeV}/c^2$	$m_{\pi^+\pi^0}$	$0.932\,\mathrm{GeV}/c^2$
Fit variables	5.2	0.939	NN	1
	4.7.1	-0.733	$\cos\Theta_H$	0.660

Table 5.6: Summary of optimized selection criteria for $B^+ \to \rho^+\gamma$.

Variable category	Section	Lower limit	Variable	Upper limit
High-energy photon selection	4.4.5	0	2^{nd} moment	0.0017
	4.4.6	0.938	$\frac{s9}{s25}$	1
	5.1.1	0.737	$\mathcal{LR}(\pi^0)$	1
		0.918	$\mathcal{LR}(\eta)$	1
Vector meson selection	4.7	$0.633\,\mathrm{GeV}/c^2$	$m_{\pi^+\pi^-}$	$0.957\,\mathrm{GeV}/c^2$
		0.008	Prob $\left(\chi^2_{\rho^0}\right)$	1
Fit variables	5.2	0.978	NN	1
	4.7.1	-0.807	$\cos\Theta_H$	0.786

Table 5.7: Summary of optimized selection criteria for $B^0 \to \rho^0\gamma$.

Variable category	Section	Lower limit	Variable	Upper limit
High-energy photon selection	4.4.5	0	2^{nd} moment	0.0019
	4.4.6	0.945	$\frac{s9}{s25}$	0.990
	5.1.1	0.767	$\mathcal{LR}(\pi^0)$	1
		0.939	$\mathcal{LR}(\eta)$	1
π^0 selection	4.5	0.413	$\cos\Theta_{\gamma\gamma}$	1
		$0.112\,\mathrm{GeV}/c^2$	$m_{\gamma\gamma}$	$1\,\mathrm{GeV}/c^2$
Vector meson selection	4.7	$0.764\,\mathrm{GeV}/c^2$	$m_{\pi^+\pi^-\pi^0}$	$0.795\,\mathrm{GeV}/c^2$
		0.000	Prob (χ^2_ω)	1
		0 cm	Flight distance	0.189 cm
Fit variables	5.2	0.776	NN	1
	4.7.1	-0.717	$\cos\Theta_H$	0.838
	4.7.2	-0.843	$\cos\Theta_D$	1

Table 5.8: Summary of optimized selection criteria for $B^0 \to \omega\gamma$.

Mode	N_{sig} in 300fb^{-1}	N_{BG} in 300fb^{-1}	FOM = $\frac{S}{\sqrt{S+B}}$	$\frac{\text{Training}}{\text{Validation}}$ FOM
$B^+ \to \rho^+\gamma$	22.0	76.6	2.22	1.09
$B^0 \to \rho^0\gamma$	14.1	45.2	1.83	1.11
$B^0 \to \omega\gamma$	6.9	19.9	1.33	1.19

Table 5.9: Summary of Bump Hunter performance. In order to assess the performance of the cuts, the number of signal and background events are counted in the m_{ES}–ΔE signal region after all cuts are applied, based on MC. The figure of merit (FOM) is calculated based on these numbers. Here, the FOM is the signal significance $S/\sqrt{(S+B)}$. This would be the final signal significance in a cut-and-count analysis without a final maximum likelihood fit.

Chapter 6

Expected Yields

After all analysis cuts are applied, the resulting signal efficiencies are 11.87±0.07%, 14.52±0.07%, and 8.44±0.05% for the $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ mode, respectively. The efficiencies after each cut for signal MC are listed in Table 6.1 to 6.3 for the three decay modes. With these cuts, the background levels are reduced to a level that the subsequent maximum likelihood fits can safely use.

Potentially dangerous peaking B backgrounds are also studied by using dedicated signal MC samples. The results of these checks are summarized in Tables 6.4 to 6.6 and are found to be negligible in the $B^0 \to \omega\gamma$ channel. For the $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ modes, the most important peaking backgrounds turn out to be $B \to K^*\gamma$ and $b \to s\gamma$ decays.

Rho Efficiency Calculation: *signal* (Run 1-5); N generated Events = 280000								
Step	CutName	LowVal	HighVal	Evts	Cands	TMC's	Rel Eff%	Abs Eff%
				147811	380855	107770	52.8	52.8
1	SkimCuts	0	0	147427	378123	107751	**99.7±0.3**	**52.7±0.1**
2	GammaisOK	0.5	1.5	147427	378123	107751	**100±0.3**	**52.7±0.1**
3	gnCrys	4	1e+03	147426	378119	107751	**100±0.3**	**52.7±0.1**
4	gSecMom	0	0.002	145660	372783	106690	**98.8±0.3**	**52±0.1**
High Energy Photon								
5	gAcceptAngle	-0.74	0.93	142679	364406	104705	**98±0.3**	**51±0.1**
6	gdistNe	25	1e+08	134258	342179	98940	**94.1±0.3**	**47.9±0.1**
7	gdistCh	25	1e+07	131764	335225	97302	**98.1±0.3**	**47.1±0.1**
	Isolation						**92.3±0.3**	**47.1±0.1**
8	gPi0Veto	0.8	1.1	123018	310586	91004	**93.4±0.3**	**43.9±0.1**
9	gEtaVeto	0.94	1.1	118690	299113	87798	**96.5±0.3**	**42.4±0.1**
	π^0/η **Veto**						**90.1±0.3**	**42.4±0.1**
Charged Tracks								
10	GTL+	0	0	117605	282056	86243	**99.1±0.3**	**42±0.1**
11	GTL-	0	0	116453	266409	84732	**99±0.3**	**41.6±0.1**
	GoodTracksLoose						**98.1±0.3**	**41.6±0.1**
12	Pi+DrcCons	0	0	107914	242289	77233	**92.7±0.3**	**38.5±0.1**
13	Pi+PID	0	0	105272	231844	75472	**97.6±0.3**	**37.6±0.1**
14	Pi-DrcCons	0	0	95947	207880	67907	**91.1±0.3**	**34.3±0.1**
15	Pi-PID	0	0	92777	197525	66103	**96.7±0.3**	**33.1±0.1**
	$\pi^{\pm}$ **selector**						**79.7±0.3**	**33.1±0.1**
π^0 selections								
16	Pi0Mass	0.12	0.15	92777	197520	66103	**100±0.3**	**33.1±0.1**
17	Pi0cosGAng	0.79	1	83623	132872	62585	**90.1±0.3**	**29.9±0.1**
ρ^0 Meson								
18	RhoMass	0.64	0.93	65282	82578	50578	**78.1±0.3**	**23.3±0.09**
19	RhoCosHel	0	0.75	60059	72013	47847	**92±0.4**	**21.4±0.09**
Vertexing								
20	BDzErr	0	0.04	58087	69496	46259	**96.7±0.4**	**20.7±0.09**
21	BDz	-1	0.4	57417	68523	45790	**98.8±0.4**	**20.5±0.09**
	$\lvert\delta z\rvert$ **combined**						**95.6±0.4**	**20.5±0.09**
22	R2All	0	0.7	57065	68116	45514	**99.4±0.4**	**20.4±0.09**
23	NN	0.8	1.1	35995	40008	30131	**63.1±0.3**	**12.9±0.07**
24	Gammas9s25	0.94	inf	35766	39756	29958	**99.4±0.5**	**12.8±0.07**
25	gPi0Conv50	0.1	0.16	35760	39749	29952	**100±0.5**	**12.8±0.07**
26	gEtaConv50	0.5	0.59	35720	39702	29918	**99.9±0.5**	**12.8±0.07**
27	Empty	-1e+03	1e+03	35720	39702	29918	**100±0.5**	**12.8±0.07**
Fit Region								
28	mB1	5.2	5.3	34964	37926	29884	**97.9±0.5**	**12.5±0.07**
29	deltaE1	-0.3	0.3	33245	35464	28760	**95.1±0.5**	**11.9±0.07**
	TruthMatched							**10.3±0.06**
Signal Region								
30	mB2	5.3	5.3	31286	32561	27662	**94.1±0.5**	**11.2±0.06**
31	deltaE2	-0.2	0.1	29195	30193	25982	**93.3±0.5**	**10.4±0.06**
	TruthMatched							**9.28±0.06**

Table 6.1: $B^+ \to \rho^+\gamma$ signal MC efficiency table.

Rho0 Efficiency Calculation: *signal* (Run 1-5); N generated Events = 328000								
Step	CutName	LowVal	HighVal	Evts	Cands	TMC's	Rel Eff%	Abs Eff%
				221965	363550	195188	67.7	67.7
1	SkimCuts	0	0	221587	361549	195158	**99.8±0.2**	**67.6±0.1**
2	GammaisOK	0.5	1.5	221587	361549	195158	**100±0.2**	**67.6±0.1**
3	gnCrys	4	1e+03	221585	361547	195158	**100±0.2**	**67.6±0.1**
4	gSecMom	0	0.0017	215504	351614	190109	**97.3±0.2**	**65.7±0.1**
High Energy Photon								
5	gAcceptAngle	-0.74	0.93	210066	342774	185764	**97.5±0.2**	**64±0.1**
6	gdistNe	25	1e+08	197769	322483	175357	**94.1±0.2**	**60.3±0.1**
7	gdistCh	25	1e+07	193914	316076	172242	**98.1±0.2**	**59.1±0.1**
	Isolation						**92.3±0.2**	**59.1±0.1**
8	gPi0Veto	0.74	1.1	178731	290537	158954	**92.2±0.2**	**54.5±0.1**
9	gEtaVeto	0.92	1.1	169798	275668	150980	**95±0.2**	**51.8±0.1**
	π^0/η **Veto**						**87.6±0.2**	**51.8±0.1**
Charged Tracks								
10	GTL+	0	0	164961	247454	145151	**97.2±0.2**	**50.3±0.1**
11	GTL-	0	0	160599	222991	139885	**97.4±0.2**	**49±0.1**
	GoodTracksLoose						**94.6±0.2**	**49±0.1**
12	Pi+DrcCons	0	0	134811	184850	114600	**83.9±0.2**	**41.1±0.1**
13	Pi+PID	0	0	126992	170366	108524	**94.2±0.3**	**38.7±0.1**
14	Pi-DrcCons	0	0	107990	143352	90671	**85±0.3**	**32.9±0.1**
15	Pi-PID	0	0	102135	132947	86122	**94.6±0.3**	**31.1±0.1**
	$\pi^{\pm}$ **selector**						**63.6±0.2**	**31.1±0.1**
ρ^0 Meson								
16	r0Mass	0.63	0.96	84438	95837	72515	**82.7±0.3**	**25.7±0.09**
17	RhoChi2Prob	0.008	1.1	79450	87390	69040	**94.1±0.3**	**24.2±0.09**
18	RhoCosHelic	0	0.75	70281	73245	62913	**88.5±0.3**	**21.4±0.08**
Vertexing								
19	BDzErr	0	0.04	69178	72123	61904	**98.4±0.4**	**21.1±0.08**
20	BDz	-1	0.4	68601	71528	61388	**99.2±0.4**	**20.9±0.08**
	$\lvert\delta z\rvert$ **combined**						**97.6±0.4**	**20.9±0.08**
21		-1e+03	1e+03	68601	71528	61388	**100±0.4**	**20.9±0.08**
22	R2All	0	0.7	67969	70868	60826	**99.1±0.4**	**20.7±0.08**
23	NN	0.85	1.1	50052	50764	45687	**73.6±0.3**	**15.3±0.07**
24	Gammas9s25	0.94	inf	49899	50609	45575	**99.7±0.4**	**15.2±0.07**
25	gPi0Conv50	0.1	0.16	49885	50595	45563	**100±0.4**	**15.2±0.07**
26	gEtaConv50	0.5	0.59	49827	50537	45508	**99.9±0.4**	**15.2±0.07**
27	Empty	-1e+03	1e+03	49827	50537	45508	**100±0.4**	**15.2±0.07**
Fit Region								
28	mB1	5.2	5.3	49302	49626	45488	**98.9±0.4**	**15±0.07**
29	deltaE1	-0.3	0.3	47628	47881	44126	**96.6±0.4**	**14.5±0.07**
	TruthMatched							**13.5±0.06**
Signal Region								
30	mB2	5.3	5.3	46985	47057	43933	**98.6±0.5**	**14.3±0.07**
31	deltaE2	-0.2	0.1	44790	44840	41948	**95.3±0.5**	**13.7±0.06**
	TruthMatched							**12.8±0.06**

Table 6.2: $B^0 \to \rho^0\gamma$ signal MC efficiency table.

Omega Efficiency Calculation: *signal* (Run 1-5); N generated Events = 328000								
Step	CutName	LowVal	HighVal	Evts	Cands	TMC's	Rel Eff%	Abs Eff%
				122519	206699	84240	37.4	37.4
1	SkimCuts	0	0	122454	206475	84231	**99.9±0.3**	**37.3±0.1**
2	GammaisOK	0.5	1.5	122454	206475	84231	**100±0.3**	**37.3±0.1**
3	gnCrys	4	1e+03	122454	206475	84231	**100±0.3**	**37.3±0.1**
4	gSecMom	0	0.0019	120441	202621	83037	**98.4±0.3**	**36.7±0.1**
High Energy Photon								
5	gAcceptAngle	-0.74	0.93	118423	199718	81713	**98.3±0.3**	**36.1±0.1**
6	gdistNe	25	1e+08	111360	187317	77234	**94±0.3**	**34±0.1**
7	gdistCh	25	1e+07	109104	183360	75869	**98±0.3**	**33.3±0.1**
	Isolation						**92.1±0.3**	**33.3±0.1**
8	gPi0Veto	0.77	1.1	101031	169244	70489	**92.6±0.3**	**30.8±0.1**
9	gEtaVeto	0.94	1.1	97202	162552	67849	**96.2±0.3**	**29.6±0.1**
	π^0/η **Veto**						**89.1±0.3**	**29.6±0.1**
Charged Tracks								
10	GTL+	0	0	92971	153179	64546	**95.6±0.3**	**28.3±0.09**
11	GTL-	0	0	88407	143499	61095	**95.1±0.3**	**27±0.09**
	GoodTracksLoose						**91±0.3**	**27±0.09**
12	Pi+PID	0	0	83927	134318	58740	**94.9±0.3**	**25.6±0.09**
13	Pi-PID	0	0	79931	126406	56511	**95.2±0.3**	**24.4±0.09**
	$\pi^{\pm}$ **selector**						**90.4±0.3**	**24.4±0.09**
14	Pi0Mass	0.12	inf	79931	126406	56511	**100±0.4**	**24.4±0.09**
15	Pi0cosGAng	0.41	1.1	72479	102751	52842	**90.7±0.3**	**22.1±0.08**
ω Meson								
16	wChi2Prob	0.00025	1.1	69724	97433	51014	**96.2±0.4**	**21.3±0.08**
17	wDisplnt	0	0.19	66626	92398	48661	**95.6±0.4**	**20.3±0.08**
18	wMass	0.76	0.8	48653	55269	38173	**73±0.3**	**14.8±0.07**
19	wCosHel	0	0.75	43804	49148	34173	**90±0.4**	**13.4±0.06**
Vertexing								
20	BDzErr	0	0.04	42769	47991	33352	**97.6±0.5**	**13±0.06**
21	BDz	-1	0.4	42409	47581	33053	**99.2±0.5**	**12.9±0.06**
	$\lvert\delta z\rvert$ **combined**						**96.8±0.5**	**12.9±0.06**
22	R2All	0	0.7	42019	47155	32748	**99.1±0.5**	**12.8±0.06**
23	NN	0.9	1.1	29644	32327	23810	**70.5±0.4**	**9.04±0.05**
24	Gammas9s25	0.95	0.99	29337	31996	23574	**99±0.6**	**8.94±0.05**
25	gPi0Conv50	0.1	0.16	29331	31990	23569	**100±0.6**	**8.94±0.05**
26	gEtaConv50	0.5	0.59	29297	31952	23539	**99.9±0.6**	**8.93±0.05**
27	DalitzOmega	-1	1	29297	31952	23539	**100±0.6**	**8.93±0.05**
Fit Region								
28	mB1	5.2	5.3	28859	31178	23538	**98.5±0.6**	**8.8±0.05**
29	deltaE1	-0.3	0.3	27667	29585	22839	**95.9±0.6**	**8.44±0.05**
	TruthMatched							**6.96±0.05**
Signal Region								
30	mB2	5.3	5.3	26505	27781	22513	**95.8±0.6**	**8.08±0.05**
31	deltaE2	-0.2	0.1	24984	26027	21417	**94.3±0.6**	**7.62±0.05**
	TruthMatched							**6.53±0.04**

Table 6.3: $B^0 \to \omega\gamma$ signal MC efficiency table.

Description	Fit region		Signal region	
	raw	300fb^{-1}	raw	300fb^{-1}
$B^+ \to \rho^+\gamma$	33480	35.87±0.21	29288	31.38±0.19
$B^0 \to K^{*0}\gamma$	2432	1.33±0.03	344	0.19±0.01
$B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$	51	1.13±0.16	5	0.11±0.05
$B^+ \to K^{*+}\gamma$	3211	18.46±0.33	994	5.72±0.18
$B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$	21	0.61±0.13	6	0.17±0.07
Sum of $B \to K^*\gamma$		19.79±0.33		5.90±0.18
$B^0 \to K^0_S\eta$	54	0.09±0.01	12	0.02±0.01
$B^+ \to K^{*+}\pi^0$, $K^{*+} \to K^+\pi^0$	155	0.00±0.00	37	0.00±0.00
$B^+ \to K^{*+}\eta$, $K^{*+} \to K^+\pi^0$	21	0.14±0.03	6	0.04±0.02
$B^+ \to K^+\pi^0$	11	0.01±0.00	4	0.00±0.00
$B^+ \to K^+\eta$	0	0.00±0.00	0	0.00±0.00
$B^0 \to \rho^0\pi^0$	564	0.01±0.00	88	0.00±0.00
$B^+ \to \rho^+\pi^0$	8615	0.98±0.01	5380	0.61±0.01
$B^+ \to \pi^+\pi^0$	39	0.04±0.01	4	0.00±0.00
$B^0 \to \rho^0\eta$	107	0.04±0.00	21	0.01±0.00
$B^+ \to \eta\pi^+$	107	0.06±0.01	1	0.00±0.00
$B^+ \to \rho^+\eta$	1256	8.29±0.24	637	4.21±0.17
Sum		9.66±0.24		4.90±0.17
$B^0 \to \rho^0\gamma$	599	0.27±0.01	138	0.06±0.01
$B^0 \to \omega\gamma$	422	0.19±0.01	79	0.04±0.00
$B^0 \to X^0_{sd}\gamma$ (excl. $B^0 \to K^{*0}\gamma$)	188	45.74±3.34	39	9.49±1.52
$B^+ \to X^+_{su}\gamma$ (excl. $B^+ \to K^{*+}\gamma$)	210	51.10±3.53	49	11.92±1.70
Sum of $B \to X_s\gamma$ (excl. $B \to K^*\gamma$)		96.84±4.86		21.41±2.28
$e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ background	138	50.72±4.32	25	9.19±1.84
$e^+e^- \to \Upsilon(4S) \to B^+B^-$ background	319	120.77±6.76	118	44.67±4.11
Sum of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ background		171.49±8.02		53.86±4.50
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	1345	1575±43	90	105±11
$e^+e^- \to c\overline{c}$	798	626±22	55	43±6
$e^+e^- \to \tau^+\tau^-$	63	65±8	5	5±2
Sum of off–resonance MC		2267±49		154±13
off–resonance data	220	2425±163	9	99±33
$S/\sqrt{S+B}$		0.72		2.03

Table 6.4: Expected yield for the $B^+ \to \rho^+\gamma$ decay mode. Everything is normalized to 300fb^{-1} and all cuts are applied. For the B background, the signal has been excluded, but not the peaking B background.

Description	Fit region		Signal region	
	raw	300fb^{-1}	raw	300fb^{-1}
$B^0 \to \rho^0\gamma$	47818	21.87±0.11	44853	20.51±0.10
$B^0 \to K^{*0}\gamma$	2513	13.71±0.27	1933	10.55±0.24
$B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$	587	12.95±0.53	464	10.24±0.48
$B^+ \to K^{*+}\gamma$	840	4.83±0.17	93	0.53±0.06
$B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$	2	0.06±0.04	1	0.03±0.03
Sum of $B \to K^*\gamma$		18.54±0.32		11.08±0.25
$B^0 \to K^{*0}\eta$	18	0.20±0.05	8	0.09±0.03
$B^0 \to K^0_S\eta$	15	0.02±0.01	5	0.01±0.00
$B^+ \to K^{*+}\eta$, $K^{*+} \to K^+\pi^0$	57	0.38±0.05	14	0.09±0.02
$B^+ \to K^+\pi^0$	0	0.00±0.00	0	0.00±0.00
$B^+ \to K^+\eta$	2	0.00±0.00	1	0.00±0.00
$B^0 \to \rho^0\pi^0$	9593	0.19±0.00	7160	0.15±0.00
$B^+ \to \pi^+\pi^0$	24	0.03±0.01	2	0.00±0.00
$B^0 \to \rho^0\eta$	1543	1.18±0.03	989	0.76±0.02
$B^+ \to \eta\pi^+$	14	0.05±0.01	0	0.00±0.00
$B^+ \to \rho^+\eta$	57	0.38±0.05	14	0.09±0.02
Sum		2.43±0.09		1.19±0.05
$B^+ \to \rho^+\gamma$	351	0.38±0.02	84	0.09±0.01
$B^0 \to \omega\gamma$	1547	0.71±0.02	1305	0.60±0.02
$B^0 \to X^0_{sd}\gamma$ (excl. $B^0 \to K^{*0}\gamma$)	95	23.12±2.37	47	11.44±1.67
$B^+ \to X^+_{su}\gamma$ (excl. $B^+ \to K^{*+}\gamma$)	82	19.95±2.20	12	2.92±0.84
Sum of $B \to X_s\gamma$ (excl. $B \to K^*\gamma$)		43.07±3.24		14.36±1.87
$e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ background	90	33.08±3.49	44	16.17±2.44
$e^+e^- \to \Upsilon(4S) \to B^+B^-$ background	63	23.85±3.00	12	4.54±1.31
Sum of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ background		56.93±4.60		20.71±2.77
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	1098	1286±39	109	128±12
$e^+e^- \to c\overline{c}$	533	418±18	36	28±5
$e^+e^- \to \tau^+\tau^-$	121	125±11	7	7±3
Sum of off–resonance MC		1830±44		163±13
off–resonance data	215	2369±162	19	209±48
$S/\sqrt{S+B}$		0.50		1.43

Table 6.5: Expected yield for the $B^0 \to \rho^0\gamma$ decay mode. Everything is normalized to 300fb^{-1} and all cuts are applied. For the B background, the signal has been excluded, but not the peaking B background.

Description	Fit region		Signal region	
	raw	300fb^{-1}	raw	300fb^{-1}
$B^0 \to \omega\gamma$	26076	11.92±0.08	23441	10.72±0.07
$B^0 \to K^{*0}\gamma$	272	1.48±0.09	51	0.28±0.04
$B^+ \to K^{*+}\gamma$	295	1.70±0.10	90	0.52±0.05
Sum of $B \to K^*\gamma$		3.18±0.13		0.80±0.07
$B^+ \to \omega\pi^0$	363	0.63±0.03	263	0.46±0.03
$B^0 \to K^0_S\eta$	7	0.01±0.00	3	0.00±0.00
$B^+ \to K^{*+}\eta$, $K^{*+} \to K^+\pi^0$	0	0.00±0.00	0	0.00±0.00
$B^+ \to K^+\pi^0$	1	0.00±0.00	0	0.00±0.00
$B^+ \to K^+\eta$	0	0.00±0.00	0	0.00±0.00
$B^+ \to \pi^+\pi^0$	4	0.00±0.00	0	0.00±0.00
$B^0 \to \rho^0\eta$	14	0.01±0.00	2	0.00±0.00
$B^+ \to \eta\pi^+$	1	0.00±0.00	0	0.00±0.00
$B^+ \to \rho^+\eta$	12	0.08±0.02	2	0.01±0.01
Sum		0.73±0.00		0.48±0.00
$B^+ \to \rho^+\gamma$	100	0.11±0.01	24	0.03±0.01
$B^0 \to \rho^0\gamma$	128	0.06±0.01	24	0.01±0.00
$B^0 \to X^0_{sd}\gamma$ (excl. $B^0 \to K^{*0}\gamma$)	39	9.49±1.52	13	3.16±0.88
$B^+ \to X^+_{su}\gamma$ (excl. $B^+ \to K^{*+}\gamma$)	32	7.79±1.38	9	2.19±0.73
Sum of $B \to X_s\gamma$ (excl. $B \to K^*\gamma$)		17.28±2.05		5.35±1.14
$e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ background	21	7.72±1.68	7	2.57±0.97
$e^+e^- \to \Upsilon(4S) \to B^+B^-$ background	25	9.46±1.89	5	1.89±0.85
Sum of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ background		17.18±2.53		4.47±1.29
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	373	437±23	36	42±7
$e^+e^- \to c\overline{c}$	197	155±11	20	16±4
$e^+e^- \to \tau^+\tau^-$	17	18±4	1	1±1
Sum of off–resonance MC		609±26		59±8
off–resonance data	72	793±94	4	44±22
$S/\sqrt{S+B}$		0.47		1.24

Table 6.6: Expected yield for the $B^0 \to \omega\gamma$ decay mode. Everything is normalized to 300fb^{-1} and all cuts are applied. For the B background, the signal has been excluded, but not the peaking B background.

Chapter 7

Signal Extraction via a Maximum Likelihood Fit

In order to maximize the signal sensitivity, the final extraction of the signal yield is performed via multidimensional unbinned maximum likelihood fits described in this chapter.

7.1 Fit Overview

The final signal is extracted via a multidimensional unbinned maximum likelihood fit. There is an individual fit model for each of the three sub-decay modes $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$. Two simultaneous fits are also performed, which constrain the efficiency-corrected signal yields of the sub-decay modes based on a theoretical assumption about the individual decay widths. The first simultaneous fit is performed only on the $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ decay modes whereas the second simultaneous fit is performed on all three decay modes.

Each probability density function (PDF) is determined from a one-dimensional fit to a dedicated Monte Carlo sample. The common dimensions used in all fitters are ΔE, the photon-energy rescaled m_{ES}, the cosine of the helicity angle of the ρ (ω) meson, and a transformation of the neural network output. Additionally, the ω-fitter uses the cosine of the Dalitz angle as a fifth dimension in the fit. The transformations

of the neural network outputs are performed in order use an analytic PDF. This also mitigates the problem of a small peak of the continuum background at one in the neural network outputs.

The cuts on each of these dimensions and the resulting signal efficiencies are summarized in Table 7.1. The expected number of events in each of the three modes are listed in Table 7.2, split up into categories used by the three individual fit models.

After all the fit region cuts are applied, the best candidate per event is selected by choosing the candidate whose meson (ρ/ω) mass is closest to the nominal meson mass.

	$B^0 \to \rho^0\gamma$	$B^+ \to \rho^+\gamma$	$B^0 \to \omega\gamma$
$\lvert\Delta E\rvert$	$< 0.3\,\mathrm{GeV}$	$< 0.3\,\mathrm{GeV}$	$< 0.3\,\mathrm{GeV}$
m_{ES}	$> 5.22\,\mathrm{GeV}/c^2$	$> 5.22\,\mathrm{GeV}/c^2$	$> 5.22\,\mathrm{GeV}/c^2$
NN	> 0.850050300	> 0.800067066	> 0.900033534
$\lvert\cos\Theta_H\rvert$	< 0.75	< 0.75	< 0.75
$\lvert\cos\Theta_D\rvert$	n/a	n/a	< 1.0
Signal efficiency	$(14.5207)\,\%$	$(11.8732)\,\%$	$(8.43506)\,\%$

Table 7.1: Fit region cuts and signal efficiencies for all three decay modes.

7.2 Likelihood Function

The likelihood functions for the fits are defined for M signal and background hypotheses and for N events as:

$$\mathcal{L} = \exp\left(-\sum_{i=1}^{M} n_i\right) \cdot \left(\prod_{j=1}^{N}\left[\sum_{i=1}^{M} n_i \mathcal{P}(\vec{x}_j;\ \vec{\alpha}_i)\right]\right)$$

Given the small correlations among the fit variables for all components (see Figure 7-1 to 7-6), the PDFs for each of the M components, $\mathcal{P}\left(\vec{x}_j;\ \vec{\alpha}_i\right)$, can be described as a product of one-dimensional PDFs over the fit variables, described below.

The PDF shapes describing the B backgrounds are fixed from MC as are the PDF shapes describing most of the signal dimensions, with the exception of the PDF describing the ΔE dimension. The ΔE PDFs for all three decay modes are also

	$B^+ \to \rho^+\gamma$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$
signal (all)	35.62	21.78	12.65
signal (truth vetoed)	5.56	1.63	2.45
$B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$	4.48	incl. in other $B^+ \to K^{*+}\gamma$	incl. in other $B^+ \to K^{*+}\gamma$
other $B^0 \to K^{*0}\gamma$	9.53	9.42	incl. in other $B\overline{B}$
other $B^+ \to K^{*+}\gamma$	13.72	4.66	incl. in other $B\overline{B}$
$B \to X_s\gamma$	50.00	incl. in other $B\overline{B}$	incl. in other $B\overline{B}$
other $B\overline{B}$	72.40	40.27	21.12
$e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}$	1236	1144	423
$e^+e^- \to c\overline{c}$	563	394	159
$e^+e^- \to \tau^+\tau^-$	54	120	18

Table 7.2: Expected number of events in the fit region for all three decay modes.

fixed, but the PDF shape parameters determined from pure signal MC are corrected by differences observed between on–resonance data and signal MC in $B \to K^*\gamma$ (see Section 8.8). All parameters related to the continuum background are floating in the fits. The continuum background distribution for the transformed neural network output is cross-checked using off–resonance data.

The fit to data determines the signal yield, continuum background yields, and depending on the decay mode the overall yield of the B backgrounds.

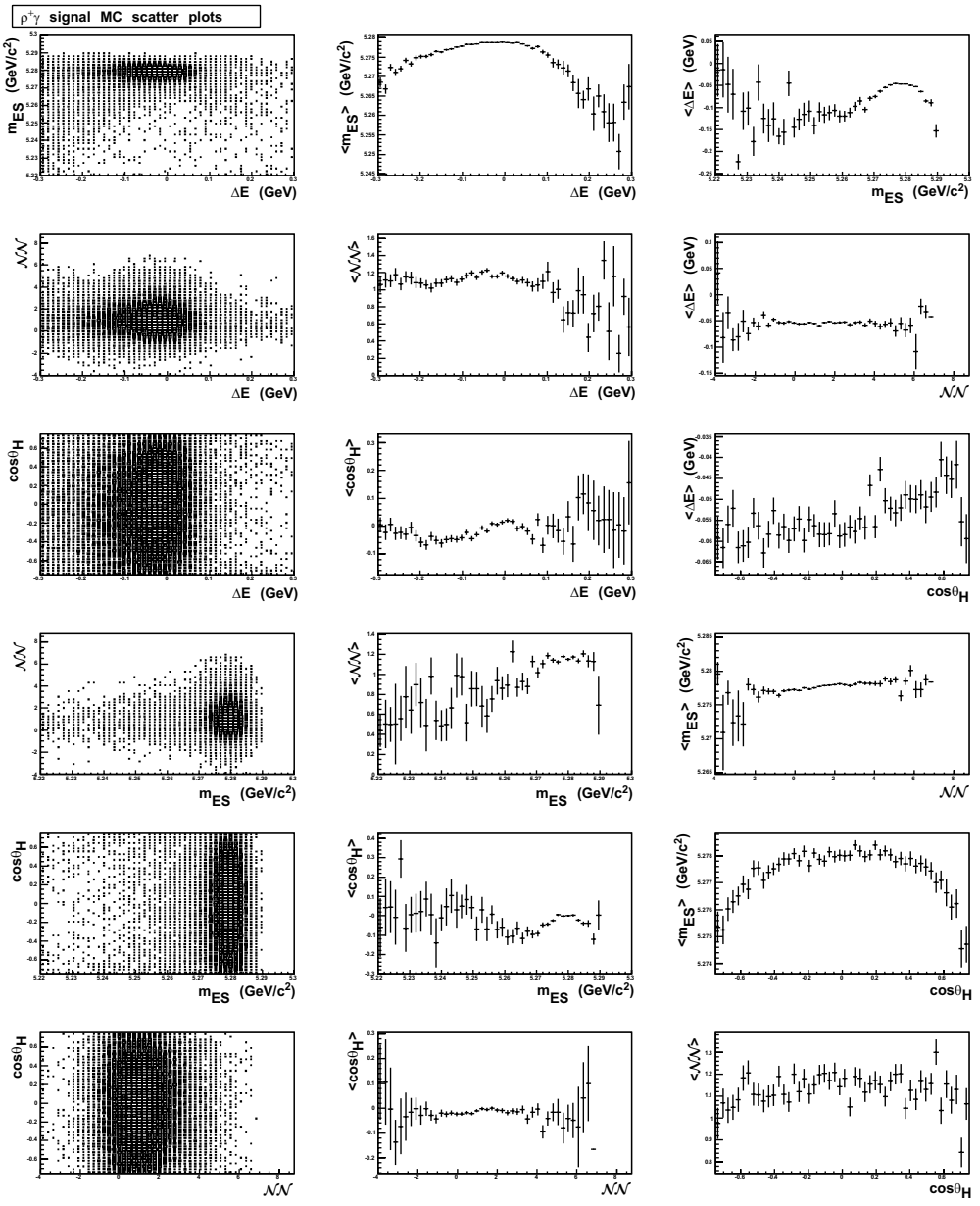


Figure 7-1: Correlations between fit dimensions for $B^+ \to \rho^+\gamma$ for signal MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

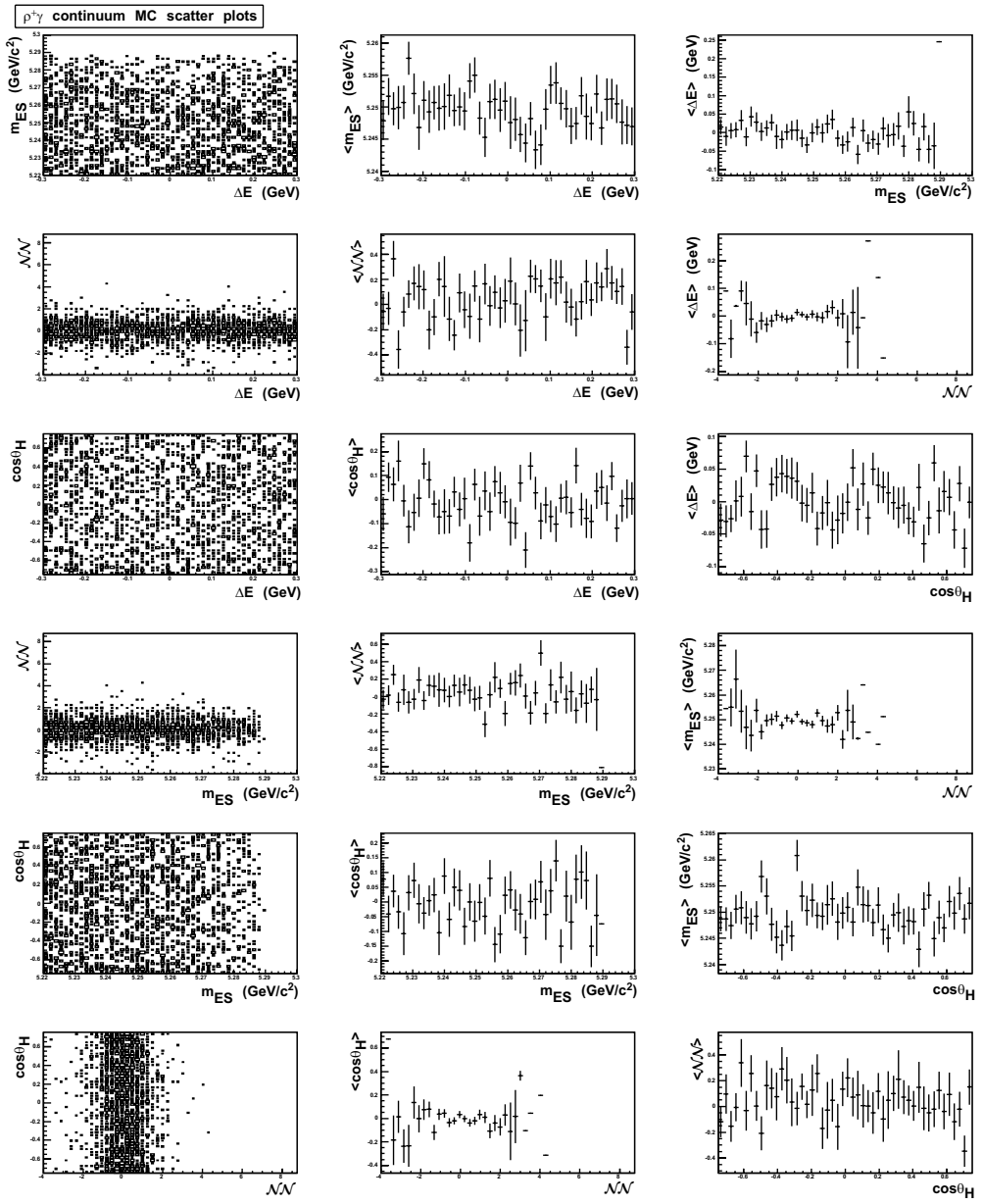


Figure 7-2: Correlations between fit dimensions for $B^+ \to \rho^+\gamma$ for continuum MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

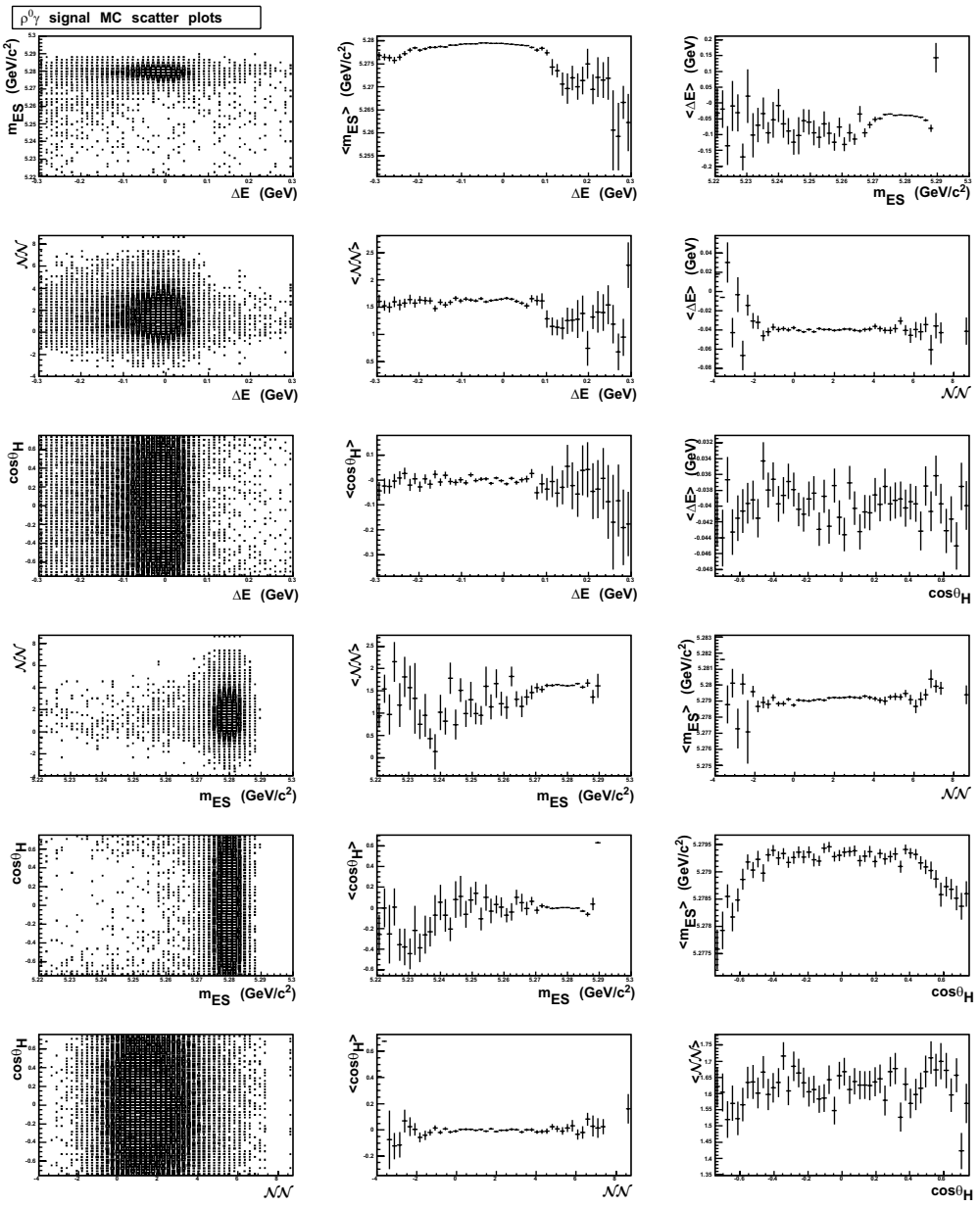


Figure 7-3: Correlations between fit dimensions for $B^0 \to \rho^0\gamma$ for signal MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

