Constraining $|V_{td}|/|V_{ts}|$ using Radiative Penguin $B \to V(K^*/\rho/\omega)\gamma$ Decays

by

Ping Tan

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

Doctor of Philosophy
(Physics)

at the

University of Wisconsin – Madison

2005

Abstract

Exclusive radiative penguin B decays, $B \rightarrow (K^{*0}/K^{*+})\gamma$ and $B \rightarrow (\rho/\omega)\gamma$, are flavor-changing neutral-current (FCNC) processes. Studies of these decays are of special interest in testing Standard Model (SM) predictions and searching for other beyond-the-SM FCNC interactions. Using 89×10^6 $B\overline{B}$ pairs from *BABAR*, we measure the branching fraction ($\mathcal{B}$), *CP*-asymmetry ($\mathcal{A}$), and isospin asymmetry (Δ_{0-}) of $B \rightarrow (K^{*0}/K^{*+})\gamma$ as follows:

$$\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = 3.92 \pm 0.20(stat.) \pm 0.24(syst.),$$
$$\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = 3.87 \pm 0.28(stat.) \pm 0.26(syst.),$$
$$\mathcal{A}(B \rightarrow K^*\gamma) = -0.013 \pm 0.36(stat.) \pm 0.10(syst.),$$
$$\Delta_{0-}(B \rightarrow K^*\gamma) = 0.050 \pm 0.045(stat.) \pm 0.028(syst.) \pm 0.024(R^{+/0}).$$

The 90% confidence intervals for the *CP*-asymmetry and the isospin-asymmetry in the $B \rightarrow K^*\gamma$ decay are given as:

$$-0.074 < \mathcal{A}(B \rightarrow K^*\gamma) < 0.049,$$
$$-0.046 < \Delta_{0-}(B \rightarrow K^*\gamma) < 0.146.$$

We also search for $B \rightarrow (\rho/\omega)\gamma$ decays using 211×10^6 $B\overline{B}$ pairs from *BABAR*. No evidence for these decays is found. We set the upper limits at 90% confidence level for these decays:

$$\mathcal{B}(B^0 \rightarrow \rho^0\gamma) < 0.4 \times 10^{-6},$$
$$\mathcal{B}(B^+ \rightarrow \rho^+\gamma) < 1.8 \times 10^{-6},$$
$$\mathcal{B}(B^0 \rightarrow \omega\gamma) < 1.0 \times 10^{-6},$$
$$\overline{\mathcal{B}}(B \rightarrow (\rho/\omega)\gamma) < 1.2 \times 10^{-6}.$$

These results are in good agreement with the SM predictions. The branching fractions of these decays are then used to constrain the ratio $|V_{td}|/|V_{ts}|$.

To my parents, for their love and support.

Contents

List of Figures

List of Tables

1 Introduction

The world view of particle physics [1, 2] is that matter is made out of three generations of leptons,

$$\left(\begin{array}{c} \nu_e \\ e \end{array} \right), \ \left(\begin{array}{c} \nu_\mu \\ \mu \end{array} \right), \ \left(\begin{array}{c} \nu_\tau \\ \tau \end{array} \right),$$

three generation of quarks,

$$\left(\begin{array}{c} u \\ d \end{array} \right), \ \left(\begin{array}{c} c \\ s \end{array} \right), \ \left(\begin{array}{c} t \\ b \end{array} \right),$$

the antiparticles of the leptons and quarks, and the mediators of interactions between them,

$$photon, \ W^\pm - bosons, \ Z^0 - boson, \ gluons, \ and \ graviton.$$

These are called elementary particles. Other particles, such as baryons and mesons, are composed of combinations of these elementary particles. For example, baryons consist of three quarks or anti-quarks and mesons consist of a quark and an anti-quark.

Particles interact with each other through four basic interactions: gravity, electromagnetic, weak, and strong interactions by exchanging graviton, photon, $W^\pm$ and Z^0 bosons, and gluons, respectively[1]. In current understanding, charged leptons involve the electro-magnetic and weak interactions, however, neutrinos involve only the weak interaction; quarks interact through electro-magnetic, weak, and strong interactions. All these particles involve the gravity interaction, which is much weaker than any other interaction and will be ignored in the following discussion.

Many theories have been proposed to describe particles and their interactions. After the 1950, it was generally assumed that quantum field theory was the form of the laws of nature [3]. A local gauge field theory was used successfully to unify the electromagnetic

---

[1]Maxwell formulated the classical theory of electromagnetism. The weak interaction was first observed in Beta decays. Yukawa proposed the strong interaction to explain the interaction between protons and nucleons in 1934. A historical review of these interactions can be found in references [1, 2].

interaction and the weak interaction among quarks and leptons, and is known as the Standard Model (SM) [4, 5, 6].

1.1 Symmetries

Since the birth of modern physics, probing the behavior of a system under various transformations has been important in understanding the underlying physical laws, and has special importance in particle physics, e.g. gauge invariance in unification of the electromagnetic interaction and the weak interaction. A symmetry transformation is a change of a system that does not change possible experimental results. Noether's theorem [7] states that for any symmetry of a system there are some corresponding conservation laws, and vice versa. Therefore, symmetries can be used as a probe of the underlying physics of a system. Symmetry transformations naturally form groups[2], "symmetry groups"; and group theories can be used to study these symmetrical transformations systematically; e.g. by studying different representations of the symmetry group.

The symmetries of particles can be divided into space-time and internal symmetries. Different groups, such as the connected Lie group, are used to characterize these symmetrical properties. For example, the most fundamental symmetry in physics is Lorentz invariance, which forms the Poincare group. Other Lie groups, such as $U(1)$ and $SU(2)$, have special importance in the construction of the SM, and are used to describe internal symmetries.

The symmetry of a system can be broken. For example, the space-inversion transforms the coordinate of a particle $\vec{r}$ to $-\vec{r}$ and was thought to be a good symmetry, called parity (P). This symmetry is broken when the weak interaction is introduced, called P-violation. Another example, the CP-transformation, converts a particle into its anti-partner with in-

---

[2]Symmetry transformation A followed by symmetry transformation B, gives another symmetry transformation C, which can be defined as $C = BA$. With this definition of multiplicity, other requirements of a group is trivial, e.g. the unit is a transformation to itself.

verse momentum and was also thought to be a good symmetry. However, it is also broken in the presence of electroweak interactions. There are several types of CP-violation. For example, a decay $i \to f$ and its CP-conjugate decay $\bar{i} \to \bar{f}$ may happen with different decay rates, represented by Γ in the following; this is the so-called "direct" CP-violation. Direct CP-violation results from different strong and weak phases in a decay with many amplitudes. The direct CP-asymmetry, defined as

$$\mathcal{A} = \frac{\Gamma(\bar{i} \to \bar{f}) - \Gamma(i \to f)}{\Gamma(\bar{i} \to \bar{f}) + \Gamma(i \to f)}, \tag{1}$$

can be used to quantify the magnitude of CP-violation. In addition to direct CP-violation, other types of CP-violation are possible [8]. The mechanism of CP-violation is of special importance in understanding our universe.

1.2 Gauge Symmetry and Gauge Fields

One of the key features in modern particle theory is gauge symmetry. In the Lagrangian formalism, the Lagrangian of matter fields $u_i(x)$ $(i = 1, 2, ..., n)$ is of the format $\mathcal{L}(u_i, \partial_\mu u_i)$. A global gauge transformation[3] takes the form, $u_i(x) \to u_i'(x) = u_i(x) + t_{ij}^k \epsilon_k u_j(x)$ $(k = 1, 2, ..., m)$, with t_{ij}^ks group generators and ϵ_ks infinitesimal parameters of the group. The Lagrangian is invariant under this gauge transformation. Here $\partial_\mu u_i \equiv \partial u_i / \partial x^\mu$ with index μ is used to indicate the space-time coordinates, and $t_{ij}^k \epsilon_k u_j(x)$ represents $\sum_{k,j} t_{ij}^k \epsilon_k u_j(x)$. This convention will be used throughout the rest of this dissertation.

In today's understanding, the Lagrangian of matter fields is also locally gauge invariant[4]. Under a group of local gauge transformations, $u_i(x) \to u_i'(x) = u_i(x) + t_{ij}^k \epsilon_k(x) u_j(x)$, a globally invariant Lagrangian can be locally non-invariant. To be locally gauge invariant,

---

[3]This is called "global" because the group element has no time-space dependence.

[4]The local gauge invariance is one of the key features in modern physics, since it ensures the calculated observables are finite

a total of m vector fields[5], A^i_μ $(i = 1, 2, ..., m)$, are introduced into the Lagrangian. The Lagrangian of these vector fields, $\mathcal{L}'$, must meet a very specific condition,

$$\frac{\partial \mathcal{L}'}{\partial F^a_{\mu\nu}} f_{bca} F^b_{\mu\nu} = 0, \ c = 1, 2, ..., m, \tag{2}$$

with f_{bca} structure constants and

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - \frac{1}{2} f_{bca}(A^b_\mu A^c_\nu - A^b_\nu A^c_\mu), \ a = 1, 2, ..., m; \ \mu \neq \nu. \tag{3}$$

These are the so-called gauge fields. Interactions between matter fields and gauge fields are then introduced in a predefined manner by making a replacement, $\partial_\mu u_i \rightarrow \partial_\mu u_i - T^k_{ij} u_j A^k_\mu$, in the Lagrangian of the matter fields, $\mathcal{L}$. The simplest form of gauge fields is the Yang-Mills field, which is

$$\mathcal{L}_{YM} = -\frac{1}{4} F_{\mu\nu}{}^k F_{\mu\nu}{}^k, \tag{4}$$

with $F_{\mu\nu}{}^k$ as defined in Eq. 3. This is the Lagrangian of the gauge fields introduced in the SM.

1.3 Spontaneous Symmetry-Breaking

Local gauge invariance has been very attractive in building theories to unify electromagnetic and weak interactions, where gauge fields are identified as mediators. These gauge fields are massless particles, which apparently contradicts the experimental fact that the mediators for the weak interaction are very massive. The Lagrangian of these gauge fields is required to take a very constrained form, as shown in Eq. 2, which eliminates the possibility of simple addition of mass terms. Instead, spontaneous symmetry-breaking is introduced to give mass to these gauge fields.

---

[5]This is equal to the number of generators in the group.

This symmetry-breaking mechanism can be best realized in the following example of a complex scalar field $\phi(x)$ described by

$$\mathcal{L} = (\partial_\mu \phi^*(x))(\partial_\mu \phi(x)) - m^2 \phi^*(x)\phi(x) - \frac{1}{4}f(\phi^*(x)\phi(x))^2, \tag{5}$$

with $f > 0$. The Lagrangian is invariant under the global $U(1)$ transformation. When $m^2 > 0$, as shown in Figure 1 a), the vacuum state of the Lagrangian is non-degenerate, and is also $U(1)$ invariant; this is an exact symmetry. When $m^2 < 0$, as shown in Figure 1 b), there are an infinite number of degenerate vacuum states $\phi_V(x) = \sqrt{2}m/\sqrt{f}$. Under a gauge transformation, any vacuum state can be transformed into a different vacuum state. The gauge invariance of vacuum states is broken; this is a spontaneous symmetry-breaking. In the spontaneous symmetry-breaking, a vacuum state must be chosen to construct any theory; e.g. $\phi(x)$ is given in terms of $\phi = (2m/\sqrt{f} + \phi_1(x) + \phi_2(x))/\sqrt{2}$ as shown in Figure 1 c). With this choice of the vacuum state, the Lagrangian in Eq. 5 takes the form,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1(x))^2 + \frac{1}{2}(\partial_\mu \phi_2(x))^2 - \frac{1}{2}m_1^2 \phi_1^2(x) - \dots, \tag{6}$$

with $m_1 = \sqrt{2}m$. So as a consequence of the spontaneous symmetry breaking, a real scalar field, $\phi_1(x)$, with mass m_1 and a massless scalar field, $\phi_2(x)$, are seen. These massless particles are called "Goldstone" bosons.
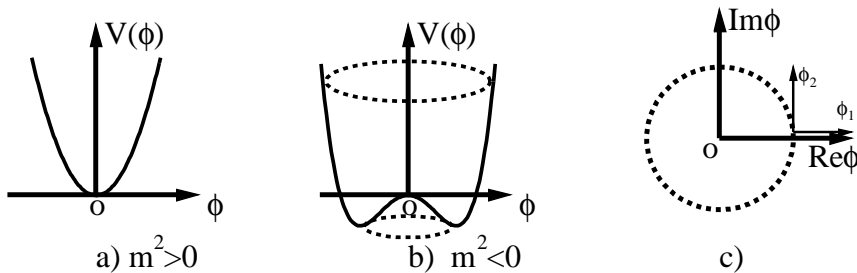
Figure 1: Vacuum states: a) non-degenerate; b) degenerate; c) fixed. Here $V(\phi) = m^2\phi^*(x)\phi(x) + \frac{1}{4}f(\phi^*(x)\phi(x))^2$, defined in Eq. 5.

Using the same mechanism, the breaking of a local gauge symmetry could give mass to vector gauge bosons. For example, for a local $U(1)$ gauge invariant scalar field,

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} + (\partial_\mu\phi^* - igA_\mu\phi^*)(\partial_\mu\phi + igA_\mu\phi) - m^2\phi^*\phi - \tfrac{1}{4}f(\phi^*\phi)^2, \tag{7}$$

after the same symmetry-breaking described above, the Lagrangian is

$$\mathcal{L} = -\tfrac{1}{4}(\partial_\mu B_\nu - \partial_\nu B_\mu)^2 - \tfrac{2g^2 m^2}{f}B_\mu B_\mu + \tfrac{1}{2}(\partial_\mu\phi_1)(\partial_\mu\phi_1) - m^2\phi^2 + \dots. \tag{8}$$

A vector gauge field, B_μ, with mass $2gm/\sqrt{f}$ is seen. The massive scalar field, $\phi_1(x)$, shows up again. The massless scalar field, $\phi_2(x)$, disappears, and is absorbed into the vector gauge field, B_μ, to give mass to the vector boson. This is the so-called Higgs mechanism of spontaneous symmetry-breaking.

1.4   The Standard Model

1.4.1   Introduction to the Standard Model

The Standard Model was first proposed by Glashow, Salam, and Weinberg to unify the electro-magnetic and weak interactions among leptons [4, 5, 6]. It is a local gauge theory. The construction of the theory starts from a global $SU(2) \times U(1)$ invariant Lagrangian,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_Y + \mathcal{L}_H,$$
$$\text{where}$$
$$\mathcal{L}_0 \equiv \sum_{j=e,\mu,\tau} i \cdot [\overline{L}_j \gamma_\mu \partial_\mu L_j + \overline{R}_j \gamma_\mu \partial_\mu R_j] \tag{9}$$
$$\mathcal{L}_Y \equiv \sum_{j=e,\mu,\tau} i \cdot [-h_j \overline{L}_j \phi R_j - h_j \overline{R}_j \,\overline{\phi} L_j]$$
$$\mathcal{L}_H \equiv \partial_\mu \overline{\phi} \partial_\mu \phi + m^2 \overline{\phi}\phi - \tfrac{1}{4}f(\overline{\phi}\phi)^2.$$

The invariant Lagrangian contains the Lagrangian of massless fermion fields ($\mathcal{L}_0$), a doublet Higgs field ($\mathcal{L}_H$), and the interaction Lagrangian between the Higgs field and the fermion fields ($\mathcal{L}_Y$), which are Yukawa terms with coupling constant h_js. In the Lagrangian of the fermion fields, $\mathcal{L}_0$, only left-handed lepton $SU(2)$-doublets (L) and right-handed lepton $SU(2)$-singlets (R) are involved.

$$L \equiv \frac{1}{2}(1 + \gamma_5)\begin{pmatrix} \nu_l \\ l \end{pmatrix}, \; R \equiv \frac{1}{2}(1 - \gamma_5)(l); \; l = e, \; \mu, \; \tau. \tag{10}$$

This combination of quark doublet or singlet is to take account of the fact that there are no right-handed neutrinos or right-handed lepton doublets observed in experiments. The Higgs field is an isotopic doublet of scalar fields, $\phi \equiv \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix}$, with f and m^2 larger than zero to ensure the symmetry breaks spontaneously. The fermion mass term, $\overline{L}R + \overline{R}L$, is not invariant due to the different transformation properties of L and R. The fermions acquire mass by virtue of Yukawa couplings. The $\mathcal{L}$ is invariant when these fields transform under the $SU(2)$ and $U(1)$ groups[6], whose coupling constants are g and g_1, respectively.

By requiring local gauge invariance, four massless gauge fields are introduced into the above Lagrangian. In the SM, these gauge fields take the form of the Yang-Mills fields shown in Eq. 4. Then making use of the spontaneous symmetry-breaking and selecting the vacuum state at $\sqrt{2}m/\sqrt{f} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, the Lagrangian $\mathcal{L}$ takes the following form[7],

$$
\begin{aligned}
&\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I, \text{ where} \\
&\mathcal{L}_0 = -\tfrac{1}{2}W_{\mu\nu}{}^*W_{\mu\nu} + \tfrac{g^2 m^2}{f}W_\mu{}^*W_\mu - \tfrac{1}{4}Z_{\mu\nu}Z_{\mu\nu} + \tfrac{m^2(g^2+g_1^2)}{2f}Z_\mu Z_\mu \\
&\quad -\tfrac{1}{4}B_{\mu\nu}B_{\mu\nu} \\
&\quad +\tfrac{1}{2}\partial_\mu\sigma\partial_\mu\sigma - m^2\sigma^2 \\
&\quad +i\sum_{l=e,\mu,\tau}[\bar{l}\gamma_\mu\partial_\mu l - \tfrac{\sqrt{2}h_l m}{\sqrt{f}}\bar{l}l + \overline{\nu}_l\gamma_\mu\tfrac{1+\gamma_5}{2}\partial_\mu\nu_l] \\
&\text{and,} \\
&\mathcal{L}_I = \{\tfrac{gg_1}{\sqrt{g^2+g_1^2}}J_\mu{}^{e,l}(x)B_\mu\} \\
&\quad +\{-\tfrac{g}{2\sqrt{2}}J_\mu{}^{+,l}(x)W_\mu{}^*(x) + h.c.\} + \{-\tfrac{1}{2}\sqrt{g^2+g_1^2}[J_\mu{}^{n,\nu}(x) + J_\mu{}^{n,l}(x)]Z_\mu(x)\} \\
&\quad -\sum_{l=e,\mu,\tau}\tfrac{1}{2}h\sigma\bar{l}l + \cdots
\end{aligned}
\tag{13}
$$

---

[6]The transformation of these fields under $SU(2)$ is

$$
\begin{aligned}
&L(x) \rightarrow \exp(-\tfrac{i}{2}g\vec{\tau}\cdot\vec{\epsilon})L(x) \text{ and } R(x) \rightarrow R(x), \\
&\phi(x) \rightarrow \exp(-\tfrac{i}{2}g\vec{\tau}\cdot\vec{\epsilon})\phi(x),
\end{aligned}
\tag{11}
$$

and under $U(1)$ as

$$
\begin{aligned}
&L(x) \rightarrow \exp(-\tfrac{i}{2}g_1 Y_L\epsilon_4)L(x) \text{ and } R(x) \rightarrow \exp(-\tfrac{i}{2}g_1 Y_R\epsilon_4)L(x), \\
&\phi(x) \rightarrow \exp(-\tfrac{i}{2}g_1\epsilon_4)\phi(x),
\end{aligned}
\tag{12}
$$

with $\vec{\tau}$, the Pauli matrix, $\tfrac{1}{2}Y_L = \begin{pmatrix} -\tfrac{1}{2} & 0 \\ 0 & -\tfrac{1}{2} \end{pmatrix}$, $\tfrac{1}{2}Y_R = -1$, and $\vec{\epsilon}$, ϵ_4, infinitesimal group parameters.

[7]A detailed derivation and the complete format of the Lagrangian is described elsewhere [9]; here only some of the terms are selected out to discuss some novel physics results in the SM.

The Lagrangian, $\mathcal{L}_0$, comprises fields that can be identified with charged intermediate vector boson fields , $W(x)$, with mass $gm/\sqrt{f}$; with a neutral intermediate vector boson field, $Z(x)$, with mass $\sqrt{g^2 + g_1^2}m/\sqrt{f}$; with an electromagnetic field, $B(x)$; with a real scalar field, $\sigma(x)$, with mass m (the Higgs boson); with lepton fields of mass $mh_l\sqrt{2/f}$, l; and with left-handed neutrinos, $\nu_l(x)$s. The Lagrangian, $\mathcal{L}_I$, describes interactions among these fields, where the $J_\mu^{e,l}(x)$, $J_\mu^{+,l}(x)$, $J_\mu^{n,\nu}(x)$, and $J_\mu^{n,l}(x)$ take the following forms:

$$
\begin{aligned}
J_\mu^{e,l}(x) &= \sum_{l=e,\mu,\tau} \overline{l}(x)\gamma_\mu l(x), \\
J_\mu^{+,l}(x) &= \sum_{l=e,\mu,\tau} \overline{\nu}_l(x)\gamma_\mu(1+\gamma_5)l(x), \\
J_\mu^{n,\nu}(x) &= \tfrac{1}{2}[\sum_{l=e,\mu,\tau} \overline{\nu}_l(x)\gamma_\mu\nu_l(x)], \\
J_\mu^{n,l}(x) &= (-\tfrac{1}{2}+\xi)[\sum_{l=e,\mu,\tau} \overline{l}(x)\gamma_\mu(1+\gamma_5)l(x)] + \xi[\sum_{l=e,\mu,\tau} \overline{l}(x)\gamma_\mu(1-\gamma_5)l(x)],
\end{aligned}
\tag{14}
$$

with $\xi = g_1^2/\sqrt{g^2 + g_1^2}$.

The $gg_1/\sqrt{g^2 + g_1^2}J_\mu^{e,l}(x)B_\mu$ is identified as electromagnetic interaction with

$$
e = -\frac{gg_1}{\sqrt{g^2 + g_1^2}}. \tag{15}
$$

The weak decays originate from interactions with charged W bosons. In the low-energy limit ($q^2 \ll m_W^2$), the propagator of the weak charged current naturally takes the form of Fermi's weak theory, leading to $G/\sqrt{2} = g^2/(8m_W^2)$. The most interesting property of the SM is the occurrence of neutral weak currents, $J_\mu^{n,\nu}(x)$ and $J_\mu^{n,l}(x)$. A number of consequences of this prediction have been subjected to experimental tests, and verified in experiments.

The SM can be extended easily to include electroweak interactions among quarks. The $SU(2) \times U(1)$ is assumed to be the gauge group of the electroweak interaction among the quarks. The quarks are then grouped as left-handed $SU(2)$ doublets,

$$
\begin{pmatrix} u \\ d \end{pmatrix}_L^\alpha \equiv \frac{1}{2}(1+\gamma_5)\begin{pmatrix} u \\ d \end{pmatrix}_L^\alpha, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L^\alpha \equiv \frac{1}{2}(1+\gamma_5)\begin{pmatrix} c \\ s \end{pmatrix}_L^\alpha, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L^\alpha \equiv \frac{1}{2}(1+\gamma_5)\begin{pmatrix} t \\ b \end{pmatrix}_L^\alpha,
$$

and right-handed $SU(2)$ singlets,

$$
\begin{aligned}
u_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)u^\alpha, & c_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)c^\alpha, & t_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)t^\alpha, \\
d_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)d^\alpha, & s_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)s^\alpha, & b_R^\alpha &\equiv \tfrac{1}{2}(1-\gamma_5)b^\alpha,
\end{aligned}
$$

with α color charges (red, blue, and green). A total of nine left-handed quark doublets and 18 right-handed quark singlets are included[8]. Similarly, the quarks obtain mass from Yukawa couplings to the Higgs boson. Electromagnetic quark currents, $J_\mu^{e,q}(x)$, charged weak quark currents, $J_\mu^{+,q}(x)$, and neutral weak quark currents, $J_\mu^{n,q}(x)$, can be identified in the same way as was done for the leptons (Eq. 13). The most remarkable feature of this model is the neutral weak quark current, which forbids flavor-changing neutral-current (FCNC) processes. For example, the decay $K_S^0 \to \mu^+ \mu^-$ is forbidden in the SM. The strong interaction among the quarks is included in the SM by choosing $SU(3)$ as the gauge group. The color states of each quark are placed in a triplet of the $SU(3)$ group. A total of eight massless gauge fields, gluons, are introduced as mediators of the strong interaction. Unlike the photon in the electroweak interaction, the gluons can interact with each other.

1.4.2 Quark Mixing and the CKM Matrix

The charged quark current, $J_\mu^{+,q}(x)$, takes a form similar to its corresponding lepton shown in Eq. 14,

$$J_\mu^{+,q}(x) = \overline{u}(x)\gamma_\mu(1+\gamma_5)d(x) + \overline{c}(x)\gamma_\mu(1+\gamma_5)s(x) + \overline{t}(x)\gamma_\mu(1+\gamma_5)b(x). \qquad (16)$$

Therefore, transitions among different generations of quarks, e.g. $s \to u$, are forbidden, violating the experimental facts, e.g. decays like $K^+ \to l\nu_l$ are observed. This discrepancy is due to the fact that the electroweak eigenstates of the quarks are not the same as the quark mass eigenstates. The electroweak eigenstates of the quarks are a mixture of the quark mass eigenstates. The mixing between the d and s quarks was first proposed by Cabibbo [10].

---

[8]The Y_L and Y_R in the $U(1)$ transformation take the form of $\frac{1}{2}Y_L = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{6} \end{pmatrix}$ and $\frac{1}{2}Y_R = \begin{pmatrix} -\frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$, respectively.
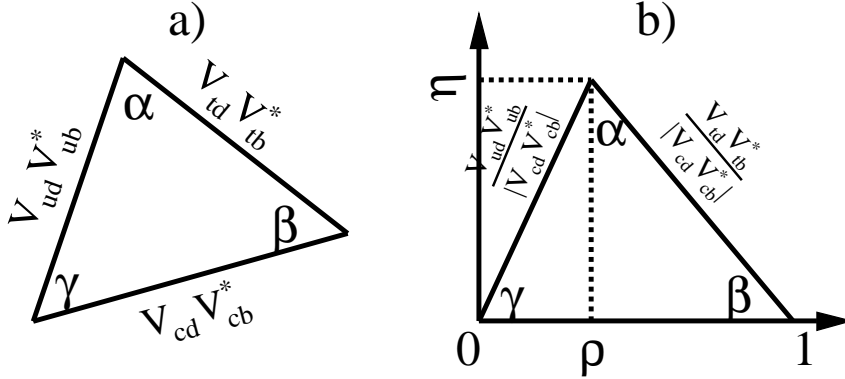
Figure 2: The Unitary Triangle (UT). In b), each side of the UT is scaled by $V_{cd}V_{cb}^*$, and one side is placed on the x-axis.

Conventionally, this mixing is assumed to happen among d, s, and b quarks, taking the form

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{17}$$

This parameterization was first proposed by Kobayashi and Maskawa [11] as a generalization case of the Cabibbo matrix, and now is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

The free parameters in the CKM matrix can be reduced by requiring its unitarity and by a redefinition of quarks' states. The total number of free parameters is four, including a complex phase term[9]. This phase term is the only source for CP-violation in the SM. It is interesting to note that mixing of only two generations of quarks would not allow CP-violation. This led to the prediction of three generations of quarks before there was any experimental evidence for the third generation.

The unitarity of the CKM matrix gives some important relations among the elements of

---

[9]For n-generation quark-doublets, the total free parameters of the mixing matrix is $(n-1)^2$: $\frac{1}{2}(n-1)n$, real parameters, normally known as "mixing angle" and $\frac{1}{2}(n-1)(n-2)$, complex phase angles.

the CKM matrix. The most important one is,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{18}$$

which forms a triangle, the Unitarity Triangle (UT), in the complex plane (See Figure 2)[10].
The UT can be rearranged in the position shown in Figure 2 b) by dividing Eq. 18 by $V_{cd}V_{cb}^*$
and placing the two vertexes of the rescaled UT at $(0, \ 0)$ and $(1, \ 0)$. The third vertex is
conventionally denoted as $(\rho, \ \eta)$ and the three inner angles are denoted by α, β, and γ. By
further defining $\lambda = |V_{us}|$, the Wolfenstein parametrization of the CKM matrix is realized[11]:

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{19}$$

The elements of the CKM matrix are free parameters in the SM and have to be measured
in experiments. Under quark mixing, the charged weak current, $J_\mu^{+,q}(x)$, in Eq. 20 is

$$J_\mu^{+,q}(x) = \overline{q}_i(x)\gamma_\mu(1 + \gamma_5)V_{ij}q_j(x) : \ i = u, \ c, \ t; \ j = d, \ s, \ b. \tag{20}$$

Therefore, in principle, all the elements can be measured by the corresponding quark decays.
For example, the nuclear beta decay, which involves the $d \rightarrow u$ weak decay, was used to
determined the $|V_{ud}|$ [12]. Currently the allowed ranges of the CKM elements [13] at 90%
Confidence Level (C.L.) are

$$\begin{pmatrix} 0.9741 & \text{to} & 0.9751 & 0.219 & \text{to} & 0.226 & 0.004 & \text{to} & 0.014 \\ 0.219 & \text{to} & 0.226 & 0.9732 & \text{to} & 0.9748 & 0.037 & \text{to} & 0.044 \\ 0.0025 & \text{to} & 0.0048 & 0.038 & \text{to} & 0.044 & 0.9990 & \text{to} & 0.9993 \end{pmatrix}, \tag{21}$$

obtained by combining experimental results and utilizing the unitarity of the CKM matrix.
Only the $|V_{td}|$ and the $|V_{ub}|$ are poorly determined.

---

[10]Other unitarity relations, such as

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0,$$
$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0,$$

can be obtained.

[11]The one actually used is a slight variation of the standard Wolfenstein parametrization: the third vertex
$(\rho, \ \eta)$ is redefined to be $(\overline{\rho} \equiv \rho(1 - \lambda^2/2), \ \overline{\eta} \equiv \eta(1 - \lambda^2/2))$, and this could improve the Wolfenstein
parametrization up to $\mathcal{O}(\lambda^6)$.

1.5 SM Limitations

The SM successfully explains all the current experimental results. However, it leaves many questions unanswered [14]:

- In the SM, there are three gauge groups, associated with three arbitrary coupling constants. The fermions are grouped into left-handed doublets and right-handed singlets. However, the SM cannot predict anything about the origins of the coupling constants. The choice of fermion representation is made purely to match the experimental facts. It cannot predict anything about the choice of the hypercharges.

- The SM includes three generation of fermions; however, it tells nothing about why this is so. It does not forbid more generations of fermions, either. The fermion masses and the mixing parameters seem to exhibit a hierarchical pattern; e.g., m_t, $m_b >> m_c$, $m_s >> m_u$, m_d. The SM cannot explain this.

- The SM predicts the existence of at least one Higgs boson. This prediction has not yet been confirmed experimentally. In addition to its existence, the mass of the Higgs boson receives quadratically divergent quantum loop corrections.

- One of the biggest puzzles in the SM is that CP-violation in the SM is too weak to explain the present matter-dominated universe.

And there are many other limitations that are not discussed here. We believe that the SM is only valid up to a cut-off energy. Beyond that cut-off energy new physics phenomena must exist. Many extensions to the SM have been proposed.

Models based on supersymmetry (SUSY) [15, 16] impose a symmetry transformation between the fermions and bosons. As a consequence of this new symmetry the SM spectrum of particles is doubled, with the introduction of new super-particles: scalars corresponding

to each SM fermion, and fermions corresponding to each boson. The quadratic divergence of the SM model Higgs mass correction is naturally avoided because contributions to it from bosonic and fermionic loops occur with opposite sign and cancel each other out.

The Higgs sector in the SUSY models is more complicated than in the SM. For example, in the Minimal Supersymmetric Standard Model (MSSM) [15], additional Higgs bosons are predicted: a light Higgs (h^0), a CP-even heavy Higgs (H^0), a CP-odd heavy Higgs (A^0), and two charged Higgs particles ($H^{\pm}$). The masses of the MSSM Higgs sector is governed by two parameters, e.g. the ratio of the two Higgs field vacuum expectation values and the mass of one of the Higgs particles. The Higgs particle h^0 behaves much like the SM Higgs, and is required to be lighter than Z-boson at tree-level. Higher order corrections allow this mass to be somewhat higher than the present experimental search limits [17, 18].

Further, SUSY must be broken at low energies because there is no experimental evidence for super-particles at such range. Several models of SUSY breaking, which predict new super-particles at masses of the order of TeV, are currently allowed within the limitations of experimental measurements [15, 16].

In SUSY models, many non-SM particles are introduced. These SUSY particles can contribute significantly to some rare SM decays, e.g. the $b \rightarrow s(d)\gamma$ decays discussed in the following. Studies of these decays can provide indirect evidence to the SUSY particles and a discussion of searching for SUSY particles using $b \rightarrow s\gamma$ decays can be found here [19].

The SUSY theories provide not only necessary and testable improvements to the SM, but also solutions to some cosmological problems. For instance, most SUSY models predict that super-particles are only pair produced, and that the lightest super-symmetric particle (LSP) is neutral and weakly interacting [20]. This LSP is a good candidate for dark matter that is found to be abundant in our Universe.

1.6 Experimental Motivation

The SM contains 19 parameters, which have to be determined from experiments[12]. Currently, many experiments have measured these parameters, to confine the SM and hopefully to give some hints for beyond-the-SM physics.

One of the predictions of the SM is that CP-violation in all particles originates from the single complex phase term in the CKM matrix. By studying CP-violation in different particle systems, this prediction can be well tested. Experimental evidence of CP-violation was first observed in the decay of kaons [21]. Large CP-asymmetries[13] were predicted in some rare b-quark decays. Measurements of these CP-parameters in the b-quark sector provide a very important test of the mechanism of CP-violation in the SM. The long life-time of B-mesons makes it experimentally possible to measure these CP-parameters. B-factories at SLAC and KEK with the $BABAR$ [22] and Belle [23] detectors, were designed specifically to collect large samples of B-mesons and make precision studies of CP-violation in the B-meson sector. With these data, other physics processes can also be studied and used to test many other SM predictions[14].

In addition to measuring the angles of the UT, the sides of the UT can be determined from direct measurements and used to check the "unitarity" of the CKM matrix. Currently only the $|V_{ub}|$ and the $|V_{td}|$ measurements have large uncertainties. From Eq. 20, we can see that measurements of these elements involve processes like $b \rightarrow u$ and $t \rightarrow d$ transitions, which are rare. The high luminosity data available in $BABAR$ make studies of these processes possible. The latest result[15] on the $|V_{ub}|$ from $BABAR$ is $|V_{ub}| =$

---

[12]The 19 parameters include nine fermion masses (quarks and leptons); three coupling constants for the electromagnetic interaction, weak interaction and strong interaction, respectively; four quark mixing parameters; the masses of the two gauge bosons and the Higgs boson.

[13]which could be related to the opening angle β of the UT.

[14]Recently, $BABAR$ has first discovered direct CP-violation in $B \rightarrow K\pi$, which is also in good agreement with the SM predication.

[15]The three errors are experimental error, the uncertainty in determining the fraction of the charmless
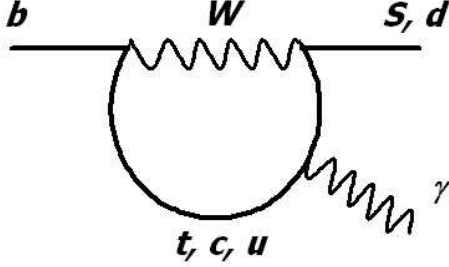
Figure 3: The lowest order of $b \to s(d)\gamma$ decays.

$(3.94\pm0.25(exp)\pm0.37(f_u)\pm0.19(theory))\times10^{-3}$, from an analysis of semileptonic charmless B-decays [24]. This dissertation focuses on performing a direct measurement of $|V_{td}|/|V_{ts}|$ using radiative penguin decays $B \to (K^*/\rho/\omega)\gamma$ (here $\rho = \rho^0/\rho^\pm$ and $K^* = K^{*0}/K^{*\pm}$, and this convention will be used throughout this disseration unless specified). This measurement of the $|V_{td}|/|V_{ts}|$ is complementary to other methods, such as using B-mixing[16].

At the quark level, the decays $B \to (K^*/\rho/\omega)\gamma$ are FCNC processes, $b \to (s/d)\gamma$. These processes are suppressed at the tree level in the SM, and the lowest order Feynman diagram is a loop diagram, as shown in Figure 3, which is dominated by the transition of the intermediate top-quark. The transition of $t \to s(d)$ is sensitive to the CKM matrix elements $V_{ts}(V_{td})$, respectively. Therefore, the decay rates of these decays can be related to these CKM elements. However, as you will see in the next section, the large theoretical error in calculating the decay rates makes this attempt meaningless. This situation can be improved by taking a ratio of the decay rates between the $B \to (K^{*0}/K^{*+})\gamma$ decays and the $B \to (\rho/\omega)\gamma$ decays. This ratio can be related to the $|V_{td}|/|V_{ts}|$ with less theoretical uncertainty than by using the decay rate of $B \to (\rho/\omega)\gamma$ decays. The following section discusses the phenomenology of this approach. Note that other physics observables in these

leptonic decay used in the analysis, and the theoretical uncertainty, respectively.

[16]A brief discussion of the B-mixing is presented in Section 2.4

decays, such as branching fractions and CP-asymmetries, can be measured and are used to test the SM or to probe for beyond-the-SM physics.

2   Phenomenology of $B \to V\gamma$ Decays

The diagram of $B \to V(\text{Vector})\gamma$ decays at quark level is shown in Figure 3, which is dominated by an electroweak penguin loop[17]. The amplitude of this electroweak transition at the quark level can be calculated to a very accurate degree using a perturbative method. However, there are a few difficulties to related the quark-level transitions to exclusive decays. First, interchanges of gluons among the quarks shown in Figure 3 can significantly modify the transition amplitude and complicate the calculation. A next-to-leading order (NLO) calculation is necessary. Second, the s or d quarks from the electro-weak transition and the light spectator quark in the B-meson fragment into the vector mesons. This process involves interactions of soft gluons. The strong coupling constant α_s has a Q^2 dependence [25],

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\log\left(Q^2/\Lambda_{QCD}^2\right)}, \tag{22}$$

with n_f the number of quark flavors and Λ_{QCD} a cut-off value in QCD ($\Lambda_{QCD} \sim 0.1$ GeV). In low Q^2 ranges, where the quarks and gluons interact strongly, the perturbative method is no longer applicable. Furthermore, there are many other possible contributions; e.g., the interchanging of gluons among the quarks in the B-meson before or after the electro-weak transition.

2.1   Effective Lagrangian

Usually the starting point of theoretical calculations of these exclusive radiative decays is from an effective Lagrangian[18] obtained from the SM by integrating out the heavy degrees of freedom (the top quarks and the $W^\pm$-bosons). With the help of Operator Product

---

[17]Here V represents K^{*0}, K^{*-}, ρ^0, ρ^-, or ω meson, and this notation will be used throughout this dissertation unless specified.

[18]The effective Lagrangian is in the form of $\mathcal{L}_I = \sum J_\mu{}^+ J_\mu{}^- + J_\mu{}^n J_\mu{}^n$ at the low Q^2, with J_μ the charged ($\pm$) and neutral (n) currents shown in Section 1.4.

Expansion (OPE) [26], the effective Lagrangian, $\mathcal{L}_{eff}$, can be written as follows:

$$\mathcal{L}_{eff} \sim -\frac{4G}{\sqrt{2}} \sum C_i(\mu) Q_i(\mu), \tag{23}$$

where μ is an arbitrary energy scale, G is the Fermi constant, $Q_i(\mu)$s are some local operators, and $C_i(\mu)$ are corresponding Wilson coefficients.

In this expansion, the transition amplitude is then separated into two distinct parts, long distance contributions (energy scale smaller than μ) contained in local operators, $Q_i(x)$, and short-distance physics contributions (energy scale larger than μ) described by the Wilson coefficients, $C_i(\mu)$. The matrix elements of the operators contain the long distance interactions and have to be determined from some non-perturbative methods, such as lattice QCD or QCD sum rules. One of many reasons that the OPE is so useful is that the Wilson coefficients are independent of any final state. Therefore, these coefficients can be treated as coupling constants. The μ-dependence of the Wilson coefficients is governed by the renormalization group expansion [26], so they can be determined at large energy scales using a perturbative method and evolved down to the relevant energy scale. In principle, the choice of the μ parameter is arbitrary, and the calculated physics quantities should be independent of it. However, a too-small μ would spoil the perturbative method to calculate the Wilson coefficients, and a too-large μ parameter would make it harder to calculate the operator elements in some non-perturbative methods. Therefore, the choice of μ is a trade off between these considerations. In B physics studies, the μ is normally chosen to have the approximately the same mass as the b quark. In reality, the theoretical calculation of the operator matrix elements and the Wilson Coefficients are stopped at some point. The μ-dependence is then introduced in physical quantities, but this dependence can be canceled out if a full calculation is performed.

With this expansion, a systematic approach to analyzing $B \rightarrow V\gamma$ decays at the NLO

in QCD was proposed by Stefan W. Bosch [27]. Following his discussion, the effective Lagrangian, considered up to dimension six, is given by

$$\mathcal{L}_{eff} = \frac{G}{\sqrt{2}} \sum_{p=u,c} \lambda_p^q \{C_1(\mu)Q_1^p(\mu) + C_2(\mu)Q_2^p(\mu) + \sum_{j=3}^{8}[C_j(\mu)Q_j(\mu)]\}, \tag{24}$$

where $\lambda_p^q = V_{pq}^* V_{pb}$; $q = s$ or d for $b \to s\gamma$ or $b \to d\gamma$, respectively. The most important operators are the four-fermion operators,

$$\begin{array}{l} Q_1^p(\mu) = (\overline{q}_\alpha \gamma_\mu (1 - \gamma_5) p_\beta)(\overline{p}_\beta \gamma^\mu (1 - \gamma_5) b_\alpha), \\ Q_2^p(\mu) = (\overline{q}_\alpha \gamma_\mu (1 - \gamma_5) p_\alpha)(\overline{p}_\beta \gamma^\mu (1 - \gamma_5) b_\beta), \end{array} \tag{25}$$

and the electromagnetic and chromomagnetic penguin operators,

$$\begin{array}{l} Q_7(\mu) = em_b/(8\pi^2) \cdot (\overline{q}_\alpha \sigma^{\mu\nu}(1 + \gamma_5)\overline{b}_\alpha)F_{\mu\nu}, \\ Q_8(\mu) = g_s m_b/(8\pi^2) \cdot (\overline{q}_\alpha \sigma^{\mu\nu}(1 + \gamma_5)T_{\alpha\beta}^a b_\beta)G_{\mu\nu}{}^a, \end{array} \tag{26}$$

where $F_{\mu\nu}$ and $G_{\mu\nu}{}^a$ are the electromagnetic and gluonic field strength tensors, respectively; $T_{\mu\nu}{}^a$ are the color $SU(3)$ group generators; α and β are color indexes, and a is the gluonic color index. Other operators related to these decays and the corresponding Wilson coefficients are described in Ref. [27, 28] and have almost no contribution to the NLO calculation discussed below.

2.2 QCD Factorization

With the effective Lagrangian shown in Eq. 24, the calculation of the decay amplitude is then turned into the calculation of the operator matrix elements. In the heavy quark limit, QCD factorization can be used to guide the calculations of the relevant operators up to the NLO of the α_s and the leading order (LO) of Λ_{QCD}/m_b. In this factorization, an operator matrix element is written as

$$< V\gamma|Q_i(\mu)|B >= \{F_V T_i^I + \int_0^1 d\xi d\nu[T_i^{II}(\xi, \nu)\Phi_B(\xi)\Phi_V(\nu)]\} \cdot \epsilon, \tag{27}$$

where ϵ is the photon-polarization 4-vector [27]. Here the F_V is a $B \to V$ transition form factor, and Φ_B and Φ_V are the leading twist light-cone distribution amplitudes of the B-meson and the vector meson, V, respectively [29]. The Φ_B and Φ_V are universal quantities, non-perturbative objects. They describe the long-distance dynamics of the matrix element, which is factorized from the the perturbative short-distance interaction expressed in the hard-scattering kernels, T_i^I and T_i^{II}.

Within this framework, the operator contributions to the LO of Λ_{QCD}/m_b and the NLO of $\mathcal{O}(\alpha_s)$ are identified, as shown in Figure 4. The LO contribution is from the magnetic penguin operator, Q_7, where the hard-scattering kernel T_7^I is a purely kinematic function and T_7^{II} is absent. The matrix element is

$$< V(k,\eta)\gamma(q,\epsilon)|Q_7|B > = -\frac{e}{2\pi^2} m_b c_V F_V \times [\epsilon^{\mu\nu\lambda\rho}\epsilon_\mu\eta_\nu k_\lambda q_\rho + i(\epsilon \cdot k\eta \cdot q - \epsilon \cdot k\eta \cdot q)], \quad (28)$$

where $c_V = 1$, except that for $V = \rho^0$, $c_V = 1/\sqrt{2}$, F_V is the form factor estimated at $q^2 = 0$, and $\epsilon^{0123} = -1$. The NLO contributions illustrated in Figure 4 b) to 4 g) have sizable corrections to the LO result, modifying the transition amplitude about 80%. The weak annihilation, shown in Figure 4 h), was estimated in QCD factorization and was found to contribute to the LO of Λ_{QCD}/m_b, which turns out to have an interesting effect on the isospin breaking for $B \to (\rho^0/\rho^+)\gamma$ decays. However, this effect in $B \to (K^{*0}/K^{*+})\gamma$ decays is CKM suppressed and has a very small contribution. Summing all the contributions, the decay amplitude of $B \to V\gamma$ can be written as following,

$$\Gamma(B \to V\gamma) = \frac{G}{\sqrt{2}} [\sum_{p=u,c} \lambda_p^q a_7^p(V\gamma)] \cdot < V(k,\eta)\gamma(q,\epsilon)|Q_7|B >, \quad (29)$$

where a_7^p ($p = u, \ c$) summarize the contributions up to the NLO of α_s and the LO of the Λ_{QCD}/m_b, including all the contributions shown in Figure 4, $q = s, \ d$.
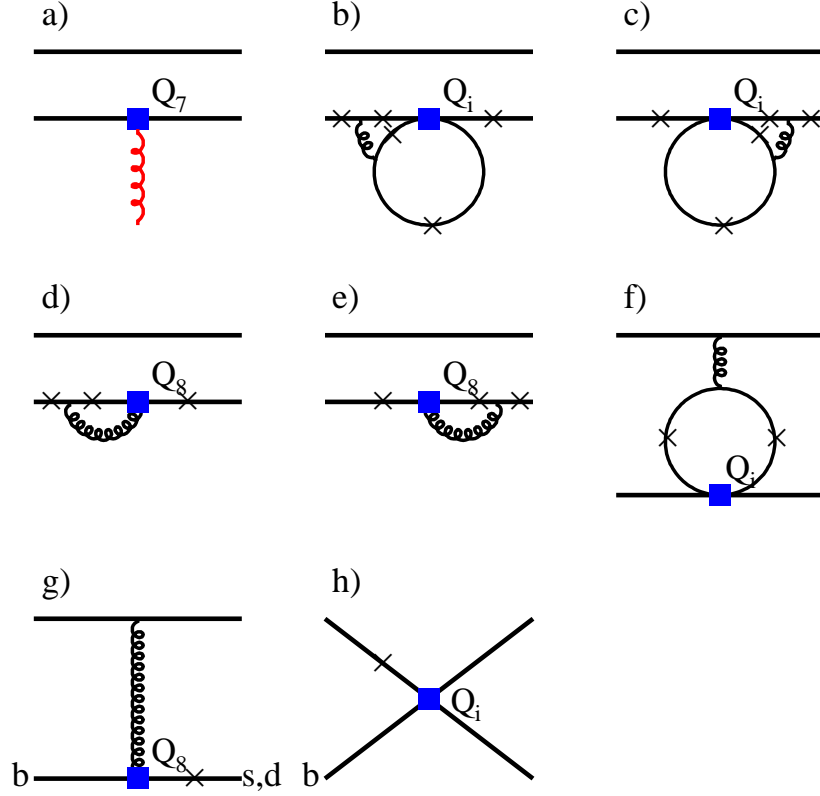
Figure 4: The contributions from the operators up to the leading order in $\mathcal{O}(\alpha_s)$ and leading order in Λ_{QCD}/m_b: a) leading order contribution in α_s; b)-c) $\mathcal{O}(\alpha_s)$ contributions to the hard-scattering kernels T_i^I from four-operators Q_i; d)-e) $\mathcal{O}(\alpha_s)$ contributions to the hard-scattering kernels T_i^I from chromomagnetic penguin operator; f)-g) $\mathcal{O}(\alpha_s)$ contributions to the hard-scattering kernels T_i^{II} from four-quark operators and chromomagnetic penguin operators; and h) weak annihilation contributions. The "$\times$" sign indicates a photon is radiated.

2.3   Physics Observables

In the following, we will follow Eq. 29, and discuss the decay amplitudes for the decays $B \to (K^*/\rho/\omega)\gamma$ in more details.

Branching fractions of $\mathbf{B \to (K^{*0}/K^{*+})\gamma}$ decays

In the $B \to (K^{*0}/K^{*+})\gamma$ decays, the QCD penguin operators $Q_3, \ldots, Q_6$ enter at $\mathcal{O}(\alpha_s)$,

but are suppressed by small Wilson coefficients. The contribution from Q_2 is zero because of its color structure. The amplitude then contains the contributions from Q_7, Q_1, and Q_8 only. The branching fraction is

$$\mathcal{B}(B \to K^*\gamma) = \tau_B \frac{G^2 \alpha m_B^3 m_b^2}{32\pi^4}(1 - \frac{m_{K^*}^2}{m_B^2})^3 |\sum_p \lambda_p^s a_7^p(K^*\gamma)|^2 |F_{K^*}|^2, \tag{30}$$

where the form factor, F_{K^*}, summarizes strong interactions among quarks inside mesons, and currently there is no reliable way to calculate it. The determination of this form factor must rely on non-perturbative methods, such as, lattice QCD calculation or QCD sum rules, which normally result in large theoretical uncertainties, about 30-50%. The NLO prediction [27] for the branching fraction of $B^0 \to K^{*0}\gamma$ decays is $\mathcal{B}(B^0 \to K^{*0}\gamma) = (7.09^{+2.47}_{-2.27}) \times 10^{-5}$, where the systematic uncertainties are dominated by the form factor calculations[19]. This theoretical prediction is already higher than the previous experimental results from CLEO [30], Belle [31], and *BABAR* [32]:

$$\begin{aligned}
\mathcal{B}(B^0 \to K^{*0}\gamma) &= (4.3 \pm 0.4) \times 10^{-5} \text{ [CLEO]}, \\
\mathcal{B}(B^0 \to K^{*0}\gamma) &= (4.09 \pm 0.21 \pm 0.19) \times 10^{-5} \text{ [Belle]}, \\
\mathcal{B}(B^0 \to K^{*0}\gamma) &= (4.23 \pm 0.40 \pm 0.22) \times 10^{-5} \text{ [\textit{BABAR}]},
\end{aligned} \tag{31}$$

and

$$\begin{aligned}
\mathcal{B}(B^- \to K^{*-}\gamma) &= (3.8 \pm 0.5) \times 10^{-5} \text{ [CLEO]}, \\
\mathcal{B}(B^- \to K^{*-}\gamma) &= (4.40 \pm 0.33 \pm 0.24) \times 10^{-5} \text{ [Belle]}, \\
\mathcal{B}(B^- \to K^{*-}\gamma) &= (3.83 \pm 0.62 \pm 0.22) \times 10^{-5} \text{ [\textit{BABAR}]}.
\end{aligned} \tag{32}$$

There are other approaches used to calculate the form factor and predict the branching fractions; e.g., using the Covariant Light-Front Quark Model (CLFQM) [33] or lattice calculation [34]. These methods make better predictions of the branching fractions, e.g. $\mathcal{B}(B \to K^*\gamma) = (3.27 \pm 0.74) \times 10^{-5}$ from the CLFQM. The discrepancy in the prediction of the branching fractions amongst these methods has not been understood. Even

---

[19]The branching fraction for the charged mode is $\mathcal{B}(B^+ \to K^{*+}\gamma) = (7.45^{+2.47}_{-2.27}) \times 10^{-5}$.

with the best prediction, a 20-30% uncertainty still remains.

Branching fractions of B $\rightarrow (\rho/\omega)\gamma$ decays

Similarly, the branching fraction of $B \rightarrow (\rho^0/\rho^+)\gamma$ decays reads,

$$\mathcal{B}(B \rightarrow \rho\gamma) = \tau_B \frac{G^2 \alpha m_B^3 m_b^2}{32\pi^4}(1 - \frac{m_\rho^2}{m_B^2})^3 | \sum_p \lambda_p^d a_7^p(\rho\gamma)|^2 c_\rho^2 |F_\rho|^2, \tag{33}$$

with $c_\rho = 1$ for $\rho^\pm$ and $c_\rho = 1/\sqrt{2}$ for ρ^0. Assuming the isospin relation between ρ^0 (isospin 1) and ω (isospin 0), to LO in the heavy-quark limit, and NLO in the α_s, both the ρ^0 meson and the ω meson in the decays $B \rightarrow V\gamma$ are produced by a $d\bar{d}$-pair. Therefore, the decay amplitude can be obtained from $B^0 \rightarrow \rho^0\gamma$ with replacements for the vector constants, mass, form factors, etc. The change of these parameters is proven to result in negligible differences between these two decays. We assume the same branching fractions for these two decay modes.

Problems similar to $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays remain here. The form factors bring large uncertainties into the predictions. An NLO prediction of the branching fraction of $B^+ \rightarrow \rho^+\gamma$ is $\mathcal{B}(B^+ \rightarrow \rho^+\gamma) = 1.58^{-0.46}_{+0.53} \times 10^{-6}$ [27]. The uncertainty of the form factor dominates the error. Another theoretical prediction of this branching is $\mathcal{B}(B^+ \rightarrow \rho^+\gamma) = (1.35 \pm 0.42) \times 10^{-6}$ [35]. These are still well below the current world best 90% confidence upper limit from *BABAR* [36]:

$$\begin{aligned}
\mathcal{B}(B^0 \rightarrow \rho^0\gamma) &< 1.2 \times 10^{-6}, \\
\mathcal{B}(B^+ \rightarrow \rho^+\gamma) &< 1.0 \times 10^{-6}, \\
\mathcal{B}(B^0 \rightarrow \omega\gamma) &< 2.1 \times 10^{-6}, \\
\mathcal{B}(B \rightarrow (\rho/\omega)\gamma) &< 1.9 \times 10^{-6}.
\end{aligned} \tag{34}$$

The branching fractions in Eq. 30 and Eq. 33 are related to the CKM elements $|V_{ts}|$ and $|V_{td}|$, respectively. In principle, those branching fractions can be used to measure these elements. However the large uncertainties in the form factors make it impossible to give any

competitive results on these quantities.

Ratio of branching fractions

The branching fractions of $B \to (\rho/\omega)\gamma$ decays can be related to the $|V_{td}|$. However, the large theoretical uncertainty in the form factor makes it meaningless for determining the $|V_{td}|$. Note that some systematic uncertainties can be canceled in the calculation of the ratio of the form factors, $\xi = F_{K*}/F_\rho$. Therefore the ratio $|V_{ts}|/|V_{td}|$ can be related to the ratio of $\mathcal{B}(B \to (\rho^0/\rho^+)\gamma)/\mathcal{B}(B \to (K^{*0}/K^{*+})\gamma)$ with much less theoretical uncertainty. In a QCD sum rule calculation, $\xi = 1.31 \pm 0.13$ (LCSR) [29], and a preliminary lattice result is $\xi = 1.1 \pm 0.1$ [34]. Following Ref. [37], we define the following CP-averaged branching fraction ratio,

$$R = \frac{\mathcal{B}(B \to \rho\gamma) + \mathcal{B}(\overline{B} \to \overline{\rho}\gamma)}{\mathcal{B}(B \to K^*\gamma) + \mathcal{B}(\overline{B} \to \overline{K^*}\gamma)}. \tag{35}$$

More explicitly, $R = R_0$ or $R_\pm$ for neutral and charged mode, respectively. From Eq. 30 and Eq. 33,

$$R = c_\rho \frac{|V_{td}|^2}{|V_{ts}|^2} \xi^{-2} r_m \frac{|a_7^c(\rho\gamma)|^2}{|a_7^c(K^*\gamma)|^2} \left| 1 + 2Re(\delta a) \frac{\overline{\eta}^2 - \overline{\rho}(1 - \overline{\rho})}{(1 - \overline{\rho})^2 + \overline{\eta}^2} \right|^2. \tag{36}$$

Here $c_\rho = 1/\sqrt{2}$ for $\rho = \rho^0$; $c_\rho = 1$ for $\rho = \rho^\pm$; $r_m = ((m_B^2 - m_\rho^2)/(m_B^2 - m_{K^*}^2))^3 = 1.023$, calculated using values in the PDG book (2004) [38]; $\delta a = [a_7^u(\rho\gamma) - a_7^c(\rho\gamma)]/a_7^c(\rho\gamma)$; $(\overline{\rho}, \overline{\eta})$ is the coordinate of the third vertex of the UT on the complex plane (Figure. 2 b).

The $Re(\delta a)$ reflects the different contributions from weak annihilations in $B \to V\gamma$ decays, which is CKM-suppressed in $B \to (K^{*0}/K^{*+})\gamma$ decays. The calculation shows that $Re(\delta a) = 0.0 \pm 0.1$ and $Re(\delta a) = -0.4 \pm 0.4$ for the neutral modes and charged modes respectively [37]. Following the same Ref. [37],

$$\frac{|a_7^c(\rho\gamma)|}{|a_7^c(K^*\gamma)|} = 1.01 \pm 0.02$$

and

$$\left| \frac{\overline{\eta}^2 - \overline{\rho}(1 - \overline{\rho})}{(1 - \overline{\rho})^2 + \overline{\eta}^2} \right| < 0.2$$

are given. From the estimation of the various constants contained in Eq. 36, using the ratio of the branching fractions between $B \to (\rho/\omega)\gamma$ and $B \to (K^{*0}/K^{*+})\gamma$ could give a measurement of the $|V_{td}|/|V_{ts}|$ to approximately 15%, ignoring experimental uncertainties. The constraint on the $|V_{td}|/|V_{ts}|$ could be interpreted as a constraint on the $|V_{td}|$, which lies within the 90% confidence interval found from a global fit to CKM data of $0.0048 < |V_{td}| < 0.014$ reported here [13].

Due to the limitation of statistics for $B \to (\rho/\omega)\gamma$ decays, an isospin-averaged branching fraction of $B \to (\rho/\omega)\gamma$ decays is more attractive experimentally than individual decays. For example, among the decay rates of the $B \to (\rho/\omega)\gamma$ decay modes, the following relation is predicted, as discussed above,

$$\Gamma(B^0 \to \rho^0\gamma) = \Gamma(B^0 \to \omega\gamma) = 0.5 \cdot \Gamma(B^+ \to \rho^+\gamma). \tag{37}$$

Applying this constraint, the isospin-averaged branching fraction [35] is

$$\overline{\mathcal{B}}(B \to (\rho/\omega)\gamma) = \frac{1}{2} \cdot \{\mathcal{B}(B^+ \to \rho^+\gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot [\mathcal{B}(B^0 \to \rho^0\gamma) + \mathcal{B}(B^0 \to \omega\gamma)]\}, \tag{38}$$

with the world average for B-meson lifetime ratio, $\tau_{B^+}/\tau_{B^0} = 1.086 \pm 0.017$ [39]. With a similar definition for the $B \to (K^{*0}/K^{*+})\gamma$ decays, the ratio $\overline{\mathcal{B}}[B \to (\rho/\omega)\gamma]/\overline{\mathcal{B}}(B \to (K^{*0}/K^{*+})\gamma)$ could be related to $|V_{td}|/|V_{ts}|$ as follows [35, 40]:

$$\frac{\overline{\mathcal{B}}[B \to (\rho/\omega)\gamma]}{\overline{\mathcal{B}}(B \to (K^{*0}/K^{*+})\gamma)} = \frac{|V_{td}|^2}{|V_{ts}|^2} \cdot r_m \cdot \zeta[1 + \Delta R], \tag{39}$$

where $\zeta = \xi^{-2}$ describes the SU(3) flavor-breaking between ρ/ω and K^*, and ΔR accounts for annihilation diagrams. However, this would introduce larger theoretical uncertainties since the weak annihilation in $B^+ \to \rho^+\gamma$ has a larger effect.

CP and isospin asymmetries

Additional observables can be derived from the amplitudes of these decays, such as direct CP-asymmetry and Isospin-asymmetry. The direct CP-asymmetry is defined as

$$\mathcal{A}(B \to V\gamma) = \frac{\Gamma(\overline{B} \to \overline{V}\gamma) - \Gamma(B \to V\gamma)}{\Gamma(\overline{B} \to \overline{V}\gamma) + \Gamma(B \to V\gamma)}. \tag{40}$$

The Isospin-asymmetry for $B \to (K^{*0}/K^{*+})\gamma$ and $B \to (\rho^0/\rho^+)\gamma$ decays is defined as

$$\Delta_{0-}(B \to V\gamma) = \frac{\Gamma(\overline{B^0} \to \overline{V^0}\gamma) - c_V\Gamma(B^- \to V^-\gamma)}{\Gamma(\overline{B^0} \to \overline{V^0}\gamma) + c_V\Gamma(B^- \to V^-\gamma)}, \tag{41}$$

where $c_V = 1$ for K^* and $c_V = 1/\sqrt{2}$ for ρ. The uncertainty due to form factors are largely reduced in Eq. 30 and Eq. 33 because of taking ratios of decay rates.

At the NLO of α_s, the SM predictions of the inclusive $b \to s\gamma$ and the $b \to d\gamma$ CP-asymmetries are [41]:

$$\begin{aligned} \mathcal{A}(b \to X_s\gamma) &= +0.0044^{+0.0024}_{-0.0014}, \\ \mathcal{A}(b \to X_d\gamma) &= -0.102^{+0.0033}_{-0.0058}, \end{aligned} \tag{42}$$

where the theoretical uncertainties are dominated by the c-penguin correction and the scale dependence (μ). Applying the quark-hadron duality, similar CP-asymmetries in exclusive decays are expected. The isospin-asymmetry is mainly introduced at the LO of Λ_{QCD}/m_b by the weak-annihilation. The SM prediction [42] of the isospin asymmetry for $B \to (K^{*0}/K^{*+})\gamma$ decays is $(+0.080^{+0.021}_{-0.032})$. In the $B \to (\rho^0/\rho^+)\gamma$ decays, the isospin asymmetry shows strong correlations with the angle α of the UT. After these decays are observed, they can be used to provide supplemental information on this angle.

2.4 Physics Discussion

As discussed above, the ratio of the form factors can be determined with much less theoretical uncertainty than the individual ones. Therefore, Eq. 36 can be used to give a theoretically cleaner result for the $|V_{td}|/|V_{ts}|$. This ratio can be also determined from

$B_{d,s}$ mixing. In Figure 5, the dominant diagrams for $B_{d,s}$ mixing are shown, where transition of top quark dominates. A NLO calculation [43] gives the following relation between the mass difference (Δm_q) and the CKM elements V_{tq},

$$\Delta m_{B_q} \propto m_{B_q} \hat{B}_q f_{B_q} |V_{tq}|^2, \quad q = s, \ d, \tag{43}$$

with f_{B_q} form factor and m_{B_q} mass. Therefore,

$$\frac{\Delta m_{B_d}}{\Delta m_{B_s}} = \frac{m_{B_d} \hat{B}_d f_{B_d}}{m_{B_s} \hat{B}_s f_{B_s}} \cdot \frac{|V_{td}|^2}{|V_{ts}|^2}. \tag{44}$$

So, $|V_{td}/V_{ts}|$ can be measured from the B_d and B_s mixing. In this equation, the ratio $\frac{m_{B_d} \hat{B}_d f_{B_d}}{m_{B_s} \hat{B}_s f_{B_s}}$ is theoretically determined within a 10% uncertainty [44].

Due to large uncertainties in the determination of the form factors, the measurements of the branching factions in the decays $B \to (\rho/\omega)\gamma$ provide less competitive results for the CKM elements. However, these results could be used to constrain the form factors, helping the choice of various physics models to perform non-perturbative calculations. Other quantities, such as the CP-asymmetry and the isospin asymmetry in these decays, are predicted theoretically at a very accurate level; e.g. less than 1% error for the CP-asymmetry in $B \to (K^{*0}/K^{*+})\gamma$ decays. Measurements of these quantities provide a very stringent test of the SM [41].
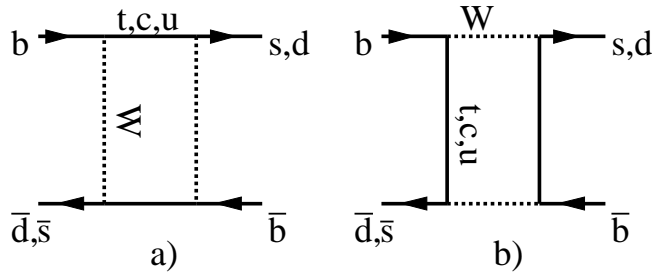
Figure 5: Box-diagram describing the $B_{s,d}$-mixing.

The decays $B \to (K^*/\rho/\omega)\gamma$ mainly occur via a loop process. Therefore these decays could be used to probe new physics beyond the SM. Some of the most attractive models beyond the SM are the two-Higgs-Doublet Models (2HDM) [45]. In these models, charged Higgs are introduced. As shown in Figure 6, in the penguin loop, new physics particles (charged Higgs $H^\pm$ and super-symmetric particle chargino ($\tilde{\chi}$)) can enter the penguin loop, and modify the corresponding Wilson coefficients. These would modify the SM predictions of the physics observables discussed above. So, a measurement of significant derivation of these observables from the SM prediction would indicate a new physics contribution. For example, the CP-asymmetry of the $B \to X_s\gamma$ decay in some supersymmetric models can be up to 10% [41]. The isospin asymmetry in the $B \to (K^{*0}/K^{*+})\gamma$ decays could have flipped sign or be greatly enhanced in some new physics models [42].
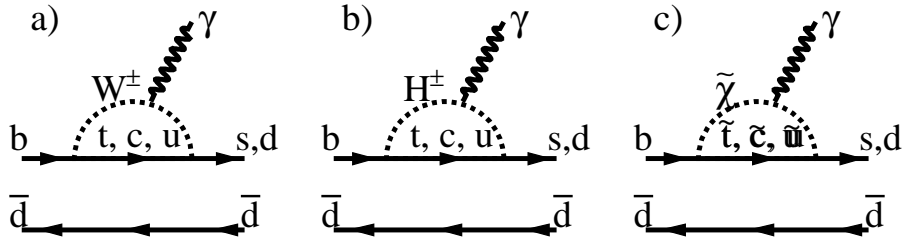
Figure 6: Possible new physics contributions to the radiative penguin decays are shown in b) and c), where the $W^\pm$-bosons are replaced by charged Higgs ($H^\pm$) or chargino ($\tilde{\chi}$).

3 The PEP-II B Factory and the *BABAR* Detector

The best way to study CP-violation in B-meson system is with an asymmetric e^+e^--collider operating at $\Upsilon(4S)$ resonance. The designs of the PEP-II B factory [46] and the *BABAR* detector [22] are optimized for such a physics study. However, it is also suitable for many other studies; e.g. precision measurements of bottom or charm physics, searching for rare B decays, or τ physics. A detailed description of the *BABAR* detector and the PEP-II B factory are given elsewhere [8, 22, 46]. A brief summary of their design and performance is presented in this chapter.

3.1 The PEP-II B Factory

The PEP-II B factory is an asymmetric e^+e^--collider designed to operate with luminosities of 3×10^{33} cm^{-2}s^{-1} at a center-of-mass energy of 10.58 GeV (the mass of the $\Upsilon(4S)$ resonance). The $\Upsilon(4S)$ resonance decays exclusively to $B^0\overline{B^0}$ and B^+B^- pairs, thus providing an ideal laboratory for the study of B mesons. In PEP-II, an electron beam of 9.0 GeV collides head-on with a positron beam of 3.1 GeV. The center of mass has a Lorentz Boost of $\beta\gamma = 0.56$ in the lab frame. This boost enables a separation of the decay vertexes of the two B mesons, making it possible to measure the time-dependent CP-asymmetry in B-meson system. This capability distinguishes PEP-II from previous B physics experiments.