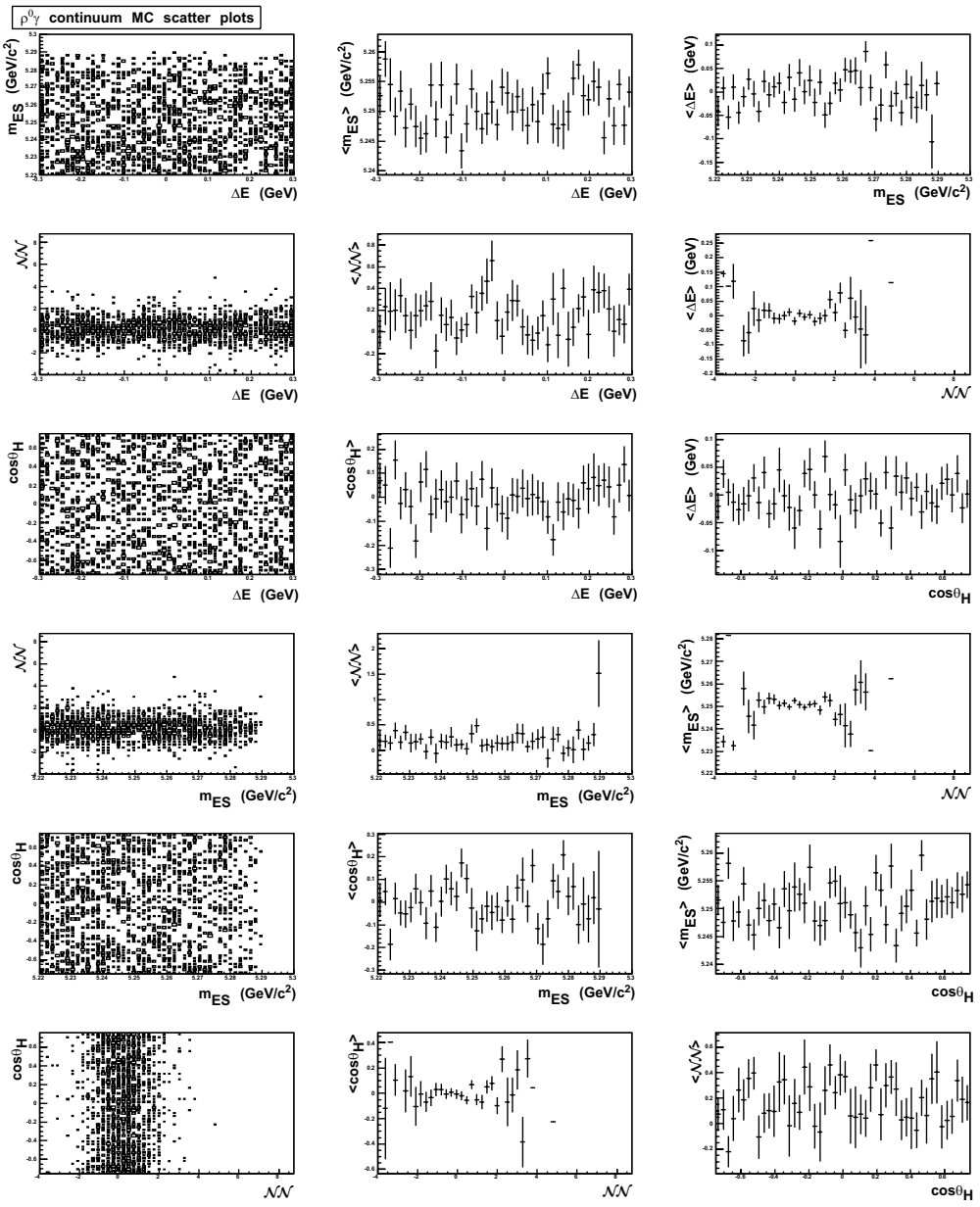


Figure 7-4: Correlations between fit dimensions for $B^0 \to \rho^0\gamma$ for continuum MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

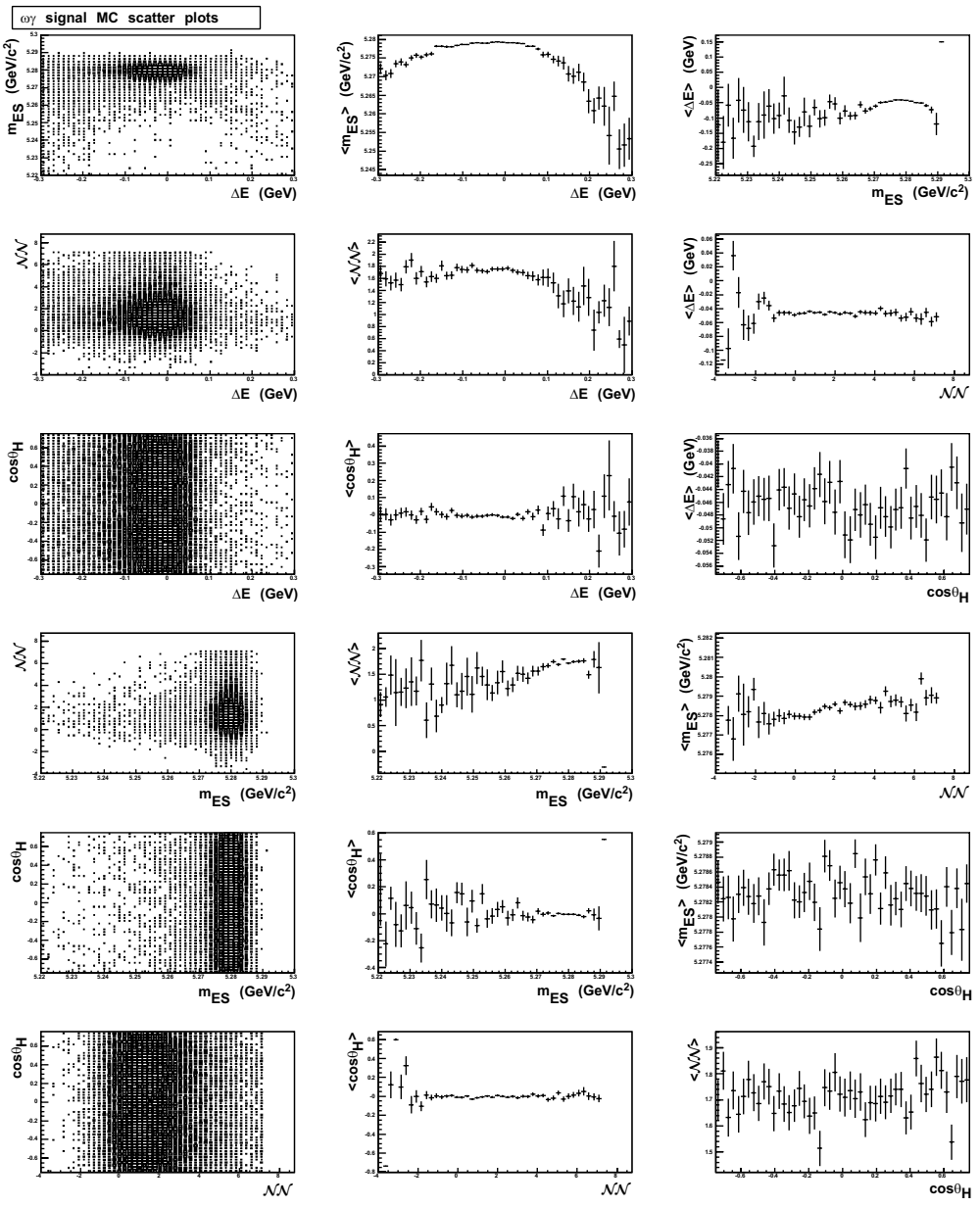


Figure 7-5: Correlations between fit dimensions for $B^0 \to \omega\gamma$ for signal MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

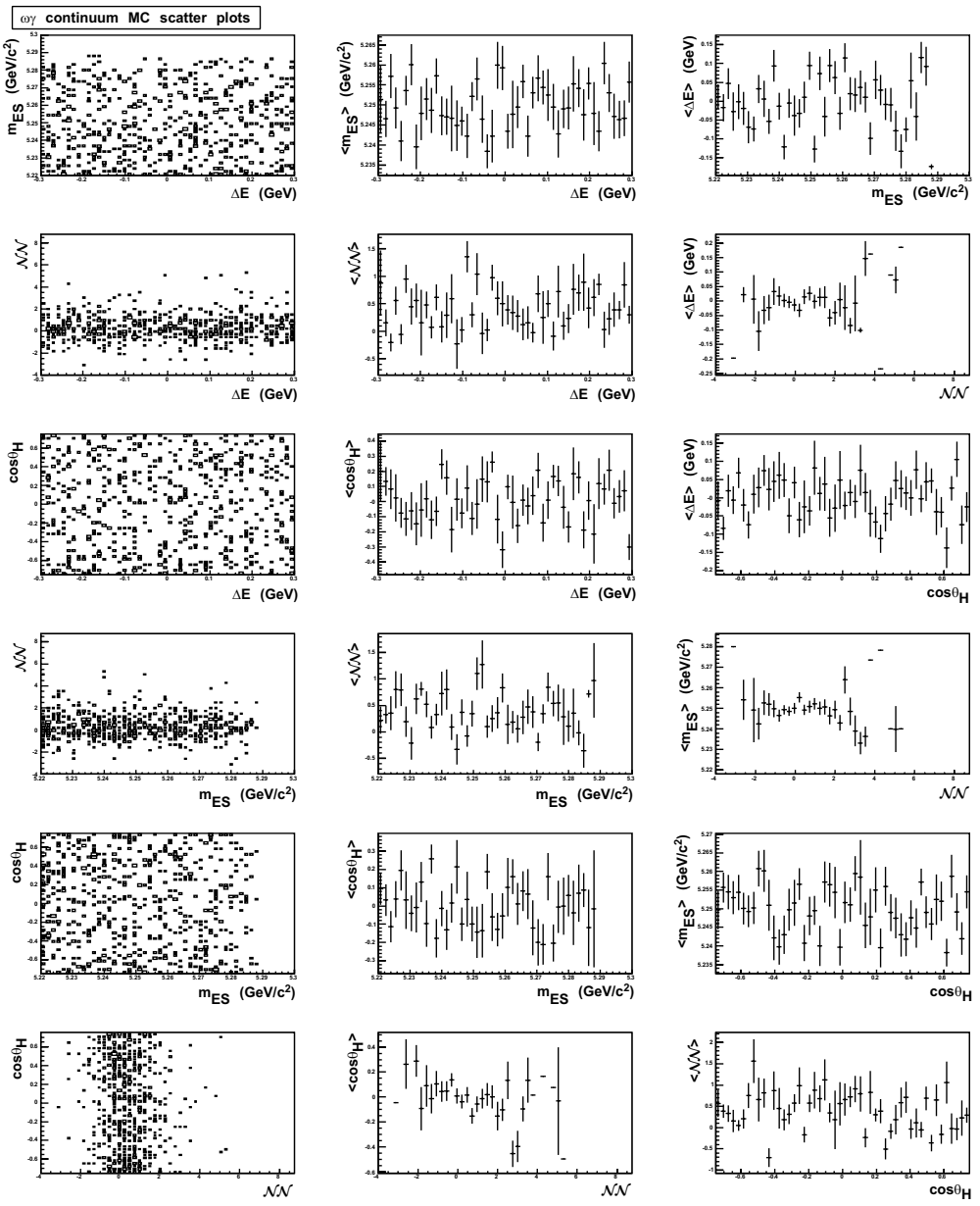


Figure 7-6: Correlations between fit dimensions for $B^0 \to \omega\gamma$ for continuum MC. All variable combinations are shown. The first column shows the scatter plots and the last two columns show the projection plots onto the X and Y axis of the scatter plot.

7.3 The $B^+ \to \rho^+\gamma$ Fit Model

The $B^+ \to \rho^+\gamma$ fit model is a fit in four dimensions and consists of six components. The transformation of the neural network output ($\mathcal{NN}$) used in this mode is (after a precut on the neural network output at 0.800067066)

$$\mathcal{NN} = \tanh^{-1}\left(\frac{(NN - 0.9)\cdot(1 - 5\cdot 10^{-8})}{0.1}\right). \tag{7.1}$$

This transformation ensures the transformed output to be limited to the range $[-4.0, 8.7522]$.

The six components of the ρ^+ fitter are:

1. signal (not requiring truth matching),
2. continuum background,
3. $B^+ \to K^{*+}\gamma$, where $K^{*+} \to K^+\pi^0$,
4. remaining $B \to K^*\gamma$ (excluding $K^{*+} \to K^+\pi^0$),
5. remaining $B \to X_s\gamma$ (excluding $B \to K^*\gamma$),
6. remaining $B\overline{B}$ background (excluding $B \to X_s\gamma$ and $B \to K^*\gamma$).

The analytical functions used for each component and each dimension are listed in Table 7.3. The analytic description of these functions are given in Appendix E.

	ΔE	m_{ES}	$\mathcal{NN}$	$\cos\Theta_H$
Signal	Cruijff	Crystal Ball	Cruijff	2^{nd}-order polynomial
Continuum	2^{nd}-order polynomial	ARGUS	Cruijff	4^{th}-order polynomial
$B\overline{B}$	Gaussian	ARGUS + Crystal Ball	Cruijff	2^{nd}-order polynomial
$B \to X_s\gamma$	Cruijff	Novosibirsk	Cruijff	2^{nd}-order polynomial
$B \to K^*\gamma$	Cruijff	Gaussian + ARGUS	Cruijff	4^{th}-order polynomial
$K^{*+} \to K^+\pi^0$	Cruijff	Crystal Ball	Cruijff	Novosibirsk

Table 7.3: PDFs used in the $B^+ \to \rho^+\gamma$ fit model.

Type of B background	other $B\overline{B}$	other $b \to s\gamma$	other $B \to K^*\gamma$	$B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$
Expected yield	$N_{B\overline{B}} = 72.40$	$N_{X_s\gamma} = 50.00$	$N_{K^*\gamma} = 23.25$	$N_{K\pi^0\gamma} = 4.48$
Total B background	$N_{tot} = N_{B\overline{B}} + N_{X_s\gamma} + N_{K^*\gamma} = 145.65$ (floating)			$N_{K\pi^0\gamma} =$ 4.48 ± 2.12 (fixed)
Fixed fractional yields (with variations)	$\frac{N_{B\overline{B}}}{N_{tot}} =$ 0.497 ± 0.058 (fixed)	$\frac{N_{X_s\gamma}}{N_{tot}} =$ 0.343 ± 0.049 (fixed)	$\frac{N_{K^*\gamma}}{N_{tot}} =$ 0.160 ± 0.033 (fixed)	$N_{K\pi^0\gamma} =$ 4.48 ± 2.12 (fixed)

Table 7.4: B background normalizations used for the $B^+ \to \rho^+\gamma$ fit model.

The $K^{*+} \to K^+\pi^0$ is separated into its own component since this background is kinematically almost the same as the signal. The only difference is the mass of one particle, the kaon, which is misidentified as a pion. Thus, the only shape differences between $B^+ \to K^{*+}\gamma$ and $B^+ \to \rho^+\gamma$ are expected to be a shift in ΔE and a skew in $\cos\Theta_H$. The yield of this component is fixed to the MC expectation of 4.5 events. It is varied by $\pm\sqrt{4.5}$ events in the on–resonance data fit and the resulting shift of the signal yield is taken as a systematic uncertainty.

The $B \to K^*\gamma$ sample has a different shape in ΔE and in the Helicity angle than the other B backgrounds. It is therefore also described by a separate component in the fit. This component and the other two B background components ($B \to X_s\gamma$ excluding $B \to K^*\gamma$, and $B\overline{B}$ excluding $B \to X_s\gamma$) have one common yield in the fit model, the sum of their three yields. The yield ratios between the three components is fixed to yield expectations derived from MC. These ratios are varied by $\pm\sqrt{N_{expected}}$ in the on–resonance data fit and the resulting shift of the signal yield is taken as a systematic uncertainty. The treatment of the individual B backgrounds, including errors on fixed yield parameters, is summarized in Table 7.4.

Figure 7-7 shows the results of the one-dimensional PDF fits to the corresponding MC samples for the $B^+ \to \rho^+\gamma$ fit model. The full statistics of each MC sample has been used and the events have been weighted individually to be normalized to 300fb^{-1}.

The shape of the transformed neural network output for the continuum background is cross-checked with off–resonance data. In order to increase the off–resonance data statistics, the ΔE window is widened to $|\Delta E| < 0.55\,\mathrm{GeV}$. The result of this cross-check can be seen in Fig. 7-8. The signal distribution is also shown for comparison.

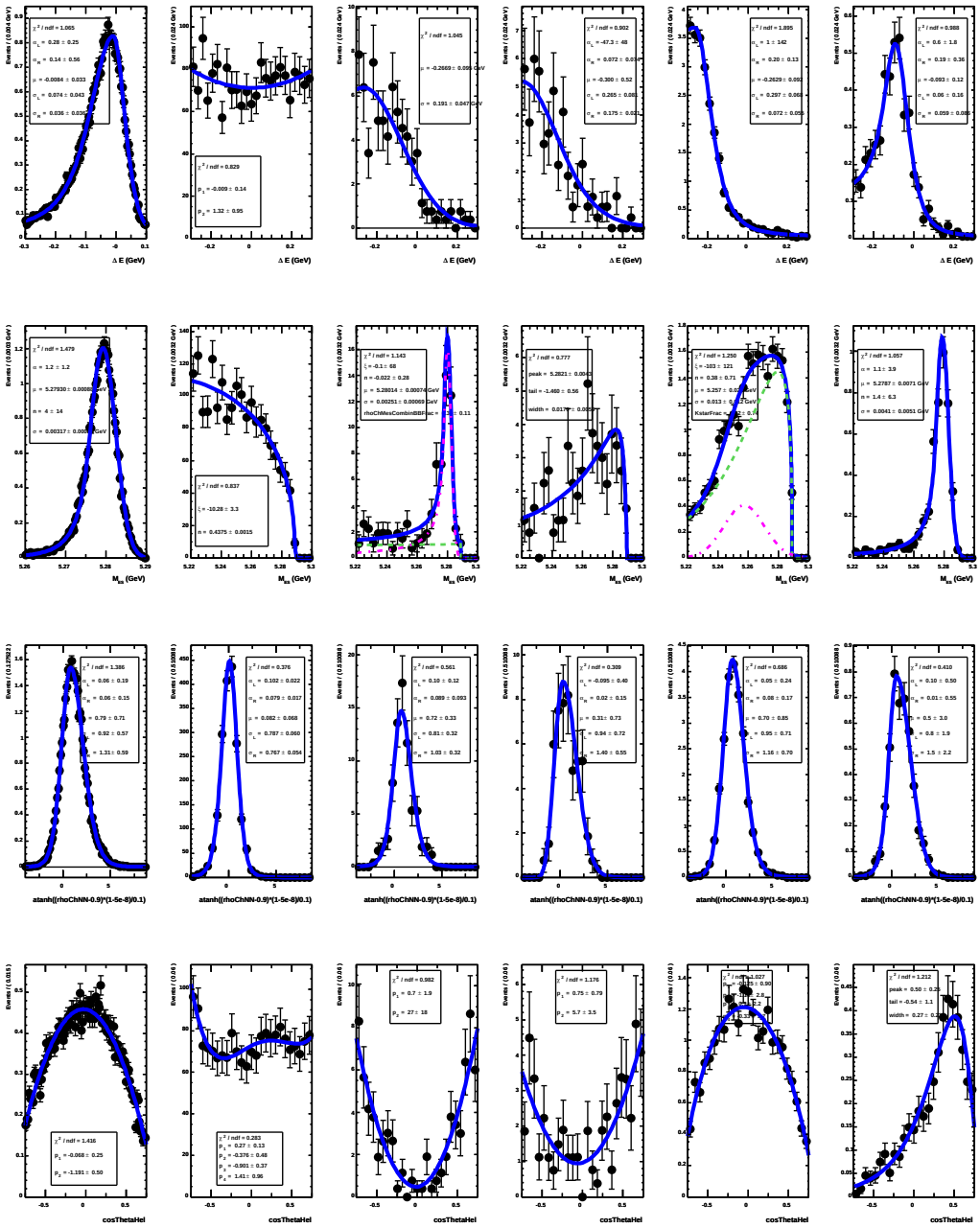


Figure 7-7: The PDFs used in the $B^+ \to \rho^+\gamma$ fit model. The dimensions are ΔE (top row), photon rescaled m_{ES} (second row), the transformed neural network output $\mathcal{NN}$ (third row) and the cosine of the ρ^+ helicity angle $\cos\Theta_H$ (last row). The six columns are depicting the six components of the fit model, signal (not requiring truth match) (first column), continuum background (second column), $B\overline{B}$ background (excluding $B \to X_s\gamma$ and $B \to K^*\gamma$) (third column), $B \to X_s\gamma$ background (excluding $B \to K^*\gamma$) (fourth column), $B \to K^*\gamma$ (excluding $K^{*+} \to K^+\pi^0$) (fifth column) and $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ peaking background (last column).

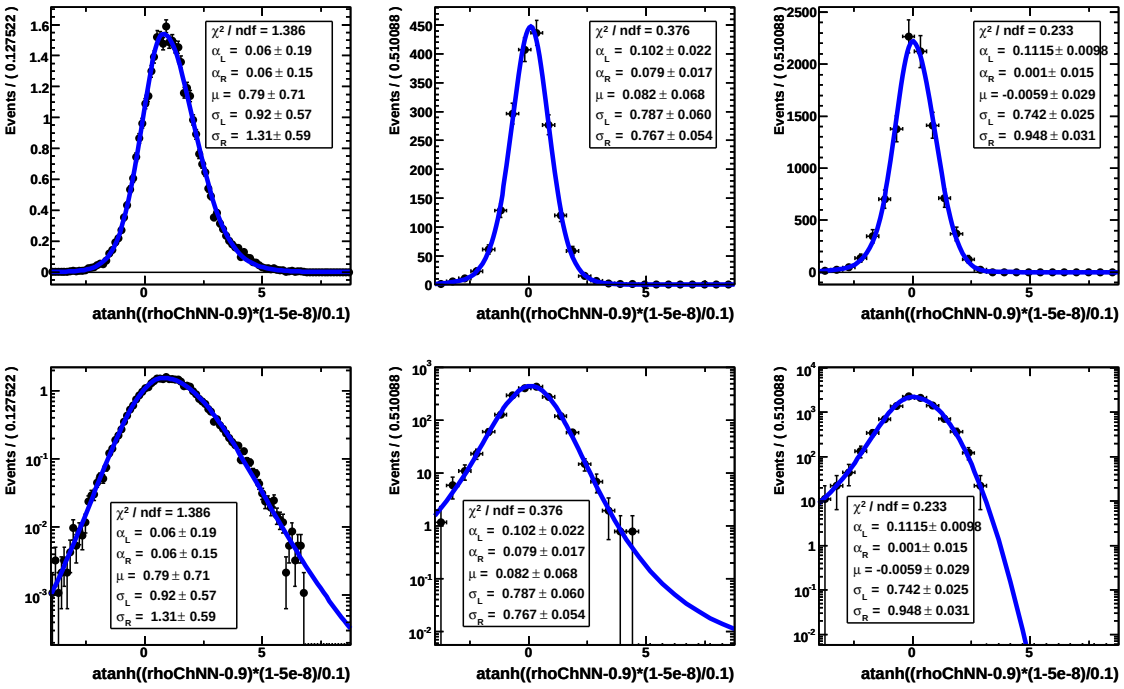


Figure 7-8: Comparison of transformed neural network shapes for the $B^+ \to \rho^+\gamma$ fit model. Shown are the shapes of the transformed neural network output for $B^+ \to \rho^+\gamma$ signal MC (left column), continuum MC (middle column) and off–resonance data (right column). In order to increase the statistics for the off–resonance data, the ΔE cut is widened for this sample to $|\Delta E| < 0.55\,\mathrm{GeV}$. The bottom row shows the same plots as the top row, only the ordinate in the bottom row has a logarithmic scale.

7.4 The $B^0 \to \rho^0\gamma$ Fit Model

The $B^0 \to \rho^0\gamma$ fit model consists of a fit in four dimensions using five components. The transformation of the neural network output used in this mode is (after a precut on the neural network output at 0.850050300)

$$\mathcal{NN} = \tanh^{-1}\left(\frac{(NN - 0.925)\cdot(1 - 5\cdot 10^{-8})}{0.075}\right). \tag{7.2}$$

This transformation ensures the transformed output to be limited to the range $[-4.0, 8.7522]$.

The five components of the ρ^0 fitter are:

1. signal (not requiring truth matching)
2. continuum background
3. $B\overline{B}$ background (excluding $B \to K^*\gamma$)
4. $B^0 \to K^{*0}\gamma$
5. $B^+ \to K^{*+}\gamma$

The analytical functions used for each component and each dimension are listed in Table 7.5. The analytic description of these functions are given in Appendix E.

	ΔE	m_{ES}	$\mathcal{NN}$	$\cos\Theta_H$
Signal	Cruijff	Crystal Ball	Cruijff	2^{nd}-order polynomial
Continuum	2^{nd}-order polynomial	ARGUS	Cruijff	2^{nd}-order polynomial
$B\overline{B}$	Cruijff	ARGUS + Crystal Ball	Cruijff	2^{nd}-order polynomial
$B^0 \to K^{*0}\gamma$	Cruijff	Crystal Ball	Cruijff	2^{nd}-order polynomial
$B^+ \to K^{*+}\gamma$	Cruijff	ARGUS + Gaussian	Cruijff + Gaussian	2^{nd}-order polynomial

Table 7.5: PDFs used in the $B^0 \to \rho^0\gamma$ fit model.

Similar to the $B^+ \to \rho^+\gamma$ fit model, the $B^0 \to K^{*0}\gamma$ background is treated with its own component. $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ are kinematically almost identical.

Type of B background	other $B\overline{B}$	$B^0 \to K^{*0}\gamma$	$B^+ \to K^{*+}\gamma$
Expected yield	$N_{B\overline{B}} = 40.27$	$N_{K^{*0}\gamma} = 9.42$	$N_{K^{*+}\gamma} = 4.66$
Total B background	$N_{tot} = N_{B\overline{B}} + N_{K^{*0}\gamma} + N_{K^{*+}\gamma} = 54.35$ (floating)		
Fixed fractional yields (with variations)	$\frac{N_{B\overline{B}}}{N_{tot}} = 0.741 \pm 0.117$ (fixed)	$\frac{N_{K^{*0}\gamma}}{N_{tot}} = 0.173 \pm 0.0564$ (fixed)	$\frac{N_{K^{*+}\gamma}}{N_{tot}} = 0.0857 \pm 0.0397$ (fixed)

Table 7.6: B background normalizations used for the $B^0 \to \rho^0\gamma$ fit model.

The only difference is the mass of the kaon in the $B^0 \to K^{*0}\gamma$ decay. In order to reconstruct a $B^0 \to K^{*0}\gamma$ event as signal, the kaon has to be misidentified as a a pion. Thus, the only shape difference between $B^0 \to K^{*0}\gamma$ and $B^0 \to \rho^0\gamma$ is expected to be a shift in ΔE.

The $B^+ \to K^{*+}\gamma$ sample has a different shape in ΔE and in the Helicity angle than the other B backgrounds. It is therefore also described by a separate component in the fit. This component, the $B^0 \to K^{*0}\gamma$ component, and the other B background component have one common yield in the fit model. The ratios between the three components is fixed to expected values as determined from MC. These ratios are varied by $\pm\sqrt{N_{expected}}$ in the on–resonance data fit and the resulting shift of the signal yield is taken as a systematic uncertainty. The treatment of the individual B backgrounds, including errors on fixed yield parameters, is summarized in Table 7.6.

Figure 7-9 shows the results of the one-dimensional PDF fits to the corresponding MC samples for the $B^0 \to \rho^0\gamma$ fit model. The full statistics of each MC sample has been used and the events have been weighted individually to be normalized to 300fb^{-1}.

The shape of the transformed neural network output for the continuum background is cross-checked with off–resonance data. In order to increase the off–resonance data statistics, the ΔE window is widened to $|\Delta E| < 0.55\,\mathrm{GeV}$. The result can be seen in Fig. 7-10. The signal distribution is also shown for comparison.

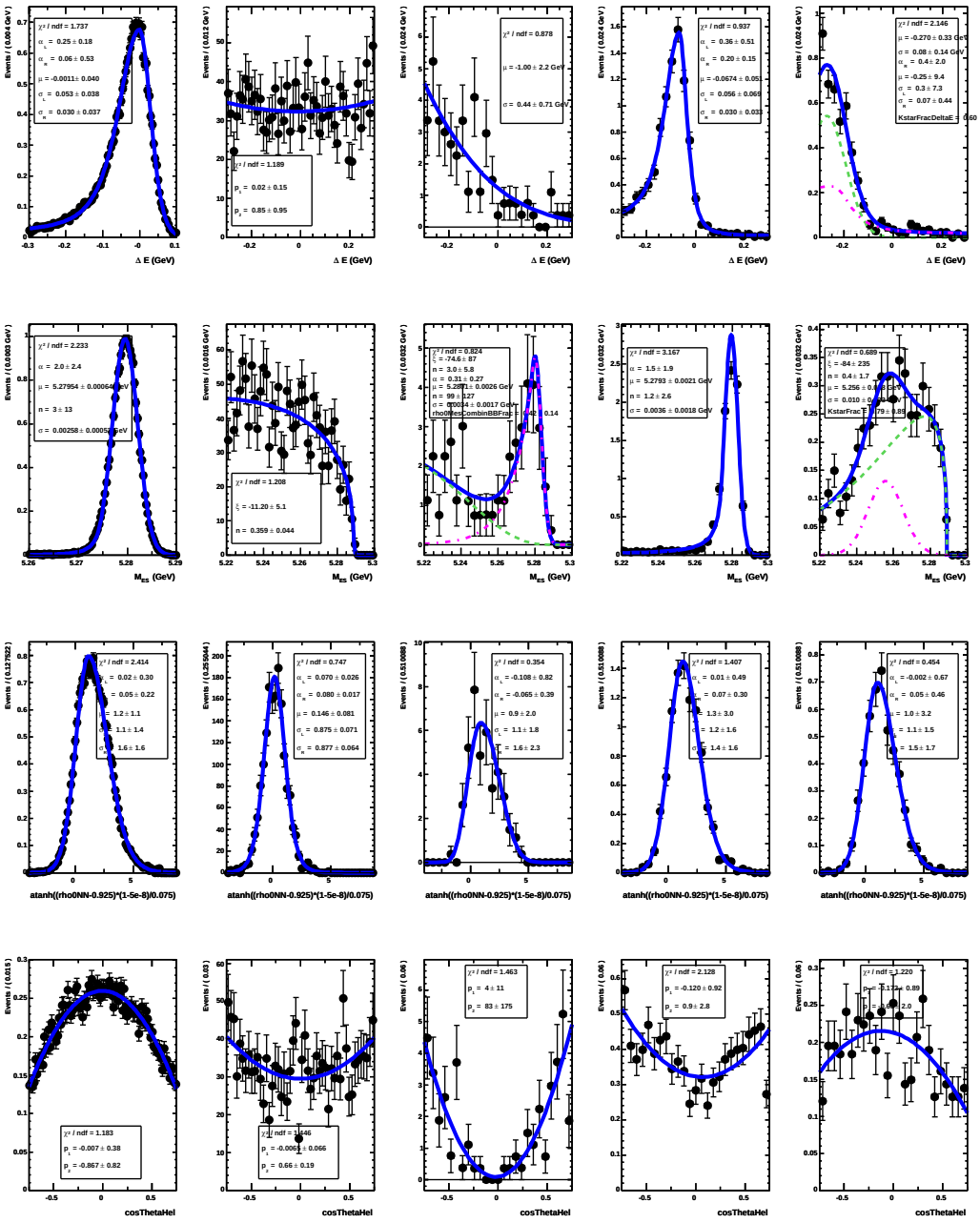


Figure 7-9: The PDFs used in the $B^0 \to \rho^0\gamma$ fit model. The dimensions are ΔE (top row), photon rescaled m_{ES} (second row), the transformed neural network output $\mathcal{NN}$ (third row) and the cosine of the ρ^0 helicity angle $\cos\Theta_H$ (last row). The five columns are depicting the five components of the fit model, signal (not requiring truth match) (first column), continuum background (second column), $B\overline{B}$ background (excluding $B \to K^*\gamma$) (third column), $B^0 \to K^{*0}\gamma$ peaking background (fourth column) and $B^+ \to K^{*+}\gamma$ background (last column).

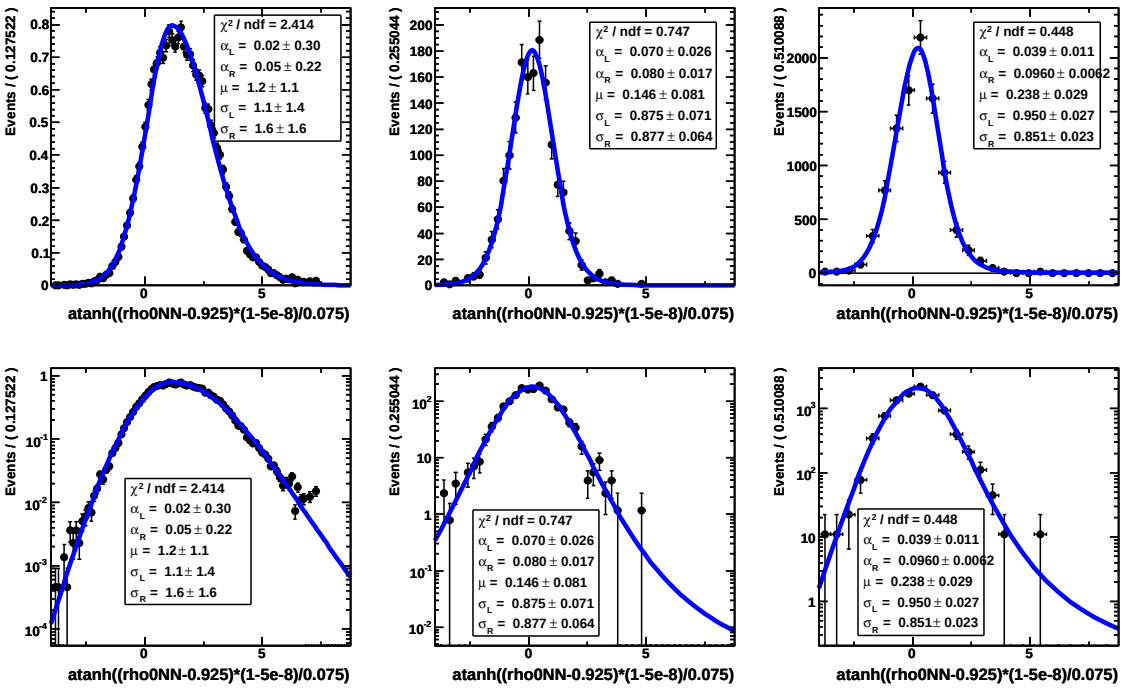


Figure 7-10: Comparison of transformed neural network shapes for the $B^0 \to \rho^0\gamma$ fit model. Shown are the shapes of the transformed neural network output $\mathcal{NN}$ for $B^0 \to \rho^0\gamma$ signal MC (left column), continuum MC (middle column) and off–resonance data (right column). In order to increase the statistics for the off–resonance data, the ΔE cut is widened for this sample to $|\Delta E| < 0.55\,\mathrm{GeV}$. The bottom row shows the same plots as the top row, only the ordinate in the bottom row has a logarithmic scale.

7.5 The $B^0 \to \omega\gamma$ Fit Model

The $B^0 \to \omega\gamma$ fit model consists of a fit in five dimensions utilizing three components. The transformation of the neural network output used in this mode is (after a precut on the neural network output at 0.900033534)

$$\mathcal{NN} = \tanh^{-1}\left(\frac{(NN - 0.95)\cdot(1 - 5\cdot 10^{-8})}{0.05}\right). \tag{7.3}$$

This transformation ensures the transformed output to be limited to the range $[-4.0, 8.7522]$.

The $B^0 \to \omega\gamma$ fit model utilizes one more dimension than the other two fitters, the cosine of the Dalitz angle. Also, the absolute value of the cosine of the ω helicity angle is used instead of the cosine of the ω helicity angle. This fit model utilizes three components:

1. signal (not requiring truth matching)
2. continuum background
3. $B\overline{B}$ background .

The analytical functions used for each component and each dimension are listed in Table 7.7. The analytic description of these functions are given in Appendix E.

	signal	continuum	$B\overline{B}$
ΔE	Cruijff	2^{nd}-order polynomial	Gaussian
m_{ES}	Crystal Ball	ARGUS	Novosibirsk
$\mathcal{NN}$	Cruijff	Cruijff	Cruijff
$\lvert\cos\Theta_H\rvert$	2^{nd}-order polynomial	2^{nd}-order polynomial	2^{nd}-order polynomial
$\cos\Theta_D$	2^{nd}-order polynomial	2^{nd}-order polynomial	2^{nd}-order polynomial

Table 7.7: PDFs used in the $B^0 \to \omega\gamma$ fit model.

The $B \to X_s\gamma$ and $B \to K^*\gamma$ background events are included in the $B\overline{B}$ background component. The expected yield for $B \to K^*\gamma$ is small though with an expected yield of less than 4 events that also do not peak in m_{ES} or ΔE. The yield of the $B\overline{B}$ background is expected to be small compared to the expected continuum

background yield. But it is significant enough such that the $B\overline{B}$ yield can be floated freely.

Figure 7-11 shows the results of the one-dimensional PDF fits to the corresponding MC samples for the $B^0 \to \omega\gamma$ fit model. The full statistics of each MC sample has been used and the events have been weighted individually to be normalized to 300fb^{-1}.

The shape of the transformed neural network output for the continuum background is cross-checked with off–resonance data. In order to increase the off–resonance data statistics, the ΔE window is widened to $|\Delta E| < 0.55\,\mathrm{GeV}$. The result can be seen in Fig. 7-12. The signal distribution is also shown for comparison.