Producing collisions since 1999, PEP-II has surpassed its original design goal and has delivered more than 250 fb^{-1} of data by 2004. During the best day, it delivered an integrated luminosity of 0.71 fb^{-1} data. The peak luminosity has reached 9.2×10^{33} cm^{-2}s^{-1}, which is more than three times of the design luminosity. More than 99% of the Bs delivered by PEP-II have been collected by the *BABAR* detector into one of the largest data sets in the world. Many topics have been studied using this data set, including the establishment of CP-violation in B-meson systems (angle β of the UT), other angles of the UT, the V_{ub}
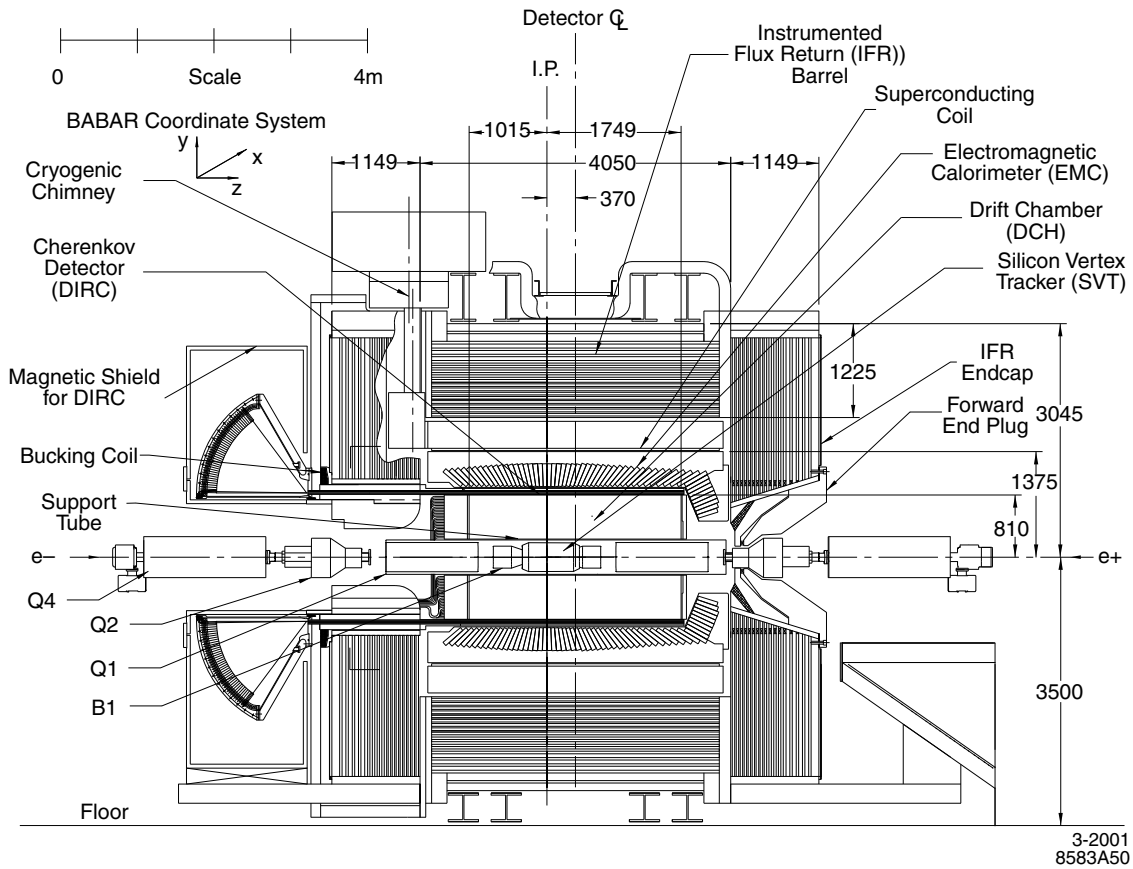
Figure 7: *BABAR* detector longitudinal section.

measurement and the direct *CP*-violation in $B \rightarrow K\pi$ decays. In the coming two years, the integrated luminosity delivered by PEP-II will reach 500 fb^{-1}, pushing *B*-physics study into new levels.

3.2   The *BABAR* Detector

The *BABAR* detector is designed to have high detection efficiency for the stable particles $e^{\pm}$, $\mu^{\pm}$, $K^{\pm}$, $\pi^{\pm}$, γ, and K^0s produced in e^+e^--collisions and to identify particle types for many physics studies. Figure 7 shows a longitudinal section of the detector. The detector surrounds the PEP-II interaction region. It is offset relative to the beam-beam interaction
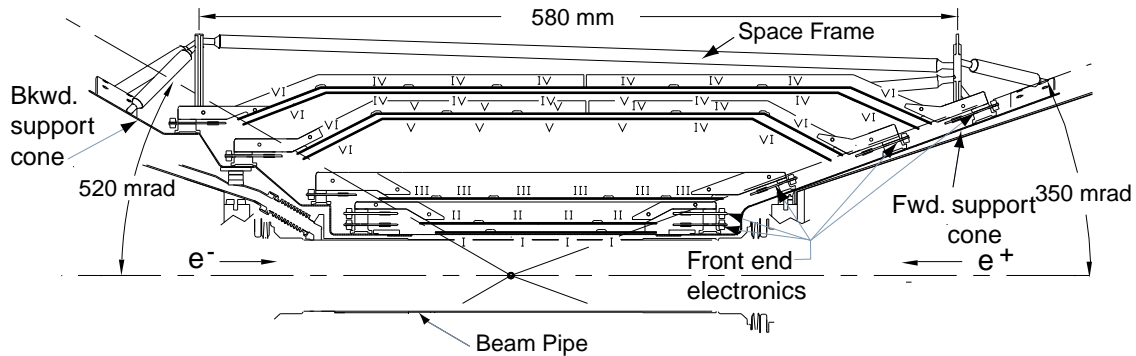
Figure 8: SVT longitudinal section.

point (IP) by 0.37 m in the direction of the lower energy beam, to maximize the geometric acceptance for the boosted $\Upsilon(4S)$ decays.

The inner detector consists of a silicon vertex tracker (SVT), a central drift chamber detector (DCH), a ring-imaging Cherenkov detector (DIRC), and a CsI electro-magnetic calorimeter (EMC). These detectors are surrounded by a superconducting solenoid that provides a magnetic field of 1.5 T. The iron flux return is instrumented for muon and neutral hadron detection. For convenience, the IP is chosen to be the origin and the beam line is chosen to be the z-axis for all measurements.

3.2.1 Silicon Vertex Tracker

The charged particle tracking system of the *BABAR* detector is made of two components, the Silicon Vertex Tracker (SVT) and the Drift Chamber (DCH). The primary physics goal of PEP-II and *BABAR* is to perform measurements of time-dependent *CP*-asymmetries in *B* decays. This requires very precise measurements of the momentum and angle of charged

particles to determine of position of the decay vertex of the B-meson[20]. The SVT is designed to satisfy these requirements.

Figure 8 shows the layout of the SVT. It is composed of five layers of double-sided silicon strip sensors, which are organized in 6, 6, 6, 16, and 18 modules for layers 1-5 respectively. The inner three layers are straight. The outer two layers are curved to increase the solid angle, and to decrease the amount of silicon. The large radii of the outer two layers is needed to link SVT tracks with DCH tracks. Strips on opposite sides of each sensor are oriented orthogonally to each other (parallel (z strips) and transverse (ϕ strips) to the beam axis), so that they provide ϕ measurements and z measurements, respectively. The SVT coverage extends down to 350 mrad and 520 mrad in polar angle from the beam line in the forward direction and in the backward direction, respectively.

SVT modules are divided electrically into two half-modules, which are read out at each end. The z strips of the sensors in the same half-module are electrically connected with wire bonds to form a signal readout strip. Signals from the z strips are brought to readout electronics using fanout circuits. The total number of readout channels is approximately 15,000. The readout electronics are mounted completely outside the active region of the detector to minimize materials within the acceptance. The SVT front-end custom electronics, A Time-Over-Threshold Machine (ATOM), samples the readout at 15 MHz and stores it in a 193-location buffer. Upon the receipt of an "accept" trigger from the data acquisition system, the output data are formatted, serialized, and delivered to the Readout Modules (ROM).

Only the SVT hits within 200 ns of an event time are included in track reconstruction. Adjacent strips are grouped into a cluster and used in the tracking reconstruction as an SVT hit (see Section 3.2.3). The SVT single-hit efficiency, which is the probability of linking an

---

[20]For example, for a fully-reconstructed B-meson, a 80 μm resolution or better on the decay vertex along the direction of the beam line must be achieved for a time-dependent CP measurement.

SVT hit with a DCH track, is around 97%, as measured on $e^+e^- \rightarrow \mu^+\mu^-$ events. For a track at a 90° incidence, the hit resolution is 15-20 μm for the three inner layers, and 20-40 μm for the outer two layers in excellent agreement with Monte Carlo (MC) simulations. For tracks with non-zero incident angles, the resolution is slightly decreased. The energy loss, dE/dx, of tracks in the SVT can be measured with a resolution of 14% for Minimum Ionization Particles (MIP). A 2σ separation between kaons and pions can be achieved up to a momentum of 500 MeV/c, and between kaons and protons up to 1 GeV/c.

3.2.2 Drift Chamber

The principal purpose of the DCH is to measure the momenta and angles of charged particles precisely. These high-precision measurements enable exclusive reconstruction of B- or D-mesons with minimum backgrounds. The DCH complements the measurements of the impact parameters and the charged track directions provided by the SVT and provides the only measurements for particles decaying or interacting outside the SVT volume[21]. The DCH is also used to supply information for the charged particle trigger, and to provide particle identification information for low-momentum charged particles using dE/dx.

Figure 9 shows a longitudinal cross-section of the DCH. The DCH is composed of 40 cylindrical layers of drift cells, providing up to 40 spatial and ionization loss measurements for charged particles. Longitudinal position information is obtained by the wires placed at small angles with respect to the z-axis. The DCH is asymmetrically located with respect to the IP. The forward and backward length are chosen in such a way that particles emitted at polar angles of 17.2° or 27.4° will traverse at least half of the 40 layers of the DCH before exiting through endplates for the forward or backward direction respectively. Instead of the commonly used argon-based drift chamber gases, a 80:20 mixture of helium:isobutane gas

---

[21]For example, K^0_S decay vertexes.
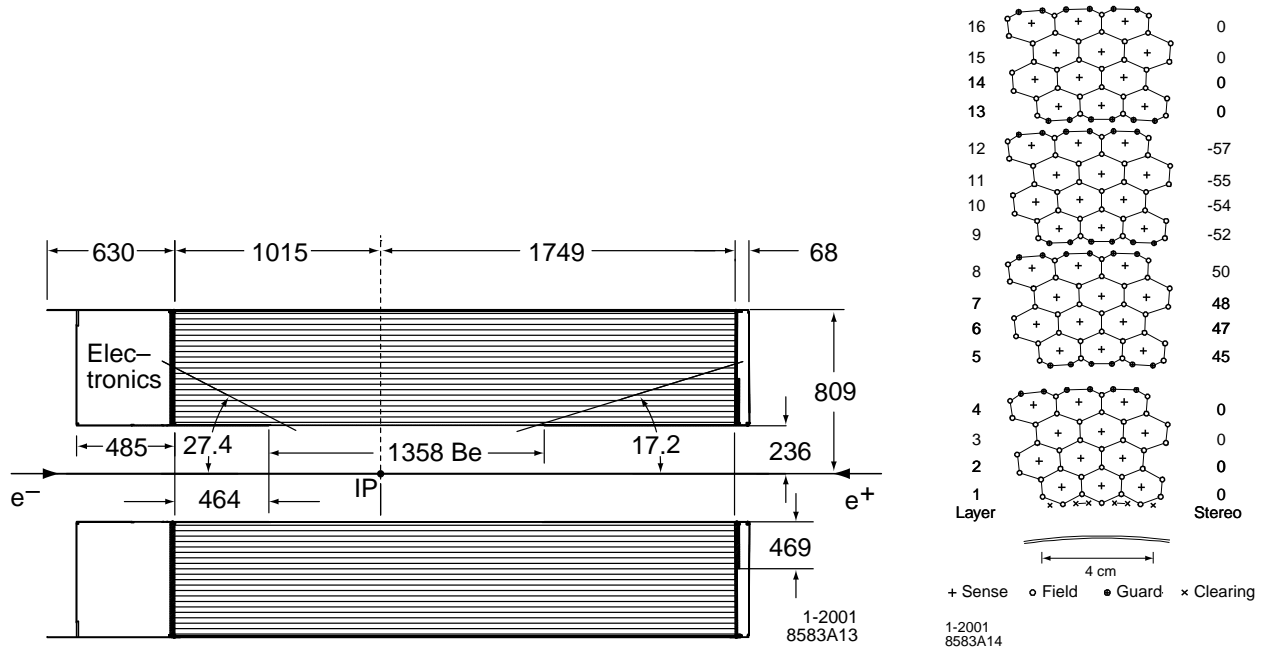
Figure 9: Left) Longitudinal section of the DCH with principal dimensions; the channel center is offset by 0.37 m from the interaction point; Right) Schematic layout of drift cells for the four innermost superlayers. The numbers on the right side of the plot gives the stereo angles (mrad) of sense wires in each layer.

is used to decrease multiple scattering. The 40 layers of drift cells are grouped into ten superlayers, with the same wire orientation and equal numbers of cells in each layer of a superlayer, as shown in Figure 9. The stereo angles of the superlayers alternate between axial (A) and stereo (U, V) pairs. The total number of drift cells is 7104.

The DCH electronics system measures the drift time and the integrated charge for every wire struck. The DCH amplifier and digitizer electronics are integrated into electronics front-end assemblies (FEAs) that are mounted on the rear endplate. The assemblies are connected to sense wires through service boards. A readout interface board in each FEA organizes the readout of the digitized data. Then data I/O and trigger I/O modules multiplex serial data from the FEAs to the readout modules.
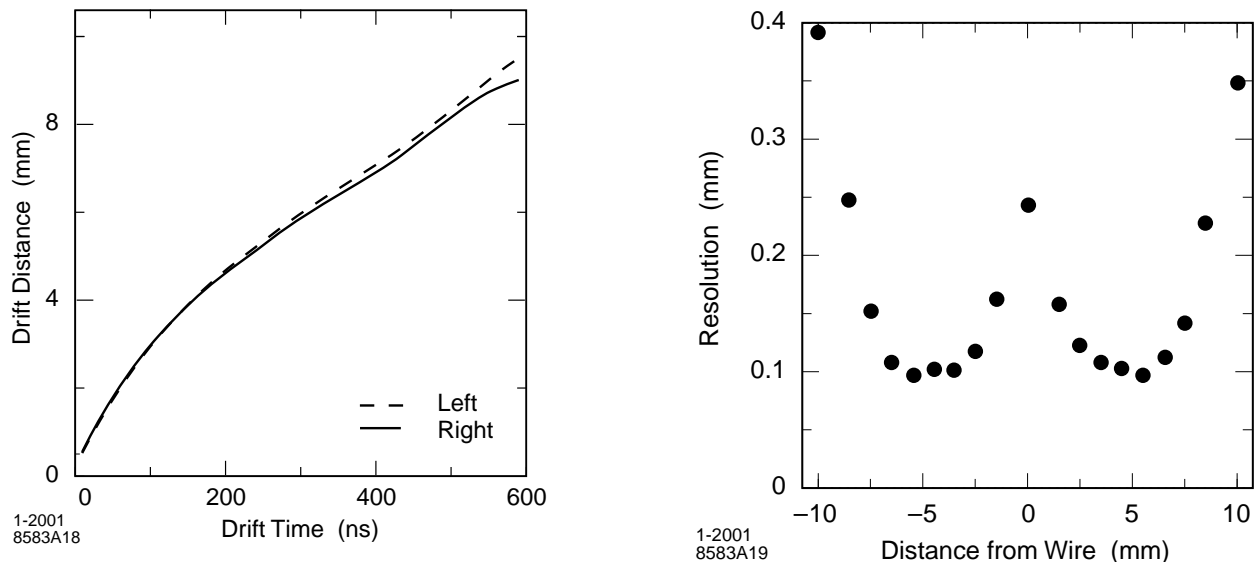
Figure 10: The drift time versus distance relation (left) and the position resolution as a function of the drift distance (right).

The DCH is operated at 1960 V, except for a brief time at 1900 V. Figure 10 shows the precise relation between the measured drift time and drift distance determined on $e^+e^- \to \mu^+\mu^-$ events, and the position resolution as a function of the drift distance, separately for the left and right side of the sense wire. The dE/dx for charged particles traversing the DCH is derived from measurements of the total charge deposited in each drift cell. A resolution of 7.5% is determined using Bhabha events. This resolution allows a π/K separation up to 700 MeV/c.

3.2.3 Performance of the Tracking System

Charged tracks are reconstructed by combining information from both the SVT and the DCH. The track finding and fitting procedures make use of a Kalman filter algorithm [47, 48]. Starting from 4-hit track segments in the DCH superlayers, tracks are selected by performing helical fits to the hits found by the track-finding algorithm. The resulting tracks are then

extrapolated into the SVT and the matching SVT track segments are added. Remaining SVT hits are passed to a stand-alone SVT track-finding algorithm. An attempt is made to combine tracks found only by one of the two tracking systems to recover tracks scattered in the material of the support tubes between the two detectors.

The absolute DCH tracking efficiency is defined as the ratio of the number of reconstructed DCH tracks to the number of tracks detected in the SVT with the requirement that they fall within the acceptance of the DCH. This efficiency is 98±1% for tracks with momentum of above 200 MeV/c and polar angle larger than 500 mrad. The SVT extends charged particle detection down to transverse momenta of 50 MeV/c, improving both tracking efficiency and resolution. The SVT single hit efficiency is around 97%. Position and angular resolutions of tracks near the IP are dominated by the SVT. The transverse decay position of a fully reconstructed B-meson is determined to a resolution of 70 μm and the resolution of the displacement of the decay vertexes on the z-axis is around 190 μm. The transverse momentum resolution, which is measured mainly by the DCH, is well represented by

$$\sigma_{p_t}/p_t = (0.13 \pm 0.01)\% \cdot p_t + (0.45 \pm 0.03)\%. \tag{45}$$

These values satisfy the design requirement and are reproduced by Monte Carlo simulations.

3.2.4 Detector of Internally Reflected Cherenkov Light

For many analyses in *BABAR*, it is essential to provide particle identification (PID) of the stable charged particles ($e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, and $p^\pm$) produced in the e^+e^- collisions to determine the underlying b-quark flavor in time-dependent *CP*-asymmetry analyses. More specifically, for the decays $B \to (K^*/\rho/\omega)\gamma$ the separation between kaons and pions is one of the keys to a successful measurement. The Detector of Internally Reflected Cherenkov Light (DIRC) is designed to identify charged kaons and pions.

The DIRC is a new kind of ring-imaging Cherenkov detector. Figure 11 shows a schematic diagram of the DIRC geometry. The main components of the DIRC are 144 fused synthetic silica bars with dimensions of 4.9 m long, 17.25 mm thin, and 35.00 mm wide. These silica bars are arranged in a 12-sided polygonal barrel of 12 silica bars each, surrounding the beam axis. When charged particles ($\beta \approx 1$) pass through these silica bars, Cherenkov photons radiated by the particle are reflected internally to the ends of the silica bars, travel through a water-filled standoff box, and are detected by an array of 10,752 photomultiplier tubes (PMT). To avoid instrumenting both ends of the silica bars with photon detectors, a mirror is placed at the end of the forward end, perpendicular to the bar axis, to reflect incident photons backward to the instrumented end. The PMTs are located at the rear of the standoff box, and grouped into 12 sectors, each containing 896 PMTs.
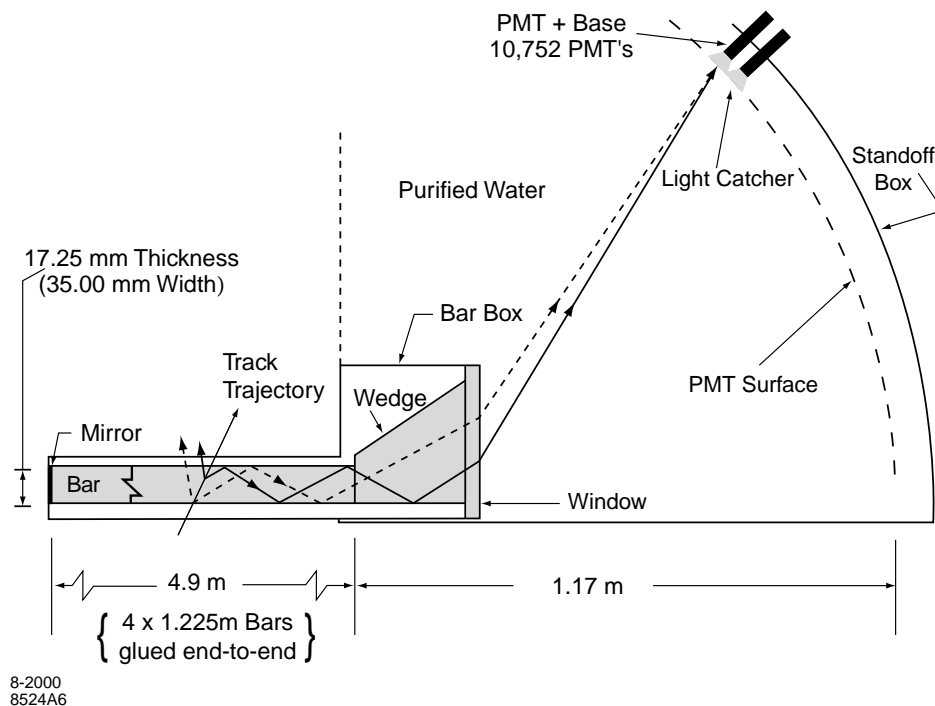
Figure 11: Schematics of the DIRC fused-silica radiator bar and imaging region.
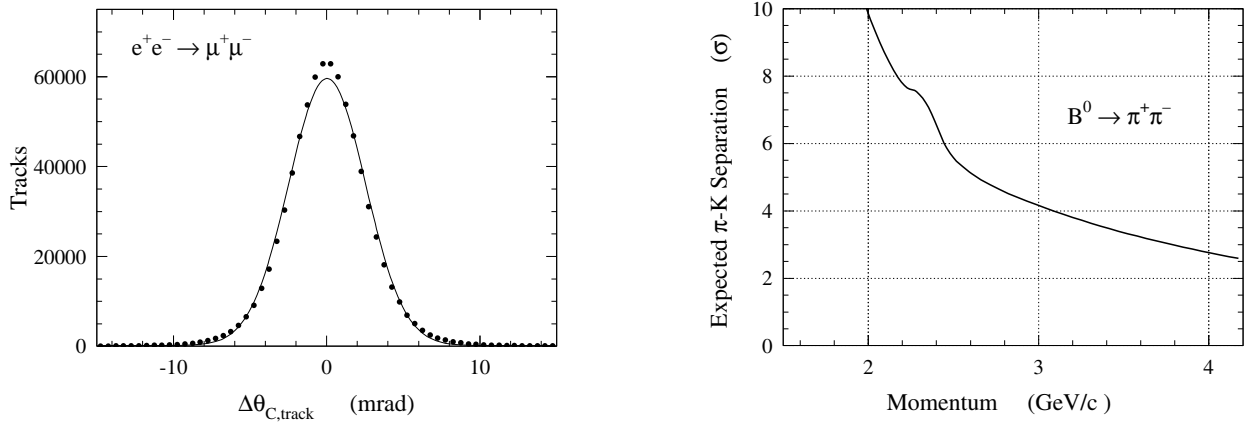
Figure 12: Left), the deferent between the measured and expected Cherenkov angle, $\Delta\theta_{c,\mathrm{track}}$, for $\mu^+\mu^-$ events; Right), expected $\pi - K$ separation in $B^0 \to \pi^+\pi^-$ events versus track momentum inferred from the measured Cherenkov angle resolution and the number of Cherenkov photons per track in dimuon events.

The DIRC is intrinsically a three-dimensional imaging device, measuring the position and arrival time of PMT signals. The emission angle of the Cherenkov photons is reconstructed from the observed space-time coordinates of the PMT signals, then transformed into Cherenkov angles (θ, ϕ with respect to the incident track direction). The measured arrival time is mainly used to suppress background hits from beam-induced backgrounds and, more importantly, to exclude photons from other tracks in the same event. The reconstruction routine currently provides a likelihood value for each of the five stable particle types (e, μ, K, π, and p) for every track passing through the the DIRC active volume. These likelihoods are calculated in an iterative process by maximizing the likelihood value for the entire event while testing different hypotheses for each track. If enough photons are found, a fit of θ and the number of the observed signal and background photons are calculated for each track.

Figure 12 shows the resolution of the track Cherenkov angle in di-muon events and

expected $\pi - K$ separation in the decays $B^0 \to \pi^+\pi^-$. The width of the fitted Gaussian distribution indicates a 2.5 mrad resolution. From the measured Cherenkov angle and the number of Cherenkov photons per track, a 4.2σ separation between kaons and pions at 3 GeV/c is expected in $B^0 \to \pi^+\pi^-$ decays. The efficiency for correctly identifying a charged kaon and the probability of wrongly identifying a pion as a kaon are determined using charged kaons and pions in $D^0 \to K^-\pi^+$ decays, which are selected kinematically from inclusive D^* decays. For a particular choice of particle-selection criteria, the mean kaon selection efficiency and pion misidentification are 96.2±0.2% and 2.1±0.1%, respectively. A further discussion of the different PID selectors used in $B \to (K^*/\rho/\omega)\gamma$ decays follows in later sections.

3.2.5   Electromagnetic Calorimeter

The electromagnetic calorimeter (EMC) is designed to measure electro-magnetic showers with excellent efficiency, energy and angular resolutions. This capability allows the detection of photons from π^0 or η decays as well as other electromagnetic and radiative processes, and
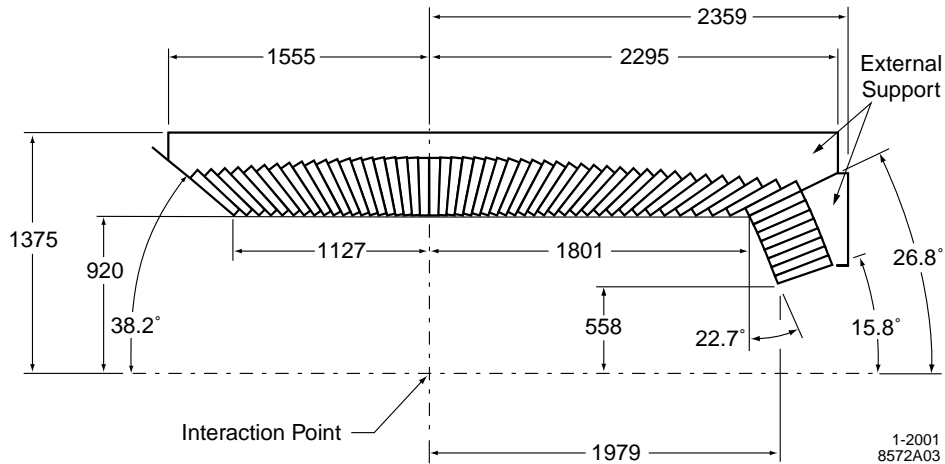
Figure 13: A longitudinal cross-section of the EMC (only the top half is shown) indicating the arrangement of the 56 crystal rings. The detector is axially symmetric about the z-axis. All dimensions are given in mm.

provides PID for electrons.

Figure 13 shows a longitudinal cross-section of the EMC, which consists of a cylindrical barrel and a conical forward endcap. The EMC has full coverage in azimuth and extends in polar angle from 15.8° to 141.8°, corresponding to a solid-angle coverage of 90% in the center-of-mass (c.m.) frame. The barrel contains 5760 thallium-doped cesium iodide (CsI(TI)) crystals, which are arranged in 48 distinct rings of 120 identical crystals. The endcap holds 820 crystals arranged in eight rings for a total of 6580 crystals. The crystals have a tapered trapezoidal cross-section, and are supported at the outer radius, to minimize the amount of material before the active detection region.

A typical electromagnetic shower spreads over many adjacent crystals, forming a cluster of energy deposits. Pattern-recognition algorithms have been developed to identify these clusters efficiently and to differentiate single clusters with more than one local energy maximum, which are referred to bumps. The position and energy of bumps are determined iteratively in these algorithms. The shower shape of a bump is used to identify a signal photon from a merged[22] π^0. Furthermore, the algorithms determine whether a bump is associated with any charged or neutral particle. For photon candidates, no association to any charged track is required.

The energy resolution of the EMC is determined from the physics processes shown in Figure 14. A fit to the energy-dependence results gives

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\text{GeV})}} \oplus (1.85 \pm 0.12)\%, \tag{46}$$

where $\oplus$ represents sum of the first term and second term in quadrature. A measurement of the angular resolution is based on π^0 (or η) decays to two photons of approximately equal

---

[22]When the open angle between the daughters of a π^0 is small, the bumps produced on the EMC largely overlap with each other, and the pattern recognition algorithms cannot differentiate them and treat them as a single bump, called a merged π^0. The second Fox-Wolfman moment of a bump tends to have a larger value if this bump is a merged π^0. This can be used to distinguish a single photon from a merged π^0.
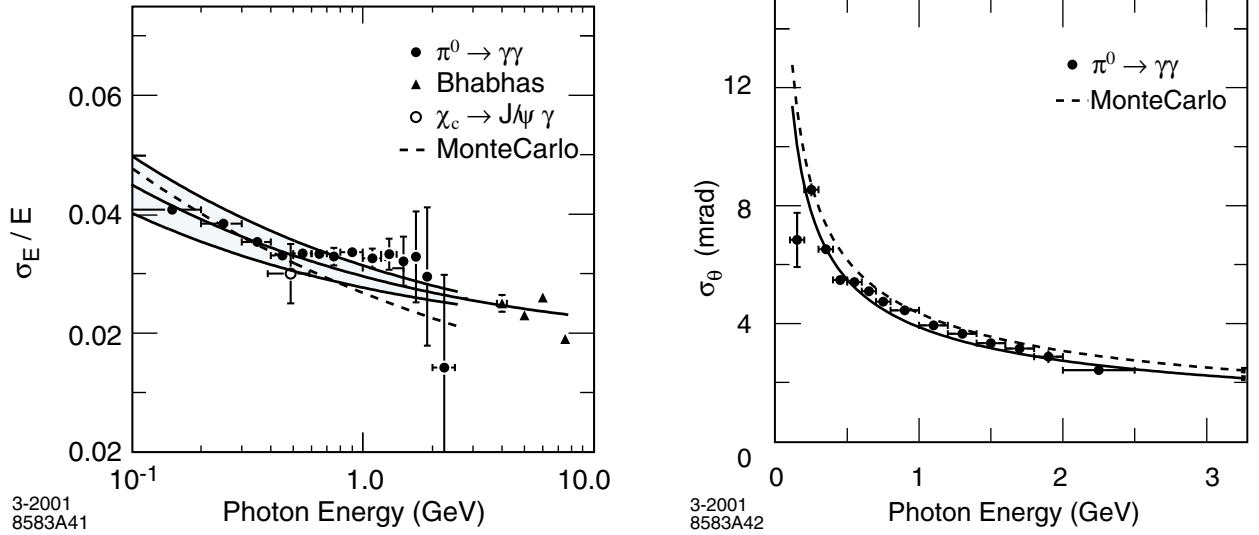
Figure 14: Left), the energy resolution for the EMC-measured photons and electrons from various processes, where the solid curve is a fit and the shaded region denotes the RMS error of the fit; Right), the angular resolution of the EMC for photons from π^0 decays.

energy. The result is shown in Figure 14. The resolution varies between about 12 mrad at low energies and 3 mrad at high energies. An empirical parameterization of the energy dependence is

$$\frac{\sigma_\theta}{\theta} = (\frac{3.87 \pm 0.07)\%}{\sqrt{E(\,\mathrm{GeV})}}) + (0.00 \pm 0.04) \ (\mathrm{mrad}), \tag{47}$$

where the first term and the second term is summed linearly. The PID of electrons relies mainly on the EMC, using shower energy, lateral shower moments, and track momentums to differentiate electrons from pions and kaons. Among those, the shower energy to the track momentum ratio (E/p) is the most important variable. In addition, the dE/dx energy loss in the DCH and the DIRC Cherenkov angle are also used. The performance of electron PID is checked using radiative Bhabha events, $e^+e^- \to e^+e^-e^+e^-$ events, and $K^0_S \to \pi^+\pi^-$ events. For a specific electron selector, the electron efficiency is 94.8% in the momentum range 0.5$< p <$2.0 GeV/c, and the probability of mis-identifying a pion as an electron is on the order of 0.3%.

3.2.6   Instrumented Flux Return

The Instrumented Flux Return (IFR) was designed to identify muons with high efficiency and good purity, and to detect neutral hadrons (primarily K_L^0 and neutrons) over a wide range of momenta and angles. The detection of muons is very important in several physics analyses, for example tagging the flavor of neutral B-meson, or studying the semileptonic and rare decays involving leptons from B and D mesons. The detection of K_L^0s allows the study of some exclusive B decays, particularly decays to some CP-eigenstates. In addition, the IFR can help in vetoing charm decays and improve the reconstruction of neutrinos by detecting K_L^0s.

The IFR uses the steel flux return of the magnet as a muon filter and hadron absorber. Single gap resistive plate chambers (RPCs) with orthogonal strip readouts were chosen as the detectors. Figure 15 shows an overview of the IFR. The steel is segmented into 18 plates, increasing in thickness from 2 cm for the inner nine plates to 10 cm for the outermost plates.
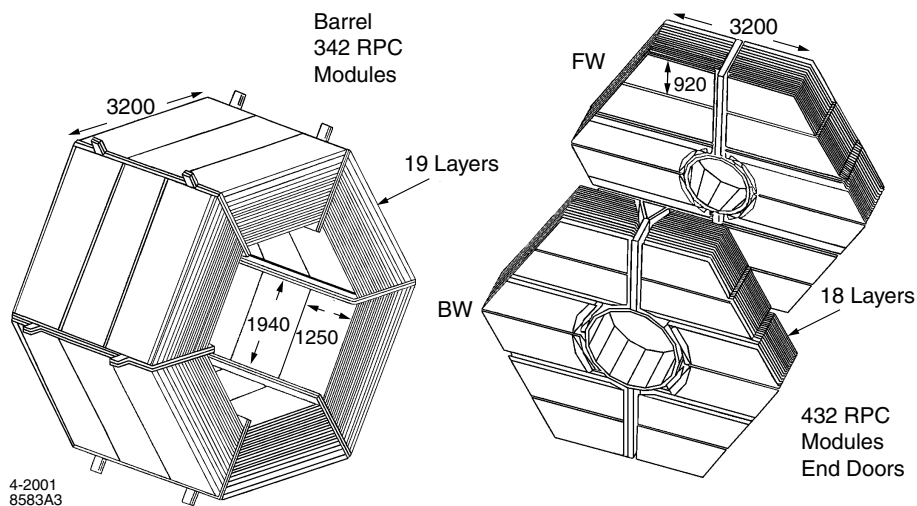
Figure 15: Overview of the IFR: Barrel sectors and forward (FW) and backward (BW) end doors: the shape of the RPC modules and their dimensions are indicated.