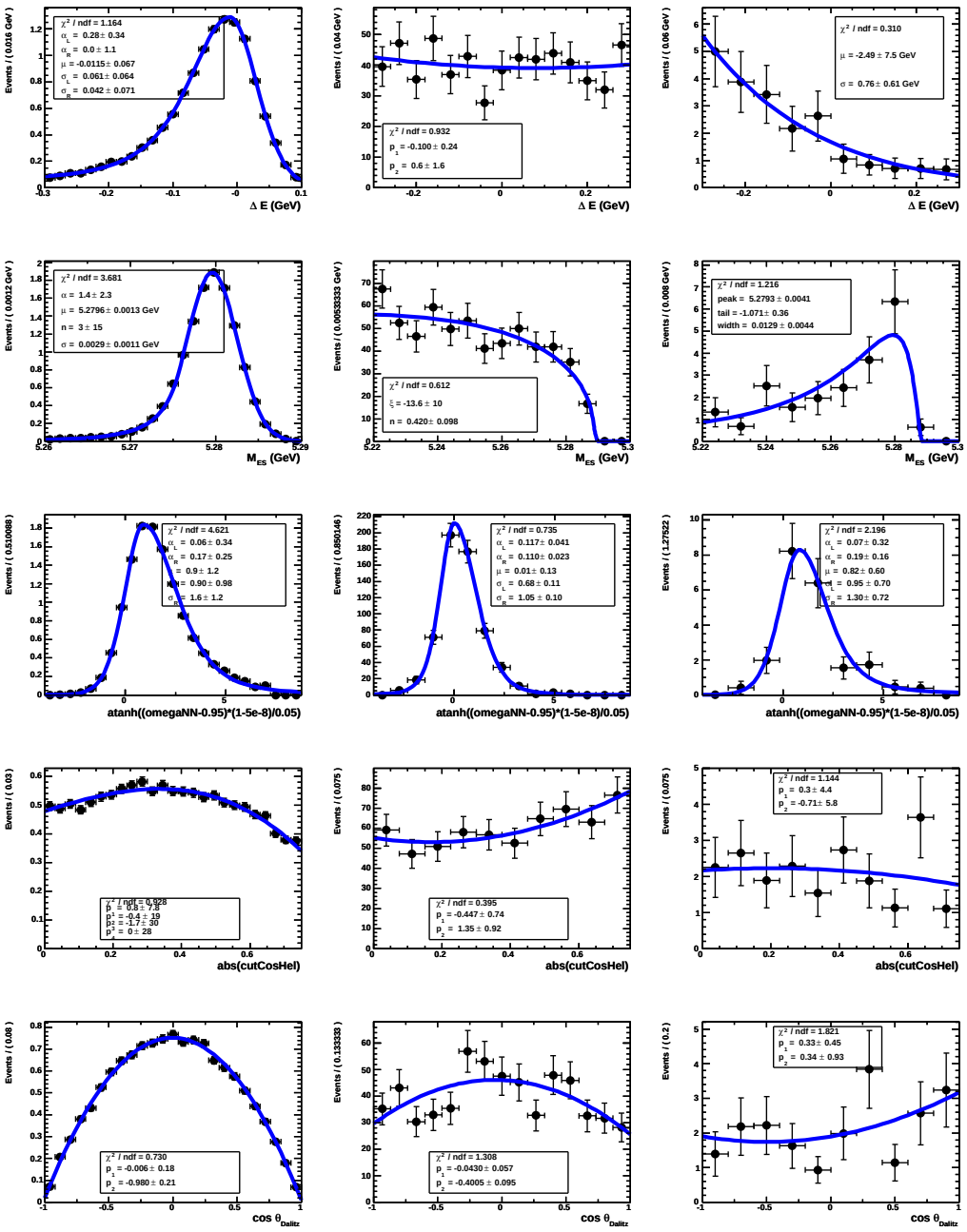


Figure 7-11: The PDFs used in the $B^0 \to \omega\gamma$ fit model. The dimensions are ΔE (top row), photon rescaled m_{ES} (second row), the transformed neural network output $\mathcal{NN}$ (third row), the absolute value of the cosine of the ω helicity $|\cos\Theta_D|$ (fourth row) and the cosine of the Dalitz angle $\cos\Theta_D$ (last row). The three columns are depicting the three components of the fitter, signal (not requiring truth match) (first column), continuum background (second column) and $B\overline{B}$ background (last column).

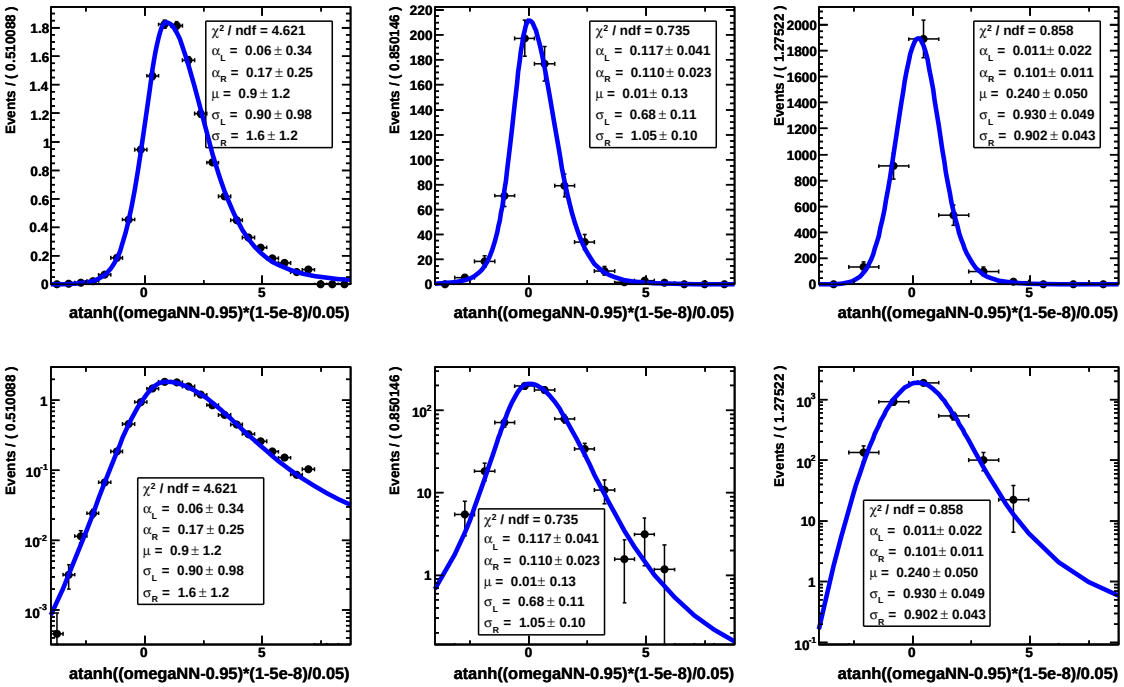


Figure 7-12: Comparison of transformed neural network shapes for the $B^0 \to \omega\gamma$ fit model. Shown are the shapes of the transformed neural network output $\mathcal{NN}$ for $B^0 \to \omega\gamma$ signal MC (left column), continuum Monte Carlo (middle column) and off–resonance data (right column). In order to increase the statistics for the off–resonance data, the ΔE cut is widened for this sample to $|\Delta E| < 0.55\,\mathrm{GeV}$. The bottom row shows the same plots as the top row, only the ordinate in the bottom row has a logarithmic scale.

7.6 The Simultaneous Fit Models

In this section, the individual decay modes are combined together in two different ways. One way is to just combine all three decay modes $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ into one unified fit model. The other way is to only combine the two ρ modes without including the ω mode. This is preferred because the ω meson is not a member of the ρ isospin triplet.

The combination of all three decay modes is done by considering the relations among them predicted by the quark model:

$$\Gamma(B^0 \to \rho^0\gamma) = \Gamma(B^0 \to \omega\gamma) = 0.5 \cdot \Gamma(B^+ \to \rho^+\gamma) \tag{7.4}$$

Following [39], the combined branching fraction is defined as

$$\mathcal{B}(B \to (\rho/\omega)\,\gamma) = \frac{1}{2} \cdot \left(\mathcal{B}(B^+ \to \rho^+\gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot \left[\mathcal{B}(B^0 \to \rho^0\gamma) + \mathcal{B}(B^0 \to \omega\gamma) \right] \right), \tag{7.5}$$

with the B meson lifetimes being $\tau_{B^+} = (1.643 \pm 0.010)$ ps and $\tau_{B^0} = (1.528 \pm 0.009)$ ps [22].

One way to combine the individual results is a least χ^2 average. The correlations of the errors amongst the three decay modes complicates the determination of the combined error. It is easier to combine the three decay modes by fitting them simultaneously. With a simultaneous fit model, the determination of the systematic errors is straightforward. Also, the simultaneous fit model has two degrees of freedom less than the sum of the three individual fit models due to the theory constraint that links to three signal yields to one (see Equation 7.5).

The"effective signal yield" for each decay mode is defined as "$N_{\text{eff}} = N_B \cdot \mathcal{B}$" where N_B is the number of B mesons and $\mathcal{B}$ is the branching fraction. It is related to the signal yield by $N_{\text{eff}} = N_{\text{sig}}/\epsilon$, where ϵ is the signal efficiency. With the combined branching fraction defined above, the simultaneous effective signal yield is defined as

follows:

$$\begin{aligned} N_{\text{eff}}^{sim} &= \tfrac{1}{2} \cdot \Big(N_{\text{eff}}(B^+ \to \rho^+\gamma) + \cdot \tfrac{\tau_{B^+}}{\tau_{B^0}} \cdot [N_{\text{eff}}(B^0 \to \rho^0\gamma) + N_{\text{eff}}(B^0 \to \omega\gamma)] \Big) \cdot \tfrac{N_B}{N_{B^0}} \\ &= \tfrac{N_{sig}(B^+ \to \rho^+\gamma)}{\epsilon(B^+ \to \rho^+\gamma)} + \cdot \tfrac{\tau_{B^+}}{\tau_{B^0}} \cdot \left[\tfrac{N_{sig}(B^0 \to \rho^0\gamma)}{\epsilon(B^0 \to \rho^0\gamma)} + \tfrac{N_{sig}(B^0 \to \omega\gamma)}{\epsilon(B^0 \to \omega\gamma)} \right]. \end{aligned} \tag{7.6}$$

The simultaneous fit model determines the combined effective signal yield directly in the likelihood fit. Each component of this fit model is identical to the ones of the corresponding individual fit models described in Section 7.3, Section 7.4 and Section 7.5. The only difference is that the signal yield for each mode is given as a function of the combined effective signal yield

$$\begin{aligned} N_{sig}\left(B^+ \to \rho^+\gamma\right) &= \tfrac{1}{2} \cdot N_{\text{eff}}^{sim} \cdot \epsilon\left(B^+ \to \rho^+\gamma\right) \\ N_{sig}\left(B^0 \to \rho^0\gamma\right) &= \tfrac{1}{4} \cdot \tfrac{\tau_{B^0}}{\tau_{B^+}} \cdot N_{\text{eff}}^{sim} \cdot \epsilon\left(B^0 \to \rho^0\gamma\right) \\ N_{sig}\left(B^0 \to \omega\gamma\right) &= \tfrac{1}{4} \cdot \tfrac{\tau_{B^0}}{\tau_{B^+}} \cdot N_{\text{eff}}^{sim} \cdot \epsilon\left(B^0 \to \omega\gamma\right). \end{aligned} \tag{7.7}$$

Here, ϵ is the corresponding signal efficiency in each decay mode, including systematic efficiency corrections related to the neural network and the π^0-reconstruction (see Sections 8.6 and 8.4). Instead of three individual signal yields floating, due to the theory constraint, only one signal yield parameter is floating, the effective signal yield N_{eff}^{sim}.

From this, Equation 7.5 can be deduced in the following way. First, note that the individual branching fractions are:

$$\begin{aligned} \mathcal{B}(B^+ \to \rho^+\gamma) &= \tfrac{N_{sig}\left(B^+ \to \rho^+\gamma\right)}{N_{B^+} \cdot \epsilon(B^+ \to \rho^+\gamma)} \\ \mathcal{B}(B^0 \to \rho^0\gamma) &= \tfrac{N_{sig}\left(B^0 \to \rho^0\gamma\right)}{N_{B^0} \cdot \epsilon(B^0 \to \rho^0\gamma)} \\ \mathcal{B}(B^0 \to \omega\gamma) &= \tfrac{N_{sig}\left(B^0 \to \omega\gamma\right)}{N_{B^0} \cdot \epsilon(B^0 \to \omega\gamma)}. \end{aligned} \tag{7.8}$$

Then, we solve each equation for $N_{\rm eff}^{sim}$:

$$\begin{aligned}
\tfrac{1}{2}\cdot N_{\rm eff}^{sim} &= \tfrac{N_{sig}\left(B^+\to\rho^+\gamma\right)}{\epsilon(B^+\to\rho^+\gamma)} \\
&= N_{B^+}\cdot\mathcal{B}(B^+\to\rho^+\gamma) \\
\tfrac{1}{4}\cdot N_{\rm eff}^{sim} &= \tfrac{N_{sig}\left(B^0\to\rho^0\gamma\right)}{\epsilon(B^0\to\rho^0\gamma)}\cdot\tfrac{\tau_{B^+}}{\tau_{B^0}} \\
&= N_{B^0}\cdot\mathcal{B}(B^0\to\rho^0\gamma)\cdot\tfrac{\tau_{B^+}}{\tau_{B^0}} \\
\tfrac{1}{4}\cdot N_{\rm eff}^{sim} &= \tfrac{N_{sig}\left(B^0\to\omega\gamma\right)}{\epsilon(B^0\to\omega\gamma)}\cdot\tfrac{\tau_{B^+}}{\tau_{B^0}} \\
&= N_{B^0}\cdot\mathcal{B}(B^0\to\omega\gamma)\cdot\tfrac{\tau_{B^+}}{\tau_{B^0}}.
\end{aligned} \tag{7.9}$$

Solving this in terms of $N_{\rm eff}^{sim}$ yields:

$$\begin{aligned}
\mathcal{B}(B^+\to\rho^+\gamma) &= \tfrac{1}{2}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^+}} \\
\mathcal{B}(B^0\to\rho^0\gamma) &= \tfrac{1}{4}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^0}}\cdot\tfrac{\tau_{B^0}}{\tau_{B^+}} \\
\mathcal{B}(B^0\to\omega\gamma) &= \tfrac{1}{4}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^0}}\cdot\tfrac{\tau_{B^0}}{\tau_{B^+}}.
\end{aligned} \tag{7.10}$$

Now, we can put these back into Equation 7.5 to get:

$$\begin{aligned}
\mathcal{B}(B\to(\rho,\omega)\gamma) &= \tfrac{1}{2}\cdot\left(\mathcal{B}(B^+\to\rho^+\gamma)+\tfrac{\tau_{B^+}}{\tau_{B^0}}\cdot[\mathcal{B}(B^0\to\rho^0\gamma)+\mathcal{B}(B^0\to\omega\gamma)]\right) \\
&= \tfrac{1}{2}\cdot\left(\tfrac{1}{2}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^+}}+\tfrac{\tau_{B^+}}{\tau_{B^0}}\cdot\left[\tfrac{1}{4}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^0}}\cdot\tfrac{\tau_{B^0}}{\tau_{B^+}}+\tfrac{1}{4}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^0}}\cdot\tfrac{\tau_{B^0}}{\tau_{B^+}}\right]\right) \\
&= \tfrac{1}{2}\cdot\left(\tfrac{1}{2}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^+}}+\tfrac{1}{2}\cdot\tfrac{N_{\rm eff}^{sim}}{N_{B^0}}\right).
\end{aligned} \tag{7.11}$$

Now, using:

$$N_{B^0}=N_{B^+}=\frac{1}{2}\cdot N_B, \tag{7.12}$$

this yields:

$$\begin{aligned}
\mathcal{B}(B\to(\rho,\omega)\gamma) &= \tfrac{1}{2}\cdot\left(\tfrac{N_{\rm eff}^{sim}}{N_B}+\tfrac{N_{\rm eff}^{sim}}{N_B}\right) \\
&= \tfrac{N_{\rm eff}^{sim}}{N_B},
\end{aligned} \tag{7.13}$$

which is just our earlier definition of the effective signal yield $N_{\rm eff}^{sim}$.

As mentioned above, also the combination amongst the two ρ modes is performed. The $B^0\to\omega\gamma$ mode is excluded in this average because the theoretical interpretation with all three modes is not straightforward and disputed amongst theoreticians. On

the contrary, the combination of the two decays $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ into a single simultaneous fit is much cleaner since the two decays are related by isospin. However, different dynamics have to be considered in the theoretical framework. For example, the W^+-annihilation diagram plays a non-zero role in the $B^+ \to \rho^+\gamma$ decay mode whereas the W^+-exchange diagram plays a non-zero role in the $B^0 \to \rho^0\gamma$ decay. Thus, isospin is not strictly conserved, but isospin violations are expected to be small [9].

The simultaneous fit model used for the two ρ modes is based on the assumption

$$\mathcal{B}(B \to \rho\gamma) = \frac{1}{2} \cdot \left(\mathcal{B}(B^+ \to \rho^+\gamma) + 2 \cdot \frac{\tau_{B^+}}{\tau_{B^0}} \cdot \mathcal{B}(B^0 \to \rho^0\gamma) \right) \tag{7.14}$$

where the B meson lifetimes are $\tau_{B^+} = 1.643 \pm 0.010$ and $\tau_{B^0} = 1.528 \pm 0.009$ [22].

7.7 "Toy MC" Studies

In order to validate the three individual fit models and the simultaneous fit model toy MC studies are performed. Toy MC studies test the self consistency of the fit model and how sensitive the fit parameters are *w.r.t.* statistical fluctuations in the data sample.

A toy MC study consists of a series of toy MC experiments. For each toy MC experiment, one generates the expected number of events for each category (allowing for Poisson fluctuations) of the fit model under consideration according to the PDF line shapes of the fit model. For this analysis, 500 of these toy MC experiments have been generated for each of the three individual fit models and the simultaneous fit model. The background events are generated according to the corresponding background PDF shapes. The signal events are sampled (allowing for Poisson fluctuations) from the fully simulated `GEANT4` signal MC samples. This allows for also testing the validity of the PDF choice for the signal component of the fit model since this PDF is not used for generating the signal events, but only for fitting them. A shift in the fit results would indicate a problem with the choice of the signal PDF.

The fully simulated signal MC samples are in fact large enough to allow for a sufficient number of independent toy MC experiments. To be precise, the number of independent experiments for each decay mode (scaled to 300fb^{-1}) are:

1. $B^+ \to \rho^+\gamma$: 933 experiments,
2. $B^0 \to \rho^0\gamma$: 2187 experiments,
3. $B^0 \to \omega\gamma$: 2187 experiments.

Besides ensuring that there is no intrinsic fit bias, signal embedded toy MC studies also test the choice of the signal PDF parametrization by not relying on the signal PDF shapes for the generation of the toy MC experiments.

Each toy MC experiment is fit with the corresponding nominal fit model. The parameters resulting from the fits are compared to the parameters used for generating the experiments. Especially interesting are the yield parameters and most important

the signal yield parameter. For a simple comparison of generated and fitted parameter values, a pull is calculated as

$$\text{pull}\,(x) = \frac{x_{\text{fit}} - x_{\text{generated}}}{\sigma_{x_{\text{fit}}}} \tag{7.15}$$

where $x_{\text{generated}}$ is the value of a certain parameter that was used to generate the data set for that specific toy experiment, x_{fit} is the value of the same parameter returned from the fit and $\sigma_{x_{\text{fit}}}$ is the error on that parameter returned by the fit. When the fit model is unbiased, it correctly describes the signal shape and correctly estimates the errors, the distribution of the pull over all 500 toy MC experiments should follow a Gaussian distribution with a mean of zero and a unit width.

The results of the signal embedded toy MC studies are showing no significant bias of the signal yield, as can be seen in Figure 7-14, Figure 7-13, and Figure 7-15 for the $B^+ \to \rho^+\gamma$ fit model, the $B^0 \to \rho^0\gamma$ fit model and the $B^0 \to \omega\gamma$ fit model, respectively.

The set of signal embedded toy MC studies for the three decay mode simultaneous fit model is performed with a slight variation. For the generation of the toy data set, the ρ^+ yield is adjusted upwards by the lifetime ration of B^+/B^0 in accordance with Equation 7.5. The results of this study can be seen in Figure 7-16.

All results from these toy MC studies are summarized in Table 7.9. No significant bias is observed in the signal yield. However, the continuum background yield is anti-correlated with the B background yield in all three individual fit models. This is of no real concern though since the signal yield is not affected by it. When comparing the sensitivity with the previous *BABAR* analysis [24], it is clear that the current analysis is about equally sensitive in the $B^0 \to \omega\gamma$ mode and about $10 - 20\%$ more sensitive in the other two modes.

Number of embedded signal events	$B^0 \to \rho^0\gamma$	$B^+ \to \rho^+\gamma$	$B^0 \to \omega\gamma$
0.0	-1.732 ± 5.008	-2.061 ± 7.71	-0.8567 ± 4.376
5.0	-	-	5.124 ± 5.09
10.0	9.552 ± 6.5	10.34 ± 9.106	10.35 ± 5.879
15.0	-	-	15.39 ± 6.561
20.0	19.35 ± 7.862	20.12 ± 10.23	19.7 ± 7.109
30.0	29.59 ± 8.892	29.75 ± 11.2	-
40.0	39.34 ± 9.767	41.01 ± 12.1	-
50.0	49.16 ± 9.201	52.59 ± 12.91	-
60.0	-	61.33 ± 13.52	-

Table 7.8: Fit results of embedding different number of signal events.

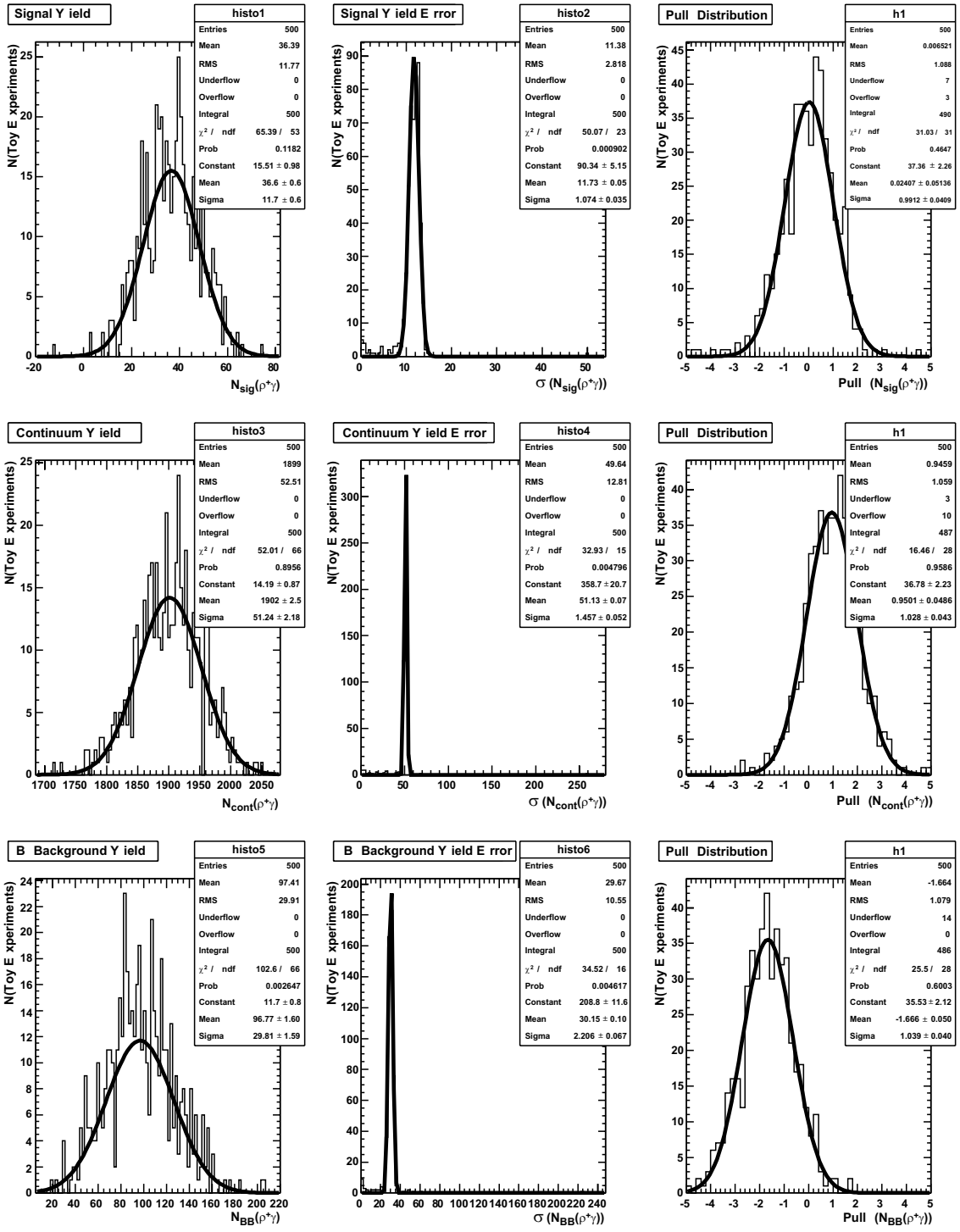


Figure 7-13: Signal embedded toy MC studies for the $B^+ \to \rho^+\gamma$ fit model. The first row shows the signal yield, the signal yield error and the pull distribution for the signal yield. The second row shows the same for the continuum background yield and the third for the B background yield.

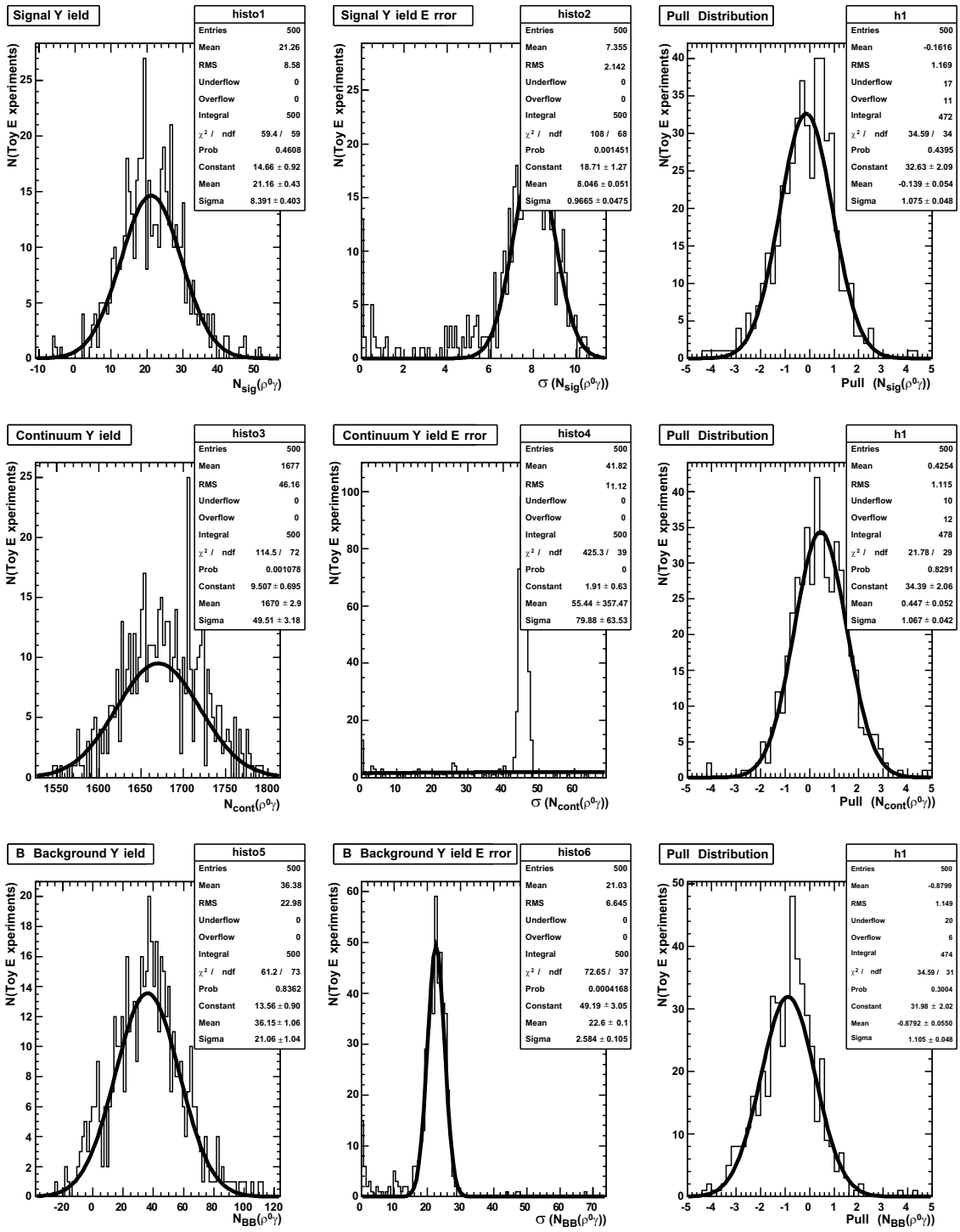


Figure 7-14: Signal embedded toy MC studies for the $B^0 \to \rho^0\gamma$ fir model. The first row shows the signal yield, the signal yield error and the pull distribution for the signal yield. The second row shows the same for the continuum background yield and the third for the B background yield.

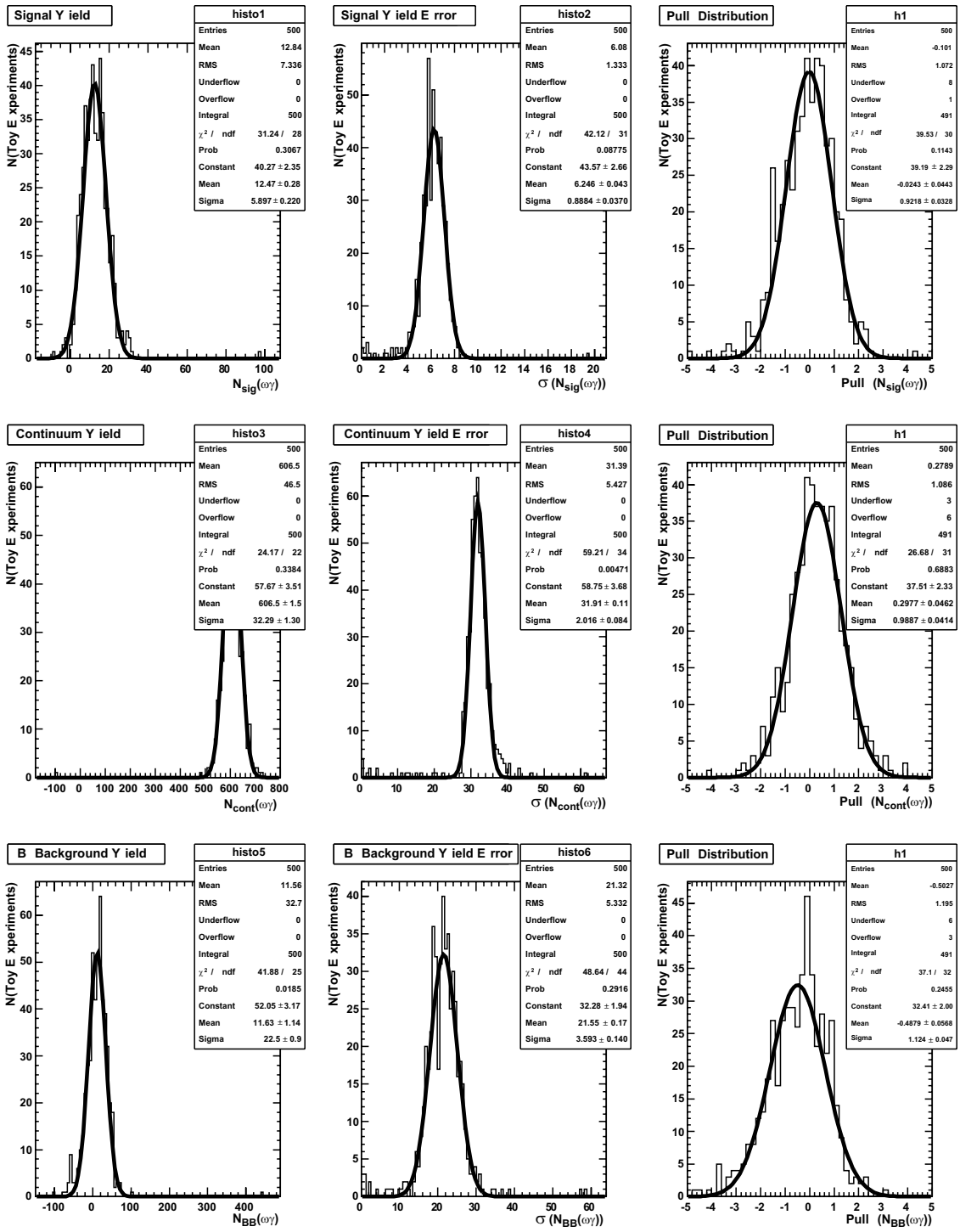


Figure 7-15: Signal embedded toy MC studies for the $B^0 \to \omega\gamma$ fit model. The first row shows the signal yield, the signal yield error and the pull distribution for the signal yield. The second row shows the same for the continuum background yield and the third for the B background yield.

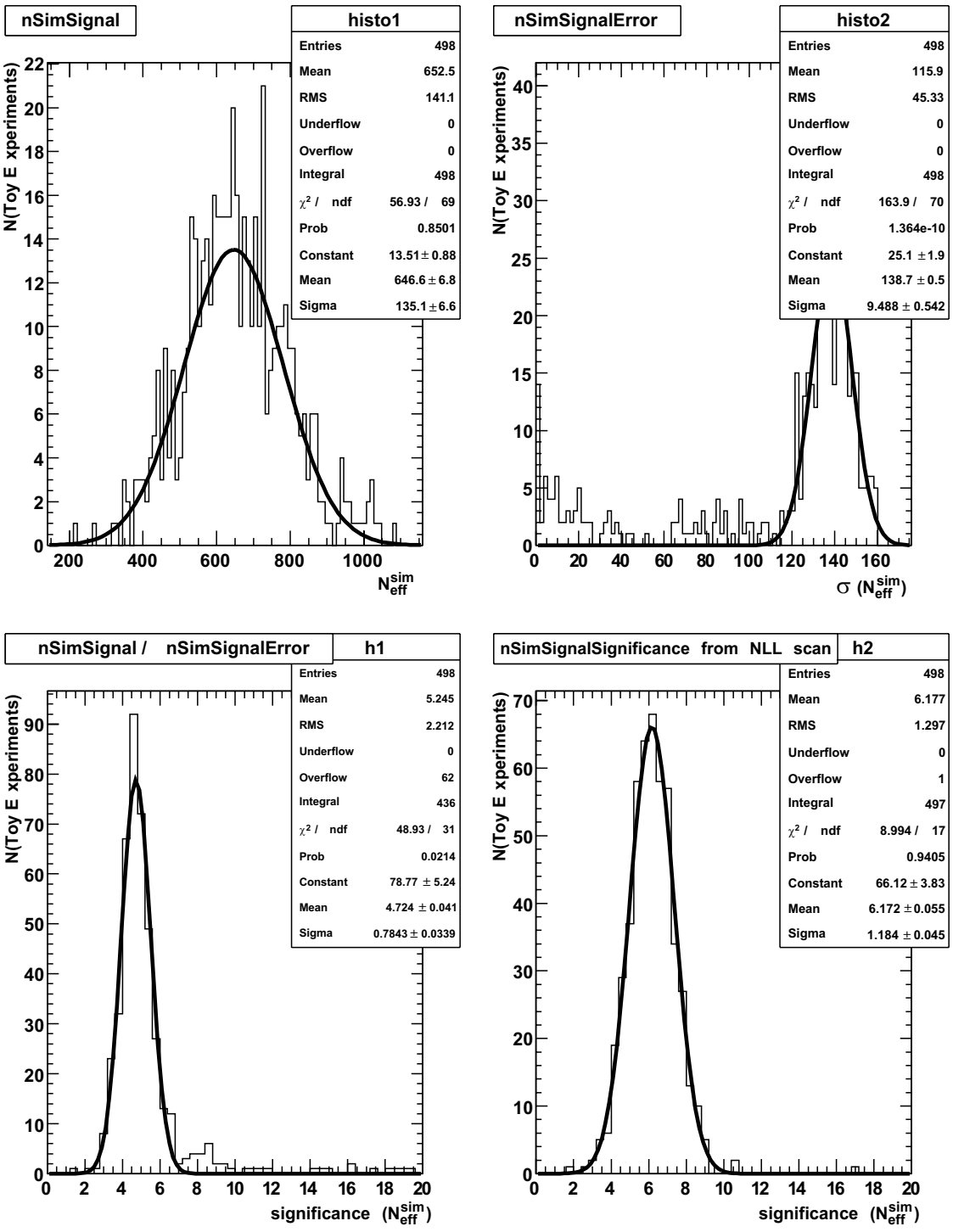


Figure 7-16: Signal embedded toy MC studies for the three decay mode simultaneous fit model. The first row shows the signal yield (left) and the error on signal yield (right). The second row shows the signal yield divided by its error (left) and the signal significance extracted from a scan of the likelihood cure for each experiment (right). The number of embedded efficiency corrected signal event was 645.

	$B^+ \to \rho^+\gamma$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$	Three mode simultaneous fit
True number of signal events	35.62	21.78	12.65	645
Assumed Branching Fraction ($\times 10^{-6}$)	1.0	0.5	0.5	1.0
Previous *BABAR* measurement @ 90% C.L. ($\times 10^{-6}$) [24]	< 1.76	< 0.36	< 0.97	< 1.16
Belle measurement ($\times 10^{-6}$) [25]	$0.55^{+0.42+0.09}_{-0.36-0.08}$	$1.25^{+0.37+0.07}_{-0.33-0.06}$	$0.56^{+0.34+0.05}_{-0.27-0.10}$	$1.32^{+0.34+0.10}_{-0.31-0.09}$
ML fit to single MC (weighted, 10% signal MC)	$31.18^{+12.4}_{-11.1}$	$23.24^{+9.02}_{-7.99}$	$15.85^{+7.41}_{-6.44}$	n/a
Significance (from likelihood scan)	3.33σ	3.66σ	2.98σ	
ML fit to single MC (weighted, 1% signal MC, 10% $K^*\gamma$ MC)	$33.97^{+12.5}_{-11.2}$	$24.86^{+8.99}_{-7.92}$	$14.92^{+7.37}_{-6.34}$	662^{+147}_{-137}
Significance (from likelihood scan)	3.69σ	4.09σ	2.81σ	6.15σ
Signal embedded toy MC study	36.6 ± 11.73	21.16 ± 8.05	12.47 ± 6.25	646.6 ± 138.7
Significance (dividing yield by error)	3.12σ	2.63σ	2.00σ	4.72σ
Corresponding significance from the old *BABAR* analysis [24] (200fb^{-1})	2.36σ	1.82σ	1.64σ	n/a
Corresponding significance from the old *BABAR* analysis [24] scaled to 300fb^{-1}	2.89σ	2.22σ	2.01σ	n/a

Table 7.9: Fit summary. Everything is assuming a data set of 300fb^{-1}. The weighted ML fits are utilizing the fully simulated MC and are using only 10% of the available signal MC (still ∼6000 events) in order to speed up the fits. Another set with only 1% of the available signal MC and 10% of the available $B \to K^*\gamma$ MC has been performed as well. The signal yield numbers for the simultaneous fit model are the efficiency corrected effective signal yields $N_{\rm eff}^{sim}$.

Chapter 8

Systematic Errors

The MC simulation is not perfect because it does not reproduce the data exactly. The choice of the functional form of the PDFs used in the fit model has an uncertainty associated with it. The size of the on–resonance data sample is not perfectly know. All these uncertainties introduce a systematic error that effects the measurements of branching fractions. Usually, the effect of these uncertainties is evaluated using data, most of the time not the actual signal decay mode but other event samples. These systematic uncertainties and their evaluations are described in this chapter.

8.1 Tracking

The systematic uncertainty of reconstructing a charged particle is determined using $e^+e^- \to \tau^+\tau^-$ events where one τ decays into a final state with three charged particles (and a neutrino) and the other τ decays into a final state with one charged particle (and two neutrinos), the so-called tau 3–1 sample. It is then determined how often one of the three charged particles coming from the same tau decay is not reconstructed, both for MC and data. The ratio of these efficiencies is not exactly one. Its deviation from unity is taken as a systematic uncertainty associated with reconstructing the trajectory of a single charged particle. This results in a systematic uncertainty of 1% per charged particle. Thus, the systematic uncertainty is 1% for $B^+ \to \rho^+\gamma$ and 2% for $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$.

8.2 Charged Particle Identification

Two aspects of the charged pion identification used in this analysis need to be checked. The efficiency of correctly identifying charged pions and the efficiency with which a charged kaon is misidentified as a charged pion. The latter is particulary important since the charged kaon misidentification rate determines largely the number of $B \to K^*\gamma$ events in the final maximum likelihood fits and could thus bias the measurement of the $B \to (\rho/\omega)\,\gamma$ signal yield. The charged pion identification efficiency directly influences the $B \to (\rho/\omega)\,\gamma$ signal reconstruction efficiencies though.

The charged particle identification performance is evaluated using a pure sample of kinematically constrained charged pions and kaons. With this kind of events, the identity of a charged particle under consideration is already known due to the kinematic constraints of a particular decay and thus the efficiency of the charged particle identification algorithm can be determined. For this analysis, a sample of background subtracted D^* decays is used. The charged particle identification algorithm (see Section 4.6.1) is applied to charged pions and kaons from D^* decay chains whose true identity is already known due to the reconstruction of all other particles in the D^* decay chain. The charged pion identification efficiencies as well as the charged kaon misidentification rates are determined separately on MC and on–resonance data as a function of the laboratory momentum. Figure 8-1 shows the performance of this algorithm.

The agreement between data and MC for the charged pion identification efficiency is generally good. Based on this study, a 2% systematic error per reconstructed charged particle is assigned to the $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ modes. The $B^0 \to \omega\gamma$ mode requires a less stringent charged pion identification for which the data MC agreement is better. The systematic uncertainty assigned to the $B^0 \to \omega\gamma$ mode is 1% per charged particle candidate.

The charged kaon misidentification rates are rather low, about 1%, and the agreement between MC and on–resonance data is rather good. The exception is the momentum range above $\approx$3 GeV/c. However, this is not relevant for this analysis since

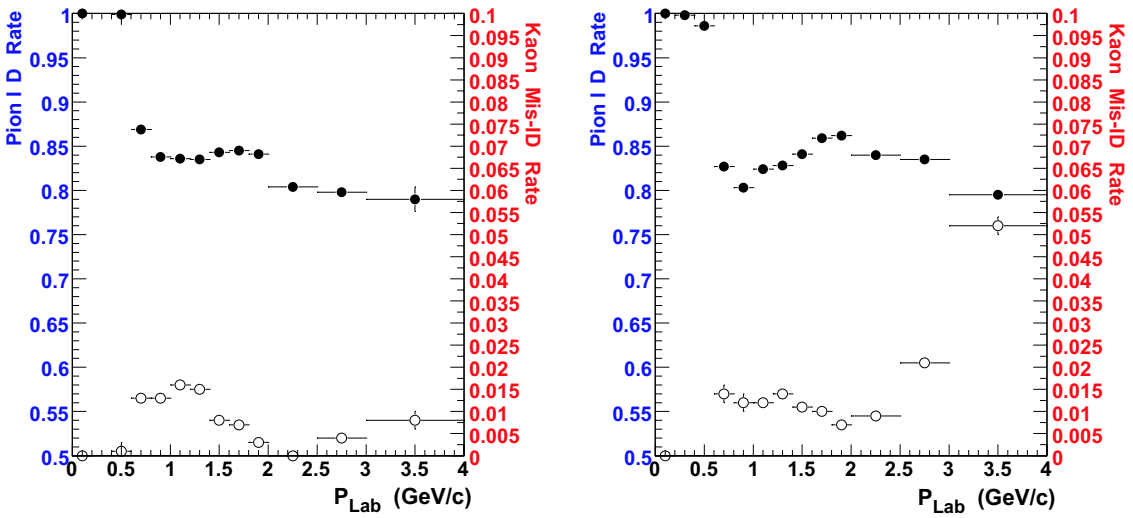


Figure 8-1: Charged particle identification performance. Generic $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ MC is shown on the left while on–resonance data is shown on the right. Filled circles are for charged pion efficiencies and use the left-hand scale. Open circles are for the charged kaon misidentification rate and use the right-hand scale. Both plots are shown as a function of the laboratory momentum of the charged particle in GeV/c.

the momenta of the reconstructed charged pions are generally less that 3 GeV/c. Based on this study, the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ yield in the $B^+ \to \rho^+\gamma$ fit model is varied by 2.12 events (out of a total of 4.48 events). The variation of the yield of the $B^0 \to K^{*0}\gamma$ component in the $B^0 \to \rho^0\gamma$ fit model is varied less since this component is part of the overall varied B background yield and not fixed as the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ yield is in the $B^+ \to \rho^+\gamma$ fit model. The variation of the $B^0 \to K^{*0}\gamma$ yield fraction is chosen to be $\frac{N_{K^{*0}\gamma}}{N_{tot}} = 0.173 \pm 0.0564$.

To summarize, the systematic uncertainties due to the charged pion identification is 2%, 4%, and 2% for the $B^+ \to \rho^+\gamma$ mode, the $B^0 \to \rho^0\gamma$ mode, and the $B^0 \to \omega\gamma$ mode, respectively.

8.3 Photon Selection

The systematic uncertainty associated with the high-energy photon consists of two parts. The photon reconstruction efficiency difference between data and MC and the efficiency difference between data and MC due to the selection cuts.

The systematic uncertainty associated with the photon reconstruction is assessed using $e^+e^- \to \mu^+\mu^-\gamma$ events. These events are selected using only requirements on the two muons:

- exactly two `GoodTracksLoose` charged particle trajectories are found,
- both of them are identified as muons,
- the missing energy of the two muons *w.r.t.* to the initial e^+e^- state is $> 1\,\mathrm{GeV}$, as determined by a kinematic fit with the constrained of the missing particle having zero mass,
- and the absolute difference between the energy missing from the two muons and the momentum missing from them be less than $0.2\,\mathrm{GeV}$.

The reconstruction efficiency is then determined as (selected $e^+e^- \to \mu^+\mu^-\gamma$ events where the photon is observed)/(all selected $e^+e^- \to \mu^+\mu^-\gamma$ events). The relative efficiency difference between data and MC is taken as the systematic error. It is found to be 1.6%.

Systematic uncertainties for the shower shape variables, second moment, $\frac{s9}{s25}$ and lateral moment are determined using $e^+e^- \to \mu^+\mu^-\gamma$ events. The $e^+e^- \to \mu^+\mu^-\gamma$ events are weighted to match the photon energy spectrum of $B \to (\rho/\omega)\,\gamma$ signal MC, which is between 1.5 to $4.5\,\mathrm{GeV}$. Figure 8-2 shows the data MC comparison for these variables. The data/MC efficiency ratios are listed in Table 8.1. The errors are added linearly, and the total error is 1.0% for the $B^+ \to \rho^+\gamma$ mode, 2.1% for the $B^0 \to \rho^0\gamma$ mode, and 0.7% for the $B^0 \to \omega\gamma$ mode. To obtain the total error on photon selection, we add these errors in quadrature to the single photon efficiency error.

Combining the two photon related systematic uncertainties yields an error of 1.9% for the $B^+ \to \rho^+\gamma$ mode, 2.6% for the $B^0 \to \rho^0\gamma$ mode, and 1.7% for the $B^0 \to \omega\gamma$ mode.

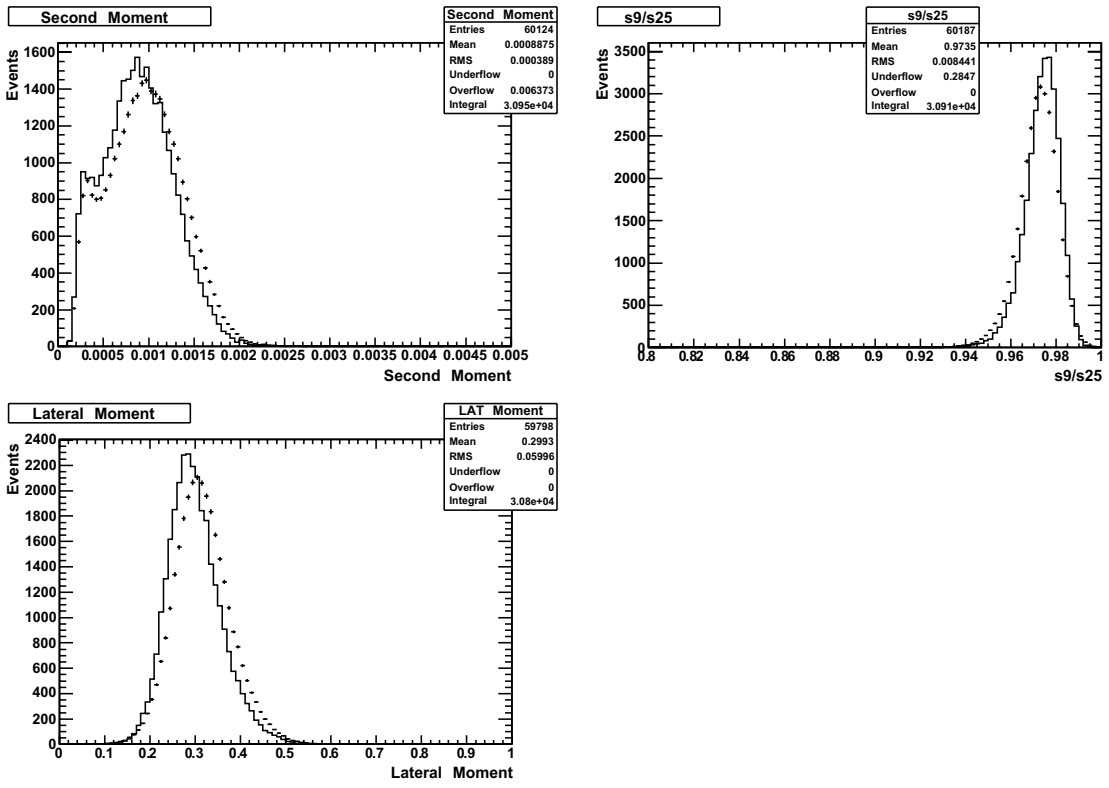


Figure 8-2: Distributions of the photon shape variables for $e^+e^- \to \mu\mu\gamma$ events. The solid histogram is MC, and the points with error bars are data.

8.4 π^0 Selection

The systematic uncertainty associated with the reconstruction and selection of the π^0s is studied using data and MC samples of the decays $\tau^+ \to \rho^+\bar{\nu}_\tau$ and $\tau^+ \to \pi^+\bar{\nu}_\tau$. These events are selected as $e^+e^- \to \tau^+\tau^-$ events where the other τ in the event is required to be reconstructed in the decay channel $\tau^- \to e^-\bar{\nu}_e\nu_\tau$. The reconstruction efficiency of $\tau^+ \to \rho^+\bar{\nu}_\tau$ is approximately proportional to the selection efficiency of the π^+ and of the π^0. The $\tau^+ \to \pi^+\bar{\nu}_\tau$ reconstruction efficiency is roughly proportional to the selection efficiency of the π^+ only. Thus, the π^0 momentum dependent double ratio

$$\frac{\frac{\tau^+\to\rho^+\bar{\nu}_\tau(\text{data})}{\tau^+\to\rho^+\bar{\nu}_\tau(\text{MC})}\,(p_{\pi^0})}{\frac{\tau^+\to\pi^+\bar{\nu}_\tau(\text{data})}{\tau^+\to\pi^+\bar{\nu}_\tau(\text{MC})}}, \tag{8.1}$$

is a measure of the π^0 reconstruction efficiency in data *vs.* MC.

Mode	Second moment	$\frac{s9}{s25}$	Lateral moment	N_{Crystals}
$\rho^+\gamma$	0.996 ± 0.006	1.000 ± 0.006	1	1
$\rho^0\gamma$	0.985 ± 0.006	1.000 ± 0.006	1	1
$\omega\gamma$	0.999 ± 0.006	1.000 ± 0.006	1	1

Table 8.1: The $\mu\mu\gamma$ data/MC efficiency ratios for the photon quality selection.

The resulting π^0 systematic error is 3%. An efficiency correction of 0.971 is applied for the $B^+ \to \rho^+\gamma$ mode and 0.968 for the $B^0 \to \omega\gamma$ mode.

8.5 π^0/η Veto

There is no large data sample of B decays with a well reconstructed high-energy photon. Thus, the systematic uncertainties for the $\mathcal{LR}$ based π^0 and η vetoes are determined by embedding a high-energy photon from $B \to (\rho/\omega)\,\gamma$ signal MC into $B \to D\pi$ MC and data samples. The high energy photon candidate is randomly chosen from the $B \to (\rho/\omega)\,\gamma$ signal MC events and is combined with photons originating from the other B decay in the $B \to D\pi$ data and MC samples. These fully reconstructed $D\pi$ events are used to form the π^0 and η likelihoods for true signal high-energy photons.

Two different $B \to D\pi$ reconstruction modes are needed, one for obtaining a clean sample of neutral B mesons and one for a sample of charged B mesons. The two samples are chosen to be $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ and $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$. The selection criteria for these two samples are listed in Appendix C.

After those selection criteria are applied, the m_{ES} data distributions for both modes are shown in Figure 8-3. The signal and sideband regions are defined as:

- Signal region:

  $(-0.05\,\text{GeV} < \Delta E < 0.05\,\text{GeV}) \times (5.27\,\text{GeV}/c^2 < m_{ES} < 5.29\,\text{GeV}/c^2)$

- Sideband region:

  $(-0.05\,\text{GeV} < \Delta E < 0.05\,\text{GeV}) \times (5.225\,\text{GeV}/c^2 < m_{ES} < 5.265\,\text{GeV}/c^2)$

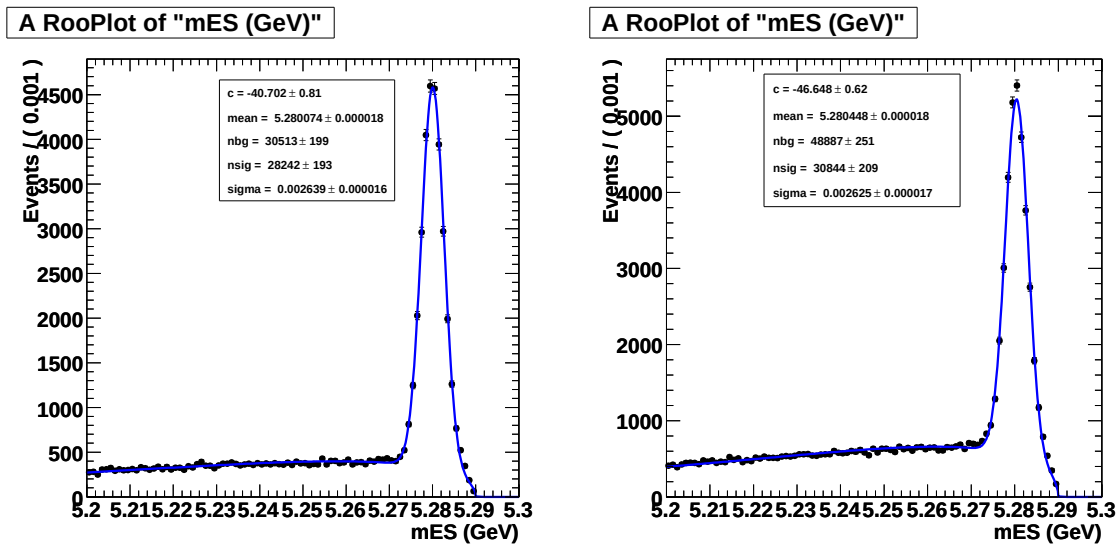


Figure 8-3: m_{ES} distributions for $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ and $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$. The charged B meson decay mode is shown on the left and the neutral B meson decay mode on the right. The lines are the results of fits to a Gaussian for signal and an ARGUS for background.

The events from the sideband region are assumed to be pure background events and are used to subtract the background events in the signal region.

The resulting $\mathcal{LR}$s are compared between the corresponding $B \to D\pi$ signal MC and background subtracted data. The results of these comparisons are shown in Figure 8-4 for the $B^0 \to D^-\pi^+$ mode and in Figure 8-5 for the $B^- \to D^0\pi^-$ mode. As can be seen in these Figures, MC and data are in very good agreement. A systematic error of 1% each for the $\mathcal{LR}\,(\pi^0)$ and $\mathcal{LR}\,(\eta)$ are assigned based on the statistical error at of the data/MC efficiency comparison. They are added linearly and thus, the resulting systematic uncertainty for the $\mathcal{LR}$ vetoes is 2%.

Finally, the systematic uncertainty of the distance cuts to the nearest neutral and charged cluster are evaluated. They result in a 2% systematic uncertainty. Combining this with the 2% systematic uncertainty for the $\mathcal{LR}$ vetoes, the total π^0 and η veto systematic uncertainty is 2.8%.

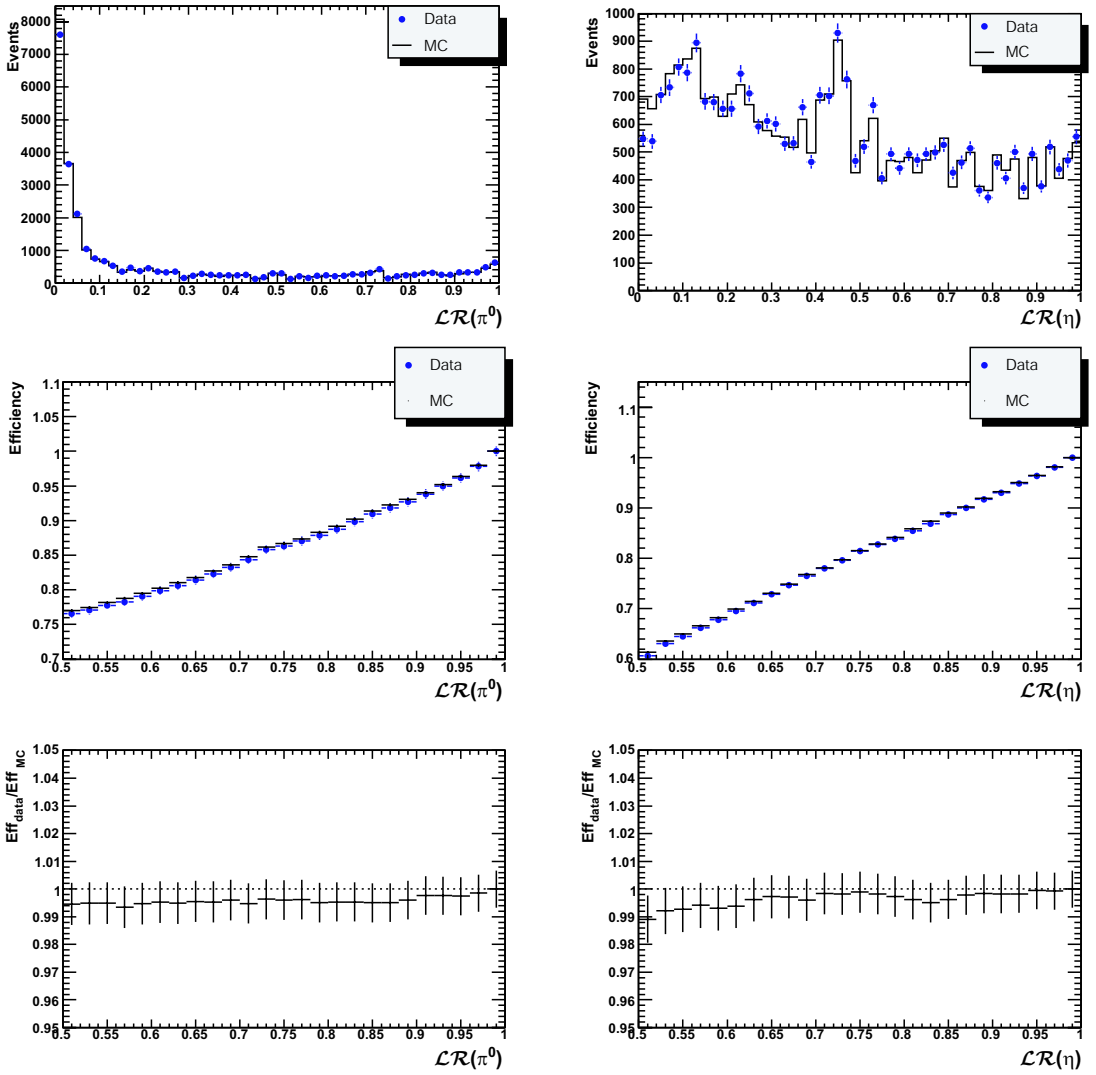


Figure 8-4: Comparison of π^0 and η vetoes between $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ data and MC. The left column shows the π^0 veto distributions and the right column shows the η veto distributions. The comparison of the $\mathcal{LR}$ distributions is shown on the top row. The efficiency comparisons as a function of the $\mathcal{LR}$ cut is shown on the middle row and the plots on the bottom row compare the ratios of data/MC efficiencies. On the top two rows, the blue dots are data entries and the black histogram/dots are MC.

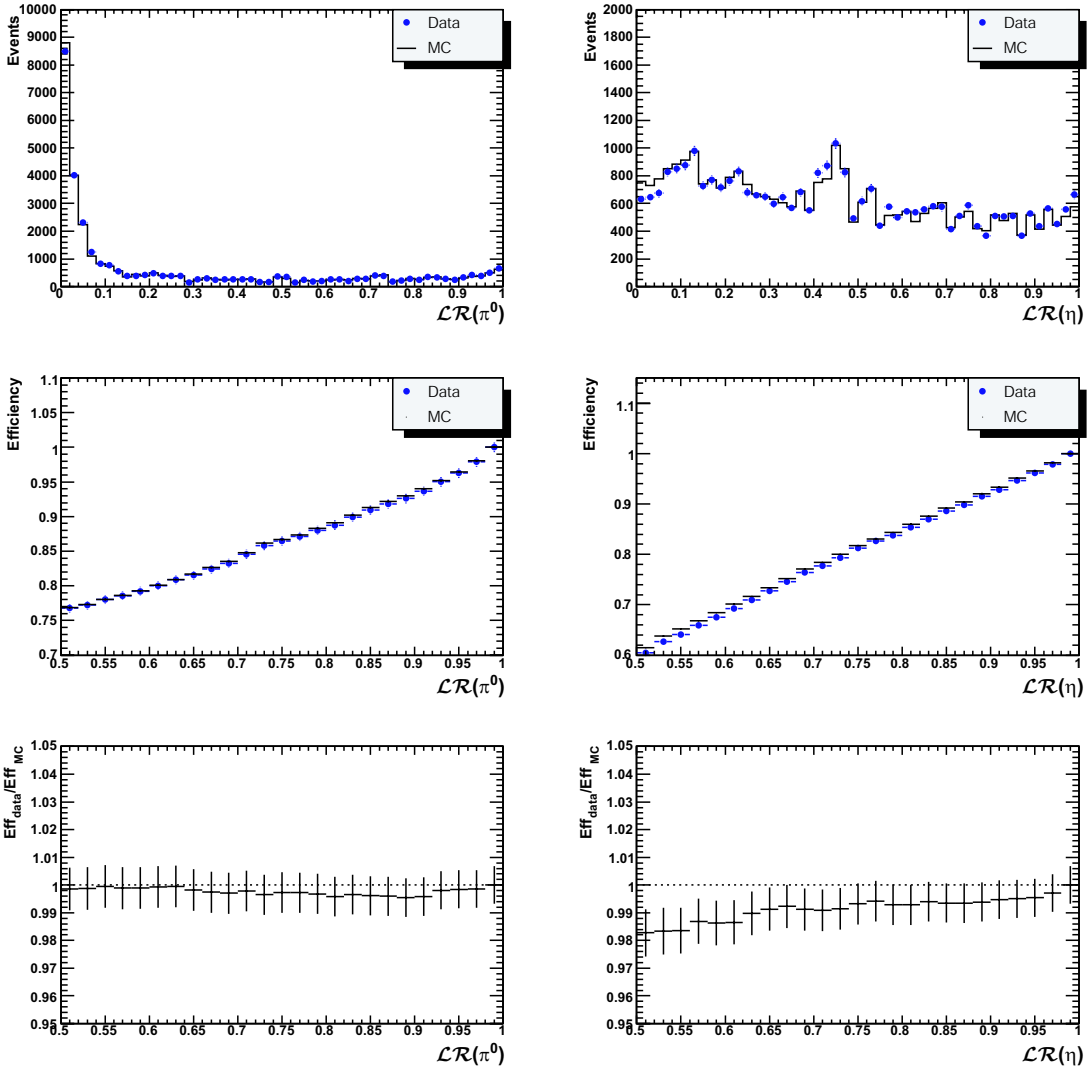


Figure 8-5: Comparison of π^0 and η vetoes between $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi-$ data and MC. The left column shows the π^0 veto distributions and the right column shows the η veto distributions. The comparison of the $\mathcal{LR}$ distributions is shown on the top row. The efficiency comparisons as a function of the $\mathcal{LR}$ cut is shown on the middle row and the plots on the bottom row compare the ratios of data/MC efficiencies. On the top two rows, the blue dots are data entries and the black histogram/dots are MC.

8.6 Neural Network

First, the outputs of all three neural networks are validated by comparing continuum background MC with off–resonance data. The comparison plots are shown in Figure 8-6.

The systematic uncertainty associated with the signal efficiency due to the cut on the neural network output is determined using a high statistics, high signal purity sample of B decays. Since most of the neural network input variables are dependent only on the rest of the event, the specific properties of the control sample is of secondary interest. The chosen samples are

- $B^- \to D^0\pi^-(D \to K\pi)$ for the $B^+ \to \rho^+\gamma$ neural network and
- $B^0 \to D^-\pi^+(D \to K\pi\pi)$ for the $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$ neural networks

(see Appendix C.1). The input variables to the neural network that are dependent on the signal side decay are computed in the control sample by treating the pion that originates directly from the B decay as the high-energy photon and the D meson as the ρ/ω meson.

The three neural networks are not retrained for this systematic evaluation, only applied.

The difference in the neural network output distributions between $B \to D\pi$ signal MC and $B \to D\pi$ data on the one hand and $B \to (\rho/\omega)\,\gamma$ signal MC and $B \to (\rho/\omega)\,\gamma$ data on the other hand is expected to be the same. Thus, the evaluation of the former difference is used for the systematic uncertainty of the latter.

In order to obtain the signal-only shape of the neural network output from these on–resonance data sets, only events from the signal region ($5.27 < m_{ES} < 5.29\,\mathrm{GeV}/c^2$) are used. A bin-by-bin background subtraction is performed using the scaled events from the lower m_{ES} sideband ($5.20 < m_{ES} < 5.27\,\mathrm{GeV}/c^2$) where the scale factor is the ratio between the integral of the ARGUS function above and below $m_{ES} = 5.27\,\mathrm{GeV}/c^2$. The comparisons of the neural network outputs of these background-subtracted $B \to D\pi$ on–resonance data samples with the corresponding $B \to D\pi$ MC are shown in Figure 8-7.

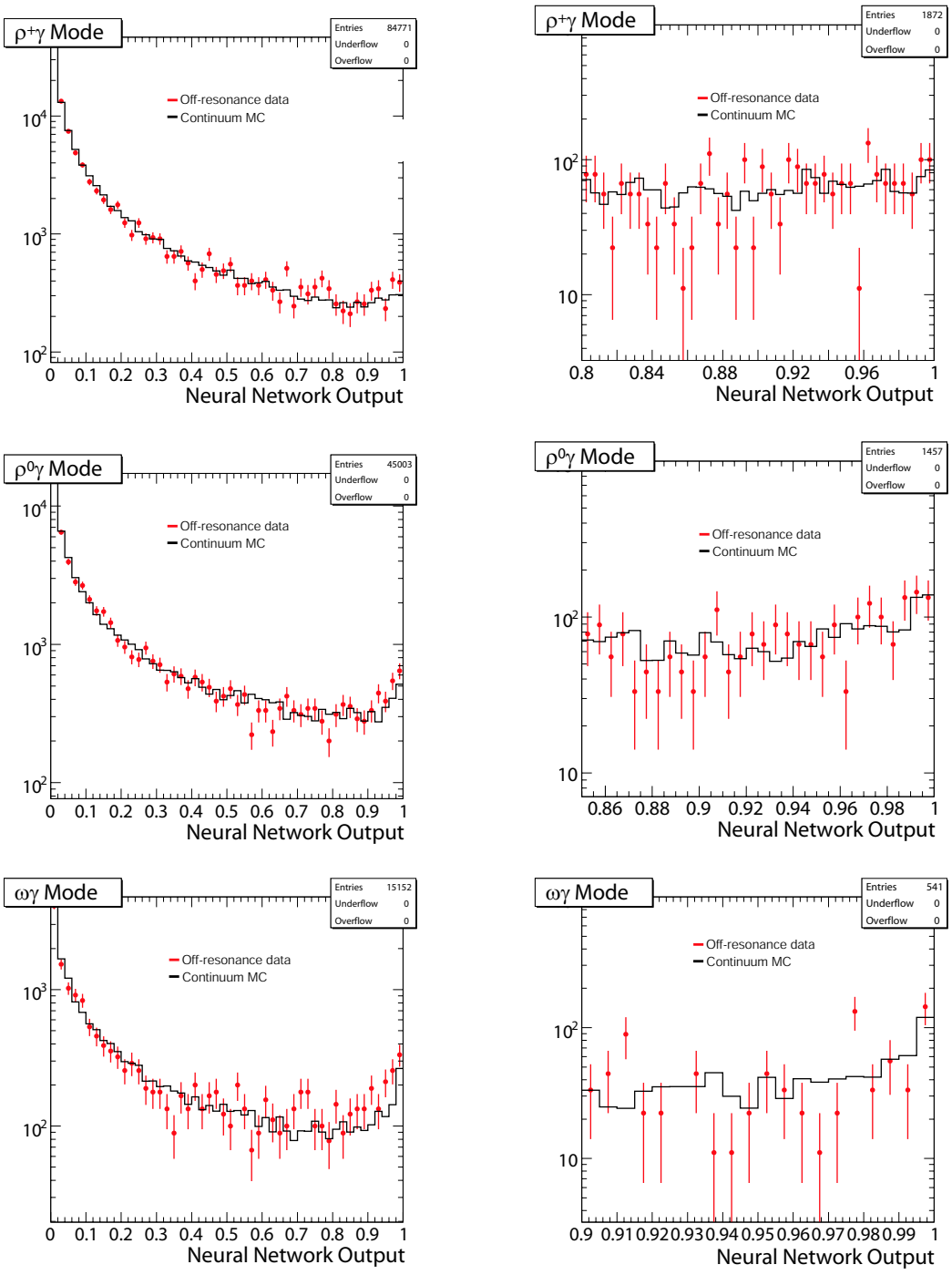


Figure 8-6: Neural network output comparison between data and MC. Comparison of the output of all three neural networks between continuum MC (black histogram) and off–resonance data (red data points). The left column shows the whole output and the right column shows the region of the output used in the final fits. All histograms are scaled to have the same area.

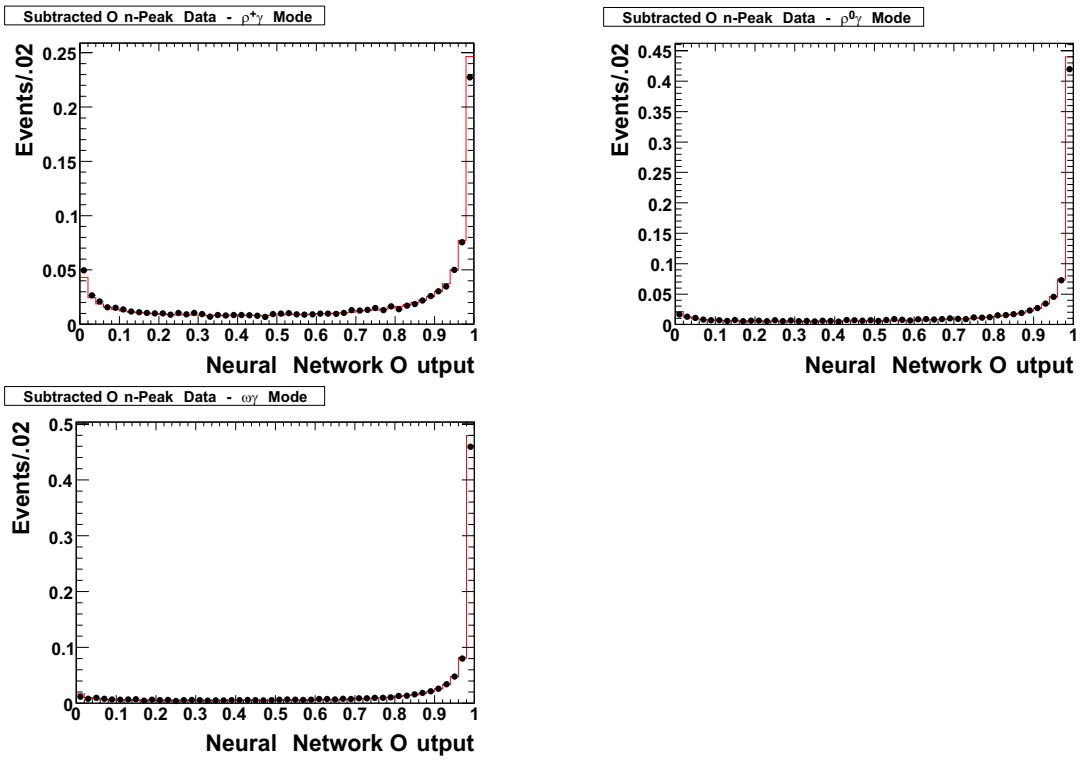


Figure 8-7: Comparison of $B \to D\pi$ neural network distributions for MC and background subtracted on–resonance data. The signal MC is shown as the red line and the on–resonance data is shown as black dots. The distributions are normalized to the same area.

The signal efficiency of the cut on the neural network output is computed for both $B \to D\pi$ signal MC and the background-subtracted $B \to D\pi$ on–resonance data control samples. Both efficiencies are shown in the plots on the left of Figure 8-8, the bin-by-bin efficiency ratio, $\epsilon_{\rm Data}/\epsilon_{\rm MC}$ is calculated and shown in the same Figure on the right. The neural network is generally slightly less efficient in on–resonance data than in MC.

The value of the efficiency ratio at the nominal neural network output cut values is taken as a correction to the signal efficiency and the statistical error of this ratio at that point is taken as the systematic uncertainty of that correction. Thus, the signal-efficiency corrections associated with the cut on the neural network output are 95.2%, 97.1%, and 92.1% for the $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively. The systematic uncertainty on these corrections are 1% for all three modes.

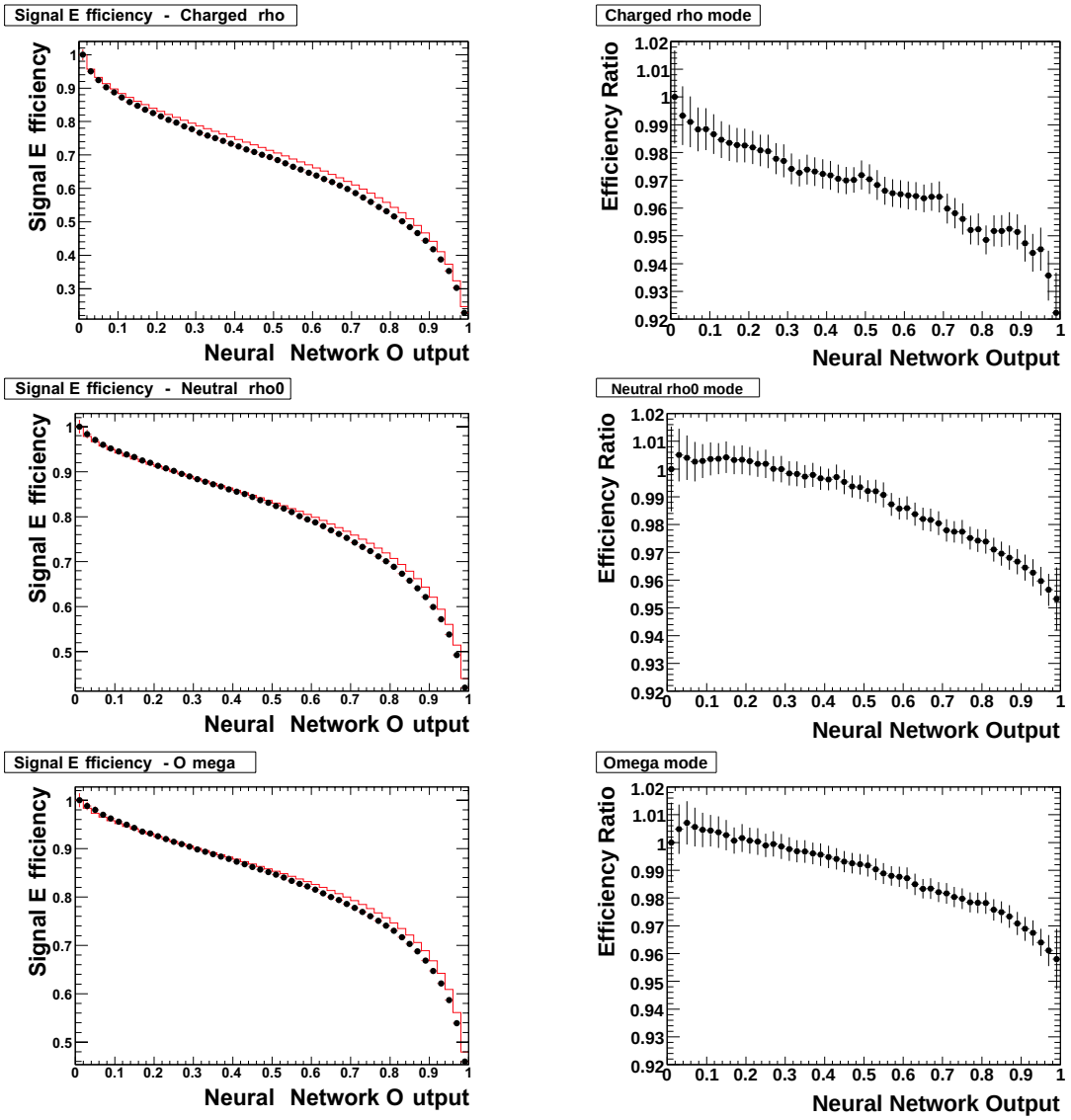


Figure 8-8: Signal efficiencies and efficiency ratios. Left: The comparison of the signal efficiency *vs.* cut value between $B \to D\pi$ signal MC (red line) and background subtracted on–resonance data (black dots). Right: The bin-by-bin $\epsilon_{\rm Data}/\epsilon_{\rm MC}$. The distributions of signal efficiency *vs.* neural net cut value are normalized to the same area. Both sets of plots have 50 bins.

8.7 Transformed Neural Network PDF Shape

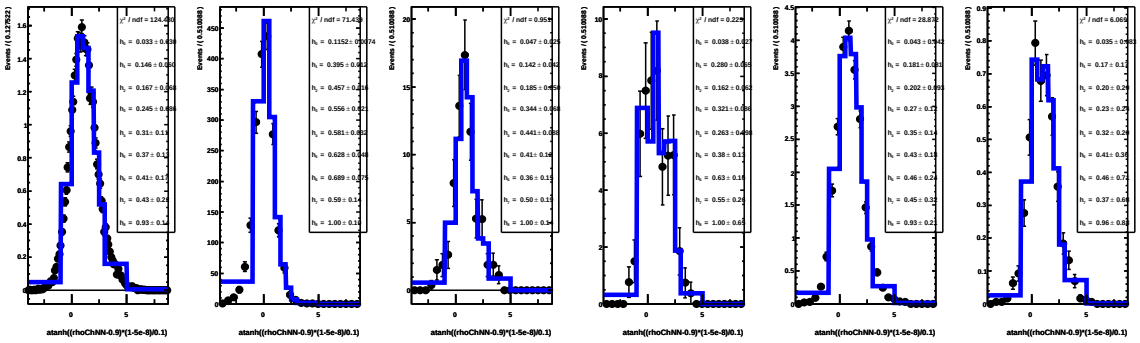
In order to study the systematic uncertainty associated with a particular choice of PDF for the transformed neural network outputs, a series of toy MC studies is performed. 500 signal MC embedded toy experiments are generated for each decay mode using the nominal fit PDFs described in Chapter 7. Then, each of these toy experiments is fitted in two different ways:

- with the nominal PDFs using the Cruijff PDF for the transformed neural network output, as described in Chapter 7;
- with the nominal PDFs for all dimensions but for the transformed neural network output. The PDFs for this dimension have been replaced with a parametric step function with 10 bins. The parameters for these PDFs are determined from individual MC samples, the same procedure as for the nominal fitter, see Figure 8-9. The continuum parameters are also floated in these fits.