The nominal gap between the steel plates is 3.5 cm in the inner layers of the barrel and 3.2 cm elsewhere. RPCs are installed in these gaps, as illustrated in Figure 15. There are 19 RPC layers in the barrel and 18 in the endcaps. Additionally there are two layers of cylindrical RPCs installed between the EMC and the magnet cryostat to detect particles exiting from EMC. The IFR detectors cover a total active area of about 2000 m^2, and contains 806 RPC modules, 57 in each of the six barrel sectors, 108 in each of the four half end doors, and 32 in the two cylindrical layers. The size and the shape of modules are matched to the steel dimensions with minimal dead spaces. Due to the size limitation on the RPC modules, two or three RPC modules are joined to form a gap-size chamber, indicated in Figure 15. A mixed gas of 56.7% argon, 38.8% freon134a, and 4.5% isobutane is used to fill in all the RPCs.

The RPC hits are read out through front-end-electronics (FEE). The reconstruction of the IFR events starts with the grouping of individual hits into one- or two-dimensional clusters. The efficiency of the RPCs has been evaluated for both normal collision data and cosmic ray muons. After the installation and commissioning of the IFR system, all RPC modules are tested with cosmic rays, and more than 90% of the RPC has an efficiency above 90%. The muon PID has been tested using muons from $\mu^+\mu^-\gamma$ events and pions from three-prong τ and $K_S^0 \to \pi^+\pi^-$ decays. For a specific muon selector, the muon detection efficiency can be close to 90% in the momentum range of 1.5$< p < 3.0$ GeV/c with a pion fake rate of about 6-8%. K_L^0 and other neutral hadrons interact in the steel of the IFR and can be identified as clusters which are not associated with any charged track. No information of the energy can be obtained, but the polar angle of particles can be determined with an angular resolution of 60 mrad. The K_L^0 efficiency increases roughly with the momentum, varying between 20% and 40% in the momentum range from 1 to 4 GeV/c.

3.2.7 *BABAR* Trigger System

The *BABAR* Trigger System (TRG) is designed as a two-level hierarchy, a Level 1 trigger (L1) implemented in hardware and a Level 3 trigger (L3) implemented in software. The L3 receives the output from L1, performs a second stage rate reduction for the backgrounds associated with the main physics sources, and identifies and flags the special categories of events needed for luminosity determination, diagnostic, and calibration purposes. During normal operation, the output rate is about 2 kHz for L1 and around 300 Hz for L3.

The L1 trigger consists of four major components: the DCH trigger (DCT), the EMC trigger (EMT), the IFR trigger (IFT) and the global trigger (GLT). The trigger decision is based on charged tracks in the DCH, showers in the EMC, and tracks detected in the IFR. The DCT, EMT, and IFT generate the trigger primitives, summarizing the position and energy of particles, which are sent to the GLT every 134 ns. The GLT processes all trigger primitives to form specific triggers and then delivers them to the Fast Control and Timing System (FCTS). The FCTS can optionally mask or prescale any of these triggers. If a valid trigger remains, an L1 accept is issued to initiate event readout. The L1 trigger efficiency for $e^+e^- \rightarrow b\overline{b}$, $c\overline{c}$, $e\overline{e}$, and $\mu\overline{\mu}$ is almost 100%, with less than 0.4% efficiency loss; the trigger efficiency for $u\overline{u}$, $d\overline{d}$, and $s\overline{s}$ events and $\tau^+\tau^-$ events is slightly less, around 94-98%.

The L3 trigger software comprises event reconstruction and classification, a set of event selection filters, and monitoring. The filters have access to the complete event data for making their decisions, including the output of the L1 trigger processors and FCTS trigger scalers. In principal, L3 operates by refining and augmenting the selection methods used in L1. For example, the L3 DCH tracking algorithm performs fast track finding and fitting to extract the helical track parameters with improved track resolutions, further rejecting background tracks. The L3 trigger efficiency is about 99% for $b\overline{b}$, $c\overline{c}$, $e\overline{e}$, and $\mu\overline{\mu}$ events, and about 92-95% for $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, and $\tau^+\tau^-$ events.

3.3 Data Set

During normal data-taking, the electron beam energy and the positron beam energy are fixed to 9.0 GeV and 3.1 GeV, respectively. In the c.m. frame of the e^+e^- system, the beam energy is at the peak of the $\Upsilon(4S)$-resonance, as shown in Figure 16 a). B-mesons are then produced via $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ decays, forming the "on-resonance" data set. For the majority of the running time, PEP-II produces on-resonance data. For a small fraction of the running time, the beam energy is adjusted to be 40 MeV below $M_{\Upsilon(4S)}$ as indicated in Figure 16 a). This data is called the off-resonance data set. In *BABAR*, 10% of the total data is off-resonance.

There are some advantages to studying B physics with a B-factory like PEP-II, compared to other hadronic colliders. The most important one is high signal-to-background ratio. The cross-sections for fermion-pair production are shown in Figure 16 b). About 28% of the hadronic pairs are $b\bar{b}$ pairs. In the off-resonance data, all but the $b\bar{b}$ pairs shown in Fig-

$e^+e^- \to$	Cross-section (nb)
$b\overline{b}$	1.05
$u\overline{u}$	1.39
$d\overline{d}$	0.35
$s\overline{s}$	0.35
$c\overline{c}$	1.30
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	$\sim$40

a)   b)
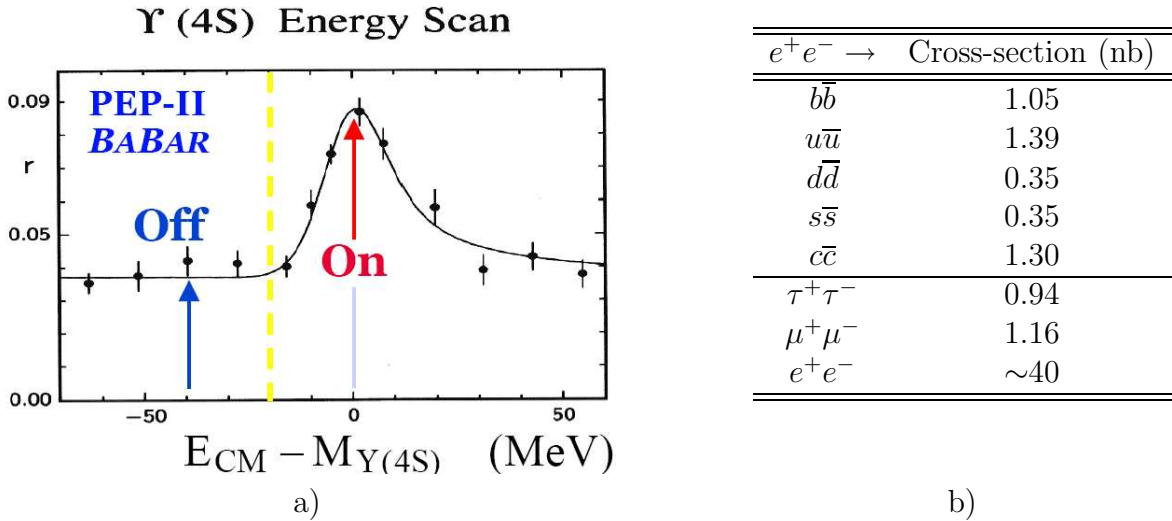
Figure 16: $\Upsilon(4S)$ energy scan is shown in a), where the red arrow and the blue arrow indicate the beam energy for "on-resonance" data-taking and "off-resonance" data-taking, respectively. Production cross-sections at $\sqrt{s} = M_{\Upsilon(4S)}$ for fermion pairs are shown in b).

ure 16 b) are produced. These processes are the major background (continuum background) for many *BABAR* analyses. The off-resonance data can be used for background studies in these analyses. Figure 17 shows the integrated PEP-II–delivered and *BABAR*-recorded luminosity since 15 October, 1999. A total 250 fb^{-1} data has been delivered by PEP-II. For the analysis of $B \to (K^{*0}/K^{*+})\gamma$ decays, we use data taken before July 2002, for the analysis of $B \to (\rho/\omega)\gamma$ decays, we use almost the entire data set.

Monte Carlo (MC) simulations are used in *BABAR* analyses for different purposes, such as event-selection optimization and background studies. The inclusive processes, such as the continuum processes and the inclusive $b\overline{b}$ process, are generated using JetSet [49]. This MC data is generally called "generic" MC; for example, generic B MC. In addition to these
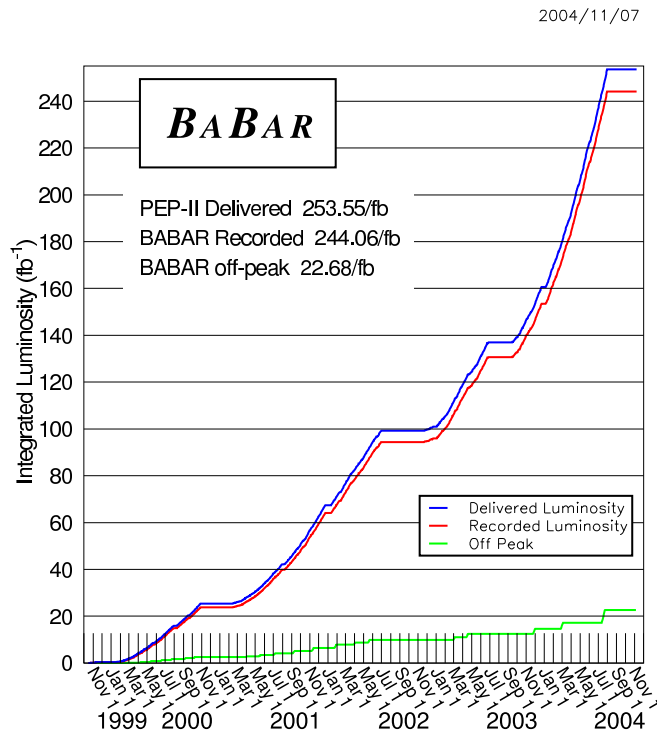
Figure 17: The integrated PEP-II-delivered and *BABAR*-recorded luminosity since 15 October, 1999. The three curves represent the PEP-II delivered (top), the *BABAR* recorded (middle), and the offpeak data (bottom).

inclusive processes, various exclusive B decays are generated using EvtGen [50]. For example, the signal MC for each $B \to (K^*/\rho/\omega)\gamma$ decays is generated using this generator. A more detailed description of the MC data sets used in *BABAR* can be found in Ref. [8].

4   Reconstruction

Reconstruction of the $B \to (K^*/\rho/\omega)\gamma$ events requires good EMC clusters (photons and merged π^0s) and well-identified charged kaons and charged pions. The information from these candidates is stored in the micro database, one of the data formats used by *BABAR* computing. These different formats are called event stores, and they are implemented to satisfy different needs, such as quality control, reconstruction, calibration, or physics analysis. Each format is optimized for that specific purpose [51].

4.1   Event Reconstruction

Figure 18 illustrates the decay chain of the modes of interest here. A total of seven exclusive decay modes will be discussed in the following sections. The reconstruction of the $B \to (K^*/\rho/\omega)\gamma$ decays is as follows. The vector meson is reconstructed from a combination of $K\pi$, $\pi\pi$, and $\pi\pi\pi$ for K^*, ρ, and ω meson, respectively. The K-meson includes charged kaons and K_S^0s, and the π-meson includes the charged and neutral pions. For the K_S^0 and π^0, only the decays of $K_S^0 \to \pi^+\pi^-$ and $\pi^0 \to \gamma\gamma$ are considered. The reconstructed vector meson $(K^*, \rho, \text{or } \omega)$ is then combined with a photon to form a B-meson candidate. This photon is called the "primary photon". In this thesis, both the decay mode and the charge-conjugate decay mode of each decay are analyzed together. This convention will be followed in the rest of this thesis unless otherwise specified.

Significant numbers of combinatorial backgrounds are present in each mode. For most background events, the primary photon is produced by π^0/η decays or Initial-State Radiation (ISR). In the c.m. frame of the e^+e^--system, the energy of the primary photon is about 2.6 GeV for the signal decays. This requirement strongly limits the combinatorial background, especially in the $e^+e^- \to b\bar{b}$ process, where the π^0/ηs are normally soft. However, in the continuum process, "$e^+e^- \to u\bar{u},\ d\bar{d},\ s\bar{s},\ c\bar{c},\ \tau^+\tau^-$", the event shape is jet-like and

$$B^0 \to (K^{*0}, \rho^0, \omega)\,\gamma \qquad\qquad B^+ \to (K^{*+}, \rho^+)\,\gamma$$

$$K^{*0} \to K^+\,\pi^-\ \ (0.67) \qquad\qquad K^{*+} \to K^+\,\pi^0\ (0.33)$$

$$K^0_S\,\pi^0\ \ (0.17) \qquad\qquad K^0_S\,\pi^+\ \ (0.33)$$

$$\pi^+\,\pi^- \qquad\qquad \pi^+\,\pi^-$$

$$\rho^0 \to \pi^+\,\pi^-\ \ \ (1.00) \qquad\qquad \rho^+ \to \pi^+\,\pi^0\ \ \ (1.00)$$

$$\omega \to \pi^+\,\pi^-\,\pi^0\ (0.89)$$
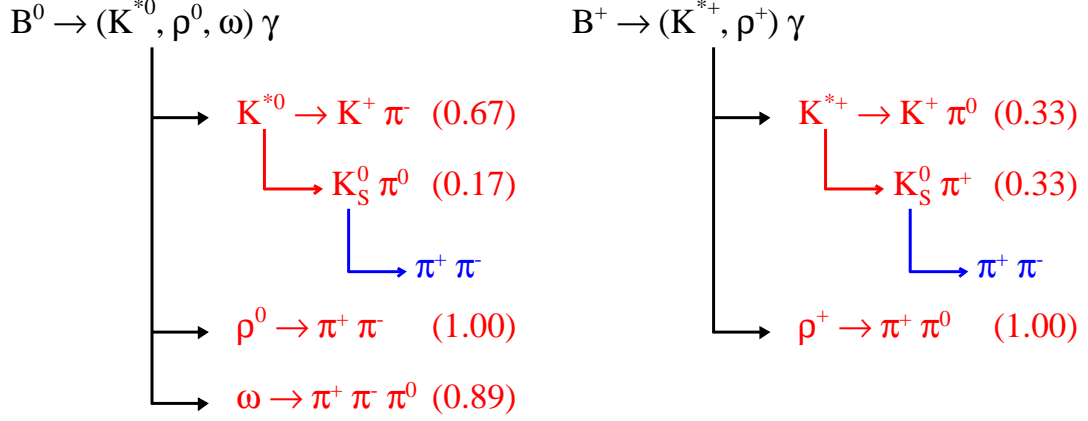
Figure 18: The decay chains of the exclusive decays discussed here. The number in the parenthesis indicates the fraction of the vector mesons decaying to the specific decay mode. The interested decay modes are those with decay fraction in parenthesis. For those modes, which contain K^0_S in the decays, only the $K^0_S \to \pi^+\pi^-$ decay is considered, and the 69% of K^0_S decays via this mode.

the π^0/ηs are boosted in the c.m. frame, therefore the π^0/ηs are relatively harder and could result in much higher combinatorial background, the "continuum background". A much more detailed discussion of the combinatorial background will be given in Section 5.1.

For many purposes in this analysis, such as estimating the signal and background yields, the following branching fractions are assumed:

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = \mathcal{B}(B^+ \to K^{*+}\gamma) = 4.0 \times 10^{-5}, \tag{48}$$

and

$$\mathcal{B}(B^0 \to \rho^0\gamma) = \mathcal{B}(B^0 \to \omega\gamma) = 0.5 \cdot \mathcal{B}(B^+ \to \rho^+\gamma) = 0.5 \times 10^{-6}. \tag{49}$$

These branching fractions can be combined with the decay fraction of the vector meson, shown in Figure 18, to give the assumed branching fraction of each exclusive decay mode.

4.1.1 Kinematic Variables - m_{ES} and ΔE^*

In the operation of the PEP-II, the beam energy is known very accurately. The produced $\Upsilon(4S)$-resonance decays exclusively to two B-mesons, which are almost at rest in the c.m. frame of the e^+e^- system. Exploiting these facts, two kinematic variables are widely used in the analysis of exclusive B decays, the beam-constrained mass, m_{ES}, and the energy difference between the beam and the reconstructed B-meson, ΔE^*. In the c.m. frame, m_{ES} and ΔE^* are defined as

$$\Delta E^* = E_\gamma^* + E_V^* - E_{\text{beam}}^*, \tag{50}$$

and

$$m_{ES} = \sqrt{E_{\text{beam}}^{*2} - (\vec{p_\gamma^*} + \vec{p_V^*})^2}. \tag{51}$$

Here the $E_{\text{beam}}^* = \frac{1}{2}\sqrt{s}$, where $\sqrt{s}$ is the total beam energy in the c.m. frame of the e^+e^- system (throughout this thesis, the (*) is used to indicate that a variable is calculated in the c.m. frame of the e^+e^- system.). Here E_γ^* (E_V^*) and $\vec{p_\gamma^*}$ ($\vec{p_V^*}$) are the energy and 3-momentum of the primary photon (the vector meson).

The variables m_{ES} and ΔE^* make optimal use of kinematic constraints and are largely uncorrelated. They can be used to separate the signal from the background in a likelihood fit. As shown in Figure 19, the signal m_{ES} and ΔE^* distributions can be well described by a Crystal-Ball function[23], a Gaussian distribution with a low-side power-law tail. The tail structure in the m_{ES} and ΔE^* results from some of the neutral candidates' energy leaking out of the EMC crystals. The signal m_{ES} distribution peaks at 5.28 GeV/c^2 with a resolution of 3 MeV/c^2. The signal ΔE^* distribution peaks around zero with a resolution of 40 MeV.

---

[23]The Crystal-Ball function [52] is defined as,

$$f(x) = C \cdot \begin{cases} e^{-(x-\langle x \rangle)^2/2\sigma^2} & , \text{ for } x > \langle x \rangle - \alpha\sigma \\ \left(\frac{n}{\alpha}\right)^n \cdot e^{-\frac{\alpha^2}{2}} / \left(\frac{\langle x \rangle - x}{\sigma} + \frac{n}{\alpha} - \alpha\right)^n & , \text{ for } x \leq \langle x \rangle - \alpha\sigma \end{cases}, \tag{52}$$

with $\langle x \rangle$ the mean, σ the width, and α and n are the parameters for the power tail.
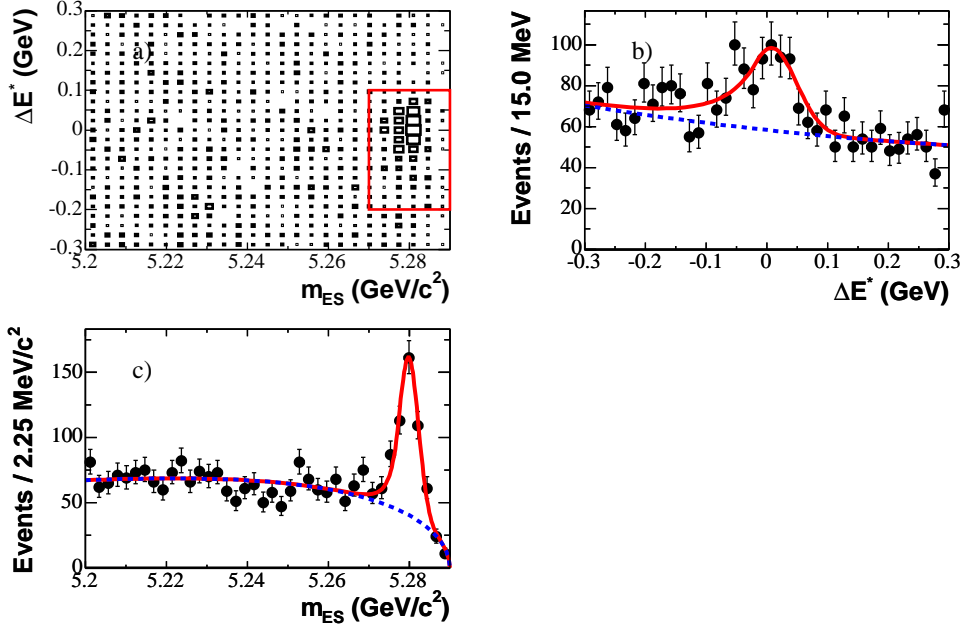
Figure 19: a) The m_{ES} and ΔE^* distribution of a $B \to (K^{*0}/K^{*+})\gamma$ Monte Carlo sample. The Monte Carlo sample corresponds to an integrated luminosity of 40 fb^{-1}, and contains continuum background and the signal decay only. A 2-dimensional fit performed with this MC data sample is projected onto plots of the c) m_{ES}, and b) ΔE^* distributions, respectively. The dashed curve indicates the background component and the solid line indicates the total fit. The fit uses all events in the m_{ES}-ΔE^* plane in a), which is the "fit region", and the small box indicates the "signal region".

The background underneath the signal peak is estimated using the events from a larger region on the m_{ES}-ΔE^* plane, and is subtracted directly in the fit. The background m_{ES} shape is well described by an ARGUS function[24], a threshold function with the threshold fixed at 5.29 GeV/c^2. The ΔE^* is modeled by a second-order Polynomial.

The resolution of the m_{ES} in the $B \to (K^*/\rho/\omega)\gamma$ decays could be improved by a process

---

[24]The ARGUS function [53] is defined as,

$$f(x) = C \cdot \frac{x}{E_{BEAM}} \cdot \sqrt{1 - \frac{x^2}{E_{BEAM}^2}} \cdot e^{-\xi\left(1 - \frac{x^2}{E_{BEAM}^2}\right)}, \quad (53)$$

with ξ the shape parameter and E_{BEAM} the threshold.

called rescaling, where the E_γ^* in Eq. 51 is scaled by a factor, k, which is determined from $k \cdot E_\gamma^* + E_V^* - E_{\text{beam}}^* = 0$. The energy leakage in the event is assumed from the detection of the primary photon only. Our studies show that this rescaling can improve the resolution of m_{ES} by 10-20% in the decays where vector mesons decay to charged particles only. This rescaled m_{ES} is used in the analysis of the $B \to (K^{*0}/K^{*+})\gamma$ decays. A further study also shows that this method increases the correlation between the m_{ES} and ΔE^* variables for some background components. So a new photon-energy calibration, which corrects the photon energy leak using the centroid position of the photon on EMC crystals, is used to improve the resolution of both m_{ES} and ΔE^* without increasing such a correlation. This calibration has been used in the analysis of $B \to (\rho/\omega)\gamma$ decays to correct the energy of the primary photons. In the following discussion, the m_{ES} will be used to represent both the m_{ES} and the rescaled m_{ES} unless otherwise specified.

4.1.2 Signal Region and Fit Region

For many purposes of the analysis, certain regions are defined, using the m_{ES} and ΔE^* variables. The most important ones are the signal region and the fit region

- **The signal region** is defined on the $\text{m}_{ES}-\Delta E^*$ plane to be within $[5.27, 5.29]$ GeV/$c^2\times$ $[-0.2, 0.1]$ GeV.

- **The fit region** covers a larger phase space than the signal region, where all events in this region can be included into a likelihood fit for final signal extraction, and is defined on the $\text{m}_{ES} - \Delta E^*$ plane to be within $[5.2, 5.29]$ GeV/$c^2\times [-0.3, 0.3]$ GeV.

The signal region and fit region are shown in Figure 19. Some other regions, such as sidebands, are also used in this analysis, and will be defined when mentioned.

4.1.3 Best Candidate Selection

Because of the presence of background particles in an event, multiple candidates of a signal decay can be reconstructed in the same event. Only one candidate is selected after all the applied selection cuts, including the background suppression cuts discussed in Section 5. The best candidate is selected as the minimum $|\Delta E^*|$.

4.2 Quality Selection Criteria

As mentioned in the beginning of this section, some quality selection criteria, including particle identification, are applied to the final decay particles or the intermediate particles after event reconstruction. These selection criteria are used mainly to reject poorly reconstructed candidates, or to ensure correct particle types. The background events are also suppressed to some extent by this selection.

4.2.1 Primary Photons

The primary photon candidate is selected from neutral EMC bumps, which are local maxima of the EMC clusters and are not associated with any charged track. The following selection criteria, which are described in details elsewhere [54], are applied

- The EMC coverage extends in polar angle from $15.8°$ to $141.8°$, $-0.774 < \cos\theta < 0.956$. To avoid both those photon candidates detected near the edge of the EMC, which most likely will be preshowered and the beam background contaminations, a further requirement $-0.74 < \cos\theta < 0.93$ is applied.

- Only neutral bumps with energy in the c.m. frame between 1.5 GeV and 3.5 GeV are considered.

- The number of crystals is required to be greater than four. However, this requirement

is redundant, since the average number of crystals for a photon within the specified energy requirement are typically about 20-30.

- The neutral bumps contain no dead, damaged, or noisy crystals. The noisy crystals are identified on a run-by-run basis. The occupancy of each crystal for the first 100 events of each run is used to determine if a crystal is noisy, by checking if it has greater than 10 MeV for half of the 100 events. The damaged crystals are identified by the online DAQ system if the event information is not completely read out.

- The neutral bumps must be at least 25 cm away from any other neutral bumps or charged tracks. This criterion is used to suppress the background from hadronic split-offs[25] and merged π^0s.

The following selection criteria are used to further suppress the merged π^0s or reject the photons from the π^0/η decays. These include the π^0/η-veto and the cut on the second-moment of the photon shower shape

- In the study of the $B \to (K^{*0}/K^{*+})\gamma$ decays, we veto the photons from $\pi^0(\eta)$ decays by requiring that the invariant mass of the candidate photon combined with any other neutral bump of laboratory energy greater than $50(250)$ MeV/c^2 to be outside the range from 115 MeV/c^2 to 155 MeV/c^2 (from 505 MeV/c^2 to 585 MeV/c^2). These vetoes effectively reject more than 90% of the background photons from the π^0/η decays.

    In the $B \to (\rho/\omega)\gamma$ decays, the expected branching fractions are two orders of magnitude smaller, and the expected background is much higher. To improve the singal significance, which is defined as $S/\sqrt{(S+B)}$ with S (B) the expected signal yield

---

[25]The "split-off" is a radiated photon candidate originating from a hadron candidate.

(background yield) in the signal region, the minimum energy requirement has been decreased to 30 MeV for the π^0-veto, and the corresponding veto ranges have been extended to [105, 155] MeV/c^2 and [500, 590] MeV/c^2 for π^0 and η-veto, respectively. These changes have been tested using Monte Carlo (MC) samples of $B \to (\rho/\omega)\gamma$ decays, and the signal significance is improved by roughly 10% for each $B \to (\rho/\omega)\gamma$ decay mode.

- The energy of a primary photon candidate spreads over tens of crystals. The shower shape can be characterized by the 2nd moment, defined as

$$2^{nd} \text{ moment} = \sum_{\text{crystal i}} \frac{E_i \cdot ((\theta_i - \theta_0)^2 + (\phi_i - \phi_0)^2)}{\sum_i E_i} \tag{54}$$

where the $\theta_i(\phi_i)$ and $\theta_0(\phi_0)$ are the centroid position of the ith crystal and the bump, respectively. The E_i is the energy of the ith crystal. The merged π^0 is distributed elliptically and has a large value of the 2nd moment. The signal photon is azimuth-symmetric around the centroid of the bump and tends to have small 2nd moment. A requirement that the 2nd moment less than 0.002 is applied to remove almost all the merged π^0s with almost no signal efficiency loss.

These selection criteria effectively reject background photons and ensure a good primary photon candidate.

4.2.2 Charged Particles

The charged particles, kaons and pions (except $\pi^\pm$ in K_S^0 decays), are selected from the track list with further requirements

- There are at least 12 DCH hits.

- The $|z_0|$ is less than 10 cm.

- The Distance of Closest Approach (DOCA) to the beam axis, d_0, is less than 1.5 cm.

- The transverse momentum of the track must be larger than 100 MeV/c.

- The tracking fit probability must be larger than zero. This probability could be zero because of some technical failures in the track reconstruction.

These requirements define a "GoodTrackLoose (GTL)" track, one of the standard track lists provided by the tracking analysis working group (AWG) for general usage in the collaboration. These selections ensure that good track parameters are measured. Since the K_s^0 has long life-time, the charged pions from K_s^0 decays may originate far from the IP; therefore, the GTL requirement is not applied. Selections with the K_s^0 candidate will be discussed in the following.

4.2.3 Charged Particle Identification

For this analysis, the most important particle identifications (PID) are kaon selection and pion selection. Between these two, pion identification has special importance in the analysis of the $B \to (\rho^0/\rho^+)\gamma$ decays, since the $B \to (K^{*0}/K^{*+})\gamma$ decays have branching fractions that are orders of magnitude higher, and $B \to (K^{*0}/K^{*+})\gamma$ decays can fake $B \to (\rho^0/\rho^+)\gamma$ decays by mis-identifying the charged kaon in a K^* decay as a pion. Kaon/pion PID [55] mainly uses the information from the DIRC. The dE/dx measurement in the SVT and the DCH also provides supplemental information. These are the only PID requirements for tracks with laboratory momenta less than 0.7 GeV/c, which is below the DIRC active range. The likelihood for each particle hypothesis (e, μ, π, K, and p) is calculated by multiplying the likelihoods from each sub-detector (SVT, DCH, and DIRC). A cut on the likelihood ratio, such as $L_K/(L_e + L_\pi + L_K + L_\mu + L_p)$, is applied to select the corresponding particle type, e.g. kaons. The performance of the selectors are studied using control samples, where
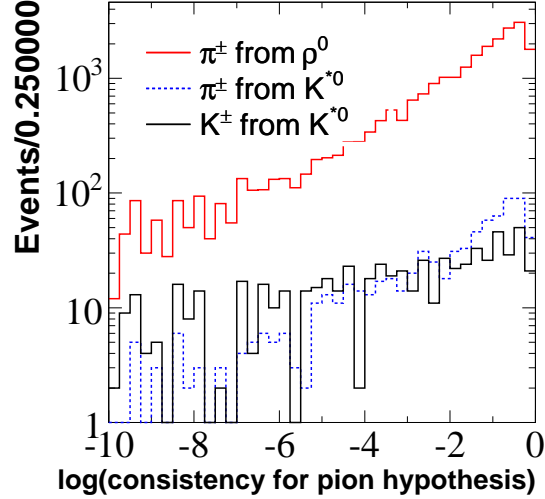
Figure 20: The number of observed photons in the DIRC for tracks with laboratory momenta greater than 0.6 GeV/c. Here we select the positive track in the $B^0 \to \rho^0 \gamma$ signal Monte Carlo and $B^0 \to K^{*0}\gamma$ background Monte Carlo. Shown are π from ρ decays (top solid line), π from K^* decays (dashed line), and true K from K^* decays that have been misidentified as π by the pion likelihood selector (bottom solid line).

the particle types can be identified kinematically. For example, the $D^* \to D\pi$ decay is used as a control sample for the study of kaon/pion PID, where the charged pion/kaon from the $D^0 \to K^-\pi^+$ decay can be identified by its charge correlation to the slow $\pi^\pm$ from $D^{*\pm}$ decay. Generally the data-MC difference shown on the control sample is considered the systematic error of the specific PID selector.

The PID selectors provided by PID AWGs fulfilled most of the PID requirements encountered in this analysis. However, a pion selector with lower kaon mis-identification rate using the available standard selectors is necessary to improve the experimental sensitivity significantly for the $B \to (\rho^0/\rho^+)\gamma$ analysis, given the orders-of-magnitude higher $B \to (K^{*0}/K^{*+})\gamma$ background. This has been achieved by further comparing the number of observed DIRC photons with the number of expected DIRC photons, after applying the
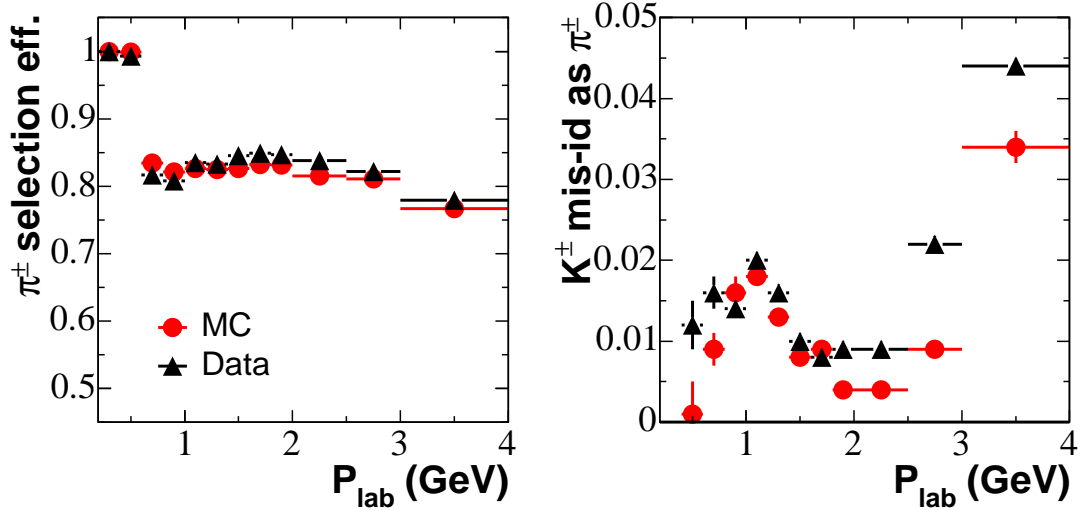
Figure 21: The performance of the pion selector used in the analysis of $B \to (\rho^0/\rho^+)\gamma$ decays. a) the pion selection efficiency, and b) the kaon mis-identification rate over the track momentum is shown. The efficiency and the mis-identification rate is obtained using a large sample of kinematically identified kaons and pions in the $D^* \to D\pi$ decays.

standard PID selectors. Figure 20 illustrates the photon consistency distribution[26] for $\pi^\pm$ or $K^\pm$ after a standard pion selector, likelihood-based pion"veryTight". A consistency of greater than 0.001 between the measured and the expected photons can suppress an extra 50% of the kaons which survived the "veryTight" selector, with a cost of a few percent loss on the pion-selection efficiency. This specially designed pion selector is used in the analysis of the $B \to (\rho^0/\rho^+)\gamma$ decays.

Figure 21 shows the selection efficiency of this pion selector and the corresponding kaon mis-identification rate over a large momentum range, on both $D^* \to D\pi$ data and MC.

---

[26]The photon consistency is defined as

$$\text{photon consistency} = 1 - \sum_{\forall i, f(i;\overline{x}) > f(x_0;\overline{x})} f(i;\overline{x}), \tag{55}$$

with x_0 the observed DIRC photon, the $\overline{x}$ expected DIRC photon with pion hypothesis, and $f(i;\overline{x})$ Possion distribution with mean $\overline{x}$.