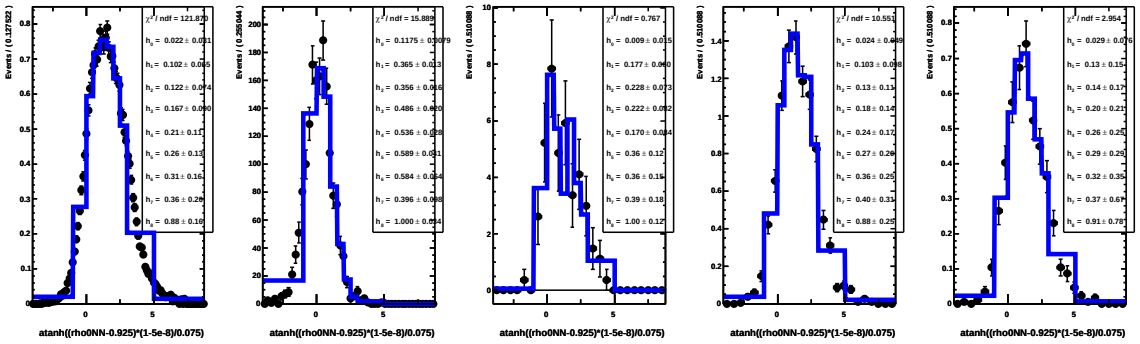
The experiment-by-experiment difference in the fitted signal yield is drawn and a Gaussian PDF has been fitted to these distributions, see Figure 8-10. These shift in signal yield are divided by the nominal signal yield and the resulting numbers are taken as the systematic uncertainties associated with the choice of PDFs for the transformed neural network output. The resulting systematic uncertainties are 0.4%, 0.3% and 2.3% for the $\rho^+\gamma$, the $\rho^0\gamma$ and the $\omega\gamma$ mode, respectively.

8.8 Signal PDF Parameter Corrections

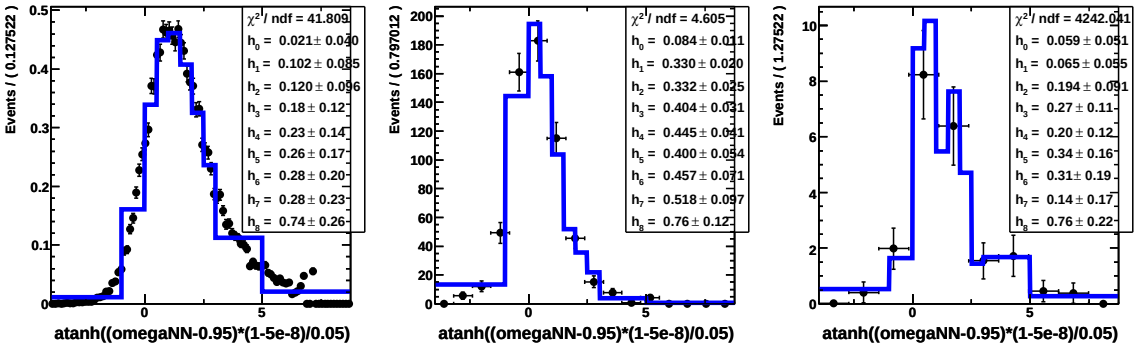
It is known that the MC is not representing the real data accurately in terms of the energy calibration of the EMC. Thus, the signal PDF in ΔE as determined from signal MC will not accurately represent the data. The $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ and $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ control samples are used to correct for this difference between data and MC as the difference is expected to be the same as between $B \to (\rho/\omega)\,\gamma$ signal MC and data.



(a) The step PDFs for the $\rho^+\gamma$ mode.

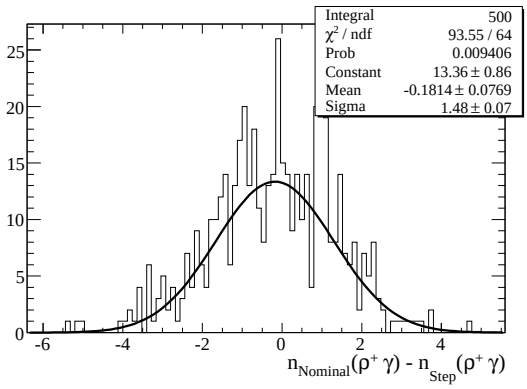


(a) The step PDFs for the $\rho^0\gamma$ mode.

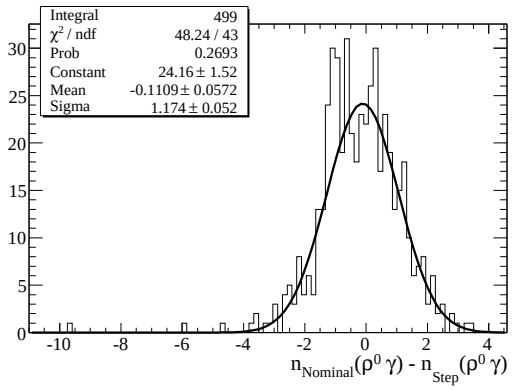


(c) The step PDFs for the $\omega\gamma$ mode.

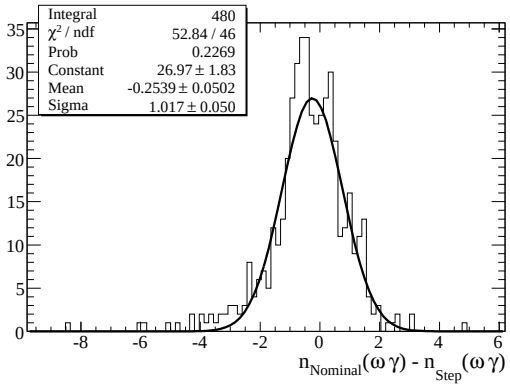
Figure 8-9: Step function PDFs used for the three decay modes. These shapes have been used for a toy study in order to evaluate the systematic errors associated with the choice of a particular PDF for the transformed neural network outputs $\mathcal{NN}$. The plots show in the first row from left to right $B^+ \to \rho^+\gamma$ signal MC, continuum MC, $B\overline{B}$ MC (excluding $B \to X_s\gamma$ and $B \to K^*\gamma$), $B \to X_s\gamma$ background (excluding $B \to K^*\gamma$), $B \to K^*\gamma$ (excluding $K^{*+} \to K^+\pi^0$) and $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ peaking background. The second row shows $B^0 \to \rho^0\gamma$ signal MC, continuum MC, $B\overline{B}$ MC (excluding $B \to K^*\gamma$), $B^0 \to K^{*0}\gamma$ peaking background and $B^+ \to K^{*+}\gamma$ background. The last row shows $B^0 \to \omega\gamma$ signal MC, continuum MC and $B\overline{B}$ MC.



(a) The difference in signal yield for the $\rho^+\gamma$ mode.



(b) The difference in signal yield for the $\rho^0\gamma$ mode.



(c) The difference in signal yield for the $\omega\gamma$ mode.

Figure 8-10: Differences in the signal yield between using the nominal fits and fits using step function PDFs for the transformed neural network output.

In order to have the same number of photons in the decay tree (high-energy photon and π^0), the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ sample is chosen to mimic the $B^0 \to \rho^0\gamma$ mode and the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ is taken to mimic the $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ modes. These control samples are reconstructed the same way as the respective $B \to \rho\gamma$ modes with three exceptions:

- one charged particle is required to be identified as a kaon instead of a pion,
- the invariant mass of the kaon and pion is required to be $0.8 < m_{K\pi} < 1.0\,\mathrm{GeV}/c^2$,
- and the neural network output is required to be above 0.8.

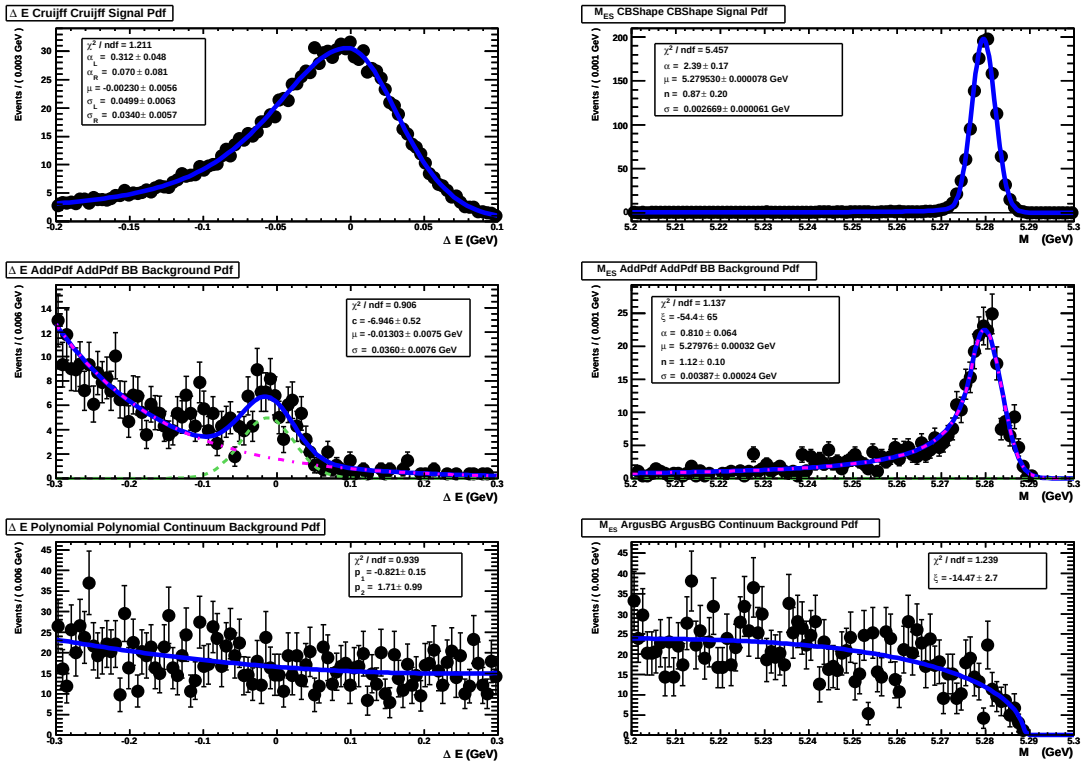


Figure 8-11: PDF shapes, as determined from MC simulation, used in $B \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) fit.

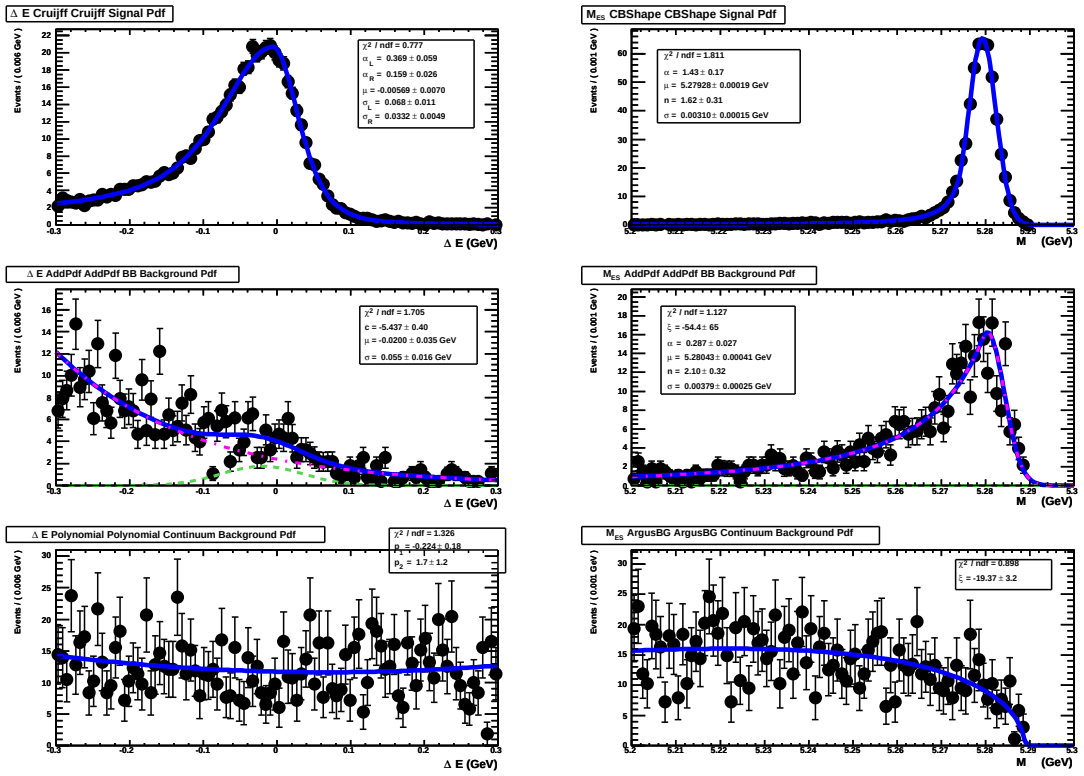


Figure 8-12: PDF shapes, as determined from MC simulation, used in $B \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) fit.

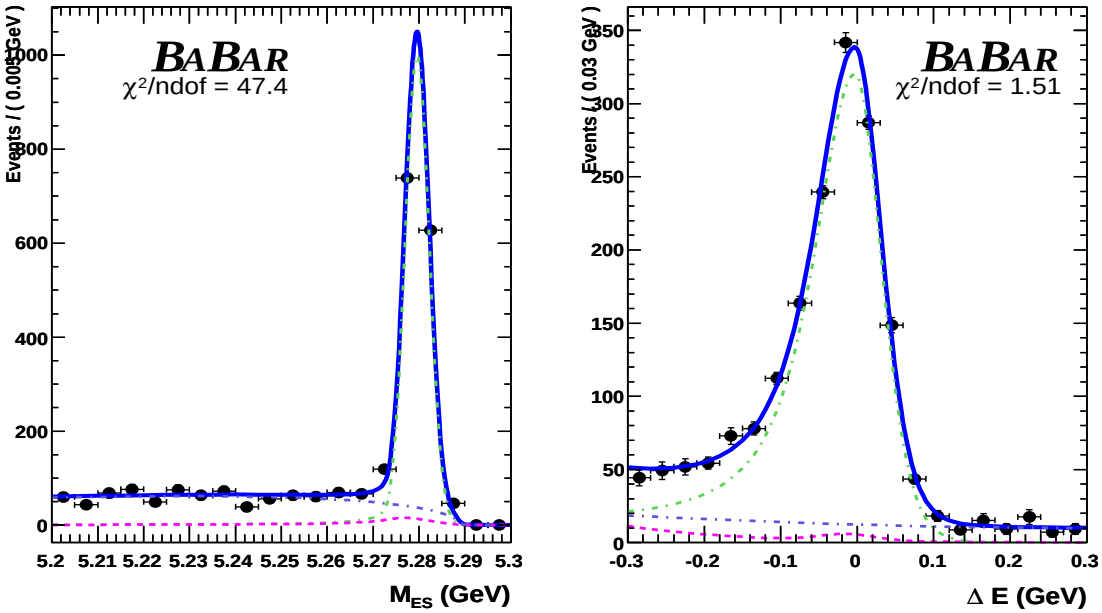


Figure 8-13: Projection plots for weighted MC $B \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) fit

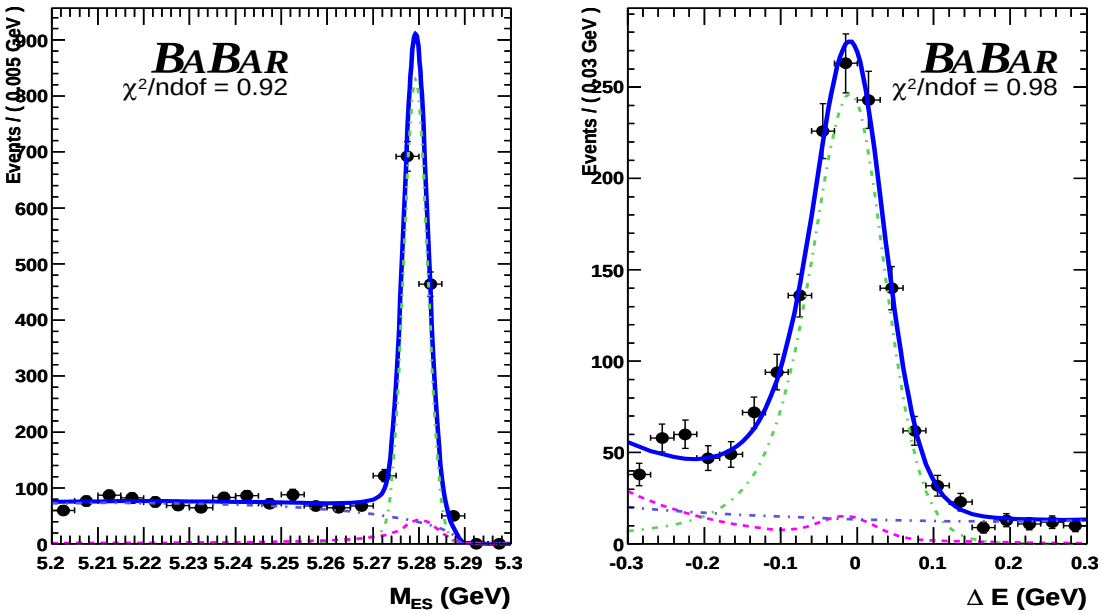


Figure 8-14: Projection plots for on-peak data $B \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) fit

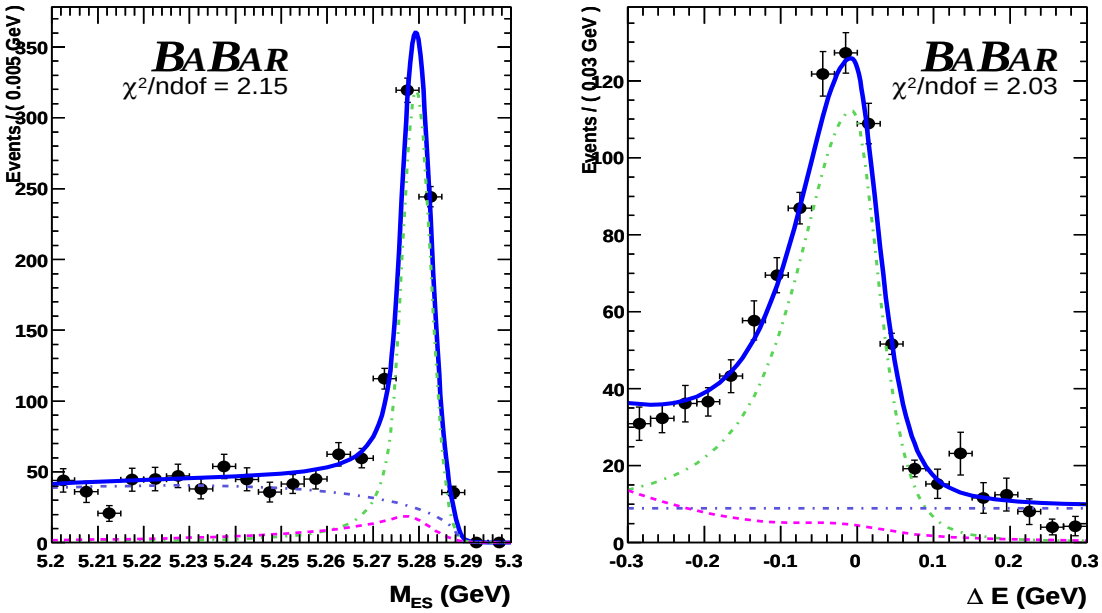


Figure 8-15: Projection plots for weighted MC $B \to K^{*+}\gamma$ ($K^{*+} \to K^{+}\pi^{0}$) fit

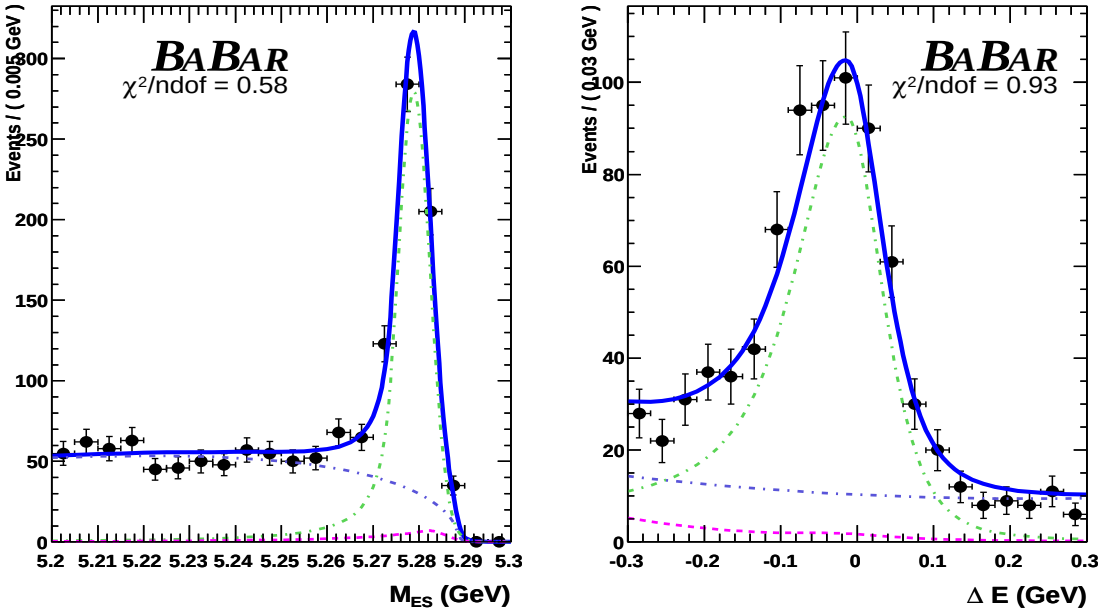


Figure 8-16: Projection plots for on-peak data $B \to K^{*+}\gamma$ ($K^{*+} \to K^{+}\pi^{0}$) fit

A two dimensional fit in m_{ES} and ΔE is performed for the $B \to K^*\gamma$ control samples in data and MC. The signal shapes are described by Crystal Ball functions, as is the B background component in m_{ES}. Cruijff functions are used to describe the signal shapes in ΔE following closely what is done for the $B \to (\rho/\omega)\,\gamma$ signal PDFs. A Gaussian on top of an exponential function is used to describe the B background component in ΔE. The continuum background is described by an ARGUS function in m_{ES} and a second-order polynomial in ΔE. The PDF shapes, as determined from MC simulation, are shown in Figures 8-11 and 8-12.

Using the results of the fits to data, the corresponding branching fractions (not corrected for data-MC efficiency differences) are measured to be (with statistical errors only):

$$\begin{aligned} \mathcal{B}[B^0 \to K^{*0}\gamma\ (K^{*0} \to K^+\pi^-)] &= (3.42^{+0.20}_{-0.16}) \times 10^{-5} \\ \mathcal{B}[B^+ \to K^{*+}\gamma\ (K^{*+} \to K^+\pi^0)] &= (4.43^{+0.20}_{-0.20}) \times 10^{-5}. \end{aligned} \tag{8.2}$$

This result can be directly compared to (and is in statistical agreement with) the previously published *BABAR* results, based on Runs 1 and 2 [40]:

$$\begin{aligned} \mathcal{B}[B^0 \to K^{*0}\gamma\ (K^{*0} \to K^+\pi^-)] &= (3.92 \pm 0.20) \times 10^{-5} \\ \mathcal{B}[B^+ \to K^{*+}\gamma\ (K^{*+} \to K^+\pi^0)] &= (4.90 \pm 0.45) \times 10^{-5}. \end{aligned} \tag{8.3}$$

Figures 8-13 through 8-16 show the m_{ES} and ΔE projection plots for data and MC for the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ mode and the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ mode.

Tables 8.2 and 8.3 summarize the resulting comparison of the PDF signal shape parameters between data and MC, together with the corrections applied to the corresponding $B \to (\rho/\omega)\,\gamma$ signal PDF shape parameters. The systematic uncertainties associated with the fixed PDF parameters are described in the following section.

		$B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$		
		MC	Data	Correction
m_{ES}	mean (MeV/c^2)	$5279.60^{+0.08}_{-0.08}$	$5279.25^{+0.11}_{-0.11}$	Not applied
	σ (MeV/c^2)	$2.63056^{+0.06620}_{-0.06466}$	$2.59178^{+0.08144}_{-0.08077}$	Not applied
ΔE	mean (MeV)	$-2.302^{+5.576}_{-5.576}$	$-8.738^{+6.329}_{-7.647}$	-6.436
	σ_L (MeV)	$49.872^{+6.284}_{-6.284}$	$50.375^{+7.051}_{-7.479}$	$\times 1.0101$
	σ_R (MeV)	$33.979^{+5.731}_{-5.731}$	$44.135^{+7.156}_{-5.510}$	$\times 1.2989$
	α_L (MeV)	$312.10^{+47.97}_{-47.97}$	$217.08^{+74.89}_{-61.17}$	$\times 0.69555$
	α_R (MeV)	$70.30^{+81.35}_{-81.35}$	$67.14^{+44.49}_{-59.35}$	$\times 0.9550$

Table 8.2: Summary of results for data and MC fits to $B \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) control sample.

		$B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$		
		MC	Data	Correction
m_{ES}	mean (MeV/c^2)	$5279.38^{+0.18}_{-0.18}$	$5278.87^{+0.21}_{-0.21}$	Not applied
	σ (MeV/c^2)	$3.2077^{+0.1473}_{-0.1431}$	$3.5356^{+0.1790}_{-0.1790}$	Not applied
ΔE	mean (MeV)	$-5.780^{+6.934}_{-6.934}$	$-14.852^{+8.364}_{-8.364}$	-9.1666
	σ_L (MeV)	$67.58^{+10.63}_{-10.63}$	$64.19^{+11.24}_{-11.24}$	$\times 0.9485$
	σ_R (MeV)	$33.28^{+4.931}_{-4.931}$	$44.97^{+0.27}_{-0.27}$	$\times 1.353$
	α_L (MeV)	$369.46^{+58.83}_{-58.83}$	$368.84^{+58.89}_{-58.89}$	$\times 0.99832$
	α_R (MeV)	$158.89^{+26.20}_{-26.20}$	$158.92^{+26.23}_{-26.23}$	$\times 1.0002$

Table 8.3: Summary of results for data and MC fits to $B \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) control sample.

8.9 Signal PDF Shape

In order to evaluate the systematic error due to fixing the shape of the signal PDFs in the fit, the fixed shape parameters are varied up and down by one sigma each. For each of these variations, a new fit on the on-peak data is performed and the relative shift in signal yield w.r.t. to the nominal signal yield is taken as a systematic uncertainty. The one-sigma errors for the PDF parameters for the $\cos\theta_{Hel}$ (and $\cos\theta_{Dalitz}$ for the $B^0 \to \omega\gamma$ mode) are taken from the fit to (un-weighted) signal MC.

The one-sigma errors for the ΔE and m_{ES} signal PDFs are taken from the $B \to K^*\gamma$ control samples, see Section 8.8. To be more specific, for the $B^0 \to \rho^0\gamma$ mode, the parameter uncertainties are taken from the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ control sample and the parameter uncertainties for the $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ mode are taken from the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ control sample.

The resulting total shift for a set of parameter variations for an individual PDF is computed using the full correlation matrix for those PDF parameters. The systematic uncertainties are 4.84%, 3.26% and 2.40% for the $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ mode, respectively.

These same variations are also performed in the simultaneous fits. The resulting systematic uncertainties of the simultaneous fit models are 3.1% for the $B \to \rho\gamma$ model and 2.6% for the $B \to (\rho/\omega)\,\gamma$ model.

8.10 B Background PDF

The fixed B background PDF parameters are varied much in the same way as the signal PDF parameters (see Section 8.9). The only difference is that here, all one-sigma errors are taken from the fits to dedicated MC. The resulting systematic uncertainties are 3.91%, 2.91%, 9.70% and 2.71% for the $\rho^+\gamma$, the $\rho^0\gamma$, the $\omega\gamma$ mode and the simultaneous fit, respectively. The large systematic uncertainty in the $\omega\gamma$ mode is mainly due to the low signal yield i.e. the systematic variation of the B-backgrounds yields a signal yield shift of only 1.06 events.

8.11 B Counting

The accuracy of counting the number of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events in the on–resonance data sample is limited to 1.1%.

8.12 $\omega \to \pi^+\pi^-\pi^0$ Branching Fraction

The branching fraction of the ω meson decaying into the reconstructed final state is known with finite precision, $\mathcal{B}(\omega \to \pi^+\pi^-\pi^0) = (89.1 \pm 0.7)\%$. Thus, an additional systematic uncertainty of 0.8% is assigned to the $B^0 \to \omega\gamma$ mode.

8.13 Simultaneous Fit Models

In addition to the systematic errors resulting from the signal PDFs (Section 8.9) and the B-background PDFs (Section 8.10) the other sources of systematic error also have to be taken into account for the simultaneous fit models. The systematic errors that are identical for all three individual modes simply factor out in the formula connecting the individual signal yields in the simultaneous fit, see Equation 7.6. Thus, those systematic uncertainties remain the same in the simultaneous fit models.

But there are also systematic errors which are mode dependent. The systematic uncertainties effect the signal efficiency. And the signal efficiency is build into the simultaneous fit models, see Section 7.6. Thus, the different systematic errors from the individual modes are taken into account by implementing them as efficiency corrections directly into the simultaneous fit models. The efficiencies are varied in three different ways:

- one at a time for each of the three modes individually,
- all three signal efficiencies corrected up at the same time and,
- all three signal efficiencies corrected down at the same time.

The maximal shift in the signal yield resulting from these three kinds of variations is taken as the systematic error. These errors are summarized in Table 8.4.

8.14 Summary of Systematic Errors

Table 8.4 summarizes the contributions to the systematic error. These are associated with the signal reconstruction efficiency, the modeling of $B\overline{B}$ backgrounds, and the choice of fixed parameters of the fit PDFs. The latter two contribute to the uncertainties on the signal yields.

Table 8.4: Fractional systematic errors (in %) of the measured branching fractions.

Source of error	$\rho^+\gamma$	$\rho^0\gamma$	$\omega\gamma$	$\rho\gamma$	$(\rho/\omega)\gamma$
Tracking efficiency	1.0	2.0	2.0	1.4	1.5
Particle identification	2.0	4.0	2.0	2.9	2.7
Photon selection	1.9	2.6	1.7	2.2	2.1
π^0 reconstruction	3.0	-	3.0	1.9	2.5
π^0 and η veto	2.8	2.8	2.8	2.8	2.8
$\mathcal{NN}$ efficiency	1.0	1.0	1.0	1.0	1.0
$\mathcal{NN}$ shape	0.4	0.3	2.3	0.4	0.7
Signal PDF shapes	4.8	3.3	2.4	3.1	2.6
B background PDFs	3.9	2.9	9.7	3.2	3.1
$B\overline{B}$ sample size	1.1	1.1	1.1	1.1	1.1
$\mathcal{B}(\omega \to \pi^+\pi^-\pi^0)$	-	-	0.8	-	0.1
Sum in quadrature	8.1	7.4	11.6	7.0	6.9

Chapter 9

Results

This chapter describes the results from the application of the five different fit models to the on–resonance data set (Section 9.1.1), their statistical and systematic significances (Section 9.1.2), the resulting branching fractions (Section 9.1.4), the determination of the isospin violation (Section 9.2), and the extraction of the CKM parameters $|V_{td}/V_{ts}|$ (Section 9.3.1).

9.1 Branching Fractions

The branching fractions are determined from the results of the fits of the five fit models to the on–resonance data set containing 347×10^6 $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events, corresponding to an integrated luminosity of $316\,\mathrm{fb}^{-1}$. The branching fractions are determined from the resulting signal yields according to

$$\mathcal{B} = \frac{N_{\mathrm{sig}}}{\epsilon_{\mathrm{sig}} \cdot N_{B\overline{B}}}, \tag{9.1}$$

where $\mathcal{B}$ is the resulting branching fraction, N_{sig} is the signal yield as determined from the fit, ϵ_{sig} is the signal efficiency as determined from signal MC corrected for differences between data and MC as determined from control samples, and $N_{B\overline{B}} = 347 \times 10^6$ is the number of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events contained in the on–resonance data sample.

9.1.1 Fit Results

Table 9.1 shows the results of the five fits and compares them to MC expectations. All fits converged successfully. In general, the number of continuum background events is ≈15% larger than the MC expectation. With these fit results, the branching fractions are calculated according to Equation 9.1

$$\begin{aligned}
\mathcal{B}\left(B^+ \to \rho^+\gamma\right) &= \left(1.10^{+0.37}_{-0.33} \pm 0.09\right) \times 10^{-6}, \\
\mathcal{B}\left(B^0 \to \rho^0\gamma\right) &= \left(0.79^{+0.22}_{-0.20} \pm 0.06\right) \times 10^{-6}, \\
\mathcal{B}\left(B^0 \to \omega\gamma\right) &= \left(0.40^{+0.24}_{-0.20} \pm 0.05\right) \times 10^{-6}, \\
\mathcal{B}\left(B \to \rho\gamma\right) &= \left(1.36^{+0.29}_{-0.27} \pm 0.10\right) \times 10^{-6}, \\
\mathcal{B}\left(B \to (\rho/\omega)\,\gamma\right) &= \left(1.25^{+0.25}_{-0.24} \pm 0.09\right) \times 10^{-6}.
\end{aligned} \tag{9.2}$$

The projections of the multi-dimensional fits to the individual fit variables are shown in Figure 9-1 to 9-6 for all three individual fits. Two sets of figures are shown, one where all components of the fit are shown (Figures 9-1, 9-3, and 9-5), and one where all background components are combined into a single curve on the plots (Figures 9-2, 9-4, and 9-6). In order to enhance the relative signal contribution in these figures, when producing a projection plot of one variable, cuts on the other variables are applied.

9.1.2 Significance

To determine the statistical significance of the fit results, a scan of the likelihood curve is performed for each fit model. This means that the signal yield is fixed to several dozen values, ranging from slightly negative numbers to about a few times the nominal signal yield. For each of these fits, the likelihood is recorded and the ratio of this likelihood L with the (maximal) likelihood of the nominal fit result L_{max} is taken. Twice the negative logarithm of this ratio is computed and drawn versus signal yield (see Figure 9-7). The statistical significance is the square root of this quantity at zero signal yield.

In the limit of Gaussian distribution of the likelihood, the quantity $\sqrt{-2\ln\frac{L\left(N_{\rm sig}\right)}{L_{max}}}$

Mode	Variable	MC expectation	Individual fit results	$B \to \rho\gamma$ simultaneous fit results	$B \to (\rho/\omega)\,\gamma$ simultaneous fit results
$B^+ \to \rho^+\gamma$	Signal yield	37.55	$42.0^{+14.0}_{-12.7}$	n/a	n/a
	Continuum background yield	1953	2226 ± 58	2226 ± 58	2226 ± 58
	B background yield	153.5	$127.3^{+38.5}_{-36.5}$	$119.8^{+37.5}_{-35.9}$	$122.8^{+37.5}_{-35.8}$
$B^0 \to \rho^0\gamma$	Signal yield	22.96	$38.7^{+10.6}_{-9.8}$	n/a	n/a
	Continuum background yield	1748	2083 ± 52	2081 ± 52	2080 ± 52
	B background yield	57.30	$39.0^{+29.0}_{-26.0}$	$46.5^{+28.5}_{-26.2}$	$49.1^{+28.5}_{-26.3}$
$B^0 \to \omega\gamma$	Signal yield	13.34	$11.0^{+6.7}_{-5.6}$		n/a
	Continuum background yield	633	749^{+38}_{-37}	n/a	751^{+38}_{-37}
	B background yield	22.26	$22.9^{+27.7}_{-26.8}$		$17.3^{+26.6}_{-26.2}$
Simultan. fits	Effective signal yield	680	n/a	$944.7^{+199.4}_{-186.9}$	871.9^{+174}_{-164}

Table 9.1: Results of the five fits. All the yields as determined from the fit of the three individual fit models and the two simultaneous fit models to the on–resonance data set of $316\,\mathrm{fb}^{-1}$ are summarized, together with the expectations from MC, scaled to $316\,\mathrm{fb}^{-1}$. The assumed branching fractions used to determine the MC expectations for the three signal modes are 1.0, 0.5, and 0.5 (in units of 10^{-6}) for the $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ mode, respectively.

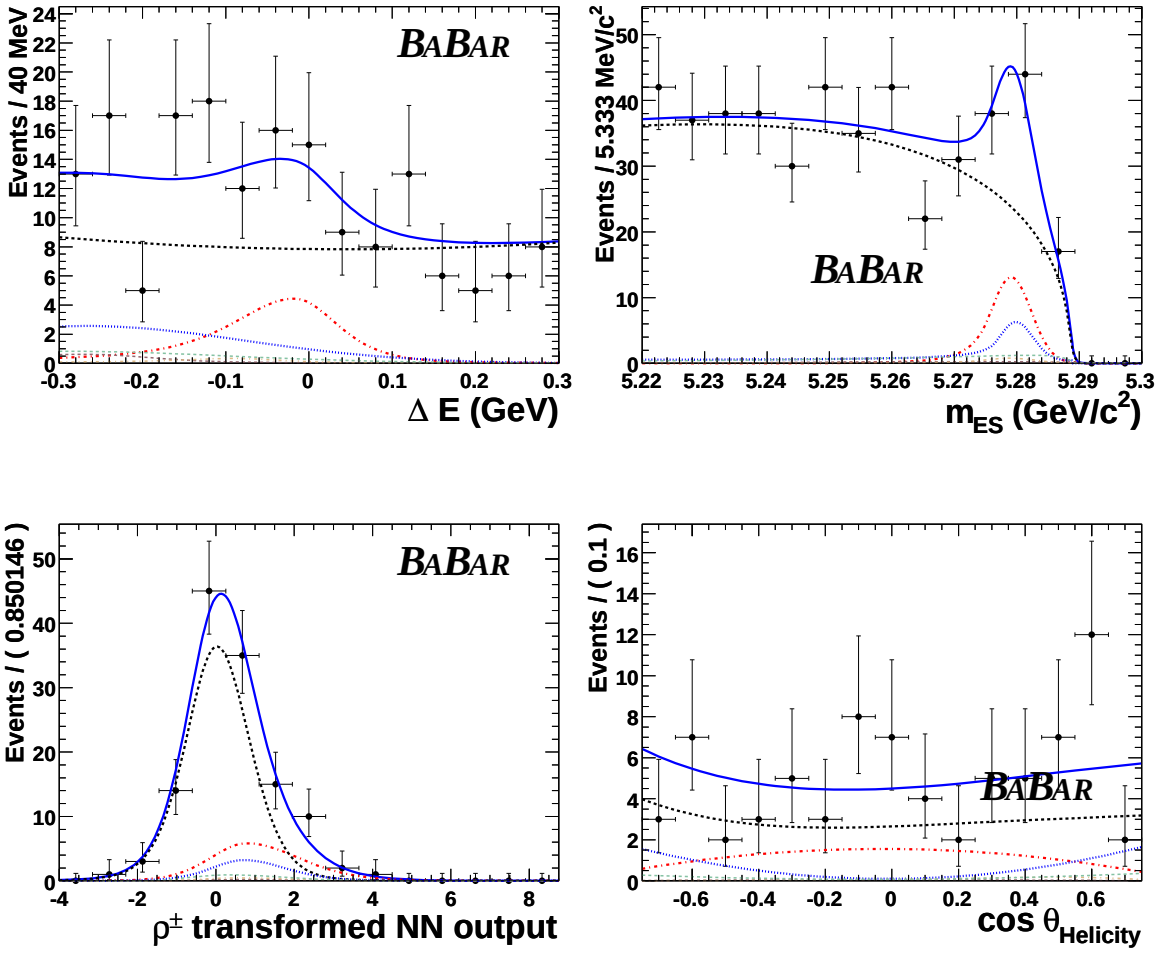


Figure 9-1: Projections for the ρ^+ mode. The solid blue curve is the full fit, the black dashed curve is the continuum background component, the light–brown curve is the $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) background, the dark–brown curve is the remaining $B \to K^*\gamma$ (excluding $K^{*+} \to K^+\pi^0$) background, the green curve is the remaining $b \to s\gamma$(excluding $B \to K^*\gamma$) background, the blue–dashed curve is all other $B\overline{B}$, and the red dashed–dotted curve is the signal component.

is the difference n of $L\left(N_{\rm sig}\right)$ *w.r.t.* L_{max} in units of the Gaussian standard deviation σ, as can be seen from comparing

$$n = \sqrt{-2\ln\frac{L\left(N_{\rm sig}\right)}{L_{max}}} \quad \Rightarrow \quad e^{-\frac{1}{2}n^2} = \frac{L\left(N_{\rm sig}\right)}{L_{max}} \tag{9.3}$$

with the Gaussian normal distribution

$$\mathcal{F}_{Gauss}(x) = C_{Gauss} \cdot e^{-\frac{(x-\langle x\rangle)^2}{2\sigma^2}} \ , \tag{9.4}$$

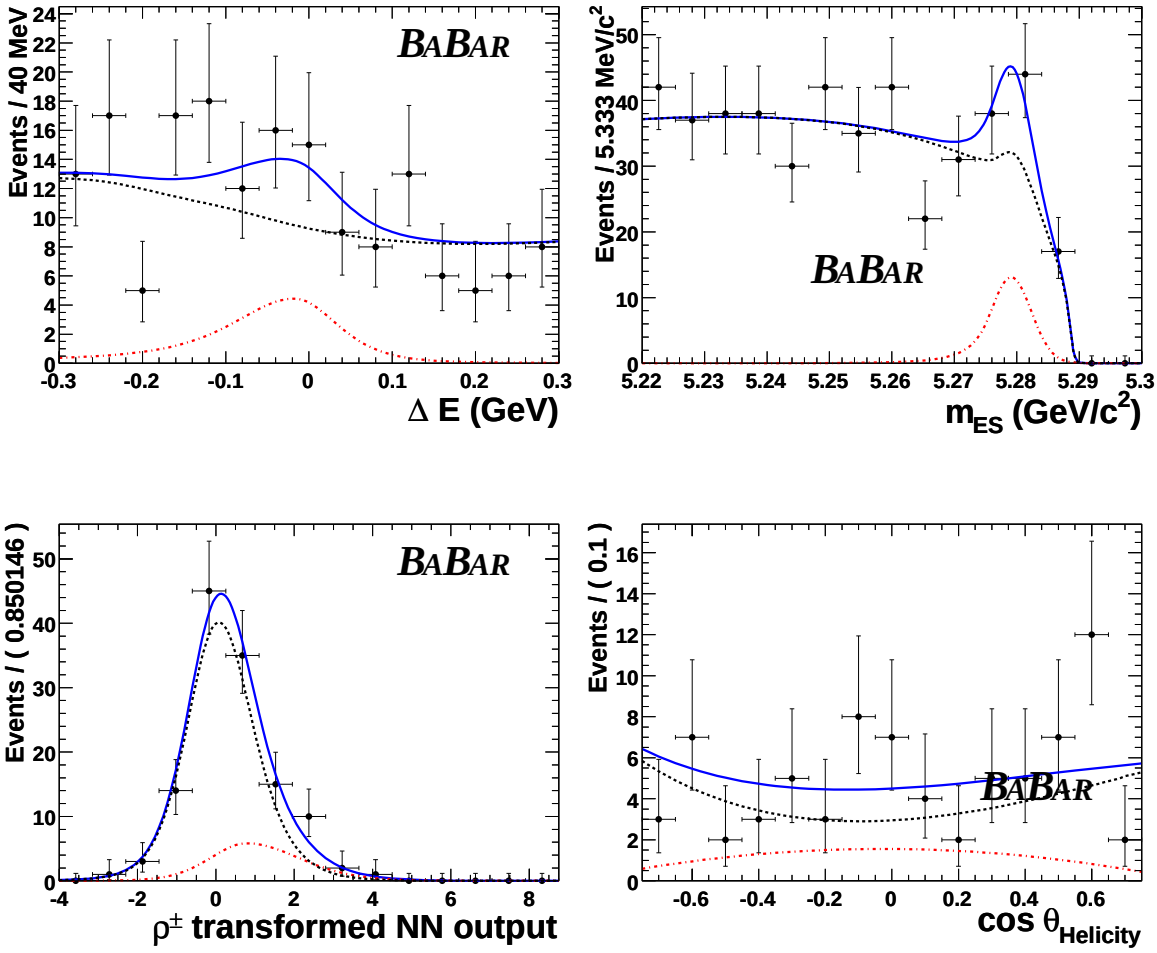


Figure 9-2: Projections for the ρ^+ mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed–dotted curve is the signal component.

where σ is the Gaussian standard deviation and $\langle x \rangle$ is the mean of the Gaussian distribution.

For the three individual mode fits, three systematic uncertainties have an effect on the signal significance, the uncertainty on the shape of the transformed neural network output (Section 8.7) and the systematic uncertainty associated with fixing the shapes of the signal PDFs (Section 8.9) and fixing the shapes and normalizations of the B background PDFs (Section 8.10). All other systematic uncertainties effect only the signal efficiency and thus do not change the significance of the fitted signal yield. The total systematic error from these three sources are 6.2% (= 2.6 events), 4.4%

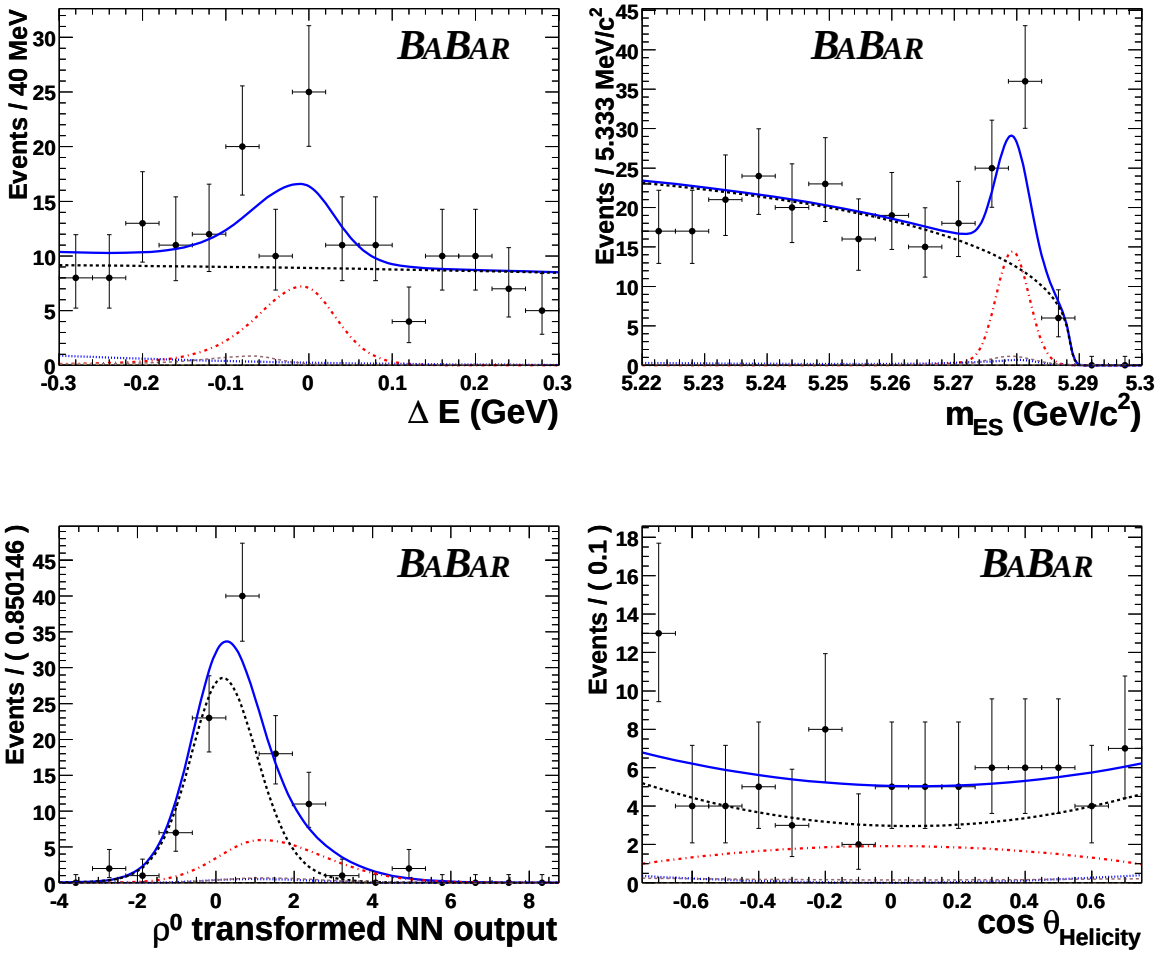


Figure 9-3: Projections for the ρ^0 mode. The solid blue curve is the full fit, the black dashed curve is the continuum background component, the dark–brown curve is the $B^0 \to K^{*0}\gamma$ background, the light–brown curve is the $B^+ \to K^{*+}\gamma$ background, the blue–dashed curve is all other $B\overline{B}$, and the red dashed–dotted curve is the signal component.

(= 1.7 events), and 10.3% (= 1.1 events) for the $B^+ \to \rho^+\gamma$ mode, the $B^0 \to \rho^0\gamma$ mode, and the $B^0 \to \omega\gamma$ mode, respectively. In order to include these systematic uncertainties into the determination of the signal significance, the likelihood curve for each individual mode is convoluted with a Gaussian where the Gaussian width is fixed to the number of events for each mode, *i.e.,* 2.6 events, 1.7 events, and 1.1 events for the $B^+ \to \rho^+\gamma$ mode, the $B^0 \to \rho^0\gamma$ mode, and the $B^0 \to \omega\gamma$ mode, respectively. The signal significance is the square root of twice the resulting convoluted negative log likelihood curve at zero signal yield. The significances are 3.8σ, 4.9σ, and 2.2σ

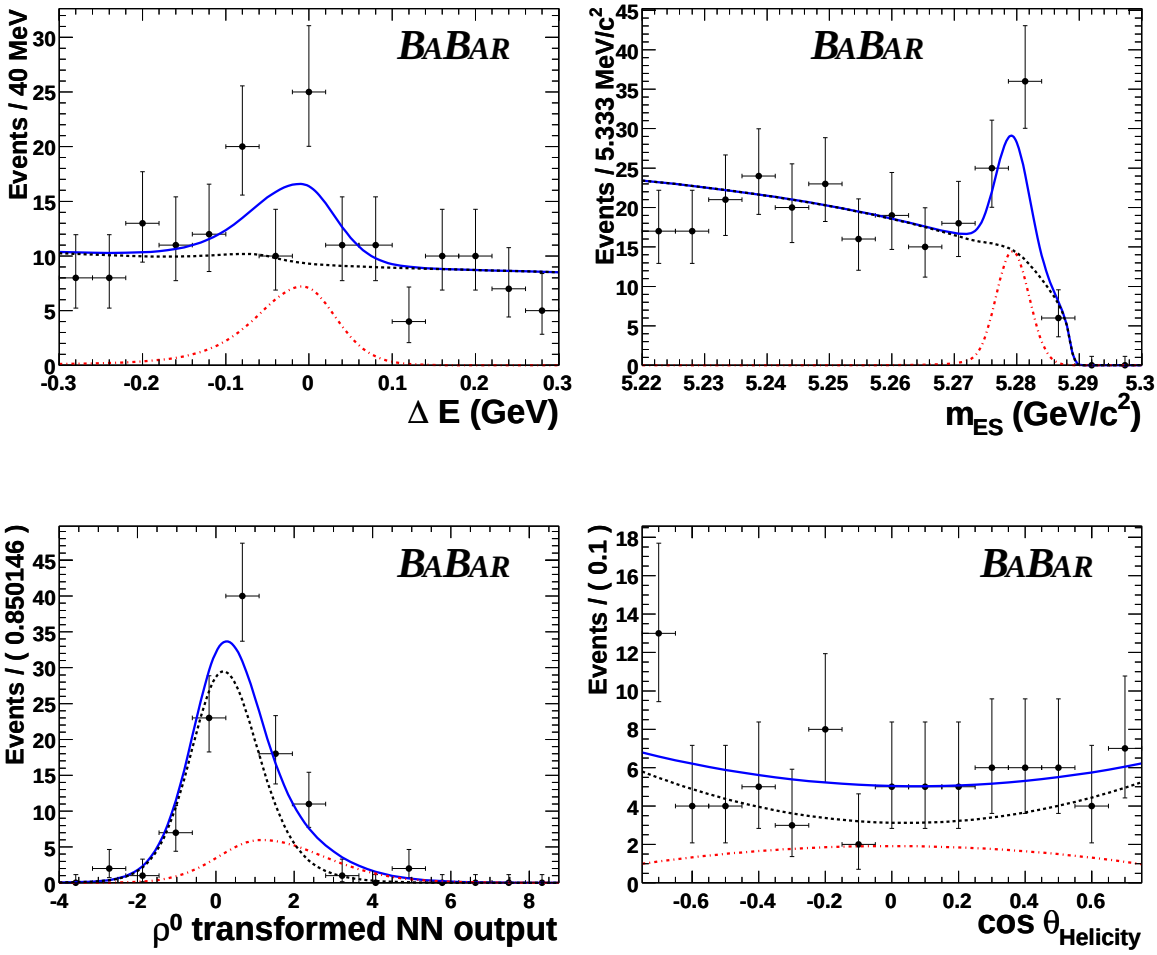


Figure 9-4: Projections for the ρ^0 mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed–dotted curve is the signal component.

for the $B^+ \to \rho^+\gamma$ mode, the $B^0 \to \rho^0\gamma$ mode, and the $B^0 \to \omega\gamma$ mode, respectively.

The systematic uncertainties for the simultaneous fits need to be included into the significance calculation in a different way since the signal efficiencies are build directly into the two simultaneous fit models.

All five likelihood curves for all five fits can be seen in Figure 9-7. Both the statistical only and the systematics included likelihood curves are shown.

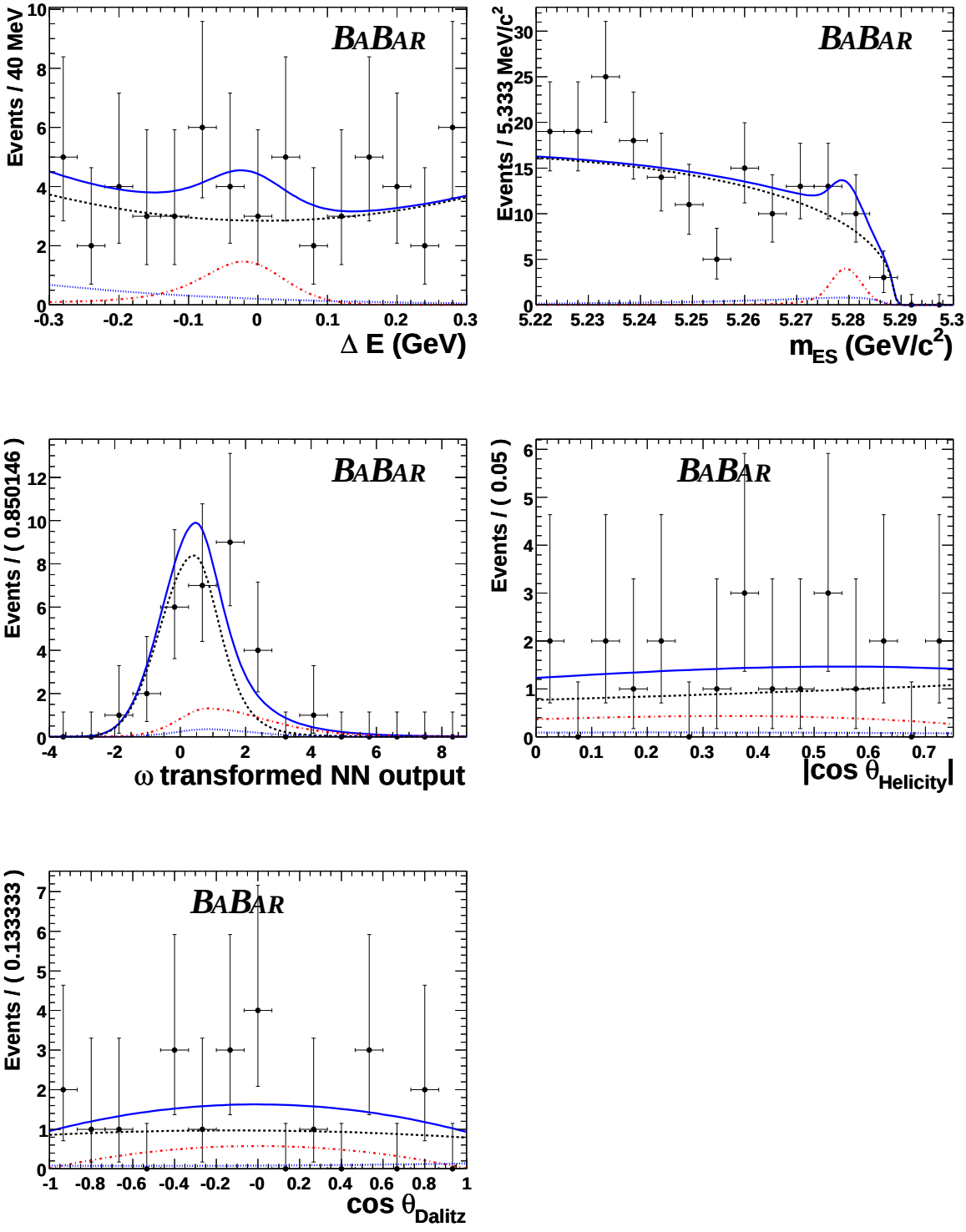


Figure 9-5: Projections for the ω mode. The blue curve is the full fit, the black dashed curve is the continuum background component, the red dashed-dotted curve is the signal component and the blue dashed curve is the $B\overline{B}$ background component.

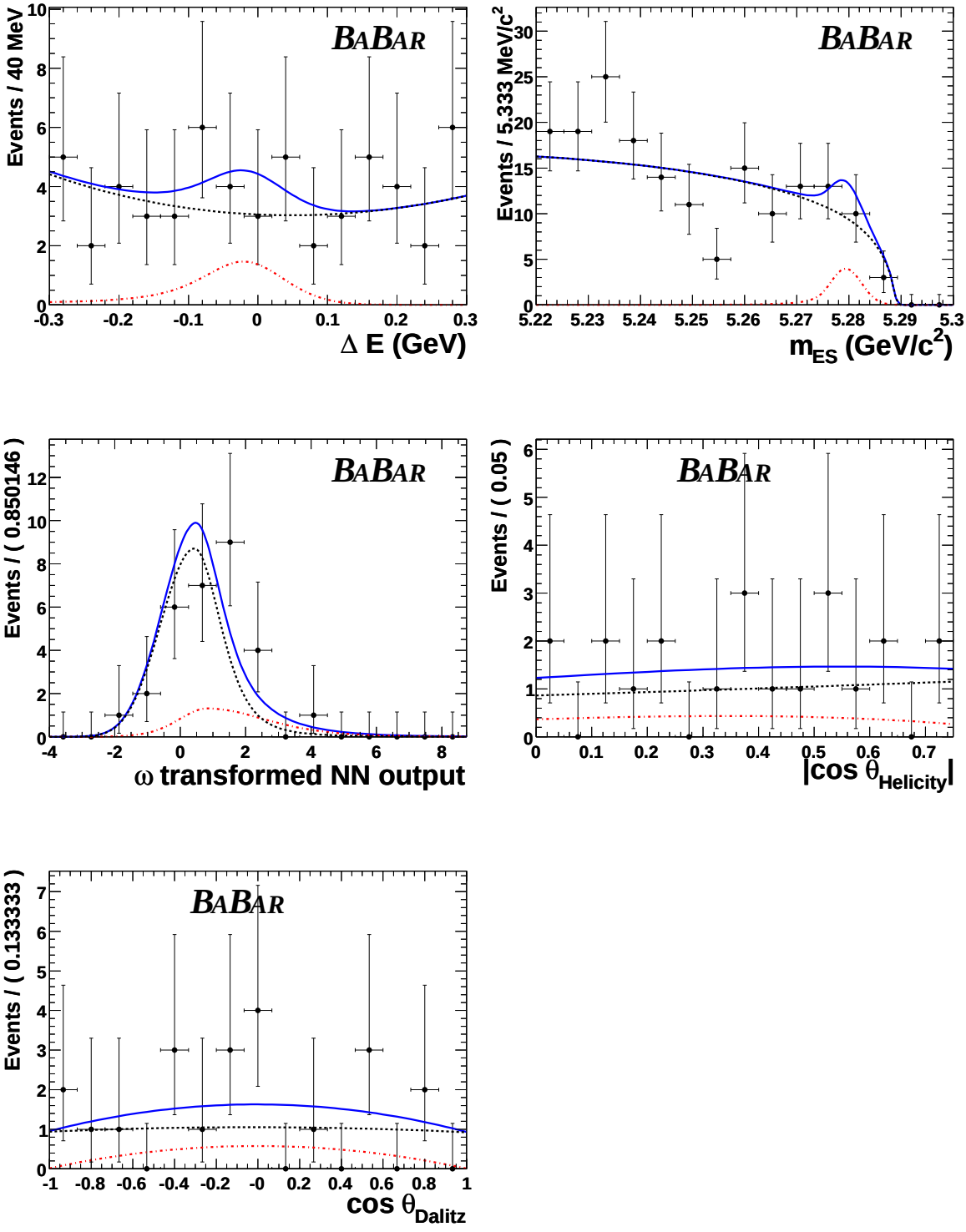
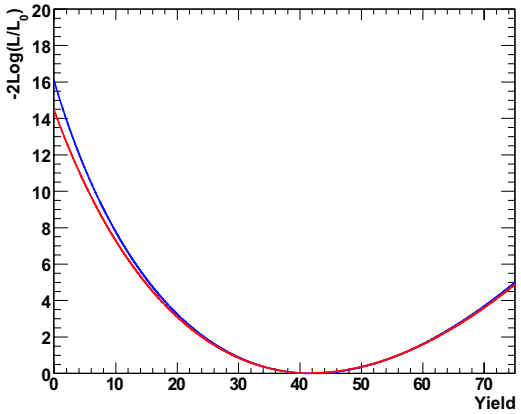
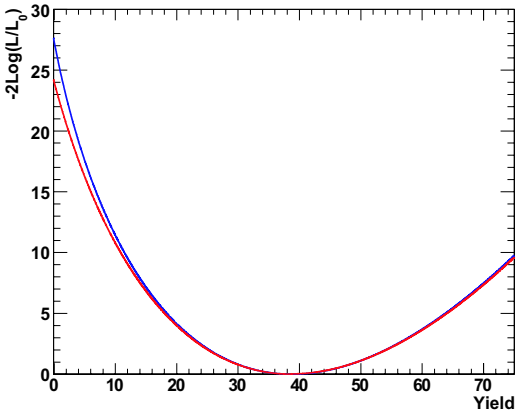


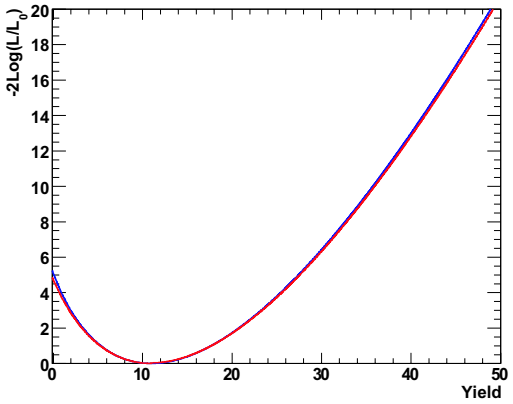
Figure 9-6: Projections for the ω mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed–dotted curve is the signal component.



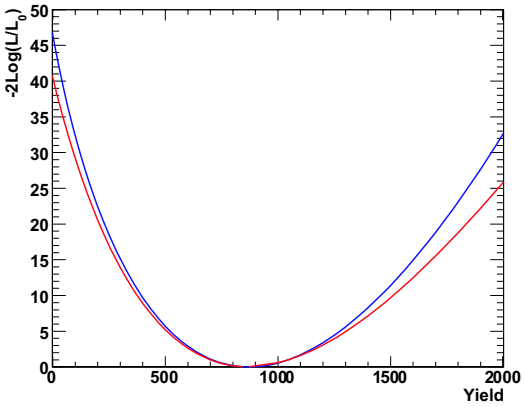
(a) The $B^+ \to \rho^+\gamma$ mode.



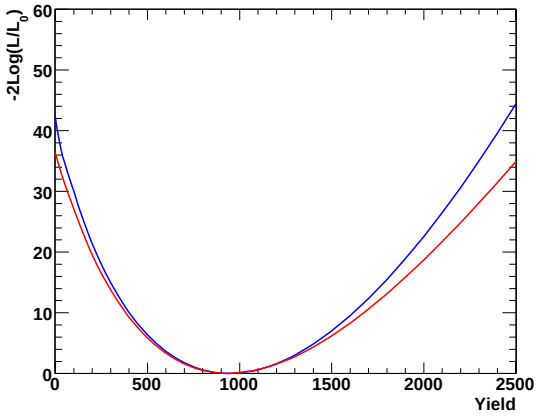
(b) The $B^0 \to \rho^0\gamma$ mode.



(c) The $B^0 \to \omega\gamma$ mode.



(d) The $B \to (\rho/\omega)\,\gamma$ simultaneous fit.



(e) The $B \to \rho\gamma$ simultaneous fit.

Figure 9-7: Likelihood curves (-2log $\mathcal{L}/\mathcal{L}_{max}$) for the fit results. The y–axis is twice the negative logarithmic likelihood, normalized to the likelihood of the nominal fit result and the x-axis shows the signal yield. The blue curve is statistics only and the red curve includes the relevant systematic errors.

9.1.3 Upper Limit

Since in the $B^0 \to \omega\gamma$ mode, no significant signal excess above zero is observed, an upper limit for the branching fraction at the 90% confidence level is calculated using a Bayesian technique. The upper limit on the branching fraction $\mathcal{B}_l$ is determined as

$$\frac{\int_0^{\mathcal{B}_l} \mathcal{L}\ d\mathcal{B}}{\int_0^{\infty} \mathcal{L}\ d\mathcal{B}} = 0.90, \tag{9.5}$$

where $\mathcal{L}$ is the likelihood curve described in the previous section and $\mathcal{B}$ is the branching fraction. A flat prior is assumed for this Bayesian calculation.

In addition to the convolution of the likelihood curve $\mathcal{L}$ with the three systematic uncertainties described in the previous section, the remaining systematic uncertainties also need to be included in the upper limit calculation. Due to the multiplicative effect on the signal efficiency of these remaining systematic uncertainties, the likelihood curve is convoluted with these errors as

$$\mathcal{L} \otimes G = \int_{-\infty}^{\infty} \mathcal{L}\left(\mathcal{B} - \mathcal{B}'\right) G\left(\mathcal{B}', \sigma\left(\mathcal{B}'\right)\right)\ d\mathcal{B}', \tag{9.6}$$

where the Gaussian G depends on the branching fraction $\mathcal{B}'$ and the width σ. σ in turn also depends on the branching fraction due to the multiplicative nature of the systematic uncertainties under consideration here, in fact, $\sigma\left(\mathcal{B}'\right) = \mathcal{B}' \cdot$ systematic error in percent.

With this method, the upper limit on the branching fraction of the decay $B^0 \to \omega\gamma$ is determined to be

$$\mathcal{B}\left(B^0 \to \omega\gamma\right) < 0.78 \times 10^{-6} \tag{9.7}$$

at the 90% confidence level.

9.1.4 Summary

A summary of all measured branching fractions, a comparison with the latest corresponding results from Belle and an indication of the expectations for these observables

Mode	$N_{\rm sig}$	$\Sigma_{\rm stat}$	Σ	$\epsilon_{\rm sig}(\%)$	$\mathcal{B}(10^{-6})$
$B^+ \to \rho^+\gamma$	$42.0^{+14.0}_{-12.7}$	4.0σ	3.8σ	11.0	$1.10^{+0.37}_{-0.33} \pm 0.09$
$B^0 \to \rho^0\gamma$	$38.7^{+10.6}_{-9.8}$	5.2σ	4.9σ	14.1	$0.79^{+0.22}_{-0.20} \pm 0.06$
$B^0 \to \omega\gamma$	$11.0^{+6.7}_{-5.6}$	2.3σ	2.2σ	7.9	$0.40^{+0.24}_{-0.20} \pm 0.05$ (< 0.78 at 90% C. L.)
Mode	$N_{\rm eff}^{sim}$	$\Sigma_{\rm stat}$	Σ	$\epsilon_{\rm sig}(\%)$	$\mathcal{B}(10^{-6})$
$B \to \rho\gamma$	$944.7^{+199.4}_{-186.9}$	6.5σ	6.0σ		$1.36^{+0.29}_{-0.27} \pm 0.10$
$B \to (\rho/\omega)\,\gamma$	871.9^{+174}_{-164}	6.8σ	6.4σ		$1.25^{+0.25}_{-0.24} \pm 0.09$

Table 9.2: Summary of the results. Listed are the signal yield $N_{\rm sig}$, the statistical significance ($\Sigma_{\rm stat}$) and the significance with systematic uncertainties included (Σ), both in standard deviations, the signal efficiency ($\epsilon_{\rm sig}$), the resulting branching fractions $\mathcal{B}$ for each mode, and the upper limit in the case of the $B^0 \to \omega\gamma$ mode. The errors on $N_{\rm sig}$ are statistical only, while for the branching fraction the first error is statistical and the second systematic.

from a recent theory calculation [9] are shown in Figure 9-8. All measured branching fractions described in this document are consistent with the theoretical expectations.

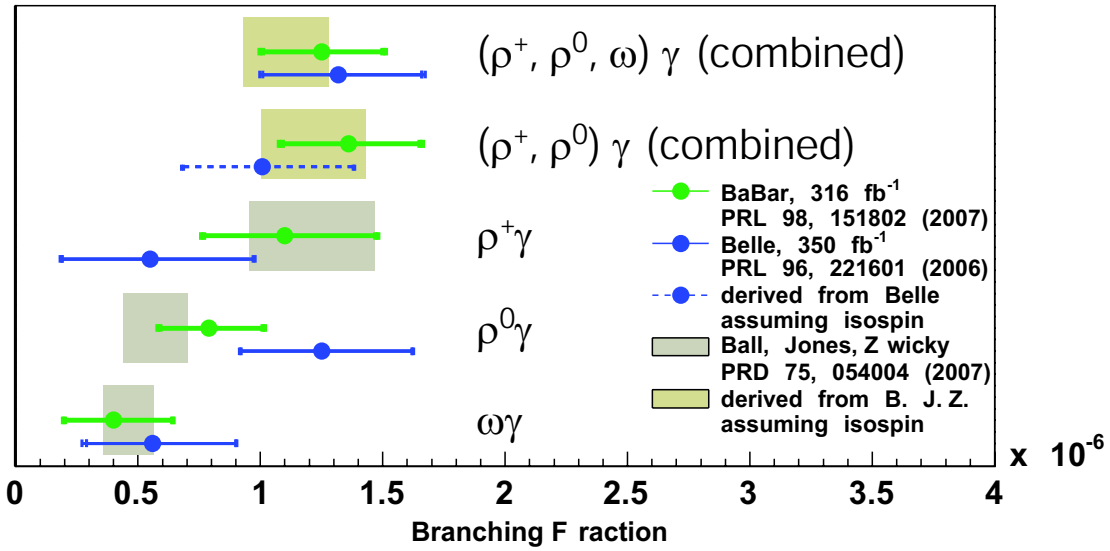


Figure 9-8: Summary of branching fraction measurements from *BABAR* and Belle.

9.2 Isospin

Since the assumption of conservation of isospin between the decays $B^+ \to \rho^+\gamma$ and $B^0 \to \rho^0\gamma$ is used in the $B \to \rho\gamma$ simultaneous fit, this assumption needs to be tested. Based on the branching fractions determined from the individual fits for these two modes, the isospin asymmetry is determined to be

$$\Delta = \frac{\Gamma(B^+ \to \rho^+\gamma)}{2\Gamma(B^0 \to \rho^0\gamma)} - 1 = \frac{\mathcal{B}(B^+ \to \rho^+\gamma)\tau_{B^0}}{2\mathcal{B}(B^0 \to \rho^0\gamma)\tau_{B^+}} - 1 = -0.35 \pm 0.27, \tag{9.8}$$

and thus consistent with zero isospin violation at the 1.3σ level.

The isospin symmetry assumption used to relate the two neutral decay modes $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$ to each other is tested by calculating the isospin asymmetry

$$A_I\,(\rho,\omega) = \frac{\Gamma(B^0 \to \omega\gamma)}{\Gamma(B^0 \to \rho^0\gamma)} - 1 = \frac{\mathcal{B}(B^0 \to \omega\gamma)}{\mathcal{B}(B^0 \to \rho^0\gamma)} - 1 = -0.49 \pm 0.63. \tag{9.9}$$

This asymmetry is consistent with zero isospin violation, but the error is too large to make a conclusive statement.

9.3 CKM Parameters

As already mentioned in the introduction, the measurements of the branching fractions discussed in the previous section can be used to extract parameters related to the CKM triangle, most notably the ratio of CKM matrix elements $|V_{td}/V_{ts}|$ (Section 9.3.1), which is related to the length of the far side of the unitarity triangle R_t.

9.3.1 $|V_{td}/V_{ts}|$

To determine $|V_{td}/V_{ts}|$, the ratio of branching fractions between $B \to (\rho/\omega)\,\gamma$ and $B \to K^*\gamma$ is needed. The following world averages are used for this calculation [22]:

- the charged branching fraction $\mathcal{B}\,(B^+ \to K^{*+}\gamma) = (4.03 \pm 0.26) \times 10^{-5}$,

- the neutral branching fraction $\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.01 \pm 0.20) \times 10^{-5}$.

With this, the ratio of branching fractions is calculated as

$$\frac{B \to (\rho/\omega)\,\gamma}{B \to K^*\gamma} = 0.030 \pm 0.006 \tag{9.10}$$

(experimental error only). Using Equation 1.8, the ratio of CKM matrix elements $|V_{td}/V_{ts}|$ is determined as

$$|V_{td}/V_{ts}|_{\rho/\omega} = 0.200^{+0.021}_{-0.020}(\text{exp.}) \pm 0.015(\text{th.}) \tag{9.11}$$

where the first error is experimental and the second theoretical. Excluding the $B^0 \to \omega\gamma$ decay from the average, the ratio of CKM matrix elements evaluates to

$$|V_{td}/V_{ts}|_{\rho} = 0.208^{+0.023}_{-0.022}(\text{exp.}) \pm 0.016(\text{th.}) \tag{9.12}$$

Combining these results with with the measurements from *Belle*[1] [25] yields the world averages

$$\begin{aligned} |V_{td}/V_{ts}|^{\text{WA}}_{\rho/\omega} &= 0.202^{+0.017}_{-0.016}(\text{exp.}) \pm 0.015(\text{th.}) \\ |V_{td}/V_{ts}|^{\text{WA}}_{\rho} &= 0.197^{+0.019}_{-0.018}(\text{exp.}) \pm 0.015(\text{th.}). \end{aligned} \tag{9.13}$$

The implications of these results to the unitarity triangle and the comparison with the measurement from B_d/B_s mixing is shown in Figure 9-9.

[1]The $B \to \rho\gamma$ branching fraction is computed using the *Belle* individual branching fractions under the assumption of isospin conservation.

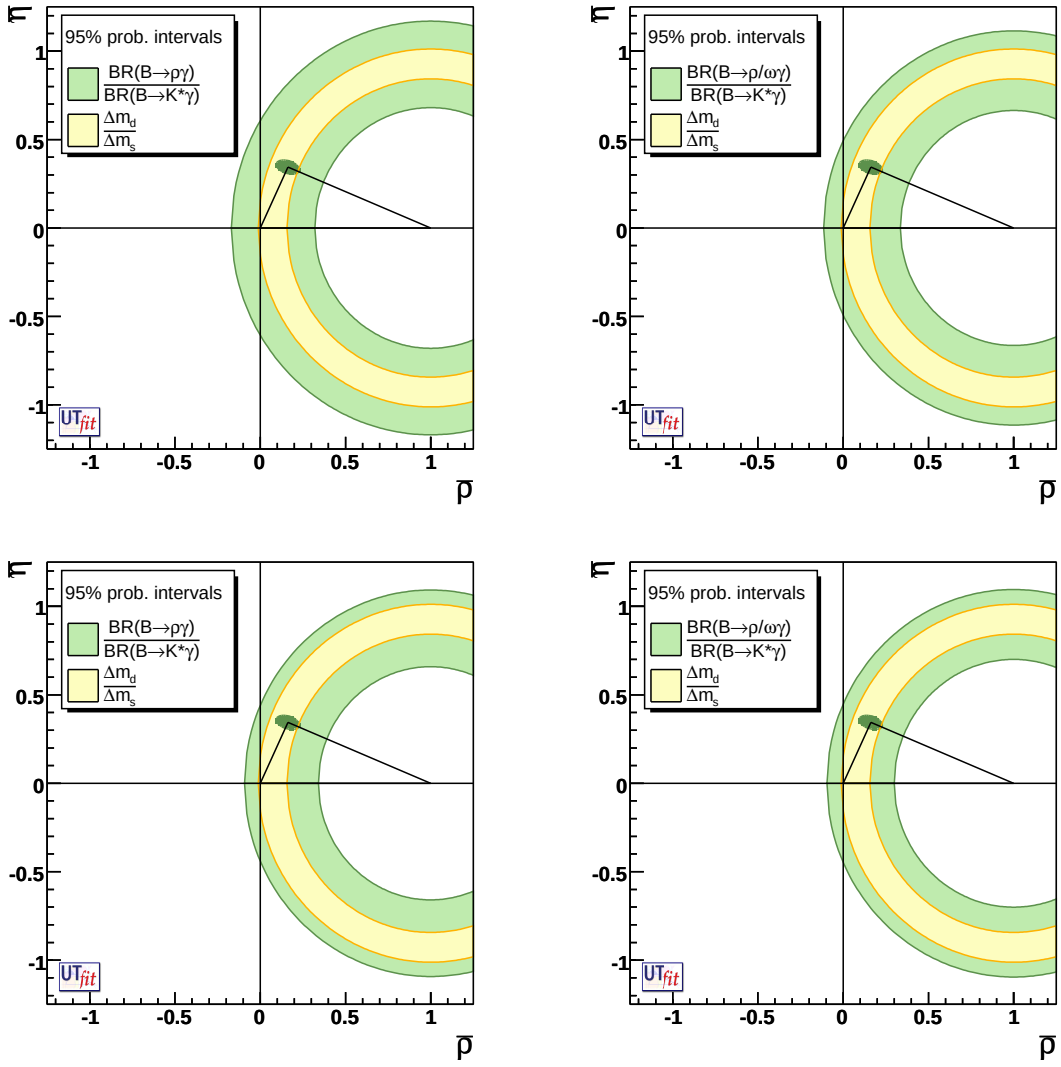


Figure 9-9: The far side of the unitarity triangle R_t. The first row shows the results from this *BABAR* measurement using $B \to \rho\gamma$ on the left and $B \to (\rho/\omega)\,\gamma$ on the right. The second row shows the average of this *BABAR* measurement with the latest *Belle* measurement [25]. All plots are produced by the UTFit group [41].