The kaon mis-identification rate is around 1% over most of the momentum range, and the corresponding pion selection efficiency is around 80%. This pion selector is applied on all the charged tracks in the $B \to (\rho^0/\rho^+)\gamma$ decays. For the charged tracks in the $B^0 \to \omega\gamma$ decay mode, the likelihood-based "tight" selector, which has a slightly higher selection efficiency (around 90%) than the "veryTight", is applied. Charged kaons in $B \to (K^{*0}/K^{*+})\gamma$ decays are selected using a "SMS KaonTight" selector, which has a selection efficiency around 85% and the corresponding pion mis-identification rate is 2-3%. The charged pions (except from K_S^0 decays) in the $B \to (K^{*0}/K^{*+})\gamma$ decays are also required to fail the "SMS KaonTight" selection. The selectors are chosen differently across each decay mode based on a dedicated comparison of the performance of available selectors to maximize the signal significance.

4.2.4   Neutral Pions

The photons used in π^0 reconstruction have a minimum energy of 30 MeV and must a have shower shape consistent with a photon. Only π^0s with reconstructed masses between 0.115 GeV/c^2 and 0.150 GeV/c^2 are used. Further, assuming that the π^0 is from the primary vertex, the interaction point calculated using all the tracks in the event, a nominal π^0 mass-constrained fit has been used improve the momentum resolution of daughter photons.

For the π^0s from the $B^+ \to K^{*+}\gamma$, $K^{*+} \to K^+\pi^0$ and $B^+ \to \rho^+\gamma$, $\rho^+ \to \pi^+\pi^0$, additional requirements are implied

- The invariant mass of the photon-pair lies inside [0.117, 0.145] GeV/c^2. The mass window is asymmetric around the nominal π^0 mass window in order to reflect the tail of the mass distribution, which results from energy leakage in the EMC.

- $\cos(\theta_{\gamma\gamma}) > 0.6$. $\theta_{\gamma\gamma}$ is the angle between the two daughter photons in the laboratory frame.

These selection criteria are used to improve the purity of the π^0 in these decay modes. The $\cos\theta_{\gamma\gamma}$ cut can effectively suppress the π^0s from the beam background at low momenta.

4.2.5 Vertexing

In many *BABAR* analyses, vertexing the decay particles is needed for various physics purposes; e.g., to determine the displacement of the two B-mesons in a time-dependent CP analysis, or to decrease the combinatorial background when combining tracks to form a particle. To fully utilize these techniques in background suppression, the following vertexings are used. The detailed description of the *BABAR* vertexing methods can be found elsewhere [56].

- **"FastVertexing"**: a fast algorithm checks if several tracks originate from same point. This is used in the reconstruction of the vector mesons, which decay to at least two charged tracks, and it is able to suppress the combinatorial background in the reconstruction. The vertexing fit χ^2 probability is required to be larger than 0.0001.

- **The "GeoKin" vertexing**, which is the standard vertexing for *BABAR* time-dependent CP-analysises to determine the displacement between the reconstructed B-meson and the rest of the event (ROE). The displacement on the beam axis (Δz) is centered about zero for both signal and continuum background events, but the width of the distribution is wider in the signal decays, reflecting the fact that the B-mesons can fly. This distribution has been used in this analysis to suppress the continuum background. The $|\Delta z| < 0.4$ cm and $|\Delta z_{\mathrm{err}}| < 0.04$ cm are required to remove the non-physical events, such as collisions in beam gas.

4.2.6 K_S^0 Selection Criteria

The K_S^0 candidates used in this analysis are reconstructed from $K_S^0 \to \pi^+\pi^-$ decay only. A positive track is combined with a negative track to make a K_S^0 candidate with the following

requirements

- The χ^2 probability of "fastvertexing" is larger than zero.

- The K_S^0 flight distance, which is defined to be the distance between the decay point of K_S^0 and the primary vertex, to be larger than 0.3 cm.

- The K_S^0 mass lies inside [0.487, 0.508] GeV/c^2.

In the $B^0 \to \omega\gamma$ decay mode, to suppress the crossfeed from $B^0 \to K^{*0}\gamma$, $K^{*0} \to K_s^0\pi^0$ decay, which has the same final states as $B^0 \to \omega\gamma$ and has a branching fraction about 10 times larger, we calculate the distance between the decay point of ω and the primary vertex and require this distance to be less than 0.3 cm. This vetoes the K_S^0 decays and suppresses $B^0 \to K^{*0}\gamma$, $K^{*0} \to K_s^0\pi^0$ decays. This is called the "K_S^0-veto". With this "K_S^0-veto", this crossfeed can be suppressed further by another 40-50% with almost no signal efficiency loss.

4.2.7 Meson Selection Criteria

For the vector mesons, the χ^2 probability of "fastvertexing" has to be larger than 0.0001 for the applicable modes; we further require the mass of K^*-, ρ-, and ω- mesons to be within [0.8, 1.0] GeV/c^2, [0.63, 0.94] GeV/c^2, and [0.764, 0.795] GeV/c^2, respectively. These selection criteria are chosen from a dedicated optimization to maximize the significance for each decay mode.

4.3 Summary

After above event quality selections, in the signal region the background-to-signal ratio is about 1:1 and 100:1 for the $B \to (K^{*0}/K^{*+})\gamma$ decays and the $B \to (\rho/\omega)\gamma$ decays, respectively. A total of about 1500 $B \to (K^{*0}/K^{*+})\gamma$ signal events are expected with

80 fb^{-1} *BABAR* data. Therefore, the analysis of the $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays is somewhat different from the analysis of the $B \rightarrow (\rho/\omega)\gamma$ decays: the former is a precision measurement and the later is still a statistically dominant analysis. Different analytical methods are used, which are discussed in the following section in details.

5   Analytical Methods

In this section, the details of the analysis for $B \to (K^*/\rho/\omega)\gamma$ decays is discussed. This discussion concentrates on background suppression and final signal extraction. To suppress backgrounds, multi-variable analysis techniques, such as Fisher discriminant and neural network, are used. For final signal extraction, a likelihood-fit method, which is suitable for extraction of small signals in presence of large background, is used.

5.1   An Overview of Backgrounds

As mentioned in Section 4.1, the "continuum" background involves the processes, "$e^+e^- \to u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, $\tau^+\tau^-$". We study the continuum background using MC simulations of these processes and off-resonance data. However the background from $e^+e^- \to b\bar{b}$ process (named as "B-background") is much more complicated. In addition to these combinatorial back-
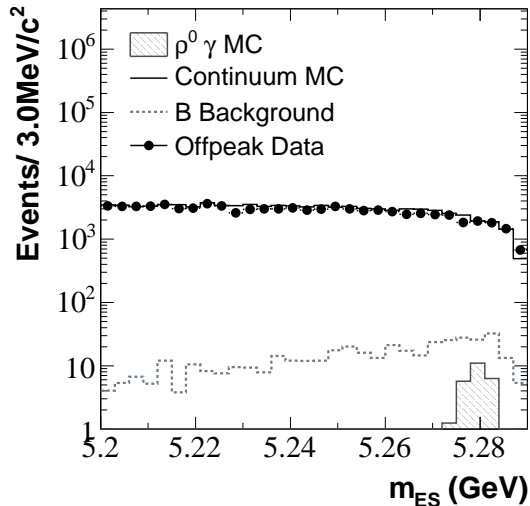
Figure 22: The m$_{ES}$ distribution of the signal, continuum background, and B-background yield in the $B^0 \to \rho^0\gamma$ decay mode at a luminosity of 200 fb^{-1}. All the quality selections discussed in Section 4 have been applied, and only the events in the fit region are used.

grounds, there exists other backgrounds, which are very signal-like.

Figure 22 shows the estimated yields of the continuum background, the B-background, and the signal at a luminosity of 200 fb^{-1} for $B^0 \to \rho^0\gamma$ decay mode, plotted as a distribution of the m$_{ES}$. The continuum background is more than two orders of magnitude higher than the signal yield, and the B-background is also very significant. Similar to other decay modes, the continuum background dominates while the B-background is not negligible.

For the $B \to (K^{*0}/K^{*+})\gamma$ decays, whose branching fractions are expected to be 10-60 times larger than the $B \to (\rho/\omega)\gamma$ decays, the background is also very different. The cross-section of the $u\overline{u}$ and $d\overline{d}$ processes is almost five times higher than the $s\overline{s}$ process [8], $\sigma_{e^+e^- \to u\overline{u},\ d\overline{d}} : \sigma_{e^+e^- \to s\overline{s}}$=1.74:0.35, and this results in 5-10 times more pions produced than kaons. Therefore the continuum background in the $B \to (\rho/\omega)\gamma$ decays are much higher than in the $B \to (K^{*0}/K^{*+})\gamma$ decays.

5.1.1   Continuum Background

In the continuum background, the primary photons mainly originate from two sources: π^0/η decays, or Initial State Radiation (ISR). Figure 23 shows the event shapes of the continuum background related to π^0/η decays and a signal decay in the c.m. frame. The event shape is spherical for the signal decay, where particles originate from two B-mesons, which are almost at rest and is jet-like for the background. The thrust axis[27] of the ROE, $\vec{T^*}$, is highly correlated with the direction of the primary photon in the continuum background, unlike in the case of the signal decay. The angle between $\vec{T^*}$ and the primary photon is defined as θ_T, whose distributions of the signal, the continuum background, and the B-background

---

[27]The thrust of an event is defined as,

$$T = \max_{|n|=1}\{\frac{\sum_i |n \cdot \vec{p_i}|}{\sum_i |\vec{p_i}|}\}, \tag{56}$$

with $|n|$ the thrust axis, named as "$\vec{T^*}$" here.
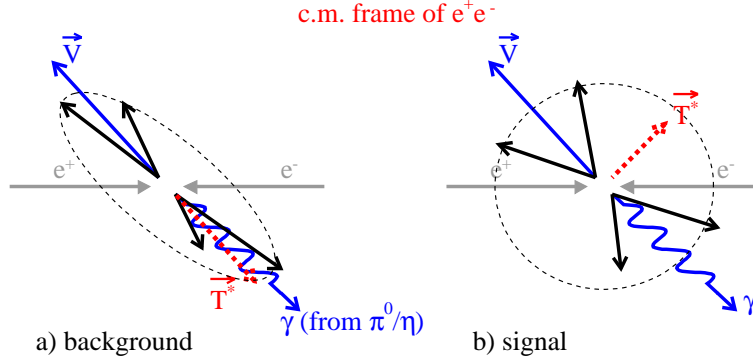
Figure 23: The event shape of, a) the continuum background (jet-like) and, b) the signal decay (spherical) in the c.m. frame of the e^+e^--system. After singling out the reconstructed signal, indicated by $\vec{V}$ and γ, the rest of event (ROE), indicated by unlabeled solid lines, are grouped together. The thrust axis of ROE, $\vec{T^*}$, is indicated by the dashed arrow.

are shown in Figure 24. For the $B^0 \to \rho^0\gamma$ signal decay and the B-background, the distribution is almost flat, but for the continuum background, it peaks around one. The $|\cos\theta_T|$ distribution gives the most significant suppression of the continuum background. Further, a likelihood fit over the m_{ES} (and ΔE^*) effectively separates the rest of this background from the signal, as discussed in Section 4.1.1.

Similarly, the event shape of the continuum background with ISR is also jet-like, but the correlation between the thrust axis and the primary photon is much weaker. The relative ratio between these two types of continuum background is different from mode to mode, but in general the one related to π^0/η decays contributes more.

5.1.2 B-Background

Some amount of B-background is also present for every decay mode. This background is not as significant as the continuum background, but the $\cos\theta_T$ (or the event shape) is not helpful in rejecting it. This is illustrated in Figure 24, where the B-background has a
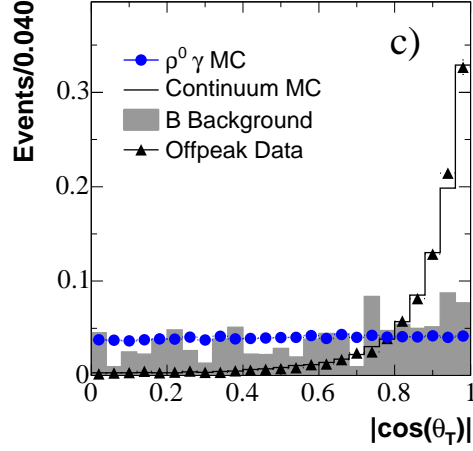
Figure 24: The $|\cos\theta_T|$ distribution for the $B^0 \to \rho^0\gamma$ signal, B-background, and continuum background.

Type	$B^0 \to \rho^0\gamma$ mode	$B^+ \to \rho^+\gamma$ mode	$B^0 \to \omega\gamma$ mode
$B \to (K^{*0}/K^{*+})\gamma$	$B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$	$B^+ \to K^{*+}\gamma, K^{*+} \to K^+\pi^0$	$B \to (K^{*0}/K^{*+})\gamma$
$B \to (\rho/\omega)(\pi^0/\eta)$	$B^0 \to \rho^0(\pi^0/\eta)$	$B^+ \to \rho^+(\pi^0/\eta)$	$B^+ \to \omega(\pi^0/\eta)$

Table 1: The peaking B-backgrounds for each $B \to (\rho/\omega)\gamma$ decay mode. Only $B \to (K^{*0}/K^{*+})\gamma$ and $B \to (\rho/\omega)(\pi^0/\eta)$ backgrounds are singled out from other B backgrounds to be considered separately.

distribution similar to the signal decay. Furthermore, the components of the B-background are much more complicated than those of the continuum background. For example, some of the B-backgrounds are very much signal-like, showing distributions of m$_{ES}$ and ΔE^* similar to the signal in the signal region; these are called "peaking B-backgrounds"[28]. These peaking B-backgrounds could significantly bias the signal yield in the fit if not properly handled.

Extensive Monte Carlo simulations, including more than 20 exclusive B-decays, have been studied to identify peaking B-backgrounds for every signal decay. Finally, we only

---

[28]More generally, the peaking B-background is related to some specific B-decays. The reconstructed event over these decays can show peaks around the signal region.

select the peaking B-background for $B \to (\rho/\omega)\gamma$ decays, shown in Table 1 and consider them separately from other B-backgrounds. Other possible peaking B- backgrounds are proven to be negligible from MC studies,

- $B \to (K^{*0}/K^{*+})\gamma$ **background in** $B \to (\rho/\omega)\gamma$ **decays.** The $B \to (K^{*0}/K^{*+})\gamma$ decays can be mistaken for $B \to (\rho^0/\rho^+)\gamma$ decays due to misidentification of the charged kaon as a pion for those decays shown in Table 1. The $B \to (K^{*0}/K^{*+})\gamma$ background in $B^0 \to \omega\gamma$ decay mode is more complicated. First, a $B^0 \to K^{*0}\gamma$, $K^{*0} \to K_s^0\pi^0$ decay has the same final states as a $B^0 \to \omega\gamma$ and can mimic the $B^0 \to \omega\gamma$ decay, but the m_ω is at the tail of the m_{K^*}. Secondly, studies show that a $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ decay could be mistaken for the $B^0 \to \omega\gamma$ decay due to mis-identification of the charged kaon as a pion, with a combination of a soft background π^0. Both these backgrounds show significant peaks around the signal region. To simplify the analysis, we do not single out any specific $B \to (K^{*0}/K^{*+})\gamma$ decay, and consider them as one peaking background.

  Many methods have been used to suppress $B \to (K^{*0}/K^{*+})\gamma$ background in $B \to (\rho/\omega)\gamma$ decay modes, such as PID (see Section 4.2.3) and K_s^0-veto (see Section 4.2.6). In $B^0 \to \omega\gamma$ decay mode, the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ background can be strongly suppressed by applying pion PID on all the charged pions.

- $B \to (\rho/\omega)(\pi^0/\eta)$ **background.** These decays can mimic $B \to (\rho/\omega)\gamma$ decays when the low-energy daughter photon in π^0/η decays is lost. Belle and *BABAR* have studied these decays recently[29], and showed that these backgrounds are also significant for

---

[29]Recent studies [57] in *BABAR* have shown that in addition to the well-measured $B^+ \to \rho^+\pi^0$ decay [58],

$$\mathcal{B}(B^+ \to \rho^+\pi^0) = (10.9 \pm 2.7) \times 10^{-6}, \tag{57}$$

there is 3.5 σ evidence for the $B^+ \to \rho^+\eta$ decay,

$$\mathcal{B}(B^+ \to \rho^+\eta) = (9.2 \pm 3.5) \times 10^{-6}, \tag{58}$$
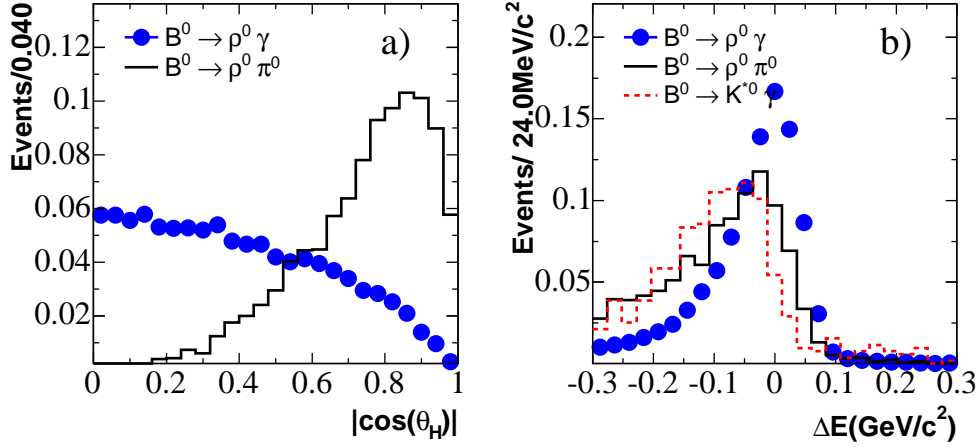
Figure 25: The a) $\cos\theta_H$ and the b) ΔE^* distribution of the signal and peaking B-background for the $B^0 \to \rho^0\gamma$ decay.

$B \to (\rho/\omega)\gamma$ decays.

These backgrounds can be strongly suppressed by a "helicity" angle cut. The helicity angle, θ_H, is defined as the angle between one daughter of the vector meson and the reconstructed B-meson in the rest frame of the vector meson with the exception in the $B^0 \to \omega\gamma$ decay; here, the θ_H is defined as the angle between the reconstructed B-meson candidate and the normal of the decay plane of the ω candidate. The θ_H is expected to follow a $\sin^2\theta$ distribution for signals and $\cos^2\theta$ for $B \to (\rho/\omega)(\pi^0/\eta)$

and over 4.3 σ evidence for the $B^0 \to \omega\eta$ decay [59],

$$\mathcal{B}(B^0 \to \omega\eta) = (4.0 \pm 1.3) \times 10^{-6}, \tag{59}$$

which we used in this analysis. For the $B^0 \to \omega\pi^0$ decay and the $B^0 \to \rho^0\eta$ decays [57], only 90% C.L. upper limits are set, which are 1.2×10^{-6} and 1.2×10^{-6}, respectively. In this analysis, we assume that

$$\mathcal{B}(B^0 \to \omega\pi^0) = \mathcal{B}(B^0 \to \rho^0\eta) = 1.0 \times 10^{-6}. \tag{60}$$

Belle has claimed the evidence for the $B^0 \to \rho^0\pi^0$ decay with a branching fraction of 5.1×10^{-6} [60], but *BABAR* only gives a upper limit of 2.9×10^{-6} [58]. A weight-averaged branching fraction is used in this analysis,

$$\mathcal{B}(B^0 \to \rho^0\pi^0) = 1.8 \times 10^{-6}. \tag{61}$$

backgrounds, as shown in Figure 25 a). The $|\cos\theta_H|$ is required to be less than $0.7(0.75)$ for $B^+ \to \rho^+\gamma$ decay and $B^0 \to \omega\gamma$ decay($B^0 \to \rho^0\gamma$ decay), to reject more than 50% of the $B \to (\rho/\omega)(\pi^0/\eta)$ backgrounds.

The $B \to (K^{*0}/K^{*+})\gamma$ background and the $B \to (\rho/\omega)(\pi^0/\eta)$ backgrounds involve either misidentification of a kaon as a pion or missing a particle. Therefore, some energy is missing, and the ΔE^* of these background events peaks around 50 MeV to 100 MeV below zero, as shown in Figure 25 b). This could help greatly the separation of the signal from these backgrounds in a fit involving the ΔE^*.

The remaining B-background for each decay mode is "combinatorial B-background", which is dominated by $B \to X_s\gamma$ decays, with $m_{X_s} > 1.1$ GeV/c^2 (here X_s represents any strange meson). In a $B \to X_s\gamma$ decay, one or more of the final decay particles are missed in the reconstruction, so the $B \to X_s\gamma$ decay could mimic a signal decay. Figure 26 shows the distribution of the combinatorial B-background on the $m_{ES} - \Delta E^*$ plane, and the projection onto the ΔE^*. A Gaussian-like peak is seen in the ΔE^* distribution, with the mean of the Gaussian outside the fit region. This is because that the $B \to X_s\gamma$ background mainly has one particle missing (e.g., a pion). The concentration of these backgrounds in one corner of the $m_{ES} - \Delta E^*$ plane can also cause significant bias on signal yields in a likelihood fit, if not properly handled.

5.2 Background Suppression Variables

For $B \to (\rho/\omega)\gamma$ decays, given the orders-of-magnitude high "background-to-signal ratio", a further rejection of the continuum background is essential. Several studies have been performed to improve the background suppression, and these studies are discussed below:
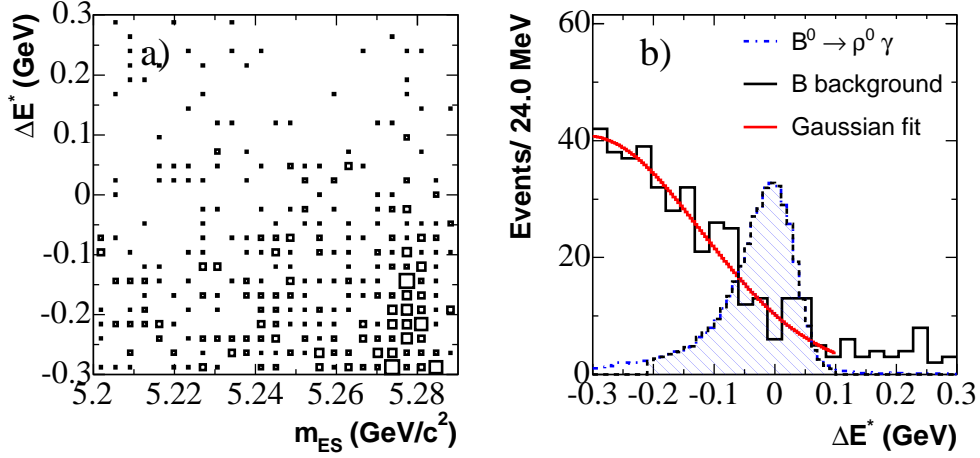
Figure 26: The a) m_{ES}-ΔE^* distribution of the combinatorial B-background, and b) the projection onto ΔE^* for the $B^0 \rightarrow \rho^0\gamma$ decay. The ΔE^* distribution in $[-0.3, 0.1]$ GeV is fitted with a Gaussian distribution. The $B^0 \rightarrow \rho^0\gamma$ signal distribution is also shown, and the signal events in signal region are in the shaded region.

- **Shape variables.** In addition to $\cos\theta_T$, some other variables are available to exploit fully the difference of the event shape between a continuum event (jet-like) and a B-decay event (spherical).

- **Variables related to signal decay kinematics.** Similar to $\cos\theta_H$, this type of variable is derived using the kinematic constraints on the signal side. These variables are effective in suppressing both the continuum background and B-backgrounds.

- **Physics content in ROE.** The B-meson decays mainly through $b \rightarrow c \rightarrow s$ via electro-weak interaction; however, the continuum background, especially the $e^+e^- \rightarrow u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ processes, mainly decay through hadronic interactions. This difference results in different multiplicities and momentum distributions of leptons, kaons, slow pions, etc. By using these particles in the ROE, some variables can be constituted to give further separation between B-decays and the continuum background.

5.2.1 Shape Variables Related to the ROE

To utilize the event shape difference fully the following variables are defined to characterize the event shape of the ROE:

- **18 energy cones.** The energy of the tracks and neutral candidates in the ROE are binned into cones with respect to the direction of the primary photon in the c.m. frame with a 10° increment in angle, as shown in Figure 27.

- **Legendre moments (L-moments).** In the c.m. frame, the L-moments [61] of an event are defined as follows:

$$
L_i = \frac{\sum_{j \text{ in ROE}} |\vec{P}_j| \times |\cos(\theta_j)|^i}{\sum_{j \text{ in ROE}} |\vec{P}_j|},
\tag{62}
$$
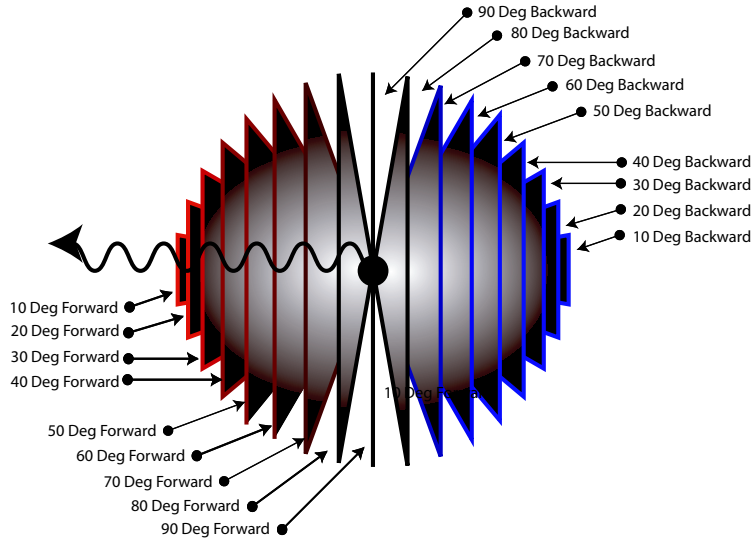
Figure 27: The 18 energy cones defined in the c.m. frame with respect to the primary photon. The cone variable is obtained by summing contribution in a cone (exclusively) from tracks and photons with momentum pointing within the cone.

here θ_j is the angle between the jth particle and a specific event axis. $\vec{P}_j$ is the momentum vector of the jth particle. Two event axes along the direction of the high-energy photon (donated by L_{jg}) and the direction of the thrust-axis $\vec{T}^*$ of the ROE (donated by L_j) can be used.

For historical reasons, 18 energy cones are used in the analysis of $B \to (K^{*0}/K^{*+})\gamma$ decays. However, in the $B \to (\rho/\omega)\gamma$ analysis, L-moments (L_1, L_2, L_3, and L_{2g}) are used. Based on the MC studies, using L-moments has been proven to give the equivalent background suppression as the 18 energy cones, but simplifying the configuration of the neural network (see Section 5.3).

To suppress the continuum background due to ISR events, the second-Wolfram moment is calculated in the photon recoil frame, R'_2. In this frame, an ISR event behaves like a two-jet continuum event, and can be effectively suppressed by the R'_2. This R'_2 variable is expected to be strongly correlated with other shape variables.

5.2.2 Decay Kinematics - θ_B, θ_H, and θ_D

In addition to $\cos\theta_H$ variable (Section 5.1.2 and Figure 25), the kinematics of $B \to (K^*/\rho/\omega)\gamma$ decays are also used to define the following variables:

1. θ_B is the angle between the reconstructed B-meson and the beam direction in the c.m. frame. The $\Upsilon(4s) \to B\overline{B}$ is a vector to scalar-scalar decay. Therefore, the $\cos\theta_B$ follows a $\sin^2\theta$ distribution for signal events, while for continuum background events and combinatorial B-background, there is no such angular constraint and the $\cos\theta_B$ is more or less flatly distributed.

2. **Dalitz angle, θ_D.** For an ω decay, θ_D is defined as the angle between the daughter π^+ and the π^0 in the rest frame of the $\pi^+\pi^-$ system. For a real $B^0 \to \omega\gamma$ signal event,

this angle θ_D follows a $\sin^2 \theta$ distribution. For the combinatorial B-background and continuum background, where lots of ωs are false, this distribution tends to be flat. However, some real ωs in the background can modify this distribution.

5.2.3 Physics Content in ROE - Tagging Variables

The *BABAR* tagging algorithm is developed to identify the underlying b-quark flavor of a tag B (ROE) at the time of its decay, by using the correlation between the charge sign of the quark and signed characteristics of the decay products. For example, in a semi-leptonic decay of B-mesons, the primary lepton has the same charge as the underlying b-quark. An algorithm is developed to identify the primary leptons from semi-leptonic B-decays, to determine the flavor of tag B. An output between minus one and one is given by this "lepton tagging" algorithm, which could be interpreted as a product of the probability of finding a primary lepton and the charge of the primary lepton. The distribution of the electron-tagging variable is shown in Figure 28; for B-decay events, including signal decays and B-backgrounds, the distribution peaks around minus one and one, which indicates a higher probability of finding semi-leptonic decays in the ROE. Similar tagging variables are also derived by exploiting other B-decays[30]. All these tagging variables are combined together using a neural network to maximize the tagging quality. The tagging algorithm has been developed and maintained by the *BABAR* "tagging" working group, the details of tagging algorithms are described elsewhere [62].

As shown in Figure 28, the distribution of tagging variables are slightly different between the B-decay events and the continuum events, due to different physics processes in the ROE.

---

[30]The "kaon tagging" uses charged kaons from $b \to c \to s$ decays; the "slowpion tagging" uses charged slow pions from D^* decay; the "kaon-slowpion tagging" uses correlations among these particles. Output similar to the "lepton tagging" is given by each of these "tagging" categories. Furthermore, the maximum track momentum in the c.m. frame ($\vec{p}^*_{max}$) is also used, which can be used to identify the B-flavor through two-body decays, such as $B \to D\pi$'s. This is the "maximum $\vec{p}^*_{max}$ tagging".
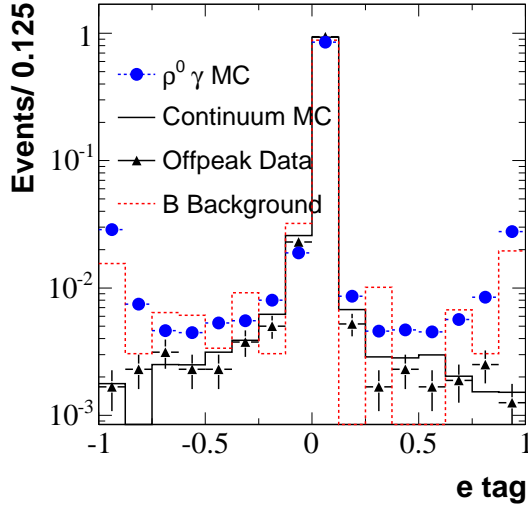
Figure 28: The distribution of the electron tag variable.

The existing kaon tagging only exploits the discriminant power carried by charged kaons and is optimized to give best separation between B-mesons and their CP-conjugated states. To utilize the strangeness produced in the ROE fully the following variables are calculated:

1. **Number of kaons.** The total number of kaons (including $K^{\pm}$ and K_s^0) in the ROE is calculated. The average number of kaons are around 1.0 in a B-decay, where the dominant decay of B-mesons is $b \to c \to s$ decay and is rarer in a continuum event, especially in $e^+e^- \to u\overline{u}$, $d\overline{d}$ processes, where there is no strangeness produced.

2. **Maximum momentum of the kaons in the c.m.frame.** For an event with at least one kaon, the maximum momentum of kaons in the c.m. frame is also calculated. This is helpful in distinguishing the strangeness produced in signal decays from $s\overline{s}$ events, where the generated strangeness is boosted, giving it higher momenta.

From Monte Carlo studies, these two variables proved to give better performance to distinguish B-decays from continuum events than the standard kaon tagging.

Variables	cont. bkg	combinatorial B bkg.	$B \to (K^{*0}/K^{*+})\gamma$ bkg.	$B \to (\rho/\omega)(\pi^0/\eta)$ bkg.
θ_T	s	$-$	$-$	$-$
energy cones/ L-moments	w	$-$	$-$	$-$
θ_H	$\sqrt{}$	$\sqrt{}$	$-$	s
θ_B, θ_D	$\sqrt{}$	$\sqrt{}$	$-$	$-$
taggings net flavor	w	$-$	$-$	$-$

Table 2: The summary of the background suppresion variables, and their suppresion power on various backgrounds shown, with "s" most significant, "w" weakly, "$-$" almost no effect, and "$\sqrt{}$" big enough to apply a cut directly. The $B \to (K^{*0}/K^{*+})\gamma$ and $B \to (\rho/\omega)(\pi^0/\eta)$ background are only considered separately from other B-background in the analysis of $B \to (\rho/\omega)\gamma$ decays.

These tagging variables are used in the analysis of $B \to (\rho/\omega)\gamma$ decays. In the analysis of $B \to (K^{*0}/K^{*+})\gamma$ decays, a simpler, but less powerful variable is used. We defined the "net flavor" content as $\sum_i |N_i^+ - N_i^-|$, where $N_i^\pm$ are the numbers of $e^\pm$, $\mu^\pm$, $K^\pm$, and slow pions of each sign identified in the event. On average, the net flavor is larger in the $B\overline{B}$ decays.

Table 2 summarizes all the background-suppression variables used in this analysis, and their suppression powers of various backgrounds. Some of these variables are correlated with each other, e.g. the shape variables. Some of these variables only provide weak separation between signal events and background events. A cut on these variables is not suitable. Therefore, to maximize the background suppression, multivariate techniques (MVA) are used to combine all these variables. An additional advantage of using MVAs is that they can can simplify the analysis.
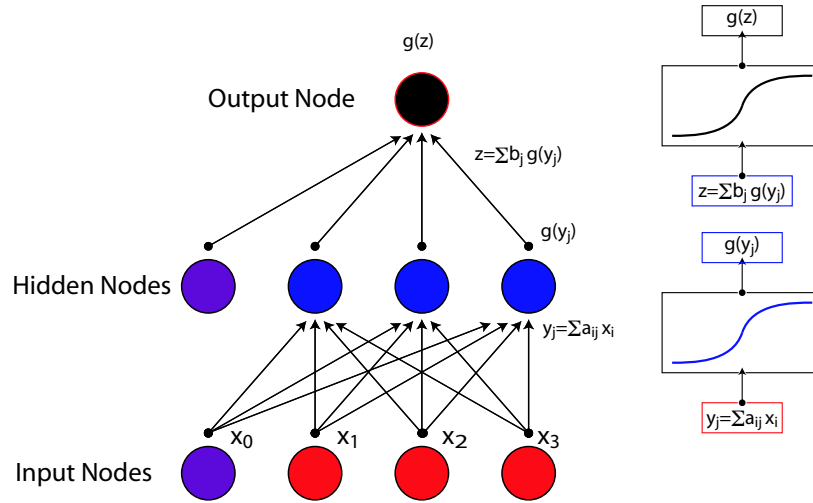
Figure 29: Basic structure of a single hidden layer NN. Input nodes (red) corresponding to event variables are passed to "hidden" nodes (blue) as linear combinations. The line combination is then transformed by an "activation function," in this case the $\tanh(x)$. Linear combinations of the hidden node outputs are then passed to the output node (black), where it is transformed once again by the activation function to give the final NN output.

5.3 Multivariate Analysis Techniques

Currently, two types of multivariate techniques are widely used in experimental particle physics, neural network [63] and linear Fisher discriminant. Both are used in this analysis.

5.3.1 Neural Network

The basic structure of an neural network (NN)[31] is depicted in Figure 29. At the bottom of the diagram, event variables x_i enter the network as input nodes. Linear combinations of these variables are sent up to the hidden nodes; the jth hidden node will receive receive the linear combination:

$$y_j = \sum a_{ij} x_j. \tag{63}$$

---

[31]The following description of NNs is based on reference [64].

Each hidden node receives a different linear combination of the input variables. This input is then transformed by an "activation" function, in this case the tanh function. The activation function determines how fast the output of the hidden layer varies as a function of the input. If the input range is small, a linear response is recovered; if it is large, a step response results. At each stage, there is a "bias" node which provides a constant output. The bias node is used to set the zero-level of the output.

A linear combination of the output from the hidden nodes $g(y_j)$,

$$z = \sum b_j g(y_j), \tag{64}$$

is then passed to the output node where it is again transformed by an activation function. The mathematical chain from input to output is shown on the right of Figure 29. An NN can, in principle, have many hidden layers and output nodes. In general, NN structures can be specified by the number of input nodes (network variables), the number of hidden layers with the number of nodes in each hidden layer, and number of output nodes. Hence, the example in Figure 29 can be described as a NN with three input nodes, one hidden layer with three nodes and one output node. In this analysis, we consider only configurations with one hidden layer.

The free parameters of an NN are the coefficients, a_{ij} and b_j, which are used to form the linear combination of input nodes to the hidden nodes, and the hidden node outputs to the output node, respectively. The optimal set of coefficients must be determined in advance through some data set (known as "training samples"), whose output for each input is already determined. This process is known as the "training" of the NN. The performance of the NN for any given set of coefficients is summarized by the sum-squared error (SSE),

$$\text{SSE}(a_{ij}, b_j) = \sum_{a=1}^{N} \left[\text{NN}(\vec{x_a}; a_{ij}, b_j) - F(\vec{x_a}) \right]^2, \tag{65}$$

here the x_a represents the vector of input variables for the ath event, $\mathrm{NN}(\vec{x_a}; a_{ij}, b_k)$ is the NN output for this vector with the coefficients a_{ij} and b_j, and $F(\vec{x_a})$ is the desired output for this vector (e.g. zero if it is a vector corresponding to a background event, one if it is a signal event). The SSE then represents a "χ^2" for the network configuration that can be minimized in a manner completely analogous to a fit, via gradient descent. The corresponding coefficients are then the optimal set of the coefficients. This "back-propagation" algorithm uses the derivatives of the SSE relative to changes in the coefficients to evaluate and adjust them iteratively to minimize the SSE. Each iteration in this process is one "cycle" of the training process.

In our application, instead of the SSE, the Mean-squared-error (MSE), which is defined as MSE = SSE/(Number of Events), is used to quantify performance of the NN. Two independent data samples are used, one for training, one for validation purposes. The output of the validation sample is already known in advance. The trained NN is then applied on the validation sample after each training cycle, and the corresponding MSE is obtained. The variation of the MSE with respect to the training cycle number for both training sample and the validation sample are then plotted, shown in Figure 30 a). The minimum of the validation distribution is then selected as the optimal set of coefficients for the specified NN, which gives the best performance on an independent sample. In practice, when the MSE of validation is decreasing very slowly or already flat, it can be thought that the performance of the current set of coefficients is very close to the optimal ones and the training can be stopped. The NN package used for this analysis is the Stuttgart Neural Network Simulator (SNNS) [65]. A ROOT-based package, RooSNNS, was then developed to configure, train and validate the NN conviniently. The MSE vs. training cycles for one of the NN used in this analysis is shown in Figure 30 a). The first few training cycles dramatically improve the performance of the NN, and the following training cycles improve the performance slowly.
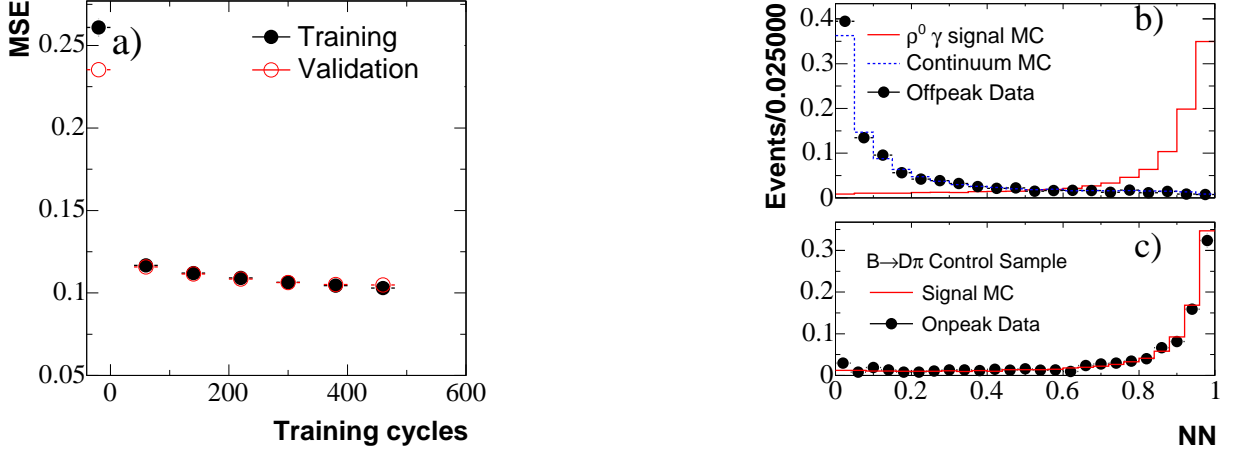
Figure 30: The a) variation of the MSE respecting to training cycles, b) and c) NN distributions of various data samples for a NN used in the $B^0 \rightarrow \rho^0\gamma$ decay mode.

All the training of the NN is performed using Monte Carlo samples. Figure 30 b) shows the distribution of the NN output for the continuum background and signal events in the validation sample, which is obtained directly from the RooSNNS. The signal peaks around one, and the background peaks around zero. A cut on this output could be applied to reject the continuum background. Since the NN is trained and validated on the MC data samples, the quality the simulations, can actually affect the performance of the NN when applied to the real data. So, various validations have been performed:

1. The NN distribution of the continuum MC is compared with the offpeak data, shown in Figure 30 b). A very good agreement between data and MC is seen.

2. The distribution of the signal NN output is validated on the $B \rightarrow D\pi$ control sample, which has distributions of the NN input variables similar to the signal decays. An identical NN is applied to $B \rightarrow D\pi$ data and MC, and the NN outputs are shown in Figure 30 c). Very good agreement is also seen.

These validations have further confirmed that the distributions of the NN output are well

simulated on the Monte Carlo samples.

5.3.2 Fisher Discriminant

The linear Fisher discriminant [66, 67] is a classification method that projects high-dimensional data onto one dimension and performs classification in this one-dimensional space. It is an MVA technique widely used in high-energy physics to simplify and optimize signal-background separation. It can be viewed as a single-output NN without any hidden layer. A process similar to the NN training is also used to obtain the Fisher coefficients. Figure 31 shows the distribution of a Fisher discriminant used in $B^0 \rightarrow \rho^0\gamma$ decay mode. The input variables are the $\cos\theta_B$ and $\cos\theta_H$. The studies on Monte Carlo simulations show that these two variables are weakly-correlated in both signal events and background events.

Compared to neural networks, the linear Fisher discriminant is much simpler and could have similar performance, if input variables are weakly-correlated, as in the example shown
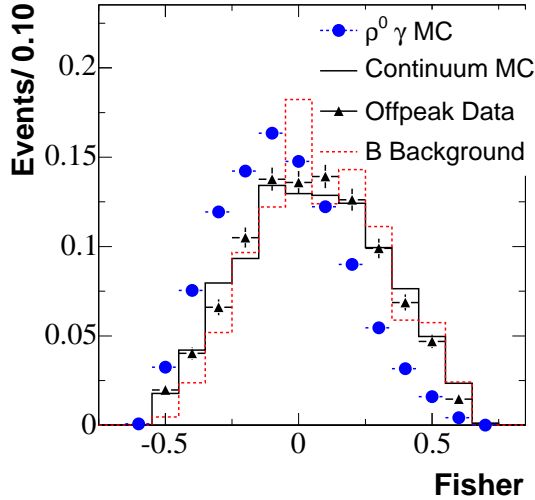
Figure 31: The Fisher distribution of the signal, the continuum background, the B-background and the off-resonance data for a Fisher used in the analysis of the $B^0 \rightarrow \rho^0\gamma$ decay mode.

above. However, this is not true for many of the shape variables and tagging variables.

5.4   Overview of the Likelihood-fit Methods

An unbinned, extended maximum likelihood-fit is used to extract the final signal yield. The likelihood function is a production over all N candidates with N_{hyp} hypothesis,

$$\mathcal{L} = \exp\left(-\sum_{i=1}^{N_{\mathrm{hyp}}} n_i\right) \cdot \left[\prod_{j=1}^{N}\left(\sum_{i=1}^{N_{\mathrm{hyp}}} n_i \mathcal{P}_i(\vec{x}_j; \ \vec{\alpha}_i)\right)\right] \ , \tag{66}$$

where n_i is the yield of each hypothesis. The $\mathcal{P}(\vec{x}_j; \vec{\alpha}_i)$ is the probability density function (PDF) of each hypothesis (component) over fit variables $\vec{x}_j$ given the set of parameters $\vec{\alpha}_i$. The fit variables normally include m_{ES}, ΔE^* for many *BABAR* analyses. One example of this kind of fits is shown in Section 4.1.1, Figure 19.

Extensive toy Monte Carlo samples are used to check the modeling and the bias of the likelihood fit. These toys are constructed as follows:

- **Pure toy Monte Carlo.** The data sample is generated using PDFs with given parameters. The pull of a free parameter p, defined as $(p^{true} - p^{fit})/\sigma_p$ with p^{true} the true value and p^{fit} the fitted value, can be used to test the bias of the likelihood fit method. The pull distribution, obtained from a large set of pure toy Monte Carlo events is a distribution with $\sigma = 1$ and mean zero for a unbiased likelihood fit.

- **Embedded toy Monte Carlo.** For this type of toy Monte Carlo, some components of the pure toy Monte Carlo are replaced with the the standard *BABAR* simulated Monte Carlo events. For example, we embed signal Monte Carlo into the PDF-generated peaking B-background, combinatorial B-background, and continuum background events to form a "signal-embedded" toy. The *BABAR*-simulated data sample resembles the real data more than a PDF-generated distribution, where only the statistical fluctuation

is considered. The correlations among the fit variables should be well preserved in the *BABAR* simulation, so by building this type of toy Monte Carlo, these correlations are introduced into the toy samples. The bias on the signal yield from ignoring these correlations in likelihood fit can be studied using this type of toy Monte Carlos.

- **Mixed toy Monte Carlo.** Additional *BABAR*-simulated Monte Carlo events are mixed with the pure toys or embedded toys described above. This type of Monte Carlo is used mainly to check the fit bias when some components are not considered in the likelihood fit.

These toys are also used to study the systematic errors related to the likelihood fits.

5.5 Design of the Analytical Methods

Background suppression using variables described above and likelihood fits to extract final signal yields is now described. An optimal analysis is developed for each decay mode separately.

5.5.1 $B \rightarrow (K^{*0}/K^{*+})\gamma$ Decays

The branching fraction of $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays is rare, on the order of 10^{-5}. However, given the high-luminosity *BABAR* data, for some of $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays the statistical errors are already starting to be comparable to the expected systematic errors. The primary experimental goal of analyzing $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays is to have the systematic error under control, while improving experimental sensitivities. This requirement determines the background treatment technique and the choice of the likelihood fits.

- **The continuum background** is suppressed by combining all the variables shown in Table 2 into a single output NN. Distinct NN is used for each decay mode to gain the

maximum performance for each decay mode. A cut is applied on the NN output to reject continuum background.

- **The combinatorial B-background** is included in the likelihood fit as a component, and separated from the signal relying on the ΔE^* distribution. In this approach, the systematic error resulting from B-backgrounds is proves to be lowest from MC study, in the range 2-3%.

  With the final choice of the likelihood fit, none of the possible "peaking" B-background can affect the signal yield more than 1%. Therefore, all the B-backgrounds are treated as combinatorial B-backgrounds.

- **The likelihood fit** is a 2-dimensional (2-D) fit over m_{ES} and ΔE^* with signal, continuum background, and combinatorial B-background. A detailed comparison among 1-D (m_{ES}), 2-D (m_{ES} and ΔE^*), and 3-D (m_{ES}, ΔE^*, and m_{K^*}) likelihood fit with signal and continuum background (including or not including the combinatorial B-background) shows that above approach is the optimal one. The experimental sensitivity is improved and the experimental systematic error, especially the one associated with the B-background, is well under control.

Table 3 summarizes NN configuration and its performance, and the final signal efficiency for each $B \to (K^{*0}/K^{*+})\gamma$ decay mode. After applying the NN to suppress the continuum background, in the signal region the signal-to-background ratio is improved by a factor of 5-10 for each decay mode.

5.5.2 $B \to (\rho/\omega)\gamma$ **Decays**

The analysis approach for the $B \to (\rho/\omega)\gamma$ decays is different from the $B \to (K^{*0}/K^{*+})\gamma$ decays because of the following facts:

Description	$K^{*0} \to K^+\pi^-$	$K^{*0} \to K^0_s\pi^0$	$K^{*+} \to K^+\pi^0$	$K^{*+} \to K^0_s\pi^+$
NN variables	$	\cos\theta_T	$, 18 energy cones, R'_2,
	$	\cos\theta_B	$, $	\cos\theta_H	$, "net flavor" and $\Delta z^{(*)}$ (only in $K^{*0} \to K^+\pi^-$)
NN hidden nodes	25	7	24	8
NN Cut	> 0.43	> 0.64	> 0.68	> 0.65
NN ϵ_{sig}	0.86	0.77	0.76	0.74
NN $\epsilon_{cont.bkg.}$	0.20	0.09	0.09	0.12
ϵ_{sig}	0.26	0.18	0.19	0.24

Table 3: The NN input variables, cut value on the NN output, NN efficiency on the signal, "NN ϵ_{sig}", and the background, "NN $\epsilon_{cont.bkg.}$", and the final signal efficiency in the fit region, "ϵ_{sig}", for each $B \to (K^{*0}/K^{*+})\gamma$ decay mode. All the efficiencies have been estimated using corresponding MC samples, which is independent from the NN training samples.

1. **No evidence for the $B \to (\rho/\omega)\gamma$ decays is found.** The theoretical predictions for branching fractions of $B \to (\rho/\omega)\gamma$ decays are about $0.5\text{-}1.0\times10^{-6}$, so this is still a statistically limited analysis.

2. **The continuum background is much more significant** than in $B \to (K^{*0}/K^{*+})\gamma$ decays, and could overwhelm the expected small-signal yields. It is essential to suppress this background, so that a significant signal could be observed.

3. **The significant peaking B-backgrounds** have comparable yields as signals in the fit region. These could significantly bias fitted signal yields and must be dealt with separately from other B-backgrounds.

Therefore, the primary goal of analyzing $B \to (\rho/\omega)\gamma$ is to improve the experimental significance over the existing data set. This can be achieved by either improving background suppressions, or using higher dimensional likelihood fit. We have improved suppression of the continuum background, using tagging variables to replace the "net flavor" used in analyzing $B \to (K^{*0}/K^{*+})\gamma$ decays (see Section 5.2.3) and π^0/η vetoes. Starting with a

two-dimensional likelihood fit (m_{ES} and ΔE^*), various multi-dimensional likelihood fits are tried to improve the significance also. We found that including the NN-output distribution into the likelihood fit after a loose cut on it improves the significance most, effectively increasing the existing data set by roughly 40-50%.

However, combining all the background suppression variables into a single NN, as in analyzing $B \to (K^{*0}/K^{*+})\gamma$ decays, will introduce various difficulties in using the NN distribution in the likelihood fit. Firstly, the combinatorial B-background is included in the likelihood fit as a component to minimize the systematic error. Only $\cos\theta_B$ and $\cos\theta_H$ distributions of this background show differences from signal decays. Therefore, the NN output of this background is neither signal-like nor background-like. The modeling of this distribution must be obtained from MC and could result in large systematic errors. Secondly, the peaking B-backgrounds are also included in the likelihood fit to obtain an unbiased fit and to minimize the systematic errors associated with it. However, the $\cos\theta_H$ distribution of $B \to (\rho/\omega)(\pi^0/\eta)$ backgrounds shown in Figure 25 is very different from the signal events. Therefore, more NN distributions have to be obtained from MC and validated over control samples. All these would complicate the analysis significantly and might overwhelm the improvement by including this variable.

In order to overcome these difficulties, the shape variables and tagging variables are combined using a single output NN. Since all these variable are related to ROE, a similar distribution is expected for signal events and all B-backgrounds, as shown in Figure 32. A loose cut is applied on the NN distribution and the remaining NN distribution is included in the likelihood fit with the assumption that NN distributions for signal and B-backgrounds are the same. In this way, only two NN distributions are involved in the likelihood fit, significantly simplifying the analysis. Some deviations of B-background NN output from the signal are also seen in Figure 32. However, this results in a very small systematic error. To
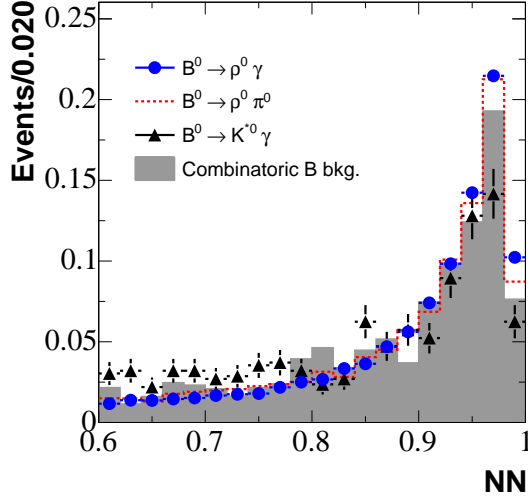
Figure 32: The NN distribution of the signal, the $B \to (K^{*0}/K^{*+})\gamma$ background, the $B^0 \to \rho^0(\pi^0/\eta)$ background, and the combinatorial B-background for the $B^0 \to \rho^0\gamma$ decays.

improve the background suppression, especially for B-backgrounds, $\cos\theta_B$, $\cos\theta_H$, and $\cos\theta_D$ (in $B^0 \to \omega\gamma$ mode) are then combined in a Fisher discriminant (shown in Figure 31) and also included in the likelihood fit. A 4-D (m_{ES}, ΔE^*, NN, and Fisher) likelihood fit is used to separate the signal from backgrounds, in which the combinatorial B-background and peaking B-backgrounds (the $B \to (K^{*0}/K^{*+})\gamma$ background and the $(\rho/\omega)(\pi^0/\eta)$ background) are included as independent components. The only exception is the $B^0 \to \omega\gamma$ decay, which only contains three components: the signal, the continuum background, and the $\omega(\pi^0/\eta)$ background.

Table 4 summarizes the NN configuration and its performance, and the final signal efficiency for each $B \to (\rho/\omega)\gamma$ decay. Application of the loose NN cuts improves the signal-to-background ratio by a factor of about eight for all three decay modes. Inclusion of the mass of vector mesons in the likelihood fit is studied in MC simulation. The modeling of PDFs for each component increased the systematic errors, but the improvement from including this

Description	$B^0 \to \rho^0\gamma$	$B^+ \to \rho^+\gamma$	$B^0 \to \omega\gamma$
NN variables		$\|\cos\theta_T\|,\ R'_2,\ L_1,\ L_2,\ L_3,\ L_{2g}$, tagging variables and $\Delta z^{(*)}$ (except in $B^+ \to \rho^+\gamma$)	
NN hidden nodes	8	7	6
NN Cut	> 0.6	> 0.6	> 0.65
NN ϵ_{sig}	0.81	0.76	0.76
NN $\epsilon_{cont.bkg.}$	0.10	0.11	0.09
Fisher		$\|\cos\theta_B\|,\ \|\cos\theta_H\|$ and $\|\cos\theta_D\|$ (in $B^0 \to \omega\gamma$)	
ϵ_{sig}	0.16	0.11	0.09

Table 4: The NN input variables, cut value on the NN output, NN efficiency on the signal, "NN ϵ_{sig}", and the background, "NN $\epsilon_{cont.bkg.}$",the Fisher variables, and the final signal efficiency in the fit region, "ϵ_{sig}", for each $B \to (\rho/\omega)\gamma$ decay mode. All the efficiencies have been estimated using corresponding MC samples, which are independent from the NN training samples.

variable is not significant. Further, a wider mass window introduces more crossfeed among $B \to (\rho/\omega)\gamma$ decays and $B \to (K^{*0}/K^{*+})\gamma$ backgrounds, which potentially introduces additional systematic uncertainties. Consequently, we choose a tighter cut on the mass window, and ignored this distribution in our likelihood fit. With the chosen mass selection, the crossfeed among $B \to (\rho/\omega)\gamma$ decays is a few precent in the fit region, and no significant peaks in the fit region are seen. Therefore, we can safely ignore the crossfeed among the $B \to (\rho/\omega)\gamma$ decays.

5.6 Likelihood Fit

5.6.1 Correlations

For each component in the likelihood fit, the PDF, $\mathcal{P}(\vec{x_j}; \vec{\alpha_i})$, is composed of a product of the individual PDFs over the variables. In this way, we ignore the correlations among the fit variables. The validity of doing so has been established by the fact that the correlations among the fit variables are weak for all components. A detailed study of the correlations has been performed over Monte Carlo simulations and the offpeak data for each decay mode.

Typically, the correlation among the fit variables is a few percent. The strongest correlation[32] is shown between the m_{ES} and ΔE^* for the combinatorial B-background at about 10%. The effect of the correlations on the signal yield has been studied by building various "signal embedded" toy Monte Carlos, shown in the Table 5.

5.6.2 Composition of the Component PDFs

Table 11 in Appendix B summarizes all the PDFs used to compose the PDF for each component, $\mathcal{P}(\vec{x_j}; \vec{\alpha_i})$, in the $B \to (K^{*0}/K^{*+})\gamma$ decays and the $B \to (\rho/\omega)\gamma$ decays, respectively. The selections of the PDFs are based on the Monte Carlo studies.

$\mathbf{B \to (K^{*0}/K^{*+})\gamma}$ decays

In the $B \to (K^{*0}/K^{*+})\gamma$ decays, to reduce systematic uncertainty, the fit parameters of the continuum background and the signal (with the exception noted later) are determined by a fit to data. For the continuum background, the ΔE^* distribution is modeled by a first-order polynomial function with the exception of $K^{*+} \to K_s^0 \pi^+$, where a second-order polynomial is used. The m_{ES} for the continuum background is modeled as an ARGUS function [53]. These are shown in Section 4.1.1 Figure 19. In the $K^{*+} \to K^+ \pi^0$ decay mode, the continuum background shape is simultaneously fit to the off-resonance data, assuming the same continuum shapes. For the B-background, the Gaussian distribution used for ΔE^* and the Novosibirsk function[33] used for m_{ES} have all shape parameters fixed to values determined from the Monte Carlo. As shown in Figure 19, the signal ΔE^* distribution is modeled as a Crystal Ball function [52], with the low-side power-law tail fixed from Monte Carlo data.

---

[32]As mentioned in Section 4.1.1, this correlation is reduced by using the m_{ES} instead of the rescaled m_{ES}.

[33]The Novosibirsk function [68] is defined as

$$f(x) = C \cdot e^{-\frac{1}{2} \cdot [\ln(1 + \tau \cdot (x - <x>) \cdot \frac{\sinh(\tau\sqrt{\ln 4})}{\sigma\tau\sqrt{\ln 4}})/\tau]^2 + \tau^2}, \tag{67}$$

with $\langle x \rangle$ the mean, σ the width, and τ the "tail" parameter.

The energy leak from the primary photon is well compensated for by the rescaled m_{ES}, so the m_{ES} distribution for signal is modeled as a Gaussian function, except for the $K^{*+} \to K^+\pi^0$ decay mode, where a Crystal Ball function, with tail parameters fixed using Monte Carlo, is used to accommodate a low-side tail due to the π^0 energy leakage. The same low-side tail in $K^{*0} \to K_s^0\pi^0$ decay mode is not separately fit due to the low statistics in this mode.

B $\to (\rho/\omega)\gamma$ decays

In the $B \to (\rho/\omega)\gamma$ decays, the m_{ES} and ΔE^* PDFs are parameterized by a Crystal Ball function for both the signal and the peaking background. The parameterization is determined from signal MC samples, except for the mean of the ΔE^* distribution, which is offset by the observed difference between data and MC samples of $B \to (K^{*0}/K^{*+})\gamma$ decays. Similar to the $B \to (K^{*0}/K^{*+})\gamma$ decays, the continuum background m_{ES} and ΔE^* distributions are parameterized by an ARGUS function and a second-order polynomial, respectively. The combinatorial B-background is described by a smooth distribution [69] (KEYS), determined from MC events in both m_{ES} and ΔE^*. The distribution of NN for signal and all the B-backgrounds is parameterized by a Crystal Ball function with the exception of $B^0 \to \omega\gamma$ decay mode, where a KEYS PDF, obtained from signal Monte Carlo, is used. The NN distribution for continuum is determined from sideband data[34], and a histogram

---

[34]The sidebands used to extract the NN and Fisher distributions of the continuum background have been defined as follows:

1. **Mass veto sideband in $B \to (\rho^0/\rho^+)\gamma$ modes.** These are formed by inverting the meson mass cut from [0.63, 0.94] GeV/c^2 to [0.5, 0.63] GeV/c^2 or [0.94, 1.10] GeV/c^2. Then, events are selected within the region [5.2, 5.26] GeV/$c^2 \times$ [-0.35, 0.6] GeV on the m_{ES}–ΔE^* plane.

2. **Mass veto sideband in $B^0 \to \omega\gamma$ mode.** These are formed by inverting the meson mass cut from [0.764, 0.795] GeV/c^2 to [0.7, 0.764] GeV/c^2 or [0.795, 0.85] GeV/c^2. Then, events are selected within the region [5.2, 5.26] GeV/$c^2 \times$[-0.35, 0.6] GeV on the m_{ES}–ΔE^* plane.

3. **Upper sideband.** These are composed by choosing events in the region [5.2, 5.3] GeV/$c^2 \times$[0.3, 0.6] GeV on the m_{ES}–ΔE^* plane without any change to the other cuts.

These sidebands were selected to be an orthonormal data set to the events in the fit region. The NN and

is used as the PDF. The distribution of the Fisher variable is parameterized by a KEYS PDF constituted from the sideband data for the continuum background, and MC events for all other hypotheses. The fit to the data determines the shape parameters of the continuum background m_{ES} and ΔE^* PDFs, as well as the signal, continuum background, and combinatorial B-background yields. All other parameters are fixed, including the peaking B-background yields.

5.6.3 Validation of Likelihood Fits

The likelihood fits are then validated using various toy Monte Carlos. In principle, a fully embedded toy Monte Carlo study, where all the components are selected from the *BABAR*-simulated events, resembles the real data most. However, because of lack of statistics in the continuum MC, we can, at most, build the "All embedded (except continuum background)" toy MC samples. This embedded MC is used to validate the modeling of B-backgrounds, and check correlations. In building toy samples, the expected yield of each component is normalized to correspond to a luminosity of 81.9 fb^{-1} and 191 fb^{-1} for $B \to (K^{*0}/K^{*+})\gamma$ decay modes and $B \to (\rho/\omega)\gamma$ decay modes, respectively, using the assumed branching fractions or cross-sections in Section 3.3, Section 4.1, and Section 5.1.2. A example of this kind of study is shown in Figure 33, where a pure toy Monte Carlo (with around 1000 independent toys) is performed, and the fitted signal yield, the error and the pull distributions are shown. Table 5 shows the mean/sigma of signal pull distributions for various toy Monte Carlo studies.

$B \to (K^{*0}/K^{*+})\gamma$ decays

Fisher are obtained by combining all the sideband data to increase the statistical samples. These distributions are then compared to the events in the fit region (with $m_{ES} < 5.26$ GeV/c^2 to exclude most of signal and B-background). The sideband data are validated to have the same NN and Fisher distribution as the the events in the fit region within the statistics.
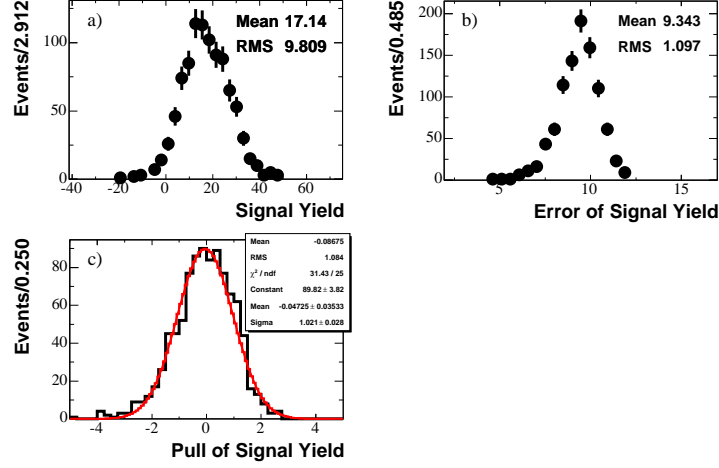
Figure 33: The pull, fitted signal yield, and the error of the signal yield from a toy Monte Carlo study in $B^0 \to \rho^0\gamma$ decay mode.

Modes	Pure	Sig. embedded	All embedded (except cont. bkg.)	Mixed
$K^{*0} \to K^+\pi^-$	-0.02/1.01	-0.54/1.06	0.35/1.02	–
$K^{*0} \to K^0_s\pi^0$	0.09/1.04	-0.67/1.12	-0.05/1.03	–
$K^{*+} \to K^+\pi^0$	-0.12/1.04	-0.37/1.05	-0.04/1.03	–
$K^{*+} \to K^0_s\pi^+$	0.02/0.94	-0.57/1.02	-0.08/0.97	–
–	–	–	–	–
$B^0 \to \rho^0\gamma$	-0.05/1.02	-0.08/1.08	0.00/0.96	–
$B^+ \to \rho^+\gamma$	-0.02/1.00	0.11/1.06	0.00/0.99	–
$B^0 \to \omega\gamma$	-0.10/0.99	-0.11/1.14	0.04/1.00	$-1.1^{+1.1}_{-1.9}$

Table 5: The pull mean/sigma of the signal yield for each decay mode, with the exception that for the mixed toy MC study in the $B^0 \to \omega\gamma$ decay, only the bias on the signal event is given, which results from ignoring the B-background in the likelihood fit.

For the $B \to (K^{*0}/K^{*+})\gamma$ decays, ignoring the correlations in signal events does introduce a slight bias on signal yields, shown in Table 5. We decide to correct this bias using the "signal embedded" toy studies. A correction factor for the signal efficiency, defined as fitted signal yield/expected signal yield, is then derived from the signal embedded toy studies, which is 0.978, 0.907, 0.963, and 0.955 for $K^{*0} \to K^+\pi^-$, $K^{*0} \to K_s^0\pi^0$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decays, respectively. The error of the fitted signal yield is considered the systematic uncertainty of this correction factor, MC statistics/fit bias, which is mainly determined by statistics of signal MCs, and is 0.9%, 3.2%, 2.4%, and 1.6% for $K^{*0} \to K^+\pi^-$, $K^{*0} \to K_s^0\pi^0$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decays, respectively.

B $\to (\rho/\omega)\gamma$ decays

For $B \to (\rho/\omega)\gamma$ decay modes, no significant bias on the signal yields is found (see Table 5). The "All embedded (except continuum background)" toy MC study bears the closest resemblance to real data. This sample can test not only the correlations among fit variables for signal and B-backgrounds, but also the hypothesis that the NN distributions of these components are the same. Additionally, the available MC samples are large enough to perform a high statistics test with this type of toy study[35]. Based on the results of the "All embedded (except continuum background)" toy MC study, no correction factor on the signal efficiency is applied. But to be conservative, the error of the fitted signal yield is taken as the systematic error of "MC statistics/fit bias/B-background modeling", which is 10.2%, 10.8%, and 5.4% on the signal efficiency for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively. In $B^0 \to \omega\gamma$ decay mode, the effect of ignoring the $B \to (K^{*0}/K^{*+})\gamma$ background and the combinatorial B-background is studied using "mixed" toy studies, and

---

[35]In BABAR, the standard MC simulations are produced in proportional to the on-resonance data with a fixed rate. When analyzing $B \to (\rho/\omega)\gamma$ decays, the available B-background MCs are already few times bigger than the available B-background MCs at the time of analyzing $B \to (K^{*0}/K^{*+})\gamma$ decays.
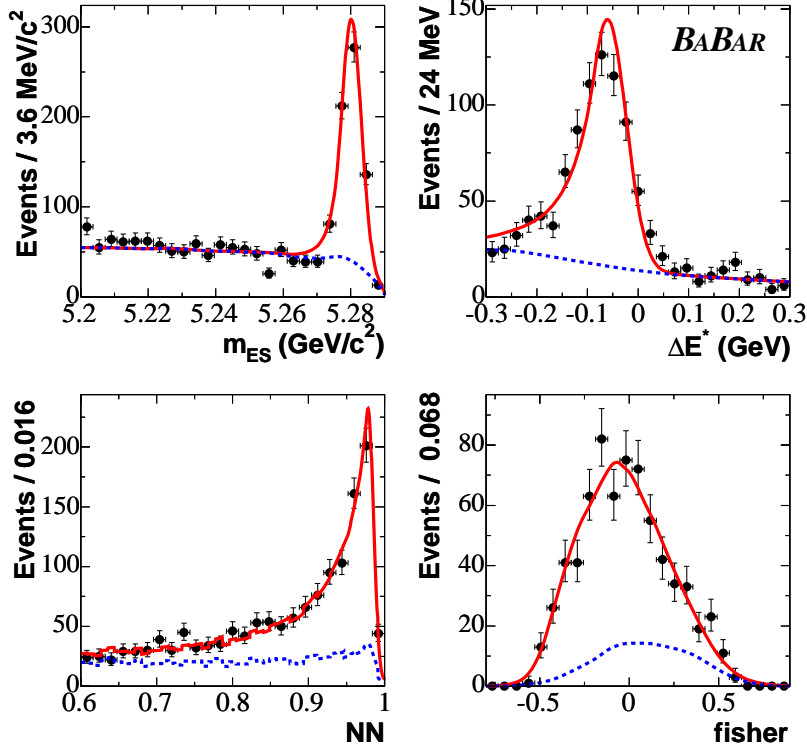
Figure 34: The likelihood fit on the $B \to (K^{*0}/K^{*+})\gamma$ control sample in $B \to (\rho^0/\rho^+)\gamma$ decay modes. The projection plots of the likelihood fit to the $B \to (K^{*0}/K^{*+})\gamma$ sample, which is obtained by releasing the PID selector in the $B \to (\rho^0/\rho^+)\gamma$ decay mode. The red solid line is the total fit and the dashed blue line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286$ GeV/c^2, $-0.20 < \Delta E^* < 0.0$ GeV, and $\mathcal{N} > 0.9$.

the bias on the signal yield is 1.1 events with a systematic error of $^{+1.1}_{-1.9}$.