Chapter 10

Conclusions

This document describes the measurements of observables in the decay $B \to (\rho/\omega)\,\gamma$ based on $316\,\mathrm{fb}^{-1}$ data collected with the *BABAR* detector. The results were recently published in Physics Review Letters [40]. The measured branching fractions are

$$\begin{aligned}
\mathcal{B}\,(B^+ \to \rho^+\gamma) &= \left(1.10^{+0.37}_{-0.33} \pm 0.09\right) \times 10^{-6}, \\
\mathcal{B}\,(B^0 \to \rho^0\gamma) &= \left(0.79^{+0.22}_{-0.20} \pm 0.06\right) \times 10^{-6}, \\
\mathcal{B}\,(B^0 \to \omega\gamma) &= \left(0.40^{+0.24}_{-0.20} \pm 0.05\right) \times 10^{-6}, \\
\mathcal{B}\,(B \to \rho\gamma) &= \left(1.36^{+0.29}_{-0.27} \pm 0.10\right) \times 10^{-6}, \\
\mathcal{B}\,(B \to (\rho/\omega)\,\gamma) &= \left(1.25^{+0.25}_{-0.24} \pm 0.09\right) \times 10^{-6},
\end{aligned} \tag{10.1}$$

where the first error is statistical and the second systematic. A Bayesian upper limit at the 90% confidence level is set on $\mathcal{B}\,(B^0 \to \omega\gamma) < 0.78 \times 10^{-6}$. These measurements comprise the first evidence of the decay $B^+ \to \rho^+\gamma$, the first evidence of the decay $B^0 \to \rho^0\gamma$ at *BABAR*, and the world's most precise measurements of the decays $B \to \rho\gamma$ and $B \to (\rho/\omega)\,\gamma$ to date.

Using the measured branching fraction of $B \to (\rho/\omega)\,\gamma$, $B^+ \to \rho^+\gamma$, and the world average of the branching fraction of the decay $B \to K^*\gamma$ [22], the ratio of CKM matrix elements $|V_{td}/V_{ts}|$ is extracted by means of Equation 1.8 as

$$\begin{aligned}
|V_{td}/V_{ts}|_{\rho/\omega} &= 0.200^{+0.021}_{-0.020}(\mathrm{exp.}) \pm 0.015(\mathrm{th.}) \\
|V_{td}/V_{ts}|_{\rho} &= 0.208^{+0.023}_{-0.022}(\mathrm{exp.}) \pm 0.016(\mathrm{th.}),
\end{aligned} \tag{10.2}$$

where the first error is experimental and the second theoretical. Combined with the *Belle* results, this yields the world averages

$$\begin{aligned} |V_{td}/V_{ts}|^{\rm WA}_{\rho/\omega} &= 0.202^{+0.017}_{-0.016}({\rm exp.}) \pm 0.015({\rm th.}) \\ |V_{td}/V_{ts}|^{\rm WA}_{\rho} &= 0.197^{+0.019}_{-0.018}({\rm exp.}) \pm 0.015({\rm th.}). \end{aligned} \tag{10.3}$$

All the above result on $|V_{td}/V_{ts}|$ are in excellent agreement with the result extracted from B_d/B_s mixing recently reported by the CDF collaboration [18]

$$|V_{td}/V_{ts}|^{\rm CDF}_{\Delta m_d/\Delta m_s} = 0.208^{+0.001}_{-0.002}({\rm exp.})^{+0.008}_{-0.006}({\rm th.}), \tag{10.4}$$

where the first error is experimental and the second theoretical. It is a remarkable success of the Standard Model that these two very different physics processes agree that beautifully.

The total experimental error of $|V_{td}/V_{ts}|^{\rm WA}_{\rho/\omega}$ is with a remarkable 8.2% already comparable to the theory error of 7.4%. In the near future, the B factories will be able to reduce the total experimental uncertainty on this quantity to a level of about 5%. However, the theoretical uncertainty will most likely not shrink at the same pace due to the already very extensive level of the calculations. The error on $|V_{td}/V_{ts}|^{\rm CDF}_{\Delta m_d/\Delta m_s}$ is already dominated by theory. This will likely not change in the foreseeable future, however future lattice QCD calculations have the potential to reduce the theory error on this quantity somewhat.

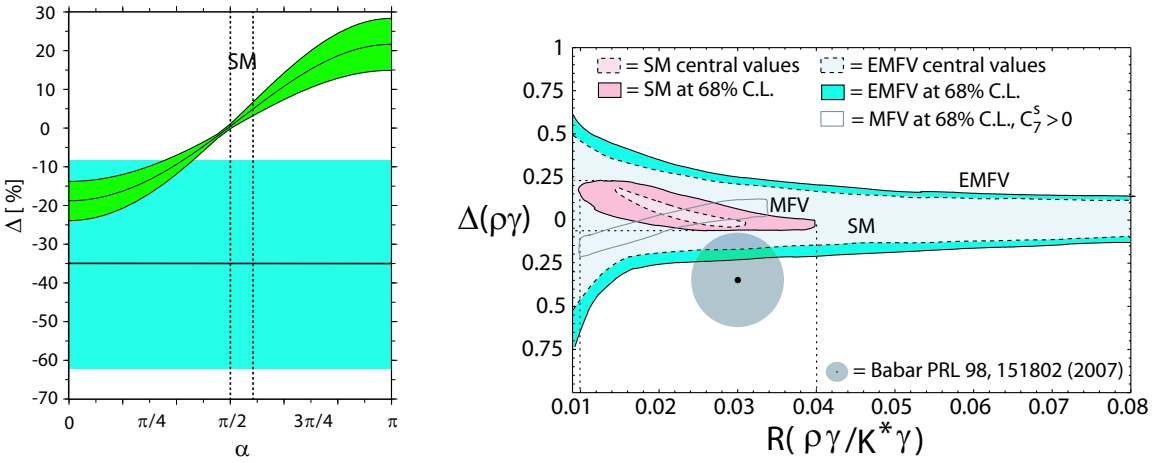
The isospin violating ratios are determined from the measured branching fractions as

$$\begin{aligned} \Delta &= \frac{\Gamma(B^+\to\rho^+\gamma)}{2\Gamma(B^0\to\rho^0\gamma)} - 1 = -0.35 \pm 0.27, \\ A_I\left(\rho,\omega\right) &= \frac{\Gamma(B^0\to\omega\gamma)}{\Gamma(B^0\to\rho^0\gamma)} - 1 = -0.49 \pm 0.63. \end{aligned} \tag{10.5}$$

These are the first measurements of isospin violation in the $B \to (\rho/\omega)\,\gamma$ system. Both are consistent with theory predictions, but due to the large uncertainties, no definite conclusions can be drawn.

As described in Section 1.1.3, the isospin violation ratio Δ is sensitive to the CKM angle α. With this measurement of Δ, the angle α can be constrained to be

less than about 68°, as can be seen in Figure 10-1(a). This very loose limit, due to the low statistical precision of the measurement of Δ has to be compared to the current SM fit result of $\alpha = \left(99.0^{+4.0}_{-9.4}\right)^\circ$ [4]. Also, Δ is a sensitive probe for new physics. One example of this can be seen in Figure 10-1(b) where again, the low statistical precision of Δ does not allow for a clear signal of physics beyond the SM. However, future measurements in the $B \to (\rho/\omega)\,\gamma$ system are expected to have smaller statistical uncertainties due to the larger data sets.



(a) (b)

Figure 10-1: Δ, the CKM angle α and new physics. The plot on the left shows the dependence of Δ on the CKM angle α in the SM (green curved band) together with this measurement of Δ (turquoise horizontal band) and the SM value for α (dotted horizontal lines) [19]. The plot on the right shows the possible impact of Δ to physics beyond the SM, explicitly to supersymmetric models with minimal flavor violation and extended minimal flavor violation [42].

With an increasing data set, future B factory measurements in the $B \to (\rho/\omega)\,\gamma$ system will also include the measurement of direct CP violation $\mathcal{A}_{\rm CP}$ in the $B^+ \to \rho^+\gamma$ decay. Current theoretical predictions depend on the computational approach, *i.e.*, calculations performed in the framework of perturbative QCD (pQCD) predict a different result than those performed using QCD factorization (QCDF) [19]:

$$\begin{aligned} \mathcal{A}_{\rm CP}\left(B^+ \to \rho^+\gamma\right) &= (17.7 \pm 15.0)\,\% \quad \text{(pQCD)} \\ \mathcal{A}_{\rm CP}\left(B^+ \to \rho^+\gamma\right) &= \left(-11.8^{+2.8}_{-2.9}\right)\,\% \quad \text{(QCDF)}. \end{aligned} \tag{10.6}$$

By the end of the decade, measurements performed at the B factories might be able to distinguish amongst these two.

In conclusion, the results presented here are all consistent with theory expectations and agree well with complementary mixing measurement in the B_d/B_s system. At the current level of precision of these measurements, no deviation from the Standard Model is observed.

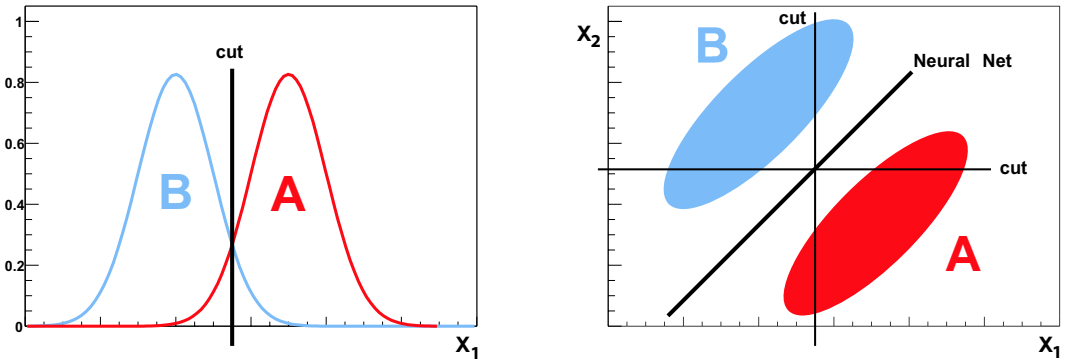
Appendix A

Introduction to Neural Networks

A neural network-based analysis has an intrinsic advantage over an analysis based on a series of cuts on individual variables. The neural network is able to consider correlations between variables and thus find a better separation between signal and background. One can, for example, consider the task of separating sample A (e.g. signal events) from sample B (e.g. background events) in the variable X_1 (see Figure A-1(a)) and the same separation of A and B in X_2. A simple cut on both variables X_1 and X_2 would not be an optimal separation of A and B. A neural net can find a better separation by considering the correlation between X_1 and X_2 (see Figure A-1(b)).

One can also imagine more complicated correlations between two variables, for example something like the situation shown in Figure A-2. The neural network is in principle able to use the complicated correlation between the different variables (here only X_1 and X_2 for simplicity) in order to find the best achievable separation between signal (A) and background (B).

A cut can be understood as a step function $\Theta(X - X')$, where events are kept if $\Theta(X - X') = 1$ and rejected if $\Theta(X - X') = 0$. The separation found by a neural net in Figure A-1(b) can be thus written as $\Theta(aX_1 + bX_2 + c)$ and the complicated



(a) Separation of two samples A and B in X_1.

(b) Two dimensional plot of variable X_1 and X_2.

Figure A-1: A one dimensional cut. (a) Separation of two samples in the variable X_1. (b) Separation of these two samples in two variables with two individual cuts and with a neural network.

shape in Figure A-2 can be understood as a combination of step functions

$$\Theta\left(\Theta\left(a_1X_1+b_1X_2+c_1\right)+\Theta\left(a_2X_1+b_2X_2+c_2\right)+\Theta\left(a_2X_1+b_2X_2+c_2\right)-2\right)\ . \tag{A.1}$$

This combination of step functions takes on the value 0 in region B and 1 in region A, thus it can be identified with the desired discrimination.

A neural net for the above example can be visualized as seen in Figure A-3. The values of two inputs X_1 and X_2 are combined at three hidden nodes Y_j, altered by weights w_{ij} and biases c_j (with $i = 1, 2$ and $j = 1, 2, 3$). The inputs y_j to these hidden nodes Y_j are computed as $y_j = \sum_i w_{ij}x_i + c_j$. This is equivalent to the argument of the step functions in example above. But the step function is replaced in a neural net with a function whose first derivative is smooth, in this case the function is the sigmoid function $\sigma\left(y_j\right) = 1/(1 + e^{-y_j})$. This so-called transfer function needs to have a smooth first derivative due to the training procedure of the neural net (see Section A.1). Also, in case of overlap regions of signal and background events, the smooth $\sigma\left(y_j\right)$ transfer function allows the neural net to assign a signal probability to each event rather than classify it absolutely like it would be done if a step function is used as the transfer function.

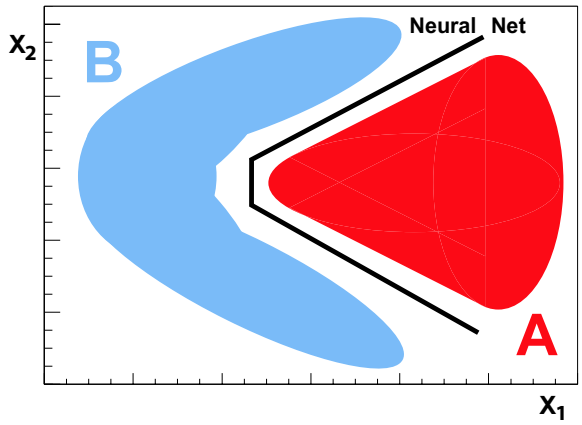


Figure A-2: A two dimensional cut compared to a neural network. A more complicated correlation between X_1 and X_2. A neural net is in principle able to find a good separation between A and B.

The outputs of the hidden nodes $g_j = \sigma(y_j)$ are then again combined with weights u_j and a single bias c_0 to the input $z = \sum_j u_j g_j + c_0$ of the last node, the output node Z. The final output of the neural net is $NN_{out} = \sigma(z)$. The complete neural net can be described in the mathematical formula

$$NN_{out} = \sigma\left(\sum_j u_j \left[\sigma\left(\sum_i w_{ij} x_i + c_j\right)\right] + c_0\right) \quad . \tag{A.2}$$

For the interested reader, [43] is a good source of information about neural networks.

A.1 Neural Network Training

The important task is now to determine the optimal combination of weights w_{ij} and u_j and biases c_j and c_0. In order to measure the performance of a given neural net, an analog to the "χ^2" for histogram fitting is used, the "sum-squared error" (SSE) is defined as:

$$SSE(w_{ij}, u_j) = \sum_{a=1}^{N} \left[NN_{out}(\vec{x}_a; w_{ij}, u_j) - F(\vec{x}_a)\right]^2 \quad , \tag{A.3}$$

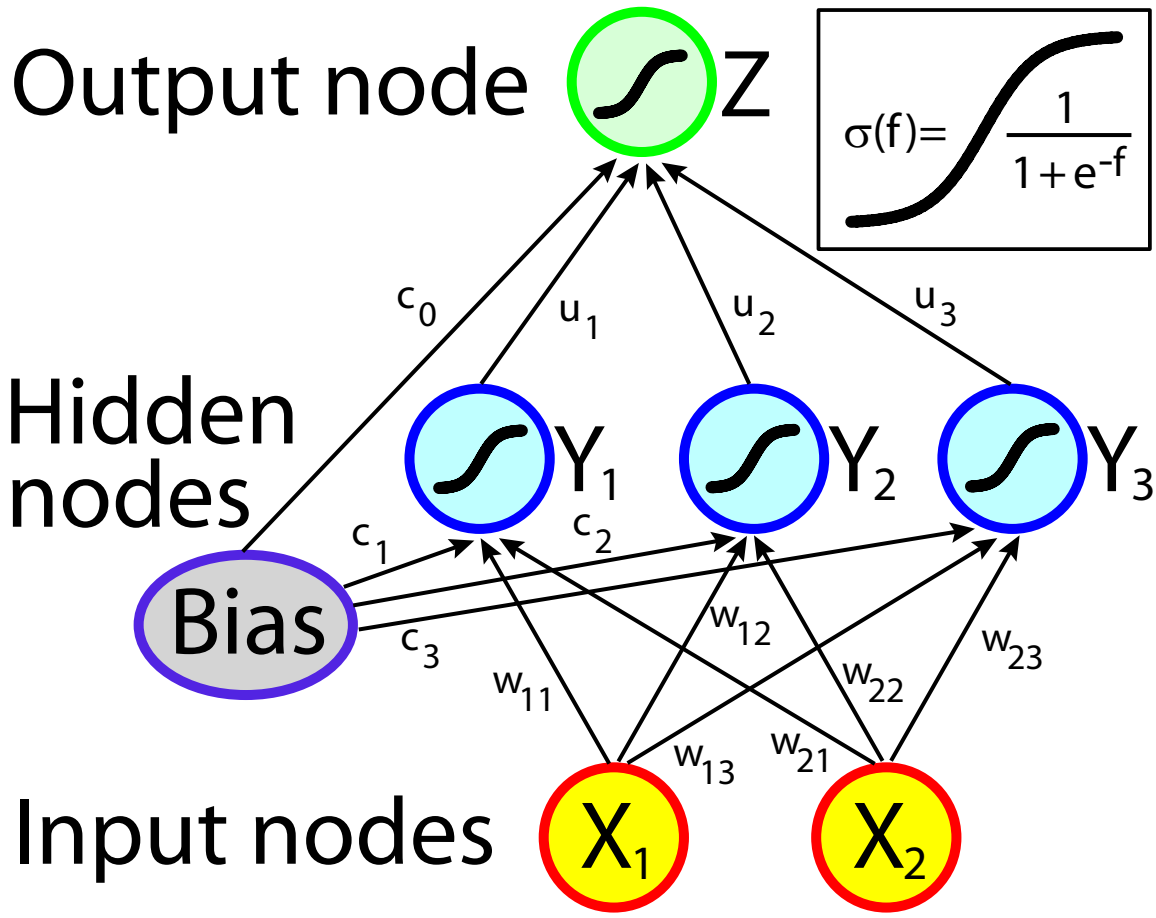


Figure A-3: Neural network visualization.

where $\vec{x}_a$ is the vector of input variables for the ath event, $NN_{out}(\vec{x}_a; w_{ij}, u_j)$ is the previously defined neural net output with the weights as parameters and $F(\vec{x}_a)$ is the desired output, e.g. 0 if the input was a background event and 1 if the input was a signal event. The SSE can be minimized in the same way as the minimization for a χ^2 fit, via gradient descent. First, the derivative of the SSE is computed relative to changes in the weights u_j, these weight coefficients are altered such that the SSE is minimized. Since the desired output of the hidden nodes is not known, the next step in the optimization is to compute the new weights u_j. With these new weights, the desired output for the hidden nodes can be computed and then, the weights from the input nodes to the hidden notes w_{ij} can be optimized in a similar way. Now, the SSE of the hidden node needs to be computed and differentiated with respect to the

weights w_{ij}. Then, the changes for these weights can be computed. This procedure is called "backpropagation", for obvious reasons.

The derivative of the SSE is needed in this procedure. Thus, the derivative of the transfer function is also needed and the step function cannot be used as a transfer function for this optimization procedure.

For the training of the neural network for this analysis, the input consists of truth matched signal MC and the same number of continuum background MC events. The continuum background sample consist of three parts, uds, $c\bar{c}$ and $\tau^+\tau^-$ events. Events from these sub samples are chosen in that number that they correspond to the same integrated luminosity. The desired neural network output for a signal event is 1 and for a continuum background event it is 0. Both samples are divided into two parts of equal size, one part is used for training the neural network and the second is used for validation. The validation with a different data sample is required in order to ensure that the neural network is not overtrained. Overtraining occurs when the neural net learns statistical fluctuations in the training sample.

The performance of the neural net is quantified by the "mean squared error" (MSE):

$$MSE = \frac{SSE}{Number\ of\ events}\ . \tag{A.4}$$

It is computed after each cycle for both the training and validation sample. If the neural net overtrains, the MSE for the validation sample does not decrease any more while the MSE for the training sample keeps decreasing after each cycle. The neural network is optimized at the minimum of the MSE for the validation sample.

Appendix B

Neural Network Input Variables

The inputs to the neural networks are examined here. The separation between signal MC and continuum MC is shown in Figures B-1 to B-1. The continuum background MC shapes of all input variables have been validated with the available off–resonance data. The shape comparison plots for this validation are shown in Figures B-7 to B-12.

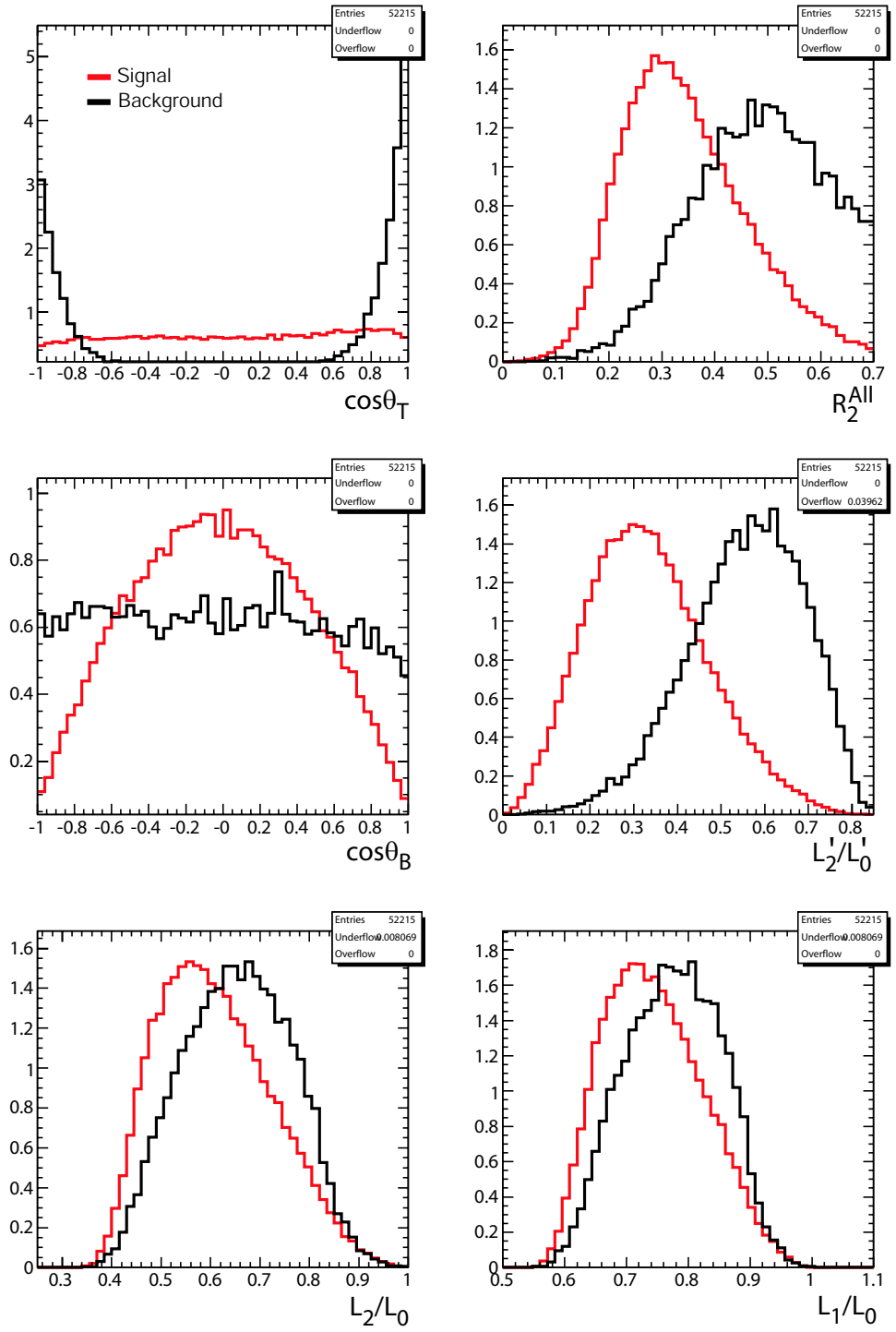


Figure B-1: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

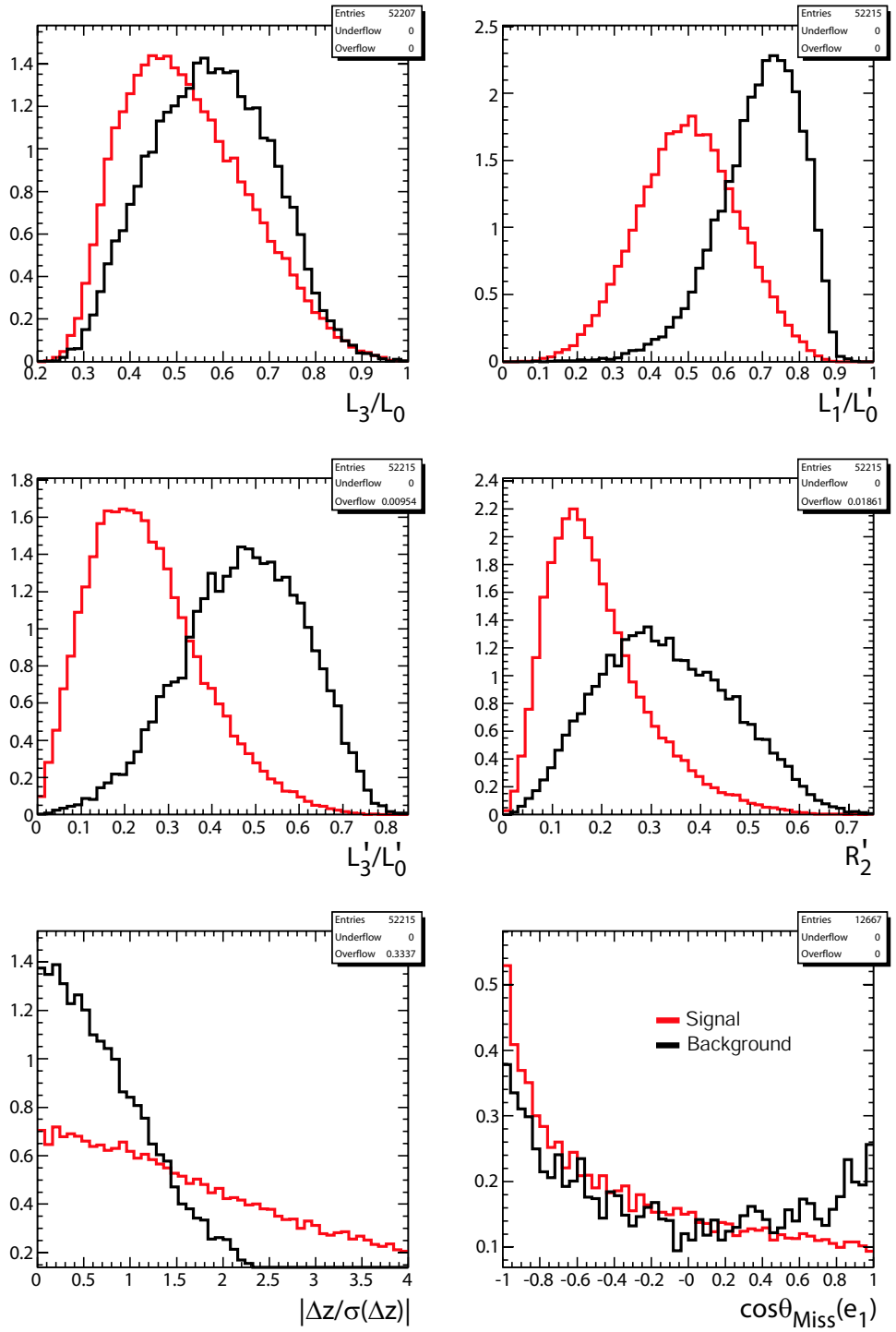


Figure B-2: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

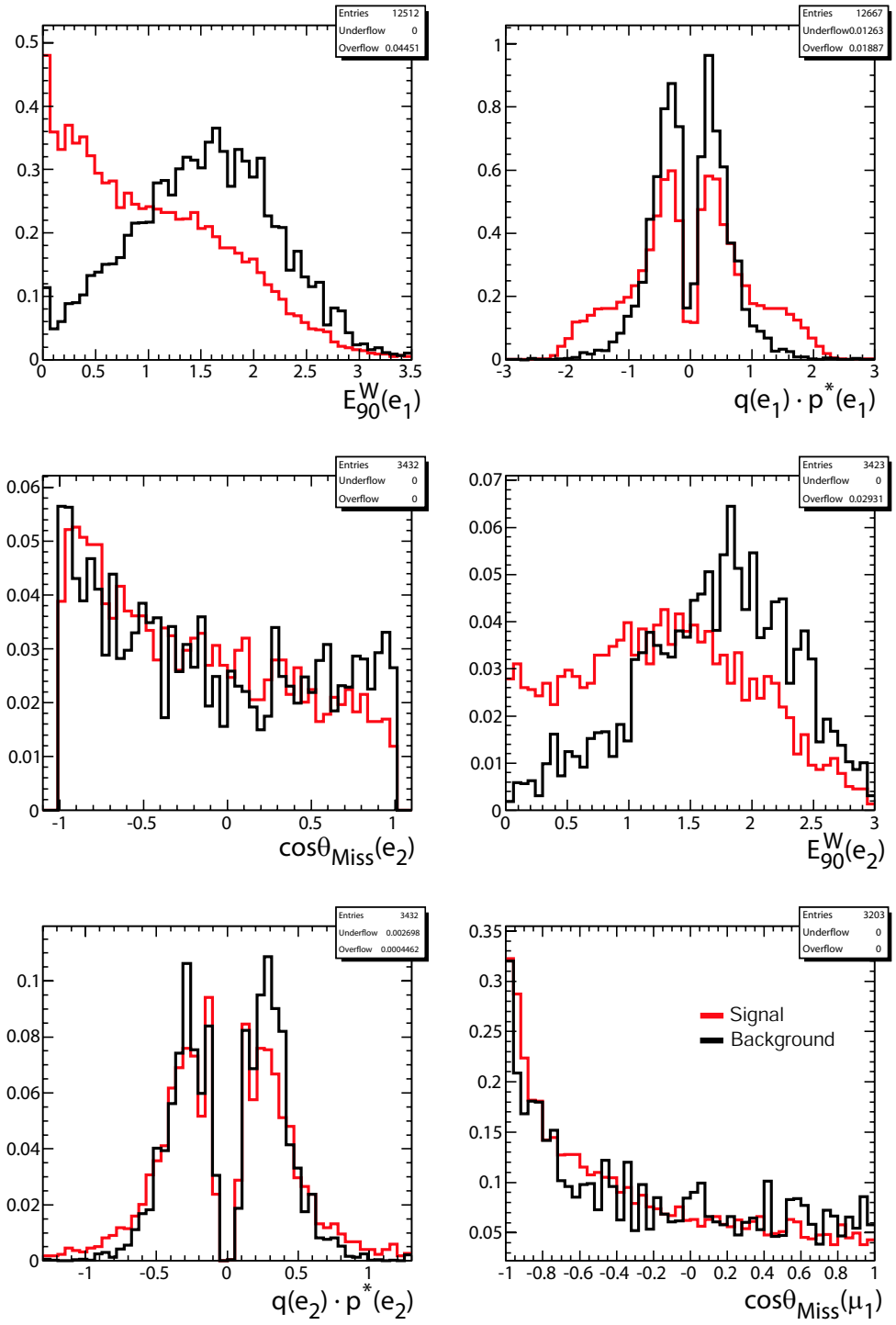


Figure B-3: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

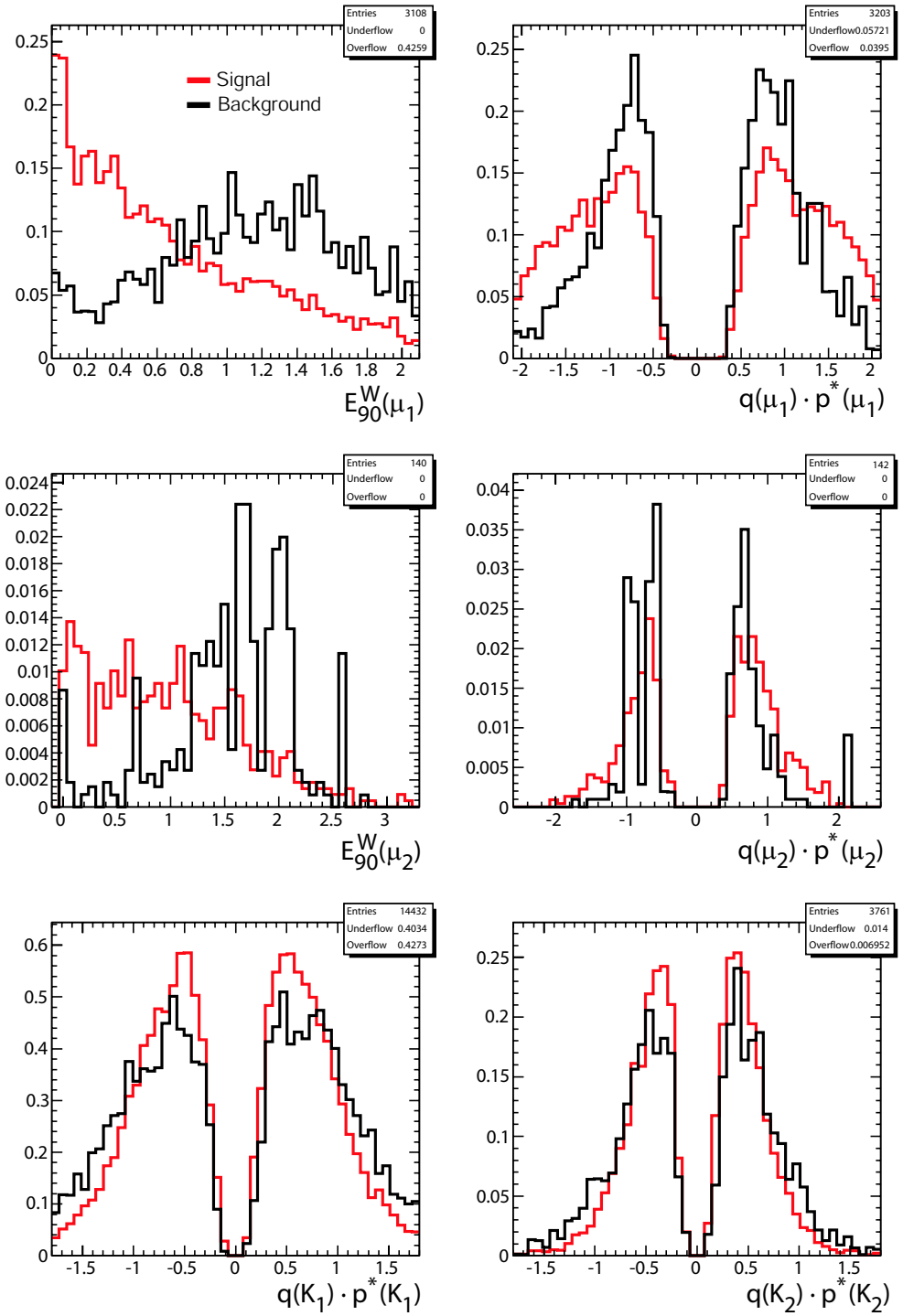


Figure B-4: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

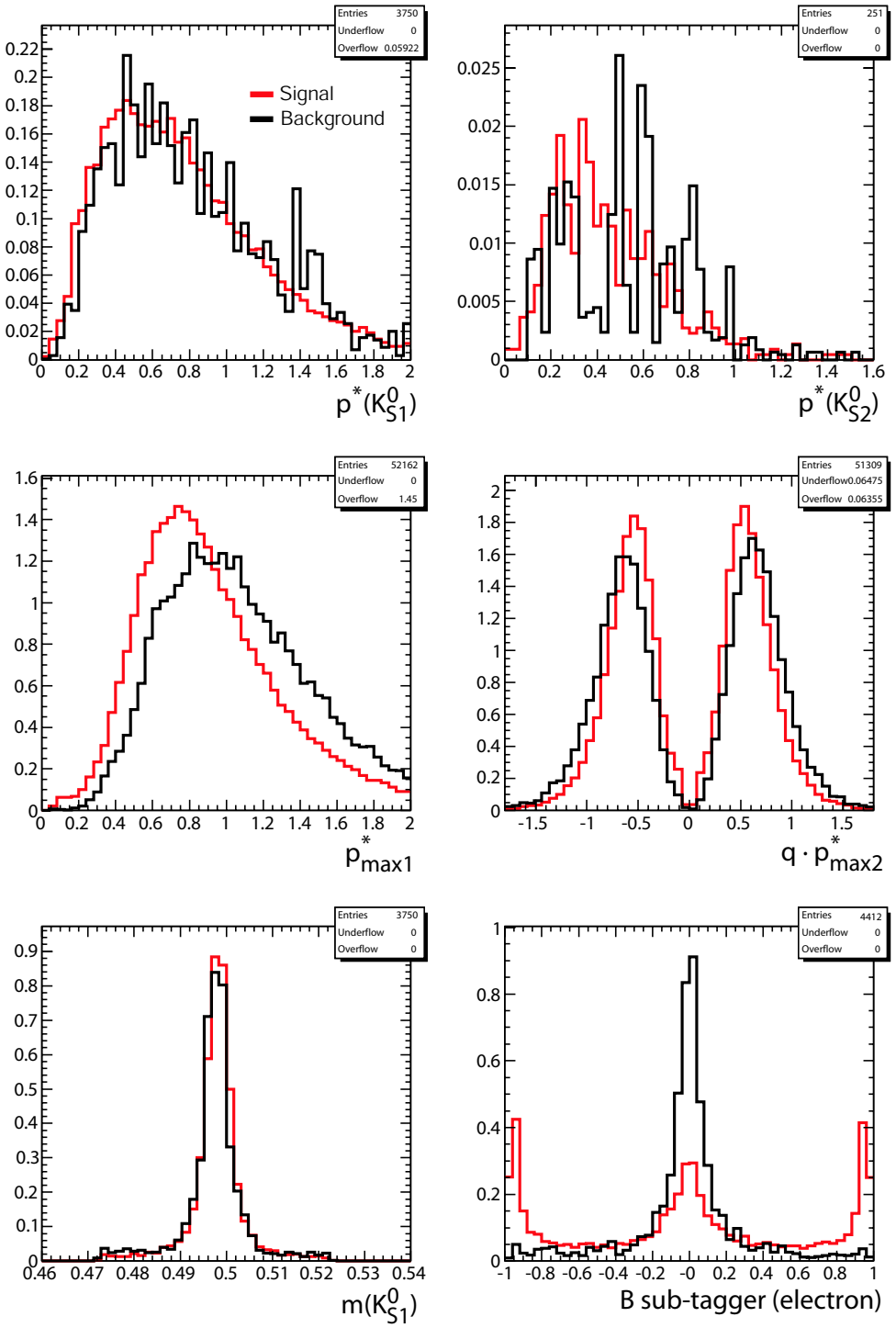


Figure B-5: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

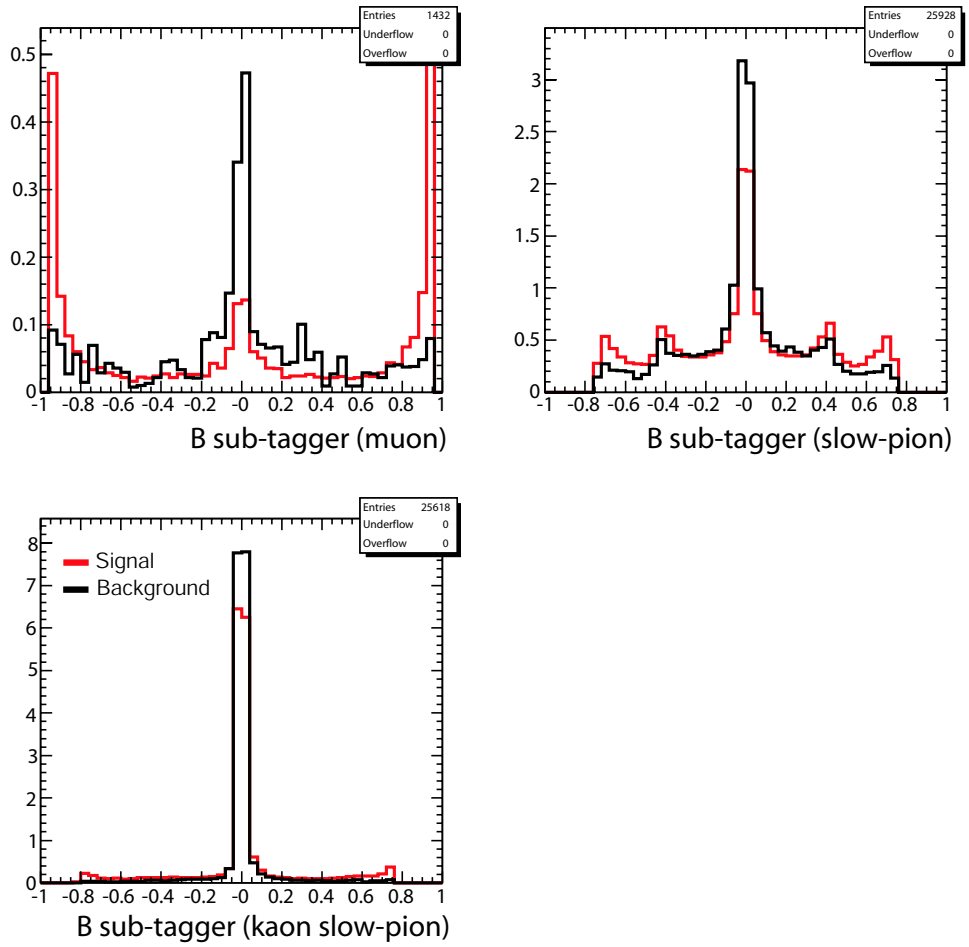


Figure B-6: Separation power of the input variables to the neural networks. Truth-matched $B^0 \to \rho^0\gamma$ signal MC is shown in red and continuum MC is shown in black. Both histograms are scaled to the same area.

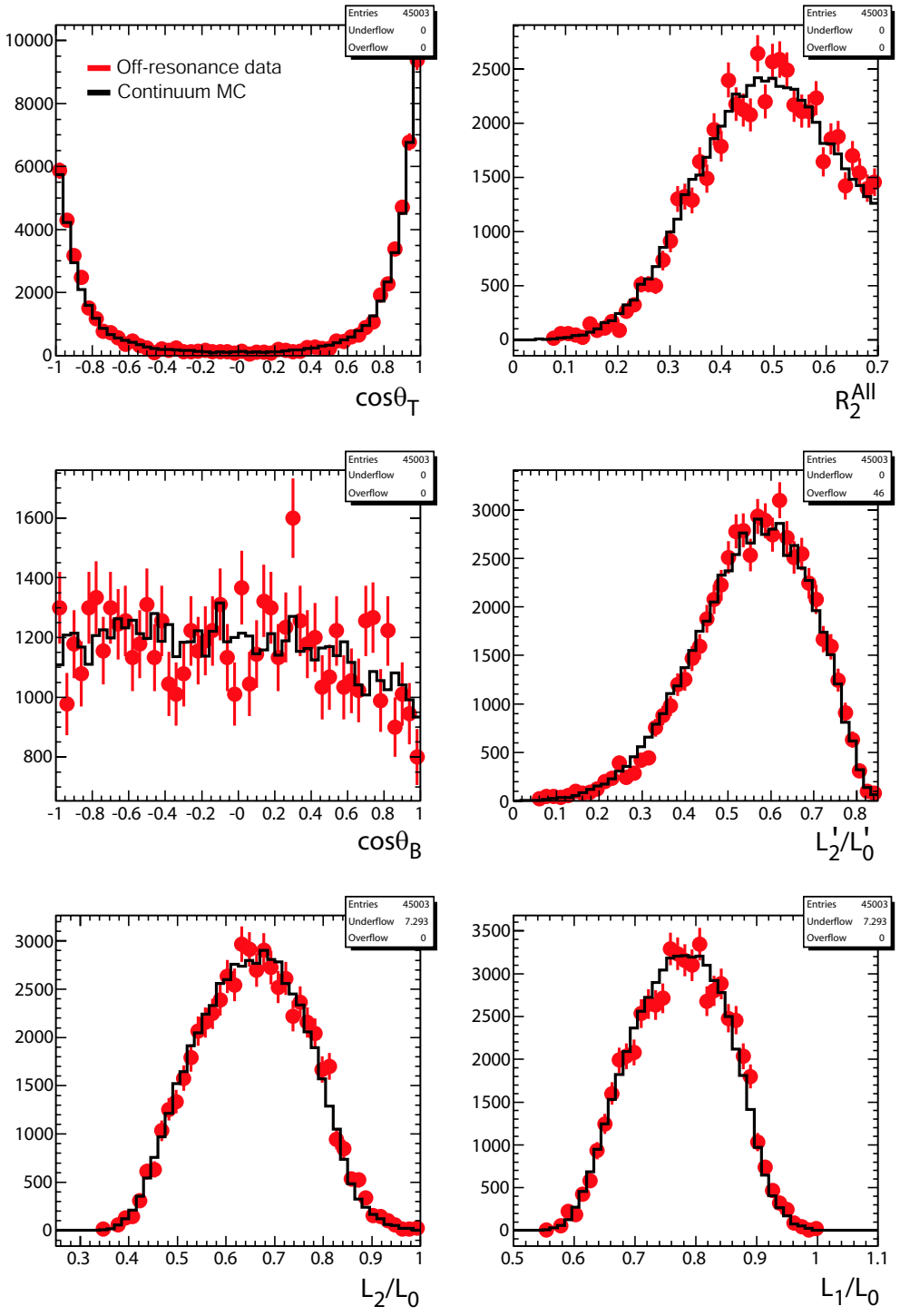


Figure B-7: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

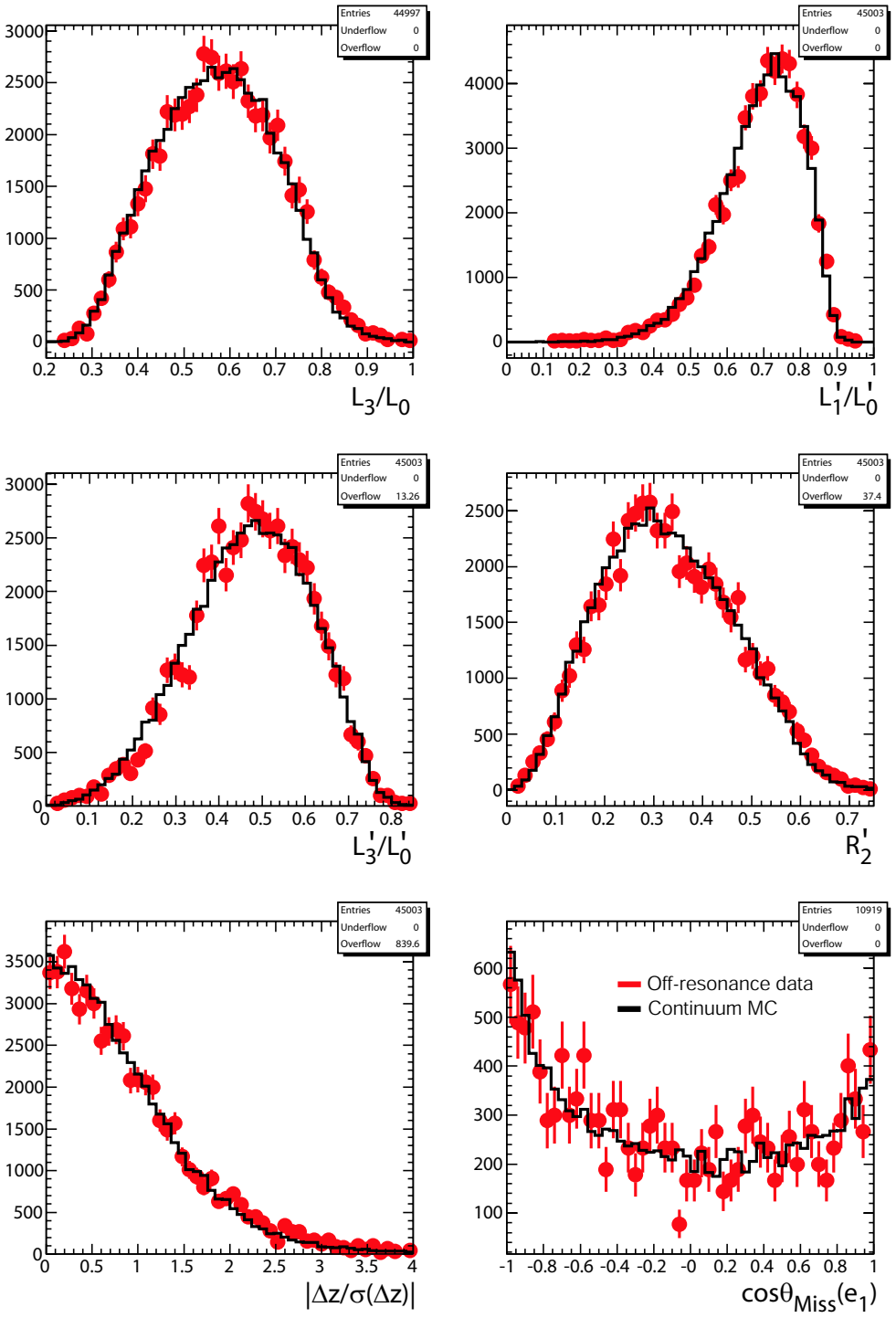


Figure B-8: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

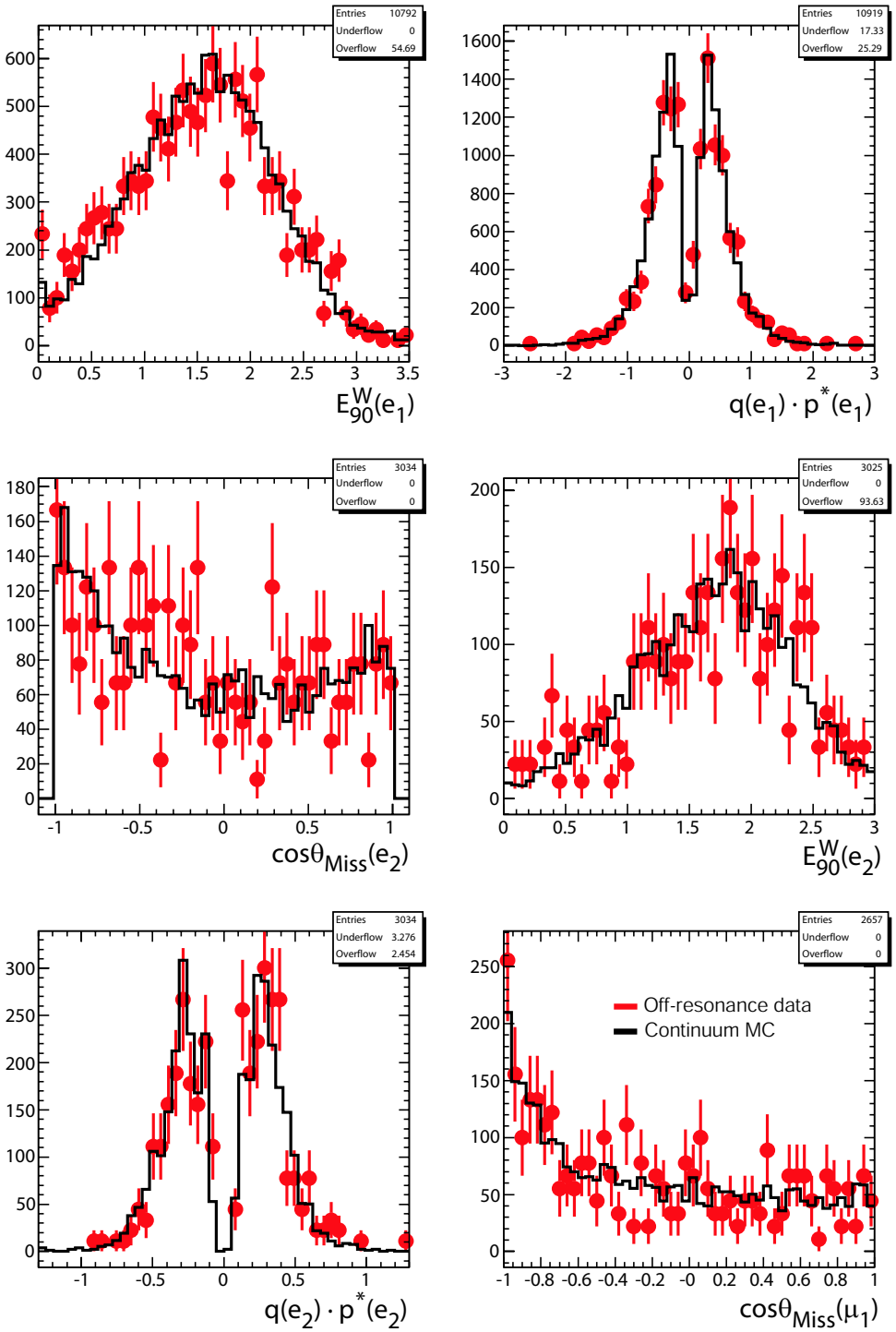


Figure B-9: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

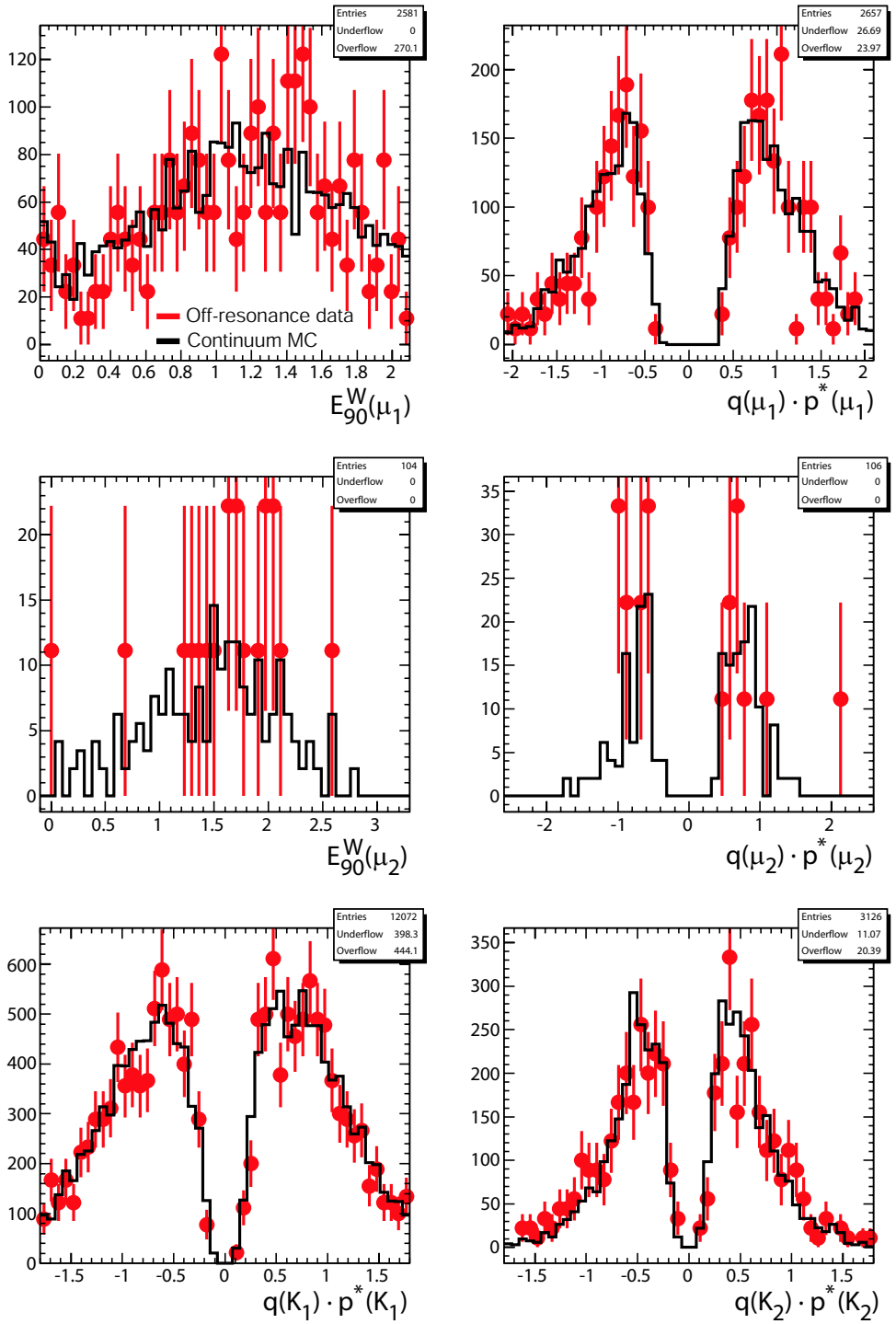


Figure B-10: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

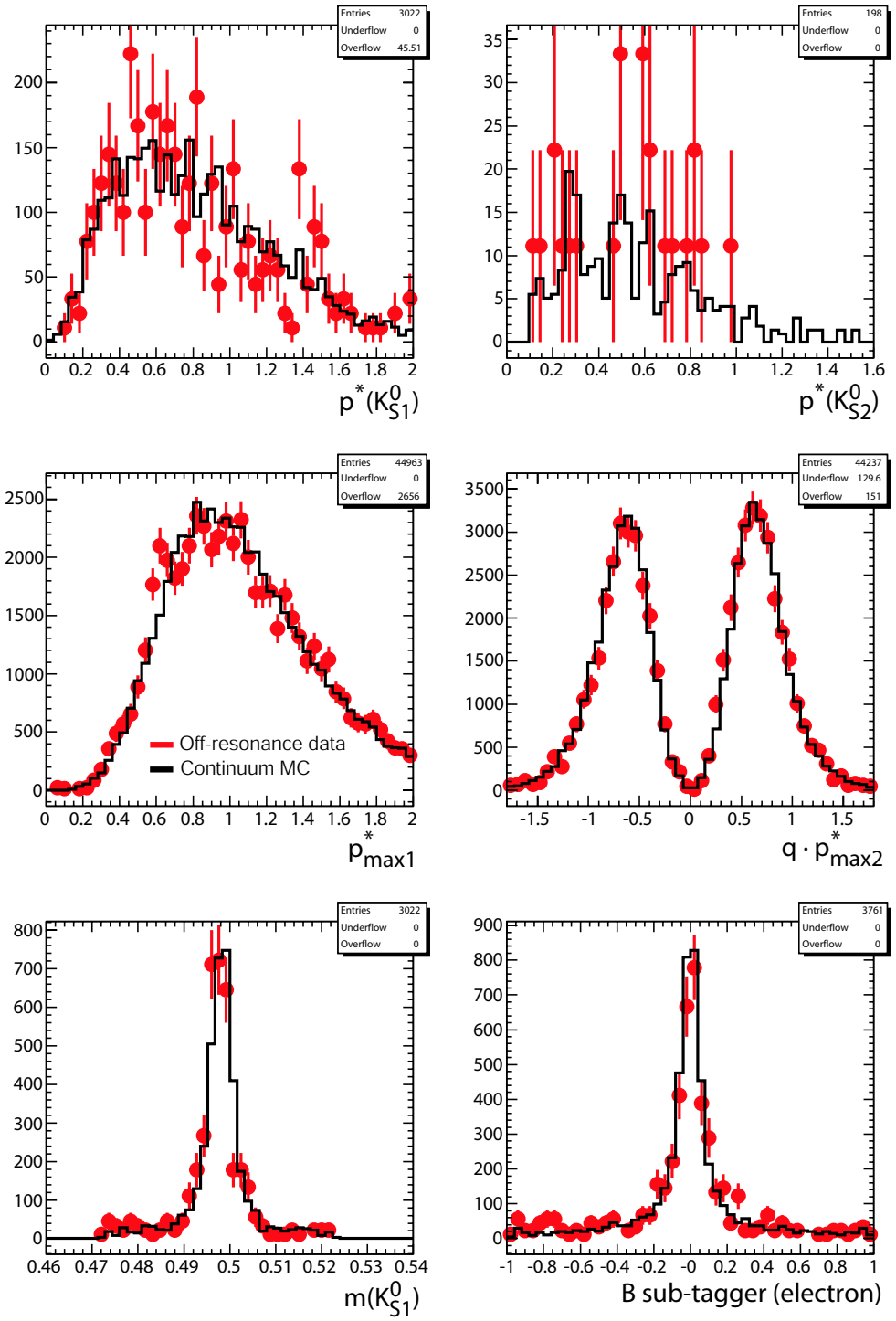


Figure B-11: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

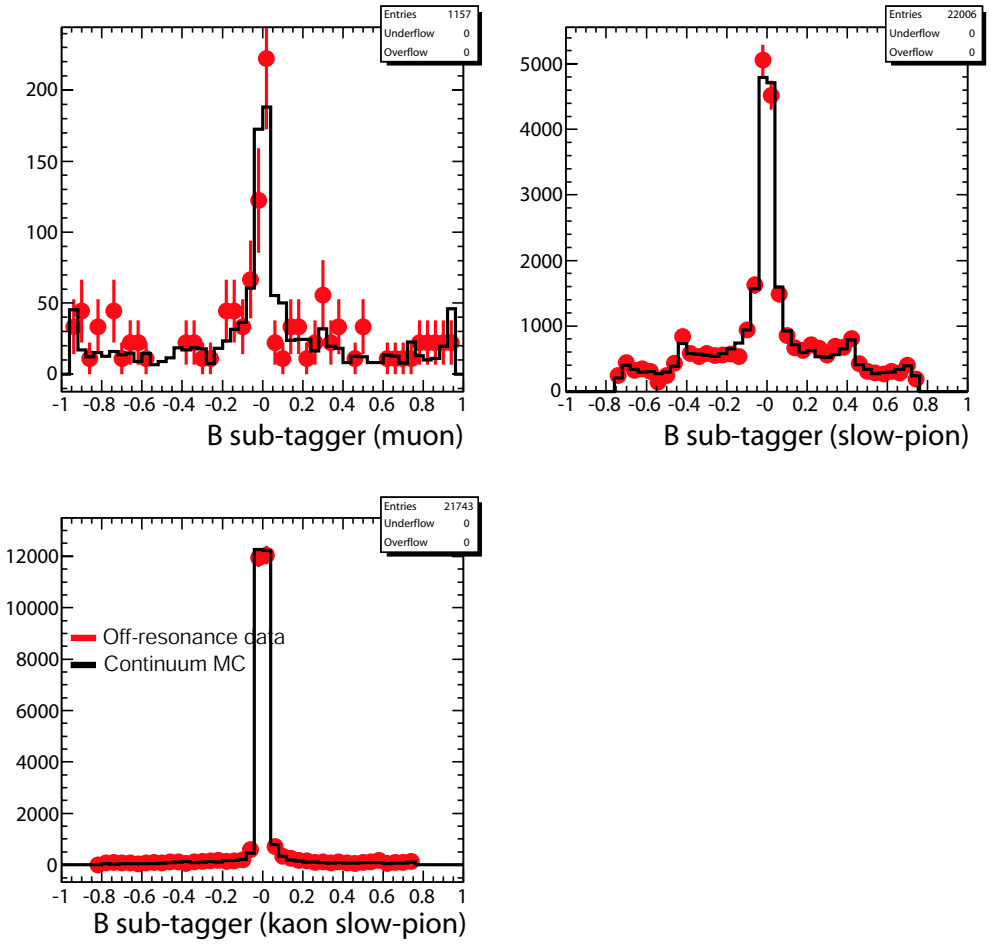


Figure B-12: Data-MC agreement of the input variables to the neural networks. Continuum MC is shown in the black histogram and off–resonance data is shown in the red data points. Both histograms are scaled to the same area.

Appendix C

Selection of Control Samples

C.1 $B \to D\pi$ for Neural Network Systematics

For the validation of the neural network, two $B \to D\pi$ samples are considered:

- $B^- \to D^0\pi^-(D \to K\pi)$ and
- $B^0 \to D^-\pi^+(D \to K\pi\pi)$.

The amount of available MC and data is shown in Table C.1.

In order to suppress background, only $B \to D\pi$ events are selected that fulfill the following requirements:

- The `GTL` list (Section 4.2.2) is used for all charged tracks except the ones from K_s^0 decays.
- A very tight likelihood based particle identification is used for selecting charged kaons.

Description	Events	Luminosity(fb^{-1})
$B^- \to D^0\pi^-, D \to K\pi$ signal MC	512,722	2,429
$B^0 \to D^-\pi^+, D \to K\pi\pi$ signal MC	1,823,330	6,361
Data		316.0

Table C.1: Dataset used in the neural net validation. Notice the luminosity of the skims shown here is only for reference.

Description	Efficiency (fit region)	Efficiency (signal region)	Expected Yields (316fb^{-1})
$B^- \to D^0\pi^-, D \to K\pi$	0.337±0.001	0.336±0.001	22408±67
$B^0 \to D^-\pi^+, D \to K\pi\pi$	0.247±0.001	0.246±0.001	22283±90

Table C.2: Signal efficiencies and expected yields for the $B \to D\pi$ control samples using the Run1-5 data set. The error on the expected yields contains the statistics error on the signal efficiency only. The uncertainties on the branching fractions are not included.

- The decay vertex of the D meson must have a χ^2 fit probability of at least 0.0001.
- The reconstructed invariant mass of the D meson must lie within $20\,\text{MeV}/c^2$ of the nominal value [22].
- Only B candidates where ΔE is within 50 MeV of zero and $m_{ES} > 5.20\,\text{GeV}/c^2$ are kept.

After these cuts are applied, there are in some cases more than one $B \to D\pi$ candidates reconstructed. Only one candidate per event is retained. This best candidate per event is chosen by selecting the one with the minimum value of $|\Delta E|$. Table C.2 shows the final signal efficiency after applying all cuts and the expected yields on the Run1-5 on–resonance data set.

Several regions in the fit variable m_{ES} are defined:

- Fit region: $5.20 < m_{ES} < 5.29\,\text{GeV}/c^2$.
- Signal region: $5.27 < m_{ES} < 5.29\,\text{GeV}/c^2$.
- Lower sideband: $5.20 < m_{ES} < 5.27\,\text{GeV}/c^2$.

As it turns out, only approximately 2% of the events in the signal region are due to other B decays, mostly $B \to D^*\pi$ where the low-energy pion or photon from the $D^* \to D\,(\pi/\gamma)$ decay is not utilized for the signal reconstruction. Thus, the signal purity in the two $B \to D\pi$ samples is very high. Also, the $B \to D^*\pi$ decays should

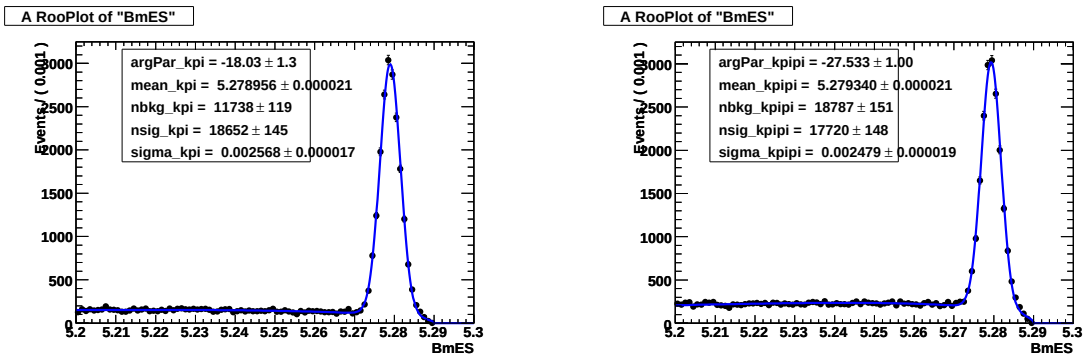


Figure C-1: The likelihood fit on the $B \to D\pi$ on–resonance data control samples. The $B^- \to D^0\pi^-(D \to K\pi)$ fit result is shown on the left and the $B^0 \to D^-\pi^+($ $D \to K\pi\pi)$ fit result is shown on the right.

Description	Events (lower sideband)	Events (signal region)	Signal yield	Background (signal region)
$B^- \to D^0\pi^-, D \to K\pi$	10130	20259	18647±145	1629±17
$B^0 \to D^-\pi^+, D \to K\pi\pi$	15784	20721	17718±148	2977±24

Table C.3: Fit results of the $B \to D\pi$ on–resonance data control samples. The number of events in the lower sideband, in the signal region, the fitted signal yields and the estimated background in the signal region for the Run1-5 data set.

have the same rest of the event as the $B \to D\pi$ decays and are thus treated as signal in the fit.

An unbinned extended maximum likelihood fit is performed in m_{ES}. The background is modeled by and ARGUS function (see Appendix E.5) and the signal peak is modeled by a Gaussian (see Appendix E.1). The results of these fits can be seen in Figure C-1

Appendix D

Particle Identification with Likelihoods

For the dE/dx information from the DCH, the likelihood is constructed using a Gaussian PDF (see Appendix E.1) that has been fitted to the pull distribution

$$\text{pull}^{DCH} = \frac{(dE/dx)^{DCH}\,(\text{measured}) - (dE/dx)^{DCH}\,(\text{expected})}{\sigma^{DCH}_{dE/dx}}. \tag{D.1}$$

The expected value of dE/dx is calculated using a parametrization of the Bethe-Bloch formula

$$(dE/dx)^{DCH}\,(\text{expected}) = \frac{x_1^{DCH} \cdot \left(x_2^{DCH} - \beta^{x_5^{DCH}} - \ln x_3^{DCH} + (\beta\gamma)^{-x_4^{DCH}} \right)}{\beta^{x_5^{DCH}}} \tag{D.2}$$

where the x_i^{DCH} are constant parameters. The error of of the dE/dx measurement in the DCH $\sigma^{DCH}_{dE/dx}$ is parameterized in terms of the polar angle θ and the transverse momentum $p_t = |\vec{p}| \cdot \sin\theta$ as

$$\sigma^{DCH}_{dE/dx} = x_6^{DCH} \cdot \left(\frac{N_{DCH}}{40} \right)^{-x_7^{DCH}} \cdot \left(\frac{|\vec{p}|}{p_t} \right)^{-x_8^{DCH}} \cdot \left(1 + \frac{x_9^{DCH}}{p_t^2} \right) \tag{D.3}$$

where again, the x_i^{DCH} are constants and N_{DCH} is the number of DCH wires that sensed the passing of the charged particle under consideration and that got also

used for reconstructing this charged particle candidate trajectory. All the constants x_i^{DCH} are determined from suitable data control samples where the particle species under investigation is already kinematically determined and thus the efficiency of the charged particle identification algorithm can be determined.

The likelihood for the dE/dx for the SVT is using a bifurcated Gaussian PDF fitted to the $(dE/dx)^{SVT}$ pull parametrization

$$\text{pull}^{SVT} = \frac{(dE/dx)^{SVT}\,(\text{measured}) - (dE/dx)^{SVT}\,(\text{expected})}{\sigma_{dE/dx}^{SVT}}. \tag{D.4}$$

Here, the expected $(dE/dx)^{SVT}$ is using a different Bethe-Bloch parametrization than the DCH which is

$$(dE/dx)^{SVT} = x_1^{SVT} \cdot \beta^{-x_2^{SVT}} \cdot \gamma_{SVT}^{x_3^{SVT}} \tag{D.5}$$

where the X_i^{SVT} are parameters and $\gamma_{SVT} = \beta/\sqrt{1-\beta^2}$ is set to 40 if it exceeds 40. The error $\sigma_{dE/dx}^{SVT}$ is parameterized in different ways

- For $N_{SVT} = 0$

$$\sigma_{dE/dx}^{SVT} = \frac{x_4^{SVT} \cdot \left(\frac{\gamma_{SVT}}{x_5^{SVT}}\right)^{-x_6^{SVT}}}{(dE/dx)^{SVT}} + x_7^{SVT} \tag{D.6}$$

- for $N_{SVT} > 0$

$$\sigma_{dE/dx}^{SVT} = \left(\frac{x_4^{SVT} \cdot \left(\frac{\gamma_{SVT}}{x_5^{SVT}}\right)^{-x_6^{SVT}}}{(dE/dx)^{SVT}} + x_7^{SVT} \right) \cdot \sqrt{\frac{5}{N_{SVT}}} \tag{D.7}$$

where the values of the parameters x_i^{SVT} for $i = 4$, 5, 6 and 7 are depending on which side of the mean $(dE/dx)^{SVT}$ the measurement lies. Also, N_{SVT} is the number of SVT planes that detected the charged particle under consideration.

Appendix E

Used Functions

E.1 The Gaussian Function

A simple symmetric Gaussian function is defined as

$$\mathcal{F}_{Gauss}(x) = C_{Gauss} \cdot e^{-\frac{(x-\langle x \rangle)^2}{2\sigma^2}} \ , \tag{E.1}$$

where $\langle x \rangle$ is the mean of the distribution, σ is its width and C_{Gauss} is the normalization constant.

E.2 The Novosibirsk Function

The Novosibirsk function is a function describing an asymmetric peak

$$\mathcal{F}_{Novo}(x) = C_{Novo} \cdot e^{-\frac{1}{2}\left(\frac{\ln^2\left(1+\tau\cdot(x-\langle x \rangle)\cdot\frac{\sinh\left(\tau\sqrt{\ln 4}\right)}{\sigma\tau\sqrt{\ln 4}}\right)}{\tau}+\tau^2\right)} \ . \tag{E.2}$$

As usual, $\langle x \rangle$ refers to the mean of the distribution and σ refers to the width. The asymmetry (w.r.t. a gaussian function) is described by the additional parameter τ, referred to as the "tail" parameter.

E.3 The Crystal Ball Function

The Crystal Ball function is a simple Gaussian at the peak center. It deviates from a Gaussian in that it has a tail attached at one side of the peak, depending on the parameter α. The tail is attached in such a way as to have continuous function and first derivative. The function is

$$\mathcal{F}_{CB}(x) = C_{CB} \cdot \begin{cases} e^{-\frac{(x-\langle x\rangle)^2}{2\sigma^2}} & \text{for} \quad x > \langle x\rangle - \alpha\sigma \\ \frac{\left(\frac{n}{\alpha}\right)^n \cdot e^{-\frac{\alpha^2}{2}}}{\left(\frac{\langle x\rangle - x}{\sigma} + \frac{n}{\alpha} - \alpha\right)^n} & \text{for} \quad x \leq \langle x\rangle - \alpha\sigma \end{cases} , \tag{E.3}$$

where $\langle x\rangle$ is the mean of the function and σ is the width. These two parameters are the same as in a usual Gaussian function. We refer to the two new parameters α and n as the "attachment point" and the "power of the tail", respectively.

E.4 The Cruijff Function

The Cruijff function is a modified Gaussian with a different width and tail on either side of the mean.

$$\mathcal{F}_{Cruijff}(x) = C_{Cruijff} \cdot \begin{cases} e^{-\frac{(x-\langle x\rangle)^2}{2\sigma_L^2 - \alpha_L(x-\langle x\rangle)^2}} & \text{for} \quad x < \langle x\rangle \\ e^{-\frac{(x-\langle x\rangle)^2}{2\sigma_R^2 - \alpha_R(x-\langle x\rangle)^2}} & \text{for} \quad x > \langle x\rangle \end{cases} , \tag{E.4}$$

E.5 The Argus Function

The Argus function was first used to describe the continuum background in m_{ES} by the ARGUS collaboration [44]

$$\mathcal{F}_{Argus}(x) = C_{Argus} \cdot \frac{x}{E_{BEAM}} \cdot \sqrt{1 - \frac{x^2}{E_{BEAM}^2}} \cdot e^{-\xi\left(1 - \frac{x^2}{E_{BEAM}^2}\right)} , \tag{E.5}$$

where ξ is the "Argus parameter" and E_{BEAM} is the "Argus endpoint". In our case, E_{BEAM} is simply half of the center-of-momentum energy ($= \frac{\sqrt{s}}{2}$).

Bibliography

[1] J. W. von Goethe, *Faust. Eine Tragödie* (Tübingen, 1808), p. 34.

[2] L. Wolfenstein, Phys. Rev. Lett. **13**, 562 (1964).

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

[4] CKMfitter Group, J. Charles *et al.*, Eur. Phys. J. **C41**, 1 (2005), hep-ph/0406184.

[5] J. R. Ellis, M. K. Gaillard, D. V. Nanopoulos, and S. Rudaz, Nucl. Phys. **B131**, 285 (1977).

[6] M. Battaglia *et al.*, (2003), hep-ph/0304132.

[7] A. Ali and A. Y. Parkhomenko, Eur. Phys. J. **C23**, 89 (2002), hep-ph/0105302.

[8] P. Ball and R. Zwicky, JHEP **04**, 046 (2006), hep-ph/0603232.

[9] P. Ball, G. W. Jones, and R. Zwicky, Phys. Rev. **D75**, 054004 (2007), hep-ph/0612081.

[10] *BABAR*, B. Aubert *et al.*, Phys. Rev. Lett. **88**, 221802 (2002), hep-ex/0112044.

[11] *BABAR*, B. Aubert *et al.*, Phys. Rev. **D66**, 032003 (2002), hep-ex/0201020.

[12] *BABAR*, B. Aubert *et al.*, Phys. Rev. Lett. **88**, 221803 (2002), hep-ex/0112045.

[13] *BABAR*, B. Aubert *et al.*, Phys. Rev. **D67**, 072002 (2003), hep-ex/0212017.

[14] *BABAR*, B. Aubert *et al.*, Phys. Rev. **D73**, 012004 (2006), hep-ex/0507054.

[15] *Belle*, N. C. Hastings *et al.*, Phys. Rev. **D67**, 052004 (2003), hep-ex/0212033.

[16] *Belle*, Y. Zheng *et al.*, Phys. Rev. **D67**, 092004 (2003), hep-ex/0211065.

[17] *Belle*, K. Abe *et al.*, Phys. Rev. **D71**, 072003 (2005), hep-ex/0408111.

[18] CDF - Run II, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 062003 (2006), hep-ex/0606027.

[19] A. Ali and A. Parkhomenko, (2006), hep-ph/0610149.

[20] S. W. Bosch and G. Buchalla, Nucl. Phys. **B621**, 459 (2002), hep-ph/0106081.

[21] Y. Li and C.-D. Lu, Phys. Rev. **D74**, 097502 (2006), hep-ph/0605220.

[22] K. Hagiwara *et al.*, Physical Review D **66**, 010001+ (2002).

[23] CLEO, T. E. Coan *et al.*, Phys. Rev. Lett. **84**, 5283 (2000), hep-ex/9912057.

[24] *BABAR*, B. Aubert *et al.*, Phys. Rev. Lett. **94**, 011801 (2005), hep-ex/0408034.

[25] K. Abe *et al.*, Phys. Rev. Lett. **96**, 221601 (2006), hep-ex/0506079.

[26] *BABAR*, e. Harrison, P. F. and e. Quinn, Helen R., 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.

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

[28] R. Santonico and R. Cardarelli, Nucl. Instrum. Meth. **187**, 377 (1981).

[29] BABAR LST, G. Cibinetto, Int. J. Mod. Phys. **A20**, 3834 (2005).

[30] BaBar LST, M. Andreotti, Prepared for 9th ICATPP Conference on Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications, Villa Erba, Como, Italy, 17-21 Oct 2005.

[31] W. Menges, (2006), physics/0609039.

[32] A. Ryd *et al.*, The evtgen package home page, World Wide Web, `http://www.slac.stanford.edu/~lange/EvtGen/`.

[33] T. Sjostrand, CERN-TH-7112-93.

[34] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[35] T. Brandt, *Electron Identifcation and Measurement of the Inclusive Semileptonic Branching Fraction of B Mesons at the BABAR Experiment*, PhD thesis, Technische Universität Dresden, 2001.

[36] University of Stuttgart and University of Tübingen, Stuttgart neural network simulator, World Wide Web, `http://www-ra.informatik.uni-tuebingen.de/SNNS/`.

[37] J. H. Friedman and N. I. Fisher, Statistics and Computing **9**, 123 (1999).

[38] *BABAR*, B. Aubert *et al.*, p. 011801 (2005), physics/0507143.

[39] A. Ali, E. Lunghi, and A. Y. Parkhomenko, Phys. Lett. **B595**, 323 (2004), hep-ph/0405075.

[40] BABAR, B. Aubert *et al.*, Phys. Rev. Lett. **98**, 151802 (2007), hep-ex/0612017.

[41] M. Ciuchini *et al.*, JHEP **07**, 013 (2001), hep-ph/0012308.

[42] A. Ali and E. Lunghi, Eur. Phys. J. **C26**, 195 (2002), hep-ph/0206242.

[43] B. H. Denby, Invited tutorial to be publ. in Proc. of 2nd Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, La Londe-les- Maures, France, Jan 13-18, 1992.

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