Offpeak data, sideband, and control samples

To validate the likelihood fit further, it is then applied to off-resonance data, sideband data, and control samples.

1. **The off-resonance data** contains no signal events. No significant non-zero signal yield is expected when performing likelihood fits on it. This has been confirmed by

the fit in each decay mode.

2. **The sideband data** in the fit region is constructed by selecting events outside the meson mass window, described in 4.2.7. This control sample has statistics comparable to the final data set. Furthermore, it contains some amount of B-backgrounds but almost no signal events. This could provide a better test of the likelihood fit than the off-resonance data set. No significant signal is present in the fit for each decay mode.

3. **In $B \to (\rho^0/\rho^+)\gamma$ decays, the $B \to (K^{*0}/K^{*+})\gamma$ backgrounds** are strongly suppressed by pion PID (Section 4.2.3). By removing the PID, a control sample for $B \to (\rho^0/\rho^+)\gamma$ decays can be built. This sample contains very large $B \to (K^{*0}/K^{*+})\gamma$ events. Therefore applying the identical fits as $B \to (\rho^0/\rho^+)\gamma$ decays (with the mean of the signal ΔE^* floating) could give a very good test of the likelihood fit in $B \to (\rho^0/\rho^+)\gamma$ decay modes. Figure 34 shows the projection plots of the $B^0 \to \rho^0\gamma$ fit on the corresponding $B \to (K^{*0}/K^{*+})\gamma$ control sample. The fitted signal yield is in good agreement with the expected $B \to (K^{*0}/K^{*+})\gamma$ yield.

After all these validations, the likelihood fit is applied to the final data set in the fit region to extract the signal yield for each decay mode. The results are discussed in Section 6.

6 Fitted Results

6.1 Fitted Signal Yields of the $B \to (K^{*0}/K^{*+})\gamma$ Decays

Figure 35 and Figure 36 show projection plots of the likelihood fits of the 81.9 fb^{-1} BABAR onpeak data for the $B^0 \to K^{*0}\gamma$ decay modes and the $B^+ \to K^{*+}\gamma$ decay modes, respectively. Significant signal peaks are seen.

There are 3304, 511, 1393, and 1174 events in the fit region for $K^{*0} \to K^+\pi^-$, $K^{*0} \to K_s^0\pi^0$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decay modes respectively. These events are included in the likelihood fits, and the corresponding signal yields are 582.6±29.7, 61.8±15.3,
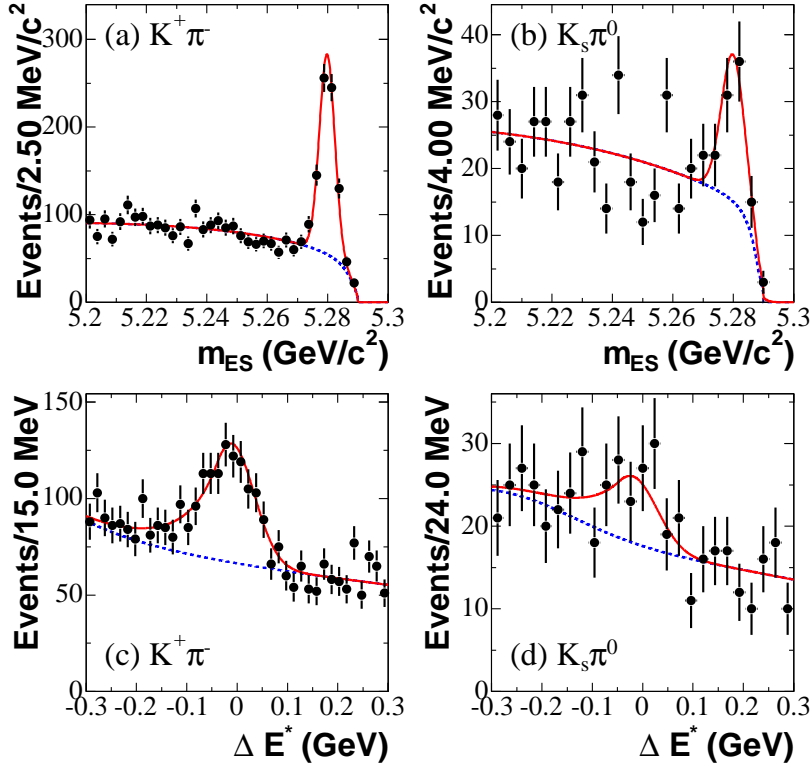
Figure 35: The projection plots of the final fits on m$_{ES}$ and ΔE^* for the $B^0 \to K^{*0}\gamma$ decay modes. The solid line is the total fit and the dashed line is the total background. No cuts on other variables were applied when making the projection plots.
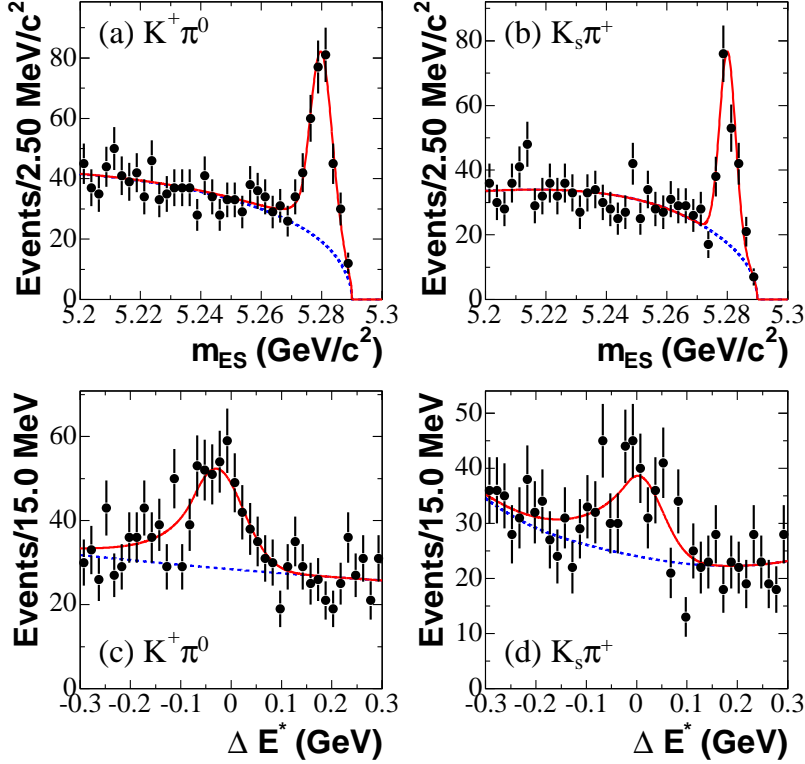
Figure 36: The projection plots of the final fits on m_{ES} and ΔE^* for the $B^+ \to K^{*+}\gamma$ decay modes. The solid line is the total fit and the dashed line is the total background. No cuts on other variables were applied when making the projection plots.

250.9±22.6, and 156.9±15.7, respectively. For $K^{*0} \to K^+\pi^-$ decay mode, the statistical precision of the signal yield is around 5%.

6.2 Fitted Signal Yields of the $B \to (\rho^0/\rho^+)\gamma$ and $B^0 \to \omega\gamma$ Decays

Figures 37, 38, and 39 show projection plots of the likelihood fits of 191.4 fb^{-1} *BABAR* data, almost all the available data by 2004, for the $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively. No significant signal peaks are seen.

There are 4367, 7071, and 1390 events in the fit region for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay mode respectively. These events are included in the likelihood fits, and
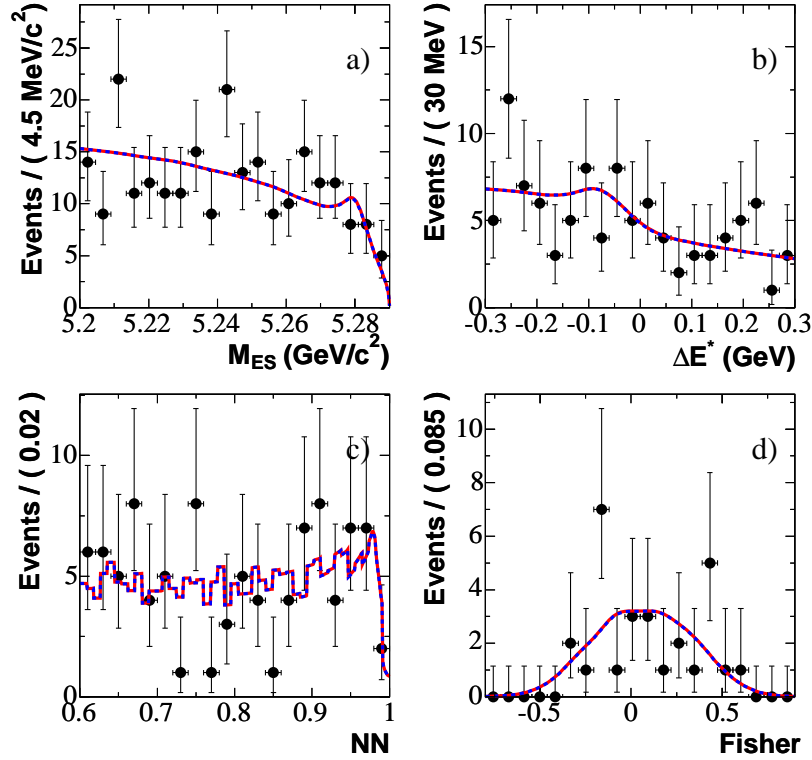
Figure 37: The projection plots of the unblinded results for $B^0 \to \rho^0\gamma$ decay mode, where a), b), c), and d) are the projection onto m$_{ES}$, ΔE^*, neural net, and Fisher, respectively. The solid line is the total fit, and the dashed line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 <$ m$_{ES} < 5.286$ GeV/c^2, $-0.10 < \Delta E^* < 0.05$ GeV and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 50%, 59%, 78%, and 49% for m$_{ES}$, ΔE^*, neural net, and Fisher projections.

the corresponding signal yields are $0.3^{+7.2}_{-5.4}$, $25.6^{+15.3}_{-13.9}$, and $8.3^{+5.7}_{-4.5}$, respectively. Ignoring the systematic uncertainties, the significance of the $B^+ \to \rho^+\gamma$ and $B \to (\rho/\omega)\gamma$ signal yields are around 2σ, however the significance of the $B^0 \to \rho^0\gamma$ signal yield is negligible. In the $B^0 \to \rho^0\gamma$ decay mode, the small bumps seen in the m$_{ES}$ and ΔE^* distribution are due to small peaking B-backgrounds. In the $B^0 \to \omega\gamma$ decay mode, the bias due to ignoring the $B \to (K^{*0}/K^{*+})\gamma$ background and the combinatorial B- background have been corrected directly in the likelihood.
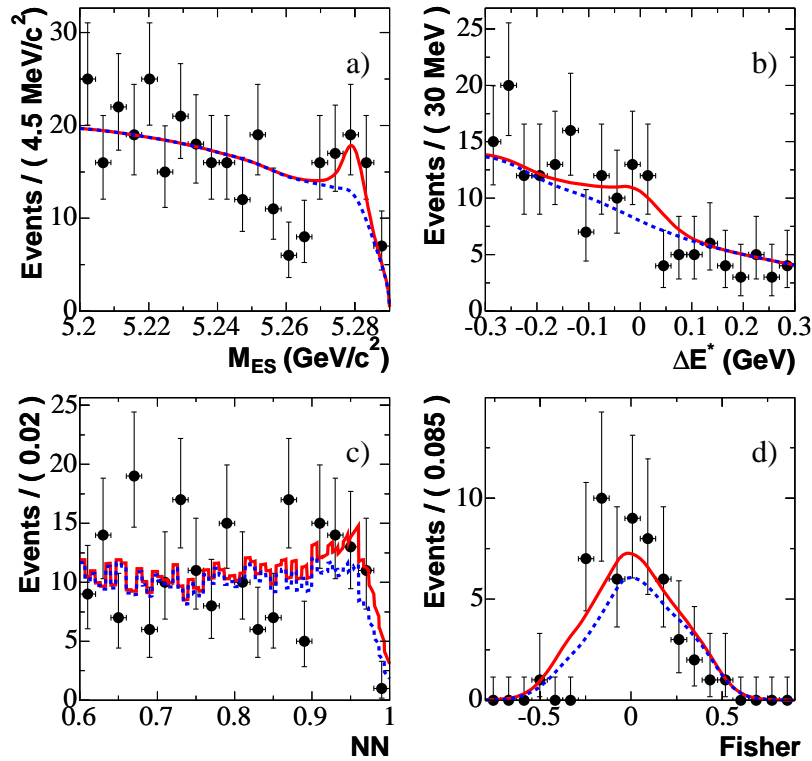
Figure 38: The projection plots of the unblinded results for $B^+ \to \rho^+\gamma$ decay mode, where a), b), c), and d) are the projection onto m_{ES}, ΔE^*, neural net, and Fisher, respectively. The solid line is the total fit, and the dashed line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286$ GeV/c^2, $-0.10 < \Delta E^* < 0.05$ GeV and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 42%, 56%, 66%, and 41% for m_{ES}, ΔE^*, neural net, and Fisher projections.
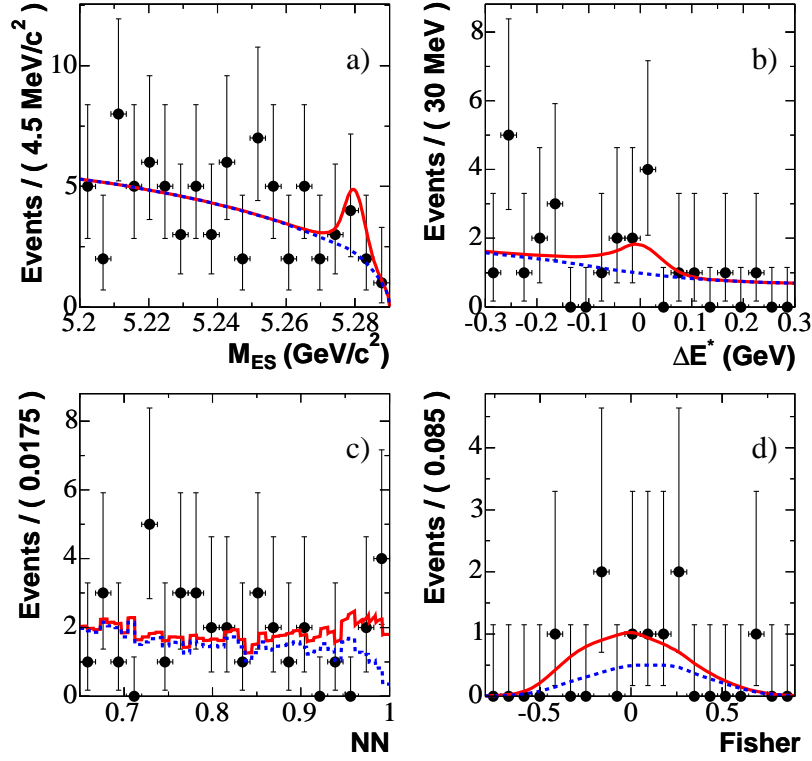
Figure 39: The projection plots of the unblinded results for $B^0 \to \omega\gamma$ decay mode, where a), b), c), and d) are the projection onto m_{ES}, ΔE^*, neural net, and Fisher, respectively. The solid line is the total fit, and the dashed line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286$ GeV/c^2, $-0.10 < \Delta E^* < 0.05$ GeV and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 46%, 58%, 70%, and 45% for m_{ES}, ΔE^*, neural net, and Fisher projections.
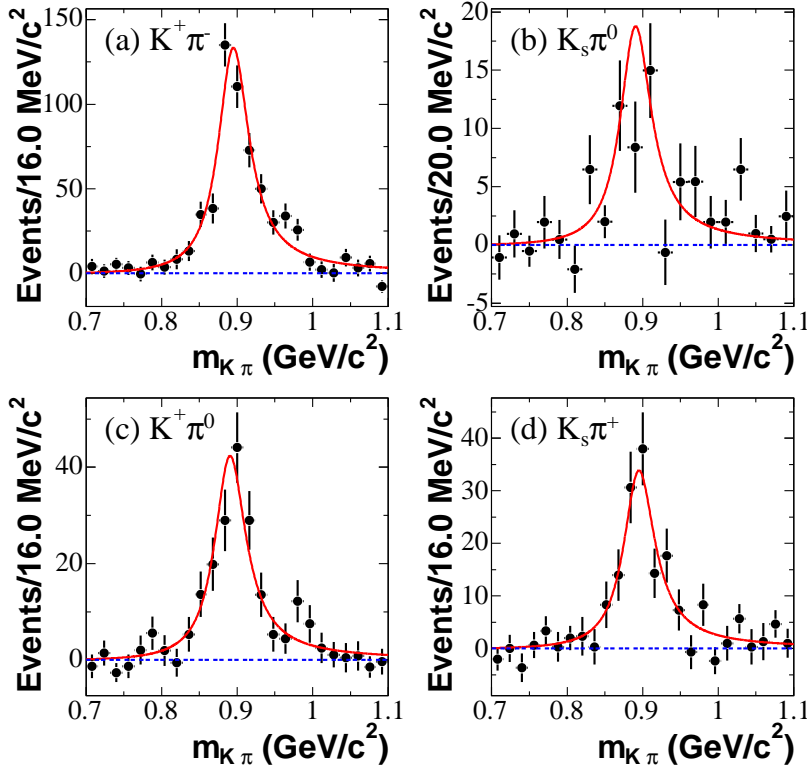
Figure 40: The background-subtracted meson mass distribution of the $B \to (K^{*0}/K^{*+})\gamma$ decays. The events in the signal region are selected to make these plots with the meson mass window extended to $[0.7, 1.1]$ GeV/c^2. These distributions are then fitted to a relativistic Breit-Wigner function, shown by the solid line. The parameters of the Breit-Wigner function is fixed by the nominal values from PDG Book [38].

6.3 Cross Checks

After obtaining the final results, the goodness of the likelihood fit is validated using a toy Monte Carlo. About 1000 pure toy Monte Carlo samples are generated using the fitted yields and the PDF parameters. The likelihood fit is then applied to these toy data samples, and a likelihood distribution is obtained. The likelihood of the final data fits and the mean of the toy MC likelihood distribution are consistent with each other within 2σ, for all the decay modes. Other projection plots, which are similar to Figures 35-39 with different selections

on fit variables are also checked. No significant deviations from the fitted curve are seen.

These checks proved that the likelihood functions are able to describe the data well.

For the $B \to (K^{*0}/K^{*+})\gamma$ decays, a further cross-check is performed to check the non-resonant contribution to these decays. The events in the signal region (see Figure 19) with a meson mass window $[0.7, 1.1]$ GeV/c^2 are projected onto the meson mass distribution $(m_{K\pi})$, and the background is subtracted using the events in $[5.2, 5.26]$ GeV/$c^2 \times [-0.2, 0.1]$ GeV, where only continuum background is present. These distributions plotted in Figure 40, show a nice $K\pi$ resonance structure. These distributions are then fitted to a relativistic p-wave Breit-Wigner function[36] with the resonance shape fixed to the nominal parameters from the Particle Data Group [38]. These fits show that the data is well modeled by the nominal resonance shape. These cross-checks have fully confirmed that the fitted $K^*\gamma$ signal yields are truly due to $B \to K^*\gamma$ events.

---

[36]The relativistic p-wave Breit-Wigner function is defined as,

$$f(m) = \frac{m \cdot \Gamma_{rel}}{\left(m^2 - m_0^2\right)^2 + \left(m_0 \cdot \Gamma_{rel}\right)^2} \ , \tag{68}$$

with

$$\Gamma_{rel} = \Gamma \left(\frac{p}{p_0}\right)^{2l+1} \left(\frac{m_0}{m}\right) \cdot \left(\frac{1 + (p_0 r)^2}{1 + (pr)^2}\right), \tag{69}$$

here m is reconstructed mass, m_0 is the peak of the resonance, p is the outgoing momentum of the daughter in the resonance rest frame, p_0 is the outgoing momentum of the daughter at the resonance peak, $l = 1$ is the angular momentum, Γ is the decay width of the resonance and $r = 3$ GeV^{-1} is the barrier radius.

7 Systematic Errors

The systematic errors in calculating the branching fractions fall in two main categories as follows:

- **Systematic errors related to the signal reconstruction efficiency.** The signal reconstruction efficiency for each decay mode was obtained from corresponding signal MC simulation to calculate the branching fraction. This efficiency is later checked against the real data. For this type of systematic uncertainty, a correction factor and a relative error on the signal efficiency is given.

- **Systematic errors related to the likelihood fit.** These are obtained from the variation of the likelihood fit, e.g. by fixing some parameters of the PDFs using Monte Carlo data or ignoring some components in modeling the data. For each source, a "fit bias" on the fitted signal yield, which is associated with an error, is then determined. The fitted signal yield is then corrected by this bias, and the corresponding error (in events) on the bias is taken as the systematic error.

7.1 Systematic Errors on Signal Reconstruction Efficiency

The number of $B\overline{B}$ pairs produced in *BABAR* is studied by the "B counting" analysis working group (AWG), using ratios of hadron events to $\mu^+\mu^-$ events between on- and off-resonance data [70]. A 1.1% uncertainty on the number of $B\overline{B}$ pairs is given. This uncertainty is common for all decay modes.

The production ratio of charged and neutral B events, $R^{+/0}$, is assumed to be 1.0 for most *BABAR* analyses. This ratio has been measured recently to be,

$$R^{+/0} \equiv \Gamma(e^+e^- \to B^+B^-)/\Gamma(e^+e^- \to B^0\bar{B}^0) = 1.006 \pm 0.048 \tag{70}$$

at $\sqrt{s} = M_{\Upsilon(4S)}$ [71], and results in an additional 2.4% systematic error on the measurements of B branching fractions. However, this error is mainly statistics dominated, and with a growing *BABAR* data set, a better determination of the production ratio can be achieved. Since the systematic error is beginning to dominate the total error for for $B \to (K^{*0}/K^{*+})\gamma$ decays, and this production ratio error is relatively significant, a 2.4% systematic error is added for each decay mode conservatively. While for $B \to (\rho/\omega)\gamma$ decays, relative to the statistical error, this error is small and is ignored.

The reconstruction efficiencies of charged tracks [72, 73], neutral candidates (including π^0 and single photons) [74] and K_s^0s [75] in MC simulations are of interest for most *BABAR* analyses and have been studied using various control samples, by different *BABAR* AWGs. Algorithms are then provided to correct the corresponding efficiencies on MCs and assign systematic errors. As shown in Table 6, the systematic errors are 0.8% and 3.0% for every track and K_s^0, respectively. The systematic error in efficiency of reconstructing a π^0 candidate is 5% and 3.0% for $B \to (K^{*0}/K^{*+})\gamma$ decays and $B \to (\rho/\omega)\gamma$ decays, respectively. The improvement in the $B \to (\rho/\omega)\gamma$ decays comes from a recent study of the π^0 systematic error [76]. The systematic error of the signal photon is derived from π^0s, whose daughter photons have almost the same energy. This systematic error is 2.5% for $B \to (K^{*0}/K^{*+})\gamma$ decays, and should be less than this for $B \to (\rho/\omega)\gamma$ decays. However this study is not fully finished, so we conservatively assign a 3.0% for each $B \to (\rho/\omega)\gamma$ decay mode.

As discussed in Section 4.2.3, PID systematic errors in MC simulations are studied by comparing efficiencies between data and MC, using the $D^* \to D\pi$ control samples. The detailed algorithms used to obtain the PID systematic error are described elsewhere [55]. The systematic error is 1.0% for the kaon selection in $B \to (K^{*0}/K^{*+})\gamma$ decays and 2% for pion selection in $B \to (\rho/\omega)\gamma$ decays, respectively. For $B \to (\rho/\omega)\gamma$ decays, the kaon mis-id rate of the pion selectors is very important since the yield of $B \to (K^{*0}/K^{*+})\gamma$ background

is estimated from MC and fixed in the likelihood fit. This mis-id rate is also checked using the same D^* control sample. A correction factor of 1.32 and 1.31 is determined for the kaon mis-id rate for $B^0 \rightarrow \rho^0\gamma$ and $B^+ \rightarrow \rho^+\gamma$ decay modes, respectively. The uncertainties on the corresponding correction factors are 0.32 and 0.31. This will be considered as part of the systematic errors of the likelihood fit.

The systematic errors of π^0/η-veto and distance cuts are studied using a "photon embedded" method [77]. A high-energy photon from MC simulation or radiative Bhabha event is embedded into generic MC samples, or off- or on-resonance data. By treating this photon as the primary photon, the efficiencies of the veto/distance cuts are compared between data and Monte Carlo simulations. These studies result in an efficiency difference between data and MC simulations of 2% and 1% for the veto and the distance cut, respectively. These are assigned as the systematic errors of the π^0/η-veto and distance cuts, respectively.

The meson mass distributions for data and Monte Carlo using large samples of ρ, K^*, and ω mesons are compared. When cutting on the mass distributions, the efficiency difference between data and Monte Carlo is found to be far less than 1%, relative to the signal efficiency. Therefore, this systematic error is ignored.

The NN distribution could result in two types of systematic errors, a systematic error on the signal efficiency when a cut is applied on the NN distribution, or a systematic error on the fitted signal yield when the NN distribution is used in a likelihood fit. These systematic errors are studied using $B \rightarrow D\pi$ control samples, shown in Appendix A. The systematic error related to signal efficiency is summarized in Table 6.

As discussed in Section 5.6.3, for $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays, ignoring the correlations between fit variables m_{ES} and ΔE^* does introduce some inefficiencies in the likelihood fit. This bias, corrected using a "signal embedded" toy Monte Carlo study, is 0.978, 0.907, 0.963, and 0.955 for $K^{*0} \rightarrow K^+\pi^-$, $K^{*0} \rightarrow K^0_s\pi^0$, $K^{*+} \rightarrow K^+\pi^0$, and $K^{*+} \rightarrow K^0_s\pi^+$ decay,

respectively. The errors on these factors are dominated by the statistics of available signal MCs, and are treated as "MC statistics/fit bias" systematic error. For $B \to (\rho/\omega)\gamma$ decays, we take no correction on the signal efficiency due to "MC statistics/fit bias", and take the statistics errors of the fitted signal yield as the systematic errors of "MC statistics/fit bias/B-background modeling", which are 10.2%, 10.8%, and 5.4% on the signal efficiency for the $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay mode, respectively. This includes the systematic error in the modeling of NN distributions for signal and B-backgrounds.

All systematic errors discussed above are summarized in Table 6. For $B \to (K^{*0}/K^{*+})\gamma$ decays the total systematic error on the signal efficiency is 5.4%, 10.7%, 9.0%, and 6.1% for $K^{*0} \to K^+\pi^-$, $K^{*0} \to K^0_s\pi^0$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K^0_s\pi^+$ decay mode, respectively. This error is already larger than the statistical error for $K^{*0} \to K^+\pi^-$ decay mode. The total error for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ is 12.1%, 10.0%, and 10.8%, respectively. This is dominated by the "MC statistics/fit bias/B-background modeling".

7.2 Systematic Errors of Likelihood Fits

The systematic errors related to likelihood fits are quite for $B \to (K^{*0}/K^{*+})\gamma$ decays and $B \to (\rho/\omega)\gamma$ decays as discussed below separately,

7.2.1 Fit Systematics in $B \to (K^{*0}/K^{*+})\gamma$ Decays

In $B \to (K^{*0}/K^{*+})\gamma$ decays, significant signals are present for all the decay modes. Therefore, almost all the parameters of the signal PDFs are determined directly from the data. Further, the parameters of the continuum background are also floating in the fits. For these parameters, no systematic error is given. We only consider the following two types of systematic errors. No bias on the signal yield is assigned.

1. Some signal parameters are still fixed from MC, e.g. the tail parameters of the Crystal-

Description	$K^{*0} \to K^+\pi^-$ Factor	$\sigma(\%)$	$K^{*0} \to K^0_s\pi^0$ Factor	$\sigma(\%)$	$K^{*+} \to K^+\pi^0$ Factor	$\sigma(\%)$	$K^{*+} \to K^0_s\pi^+$ Factor	$\sigma(\%)$	$B^0 \to \rho^0\gamma$ Factor	$\sigma(\%)$	$B^0 \to \omega\gamma$ Factor	$\sigma(\%)$	$B^+ \to \rho^+\gamma$ Factor	$\sigma(\%)$
ϵ_{MC} (before corr.)	0.257		0.179		0.189		0.244		0.160		0.090		0.137	
B Counting	1.00	1.1	1.00	1.1	1.00	1.1	1.00	1.1	1.00	1.1	1.00	1.1	1.00	1.1
R^{+0}	1.00	2.4	1.00	2.4	1.00	2.4	1.00	2.4						
Tracking Eff	0.99	1.6			1.00	0.8	0.99	0.8	0.99	1.6	0.99	1.6	1.00	0.8
PID	0.98	1.0			0.98	1.0	0.97	1.0	1.00	2.0	1.00	2.0	1.00	2.0
π^0 eff.			0.98	5.1	0.98	5.1					0.97	3.0	0.97	3.0
Photon eff.	1.00	2.5	1.00	2.5	1.00	2.5	1.00	2.5	1.00	3.0	1.00	3.0	1.00	3.0
Distance cut	1.00	2.0	1.00	2.0	1.00	2.0	1.00	2.0	1.00	2.0	1.00	2.0	1.00	2.0
$\pi^0(\eta)$ veto	1.00	1.0	1.00	1.0	1.00	1.0	1.00	1.0	1.00	1.0	1.00	1.0	1.00	1.0
K_s			0.98	3.0			0.98	3.0						
meson mass	1.00	<1.0	1.00	<1.0	1.00	<1.0	1.00	<1.0	1.00	<1.0	1.00	<1.0	1.00	<1.0
NN	1.00	3.0	1.00	3.5	1.00	2.7	1.00	2.8	1.00	4.6	1.00	4.6	1.00	1.8
MC statistics/ fit bias/B bkg. modeling	0.98	0.9	0.91	3.2	0.96	2.4	0.96	1.6	1.00	10.2	1.00	5.4	1.00	8.1
Total	0.95	5.4	0.85	10.7	0.92	9.0	0.91	6.1	0.99	12.1	0.95	10.0	0.97	10.8
ϵ_{MC} (after corr.)	0.244	5.4	0.153	10.7	0.174	9.0	0.221	6.1	0.158	12.1	0.132	10.0	0.083	10.8

Table 6: The systematic uncertainties of $\mathcal{B}$ for each decay mode. For each type of systematic uncertainty related to the signal efficiency, a correction factor on the signal efficiency and the relative error on the signal efficiency are calculated, as described in Section 7.1. The "meson mass" cut efficiency has a systematic error on the signal efficiency far below 1%, and ignored in calculating the total systematic error. The "π^0" and "photon" systematic errors are correlated, therefore they are summed in linear with each other before added in quarture with other systematic errors.

Ball functions in the signal PDF. We vary the fixed values by $\pm 1\sigma$ around the fixed values, which are obtained from a fit to the signal MC sample, and fit the final data set again. The resulting changes in the signal yields are treated as the corresponding systematic errors.

2. The combinatorial B-background PDF is fixed from Monte Carlo sample. We extract a different B-background shape from the dominant component, $B \to X_s\gamma$. Using this shape, we perform the final fit again. The variation of the signal yield is considered as the systematic error of the B-background modeling.

In this way, the total systematic error associated with the fit is ± 12.8, ± 4.5, ± 6.7, and ± 2.2, for for $K^{*0} \to K^+\pi^-$, $K^{*0} \to K_s^0\pi^0$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decay modes, respectively.

7.2.2 Fit Systematics in $B \to (\rho/\omega)\gamma$ Decays

Compared to $B \to (K^{*0}/K^{*+})\gamma$ decays, where most of the signal parameters are determined from the fit itself, all the signal parameters in each $B \to (\rho/\omega)\gamma$ decay are determined from the MC and fixed in the final likelihood fit. The only exception is the mean of ΔE^*, which is further offset by the difference seen between data and MC in $B \to (K^{*0}/K^{*+})\gamma$ decays. Only the yields of the signal, continuum background, and combinatorial B-background (if applicable), and the parameters of m_{ES} and ΔE^* of the continuum background are allowed to float in the likelihood fit. We consider the following fit systematics,

1. **Fixing parameters of the signal PDF.** We vary each of the fixed parameter by $\pm 1\sigma$ of the uncertainties obtained from fitting to the signal Monte Carlo, and fit to the final data again. The resulting variation of the signal yield is then considered as the systematic error of fixing the corresponding parameter. Due to presence of a small

signal, an asymmetric error is calculated. All the systematic errors are then summed in quadrature to obtain the total systematic error, and is labeled "fixing parameters of the signal PDF". This error is $^{+0.22}_{-0.02}$, $^{+1.54}_{-0.99}$, and $^{+0.15}_{-0.16}$ for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay mode, respectively.

2. **Binning/statistics of the histogram PDF.** In the likelihood fit, the NN distribution of the continuum background is obtained from the sideband data and fixed in the likelihood fit. This distribution is irregular, and a binned histogram is directly used in the likelihood fit.

   The binning of the histogram is chosen so that the statistics error of almost every bin is less than 10%. The default bins of the histogram PDF are 45, 60, and 45 for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$, respectively. To study the systematic error of the binning, the bins of the histogram are varied within 25, 35, 45, 55, and 65 for $B^0 \to \rho^0\gamma$ and $B^0 \to \omega\gamma$ decay modes (30, 45, 60, 75, and 90 for $B^+ \to \rho^+\gamma$ decay mode), and used in the final fit. The maximum variations of the signal yield are taken as the systematic errors of "binning of the histogram PDF", which are $^{+0.20}_{-0.70}$, $^{+0.00}_{-0.70}$, and $^{+0.50}_{-0.00}$ for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay mode, respectively.

   To check the systematic error resulting from the statistics of the sideband data, a 5-th Polynomial is fitted to the histogram PDF to smooth the distribution. The smoothed PDF is used in the likelihood fit, replacing the histogram PDF. The resulting variations of the signal yield are taken as the systematic errors of "the statistics of the histogram PDF", which are $^{+0.40}_{-0.40}$ for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay mode, respectively.

3. **The Peaking B-backgrounds.** The yields and the shape of the peaking B-backgrounds are both fixed from Monte Carlo. We only consider the systematic error associated

with fixing the yields of the peaking B-background.

The $B \to (K^{*0}/K^{*+})\gamma$ background yield and its error for the $B^0 \to \rho^0\gamma$ and $B^+ \to \rho^+\gamma$ decay modes are estimated using the measured $B \to (K^{*0}/K^{*+})\gamma$ branching fractions presented in this thesis. The statistics of the $B \to (K^{*0}/K^{*+})\gamma$ MC and the uncertainties of the kaon mis-identification rate are described in Section 7.1. One-sigma variations of the yield are made, and the corresponding changes of the fitted signal yield are taken as the systematic errors. The resulting systematic error for $B^0 \to \rho^0\gamma$ and $B^+ \to \rho^+\gamma$ is $^{+1.59}_{-1.29}$ and $^{+0.92}_{-0.81}$, respectively. For $B^0 \to \omega\gamma$ decay mode, we do not consider $B \to (K^{*0}/K^{*+})\gamma$ background in the fit. Instead, a "$B \to (K^{*0}/K^{*+})\gamma$ background mixed" toy Monte Carlo is used to estimate the bias on the signal yield. A bias of $+1.05$ events is found. We apply this bias on the signal yield and choose the corresponding error on this correction, -1.05, as the systematic error.

The $(\rho/\omega)(\pi^0/\eta)$ backgrounds are slightly different from the $B \to (K^{*0}/K^{*+})\gamma$ background, since they contain two types of decays, and some of the decays have not been observed yet. We vary the branching fraction of each decay mode by one standard deviation for a measured decay, or from zero to the current upper limit for a decay which has only an upper limit (see Section 5.1.2). The exception is for $B \to \rho^0\pi^0$ decay. We vary the branching fraction from zero to 5.1×10^{-6}, the recent Belle result of this decay mode. Combining the variation on the branching fraction with the statistical errors of the Monte Carlo samples, the variations of the yield are calculated. The corresponding changes of the fitted signal yield, when making these variations, are considered to be the corresponding systematic errors. This gives $^{+0.56}_{-0.64}$, $^{+1.31}_{-1.20}$, and $^{+0.26}_{-0.19}$ for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$, respectively.

4. **The combinatorial B-background.** The systematic uncertainty of the combina-

torial B-background has been considered as part of the "MC statistics/fit bias/B-background modeling" for $B^0 \to \rho^0\gamma$ and $B^+ \to \rho^+\gamma$ decay modes. For the $B^0 \to \omega\gamma$ decay, a "combinatorial B" mixed-MC study was performed (Section 5.6.3), and a bias of $+1.5$ events is seen, when this background is not considered in the likelihood fit. We decided not to correct the bias, but to take the full bias as the systematic error. This results a systematic error of $^{+0.0}_{-1.5}$.

5. **The best candidate selection.** For events which contain more than one candidate, we select the best candidate, which is defined as the one with the minimum $|\Delta E^*|$. Before the best candidate selection, the average multiplicity of the signal events is 1.03, 1.15, and 1.14 for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively. The relatively high multiplicity in the $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ decay mode results from the high multiplicity of the background π^0 in an event. For the continuum background, a second order polynomial is used to model the background ΔE^* distribution. We also check the possible effect of best candidate selection on modeling of the continuum background, where an artificial bump could be created by the best candidate selection. We only consider this for $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ decay modes.

First, we look at events in the $m_{ES} < 5.26$ GeV/c^2, $|\Delta E^*| < 0.3$ GeV sideband, in which more than one candidate is found. The ΔE^* distribution of the best candidates is fit with a second-order polynomial, which is fixed to be the same as that used in the unblinded fit, and a Gaussian. Then, we use the Gaussian and the same second-order polynomial to fit all events in the sideband. The size of the Gaussian component is $(1.3 \pm 1.1)\%$ and $(0.6 \pm 2.5)\%$ of the continuum events for $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ decay modes, respectively. This is consistent with no peaking structure.

For $B^+ \to \rho^+\gamma$ decay mode, we perform the final fit, replacing the continuum back-

Description	$B^0 \to \rho^0\gamma$ Sig. bias	error	$B^+ \to \rho^+\gamma$ Sig. bias	error	$B^0 \to \omega\gamma$ Sig. bias	error
Fixing of sig. PDFs	0.0	$+0.22$ -0.02	0.0	$+1.54$ -0.99	0.0	$+0.15$ -0.16
Binning of histogram PDFs	0.0	$+0.20$ -0.70	0.0	$+0.00$ -0.70	0.0	$+0.50$ -0.00
Statistics of histogram PDFs	0.0	$+0.40$ -0.40	0.0	$+0.40$ -0.40	0.0	$+0.40$ -0.40
$B \to (K^{*0}/K^{*+})\gamma$ bkg.	0.0	$+1.59$ -1.29	0.0	$+0.92$ -0.81	-1.05	$+1.05$ -1.05
$(\rho/\omega)(\pi/\eta)$ bkg.	0.0	$+0.56$ -0.64	0.0	$+1.31$ -1.20	0.0	$+0.26$ -0.19
B bkg.	n/a		n/a		0.0	$+0.00$ -1.50
Best Candidate Selection	n/a		0.0	$+0.0$ -1.2	n/a	
Total	0.0	$+1.7$ -1.6	0.0	$+2.2$ -2.2	-1.05	$+1.3$ -1.9

Table 7: The summary of all systematic uncertainties associated with the likelihood fit for each decay mode. For each type of systematic uncertainty, a bias on the signal yield and the error (in events) on the signal bias are calculated. (here "bkg." represents "background".)

ground ΔE^* PDF with a sum of a Gaussian and a second-order polynomial, with the normalization of the Gaussian fixed to 1.3% of the total background yield. This leads to a 1.2 event reduction in the signal yield. To be conservative, we take this as an additional asymmetric uncertainty on the signal yield of -1.2 events in $B^+ \to \rho^+\gamma$ decay mode. For $B^0 \to \omega\gamma$ decay mode, the same procedure is followed and the effect is estimated to be on the order of 0.1 events. Therefore, we ignore this effect for the $B^0 \to \omega\gamma$ decay mode.

6. **Systematic error of the signal NN distribution.** The systematic effect of the NN distribution is discussed in Appendix A. Because the effect on the signal yield is estimated to be less than 2% of the signal yield, we decide to ignore this systematic error.

Table 7 summarizes the fit systematic errors discussed above. The total fit systematic error is then calculated to be $^{+1.7}_{-1.6}$, $^{+2.2}_{-2.2}$, and $^{+1.3}_{-1.9}$ for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$,

respectively. A bias of -1.05 is then applied on the $B^0 \to \omega\gamma$ fitted signal yield, due to the effect of the $B \to (K^{*0}/K^{*+})\gamma$ background.

8 Physics Results

In the following, the branching fractions of $B \to (K^{*0}/K^{*+})\gamma$ and $B \to (\rho/\omega)\gamma$ decays are given first. The ratio of the branching fractions is then used to constrain $|V_{td}|/|V_{ts}|$. Finally the CP-asymmetry and isospin-asymmetry of $B \to (K^{*0}/K^{*+})\gamma$ decays are discussed.

8.1 $B \to (K^{*0}/K^{*+})\gamma$ Branching Fractions

The branching fraction of the $B^0 \to K^{*0}\gamma$ decay ($B^+ \to K^{*+}\gamma$ decay) was measured exclusively from both $K^{*0} \to K^+\pi^-$ and $K^{*0} \to K_s^0\pi^0$ decay modes ($K^{*+} \to K^+\pi^0$ and $K^{*+} \to K_s^0\pi^+$ decay modes), respectively. Individual results are combined to give a more precise measurement. A χ^2 method is used to combine the individual results. A detailed discussion is presented in Appendix C.

8.2 Results

The branching fraction, the signal efficiency, ϵ, and the fitted signal yield, N_S, of each mode are summarized in Table 8. Following the procedure described in Appendix C, the combined branching fractions for $B^0 \to K^{*0}\gamma$ and $B^+ \to K^{*+}\gamma$ decays are calculated, and shown in Table 8. Figure 41 shows a comparison of these results with the theoretical pre-

Mode	$\epsilon(\%)$	N_S	$\mathcal{B}(\times 10^{-5})$	Combined $\mathcal{B}(\times 10^{-5})$
$K^{*0} \to K^+\pi^-$	24.4±1.4	582.6±29.7	3.92±0.20±0.23	} 3.92±0.20±0.24
$K^{*0} \to K_s^0\pi^0$	15.3±1.9	61.8±15.3	4.02±0.99±0.51	
$K^{*0} \to K^+\pi^-$	17.4±1.6	250.9±22.6	4.90±0.45±0.46	} 3.87±0.28±0.26
$K^{*+} \to K_s^0\pi^+$	22.1±1.4	156.9±15.7	3.52±0.35±0.22	

Table 8: The signal efficiency ϵ, the fitted signal yield N_S and the branching fraction $\mathcal{B}$ of each decay mode are shown. The combined branching fractions for $B^0 \to K^{*0}\gamma$ or $B^+ \to K^{*+}\gamma$ are also shown. Errors are statistical and systematic, respectively, with the exception of ϵ, which has only a systematic error, and N_S, which has only a statistical error. The detailed description of the systematic errors is found in Table 6 and Section 7.2.
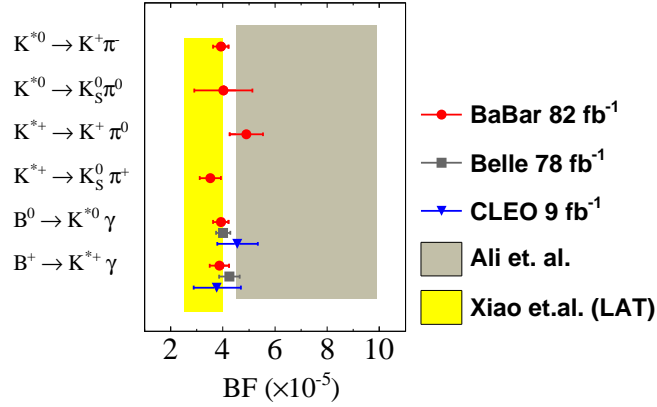
Figure 41: Summary of branching fractions for the $B \to (K^{*0}/K^{*+})\gamma$ decays. The *BABAR* result is compared to experimental results from CLEO [30] and Belle [31], and the theoretical predictions from Ali *et al.* [40] and Xiao *et al.* [78].

dictions from several models and with the results from Belle and CLEO. These results are in good agreement with the theoretical predictions. However, the experimental results are already much more precise than any of the theoretical predications. As discussed in Section 2, the theoretical uncertainties are dominated by the determination of the form factors. The measurements presented here indicate smaller form factors than the ones predicted by QCD sum rules, while they are in good agreement with the prediction of the lattice QCD calculation. These results, with the *CP*-asymmetry and isospin asymmetry discussed later, can be used to constrain the Higgs mass in 2HDM. A detailed discussion can be found in Ref. [78].

8.3 $B \to (\rho/\omega)\gamma$ Branching Fractions

Unlike the $B \to (K^{*0}/K^{*+})\gamma$ decays, no signals are seen for individual $B \to (\rho/\omega)\gamma$ decays. In order to search for evidence of $b \to d\gamma$ decays, we combine individual decay modes to improve statistics.

8.3.1 Combining Technique

The combined branching fraction for $B \to (\rho/\omega)\gamma$ is,

$$\overline{\mathcal{B}}(B \to (\rho,\omega)\gamma) = \frac{1}{2} \cdot \{\mathcal{B}(B^+ \to \rho^+\gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot [\mathcal{B}(B^0 \to \rho^0\gamma) + \mathcal{B}(B^0 \to \omega\gamma)]\}. \qquad (71)$$

To avoid complications in calculating the covariance matrix, we combine the branching fraction of $B \to (\rho/\omega)\gamma$ decay modes in a simultaneous fit. We define the following "efficiency-corrected yield":

$$N_{\text{eff}} = \frac{N(B^+ \to \rho^+\gamma)}{\epsilon_{B^+ \to \rho^+\gamma}} + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot [\frac{N(B^0 \to \rho^0\gamma)}{\epsilon_{B^0 \to \rho^0\gamma}} + \frac{N(B^0 \to \omega\gamma)}{\epsilon_{B^0 \to \omega\gamma}}],$$

where Ns and ϵs are the fitted signal yield and the signal efficiency for each decay mode. The decay rates are constrained as follows:

$$\Gamma(B^0 \to \rho^0\gamma) = \Gamma(B^0 \to \omega\gamma) = 0.5 \cdot \Gamma(B^+ \to \rho^+\gamma).$$

The efficiency-corrected yield is related to the $N_{B\overline{B}}$ pairs by $N_{\text{eff}} = 2 \cdot N_{B\overline{B}} \cdot \mathcal{B}(B \to (\rho,\omega)\gamma)$. We determine the efficiency-corrected yield directly from the likelihood fit. Each component of this fit is identical to the corresponding individual fit described in Section 5.6, with the only difference being that the signal yield for each component is given as a function of the combined efficiency-corrected yield; $0.5 \cdot N_{\text{eff}} \cdot \epsilon_{B^+ \to \rho^+\gamma}$, $0.25 \cdot N_{\text{eff}} \cdot \tau_{B^0}/\tau_{B^+} \cdot \epsilon_{B^0 \to \rho^0\gamma}$, and $0.25 \cdot N_{\text{eff}} \cdot \tau_{B^0}/\tau_{B^+} \cdot \epsilon_{B^0 \to \omega\gamma}$ for $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$ modes, respectively. The bias on the signal yield in the $B^0 \to \omega\gamma$ decay mode, and the corrections on signal efficiencies have already been applied.

8.3.2 Combined Results

The combined likelihood fit was validated using all the types of Monte Carlo samples described in Section 5.6. With the assumed branching faction $\mathcal{B}(B^+ \to \rho^+\gamma) = 1.0 \times 10^{-6}$, we expected more than 3σ significance on the combined branching fraction. Figure 42
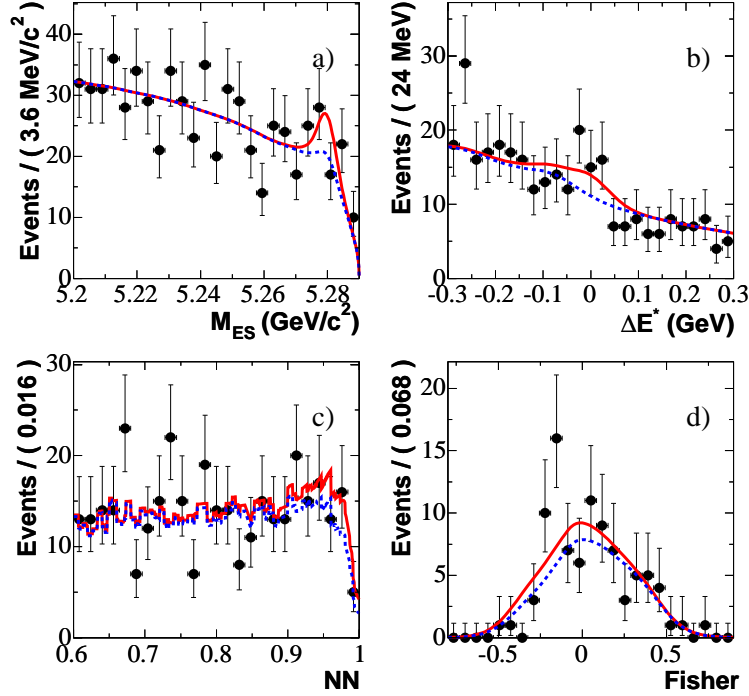
Figure 42: The projection plots of the combined likelihood fit, where a), b), c), and d) are the projection onto m_{ES}, ΔE^*, neural net, and Fisher, respectively. The solid line is the total fit and the dashed line is the total background. The plots are projections with the following cuts applied, if that variable is not being plotted: $5.272 < m_{ES} < 5.286$ GeV/c^2, $-0.10 < \Delta E^* < 0.05$ GeV and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 45%, 57%, 70%, and 44% for m_{ES}, ΔE^*, NN, and Fisher projections, respectively.

shows the projection plots of the combined fit, and the efficiency-corrected yield is 269^{+126}_{-120}. The yields of the combinatorial B-background and the continuum background have been compared with the individual fits, and they are in good agreement with each other.

Figure 43 shows the likelihood function with respect to this efficiency-corrected yield, N_{eff}. We then take the significance as $\sqrt{-2\ln \mathcal{L}_0/\mathcal{L}_{\max}}$, where $\mathcal{L}_0$ is the value of the likelihood for the null signal hypothesis. Considering statistics only, a 2.6σ significance is found. Unfortunately, we do not see evidence for $b \to d\gamma$ decays.

Figure 43: Likelihood function of the combined fit. The dashed horizontal line indicates the $-ln(\mathcal{L}/\mathcal{L}_{\max}) = 0.82$ and the solid horizontal line indicates the $-ln(\mathcal{L}/\mathcal{L}_{\max}) = 1.34$.

8.3.3 Limit Setting

Since there is no evidence for $b \to d\gamma$ decays, we set an upper limit for each decay mode at the 90% Confidence Level (C.L.). We define the 90% C.L. on the efficiency-corrected signal yield for each mode as
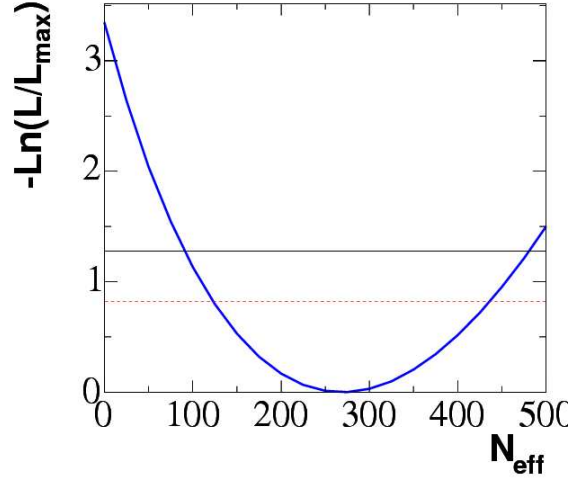
$$N_{\text{eff}}^{90} = \frac{N_{\text{sig}}}{\epsilon} + 1.28 \times \frac{\sigma_{\text{stat}}(N_{\text{sig}})}{\epsilon} + 1.28 \times \sigma_{\text{syst}}(N_{\text{sig}}/\epsilon) \,,$$

where N_{sig} is the fitted signal yield, $\sigma_{\text{stat}}(N_{\text{sig}})$ is the statistical error on the signal yield, ϵ is the efficiency and $\sigma_{\text{syst}}(N_{\text{sig}}/\epsilon)$ is the combined systematic uncertainty on the effective signal yield,

$$\sigma_{\text{syst}}(N_{\text{sig}}/\epsilon) = \frac{N_{\text{sig}}}{\epsilon} \sqrt{\left(\frac{\sigma_{\text{syst}}(\epsilon)}{\epsilon}\right)^2 + \left(\frac{\sigma_{\text{syst}}(N_{\text{sig}})}{N_{\text{sig}}}\right)^2} \,.$$

In practice, the value of $N_{\text{sig}} + 1.28 \times \sigma_{\text{stat}}(N_{\text{sig}})$ is obtained by scanning the likelihood function, $\mathcal{L}$, with respect to the signal yield to find the value at which $-2\ln \mathcal{L}/\mathcal{L}_{\max} = 1.28^2 = 1.64$, where $\mathcal{L}_{\max}$ is the maximum value of the likelihood. The values of $\sigma_{\text{syst}}(N_{\text{sig}}/\epsilon)$ are $^{+10.9}_{-9.9}$, $^{+26.8}_{-26.1}$, and $^{+18.7}_{-23.6}$ events for $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively.

Description	N_{sig}	fit sys.	ϵ	$\mathcal{B}(\pm\text{stat.} \pm \text{sys.})$ $(\times 10^{-6})$	90% limit $(\times 10^{-6})$	95% limit $(\times 10^{-6})$
$B^0 \to \rho^0\gamma$	$0.3^{+7.2}_{-5.4}$	$^{+1.7}_{-1.6}$	0.158 ± 0.019	$0.01^{+0.22+0.05}_{-0.16-0.05}$	<0.36	<0.44
$B^+ \to \rho^+\gamma$	$25.6^{+15.3}_{-13.9}$	$^{+2.2}_{-2.2}$	0.132 ± 0.014	$0.92^{+0.55+0.13}_{-0.50-0.13}$	<1.76	<1.99
$B^0 \to \omega\gamma$	$8.3^{+5.7}_{-4.5}$	$^{+1.3}_{-1.9}$	0.086 ± 0.009	$0.46^{+0.31+0.08}_{-0.25-0.12}$	<0.97	<1.08
Combined result	269^{+126}_{-120}	$^{+40.2}_{-44.5}$	n/a	$0.64^{+0.32+0.10}_{-0.28-0.10}$	<1.16	<1.31

Table 9: Summary of the fit yield, the systematic error, the signal efficiency, the branching fraction, the 90% upper limit and the 95% upper limit for each decay mode, and the combined results.

Using the values of N_{eff}^{90} found with the above method we set the 90% C.L. upper limit of 0.36×10^{-6}, 1.76×10^{-6}, and 0.97×10^{-6} for the $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$ decay modes, respectively.

As described above, we take the significance to be $\sqrt{-2\ln \mathcal{L}_0/\mathcal{L}_{\text{max}}}$, where $\mathcal{L}_0$ is the value of the likelihood for the null signal hypothesis. We translate the likelihood curve of the efficiency-corrected signal yield by $-\sigma_{\text{syst}}(N_{\text{sig}}/\epsilon)$ to account for the systematic error. We find the significances for $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ are 1.9σ and 1.5σ respectively.

We combine the three decay modes together with a simultaneous fit by considering $\Gamma(B^0 \to \rho^0\gamma) = \Gamma(B^0 \to \omega\gamma) = 0.5 \cdot \Gamma(B^+ \to \rho^+\gamma)$, to obtain a combined effective signal yield, N_{eff}. We distinguish the correlated and uncorrelated systematic uncertainty among three the decay modes. These systematic uncertainties are described in Appendix D and summarized in Table 12. The total systematic error[37] on N_{eff} is $^{+40.2}_{-44.5}$. The significance of the combined results is 2.1σ. We then determine the 90% C.L. as was done for the individual modes by scanning the likelihood with respect to N_{eff}. We find the $\overline{\mathcal{B}}(B \to (\rho/\omega)\gamma) < 1.16 \times 10^{-6}$ at 90% C.L.

---

[37]The systematic error in the number of $B\overline{B}$-pairs have been included already in this error.

Table 9 summarizes the results for the $B \to (\rho/\omega)\gamma$ decays. These results are plotted in Figure 44, and compared with the theoretical predictions and other experimental results. The experimental results are in good agreement with each other, and are slightly below the theoretical predictions.

8.4 Constraint on $|V_{td}|/|V_{ts}|$

To constrain $|V_{td}|/|V_{ts}|$, we use the isospin-averaged branching fractions. Similar to Eq. 71, we define the isospin-averaged branching fraction for $B \to (K^{*0}/K^{*+})\gamma$ decays:

$$\overline{\mathcal{B}}(B \to (K^{*0}/K^{*+})\gamma) = \frac{1}{2} \cdot \{\mathcal{B}(B^+ \to K^{*+}\gamma) + \frac{\tau_{B^+}}{\tau_{B^0}}\mathcal{B}(B^0 \to K^{*0}\gamma)\}. \tag{72}$$

Using our measured values of $B \to (K^{*0}/K^{*+})\gamma$ branching fractions, we can constrain the ratio of $\overline{\mathcal{B}}[B \to (\rho/\omega)\gamma]/\overline{\mathcal{B}}(B \to (K^{*0}/K^{*+})\gamma) < 0.030$. This in turn is used to bound
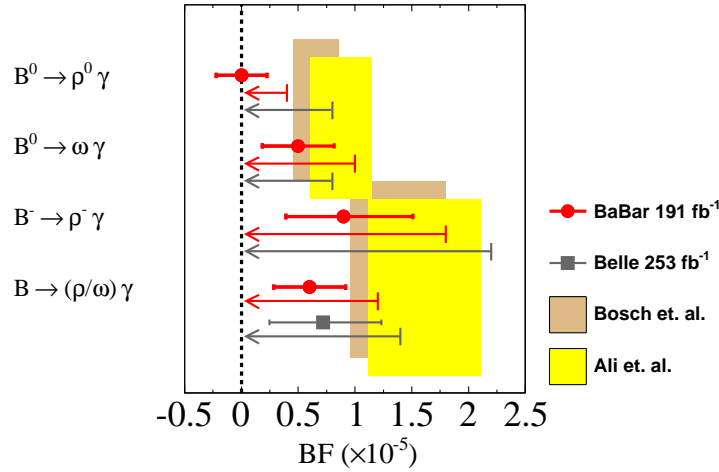

Figure 44: Summary of branching fractions for the $B \to (\rho/\omega)\gamma$ decays (The arrows indicate the 90% C.L. upper limit). The *BABAR* result is compared to experimental result from Belle [79], and the theoretical predictions from Bosch *et al.* [27] and Ali *et al.* [35].
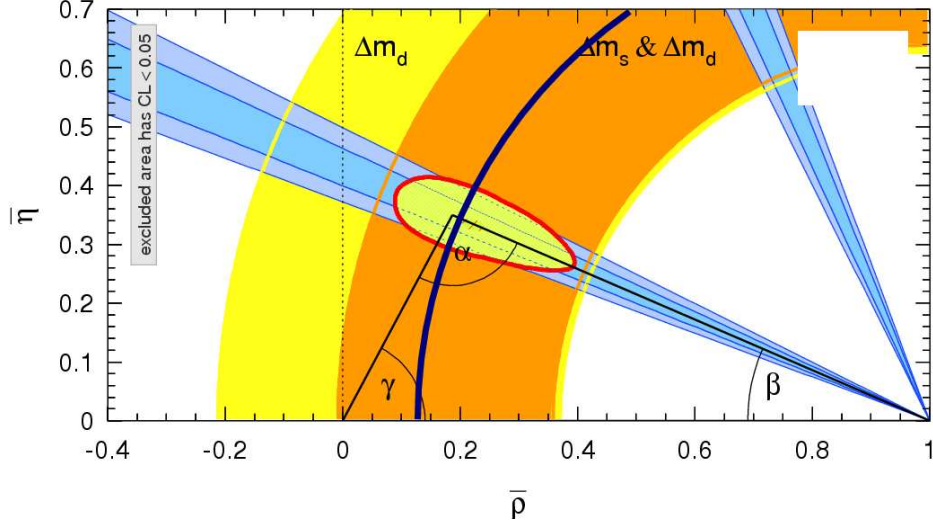
Figure 45: Constraint on the $|V_{td}|/|V_{ts}|$ from this result. The 95% C.L. limit on $|V_{td}|/|V_{ts}|$ is converted to a constraint on the $\overline{\rho}$ and $\overline{\eta}$ parameters of the Unitarity Triangle, indicated by the bold solid arc, where the left-side of the SM-allowed region (inside region of the irregular contour) is excluded.

$|V_{td}/V_{ts}|$, using the relationship [35, 40]:

$$\frac{\overline{\mathcal{B}}[B \to (\rho/\omega)\gamma]}{\overline{\mathcal{B}}(B \to (K^{*0}/K^{*+})\gamma)} = \left|\frac{V_{td}}{V_{ts}}\right|^2 \cdot r_m \cdot \zeta^2[1 + \Delta R]. \tag{73}$$

Both ζ and ΔR must be taken from theory. Several different values of ζ and ΔR have been published [35, 40, 80]. Following [35], we choose the values $\zeta = 0.85 \pm 0.10$, and $\Delta R = 0.10 \pm 0.10$, as the average over the values given for the three modes. We find the limit $|V_{td}/V_{ts}| < 0.19$ at 90% C.L. by ignoring any theoretical uncertainty. Allowing ζ and ΔR to vary within their theoretical uncertainties the upper limit changes by ±0.3. Theoretical improvement on the calculation of the form factor is needed.

The limit on the $|V_{td}|/|V_{ts}|$ can be used to constrain the $\overline{\rho}$ and $\overline{\eta}$ parameters of the Unitarity Triangle, $|V_{td}|/|V_{ts}| = \lambda \cdot \sqrt{(1 - \overline{\rho})^2 + \overline{\eta}^2}$, where $\lambda \sim 0.226$. At the 95% C.L., our result constrains $|V_{td}|/|V_{ts}| < 0.20$. Figure 45 shows this constraint on the $\overline{\rho} - \overline{\eta}$ plane.

Currently, $|V_{td}|$ is measured using the B_d mixing. The world average of the mixing

parameter Δm_d is 0.502±0.006 ps^{-1} [39]. This result is converted to constraints on the $\bar{\rho}$ and $\bar{\eta}$ parameters, shown in Figure. 45. The large uncertainty results from the large theoretical error in Eq. 43. Using the current world 95% C.L. limit on the B_s mixing parameter, $\Delta m_s > 14.5$ ps^{-1} [39], and Eq. 44, this constraint has been significantly improved, shown in the same Figure 45.

The limit on $|V_{td}|/|V_{ts}|$ can be interpreted as a limit on $|V_{td}| < 0.008$ at 90% C.L. by assuming $V_{ts} = V_{cb} = (41.3 \pm 1.5) \times 10^{-3}$ [38]. The upper limit changes by ±0.001 when varying ζ and ΔR within their theoretical uncertainties. This lies within the 90% confidence interval found from a global fit to CKM data of $0.0048 < |V_{td}| < 0.014$ reported in Ref. [38]. Other ways of interpreting our results can be seen in Ref. [37]. If the individual branching fractions are used to constrain the $|V_{td}|/|V_{ub}|$, a theoretically cleaner result can be obtained. Furthermore, a ratio between $B \to (\rho/\omega)\gamma$ decays and $B \to \rho l \nu$ decays can be used to constrain the $|V_{td}|/|V_{ub}|$, where form-factor uncertainty can be further reduced to give a theoretically cleaner way to extract the $|V_{td}|$.

8.5 CP- and Isospin- Asymmetry in $B \to (K^{*0}/K^{*+})\gamma$ Decays

The branching fractions of the $B \to (K^{*0}/K^{*+})\gamma$ decays are measured very precisely. However, the large theoretical errors in the SM predictions limit their sensitivity to probe new physics. Due to cancellation of theoretical uncertainties in a ratio of decay rates, the CP-asymmetry and the isospin-asymmetry in these decays are well calculated and have excellent beyond-the-SM sensitivity. The results of CP-asymmetry and isospin-asymmetry are presented in the following.

8.5.1 CP-Asymmetry

The $\mathcal{A}_{CP}$ is defined in Eq. 40, and related to the observed signal events as following,

$$\mathcal{A}(B \to V\gamma) = \frac{1}{1 - 2\eta} \cdot \frac{N(\overline{B} \to \overline{V}\gamma) - N(B \to V\gamma)}{N(\overline{B} \to \overline{V}\gamma) + N(B \to V\gamma)}, \tag{74}$$

where η is the mis-tag rate (the fraction of $\overline{B} \to \overline{V}\gamma$ mis-reconstructed as $B \to V\gamma$). Here we have already assumed that the signal efficiencies between the $\overline{B} \to \overline{V}\gamma$ and its CP-conjugate decay are the same. Any signal efficiency difference between them is treated as a systematic error.

We search for CP-violation using $K^{*0} \to K^+\pi^-$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decay modes. We identify the $\overline{B} \to \overline{V}\gamma$ decay and its CP-conjugated decay using the correlation of the charge of the final stable particles ($K^\pm$ and $\pi^\pm$) to the charge of the b-quark. From a Monte Carlo study, the mis-tag rate for each decay mode is found to be less than 0.1%. So, we will ignore the mis-tag rate in the following.

We extract the $\mathcal{A}_{CP}$ using a simultaneous fit over each decay mode and its CP-conjugate decay mode. Two components of the simultaneous PDF are identical to the one we used in Section 5.6. Two more free parameters are introduced, the asymmetry of the signal $\mathcal{A}_{CP,Sig}$ and the asymmetry of the continuum background $\mathcal{A}_{CP,Cont}$. The asymmetry of the combinatorial B-background is assumed to be the same as that of the continuum background. The fitted CP-asymmetries are:

$$\begin{aligned}
\mathcal{A}(K^{*0} \to K^+\pi^-) &= -0.069 \pm 0.046 \pm 0.011, \\
\mathcal{A}(K^{*+} \to K^+\pi^0) &= +0.084 \pm 0.075 \pm 0.007, \\
\mathcal{A}(K^{*+} \to K_s^0\pi^+) &= +0.084 \pm 0.075 \pm 0.007.
\end{aligned} \tag{75}$$

The first error is the statistical error and the second error is the systematic error, which mainly originates from the asymmetries in the charged particle identification. The systematic asymmetry errors in particle identification are 1.0%, 0.5%, and 0.5% for $K^{*0} \to K^+\pi^-$, $K^{*+} \to K^+\pi^0$, and $K^{*+} \to K_s^0\pi^+$ decay modes, respectively. The *BABAR* detector also

shows a slight asymmetry in the detection of the charged tracks. This has been studied by the tracking AWG, and a systematic error of 0.25% is assigned for each track [72]. In the low-momentum range, charged kaons interact with material with different cross-sections from their *CP*-states [38]. This results in an additional 0.15-0.30% systematic error on the *CP*-asymmetry. Finally, a 0.25% systematic error is assigned for each decay mode to account for the assumption that the combinatorial *B*-background has same *CP*-asymmetry as the continuum background. The fitted *CP*-asymmetry of the continuum background is consistent with zero within statistics. Following the technique in Appendix C, the individual results are combined to give

$$\mathcal{A}_{CP}(B \to K^*\gamma) = -0.013 \pm 0.036 \pm 0.010.$$

This gives 90% C.L. interval of

$$-0.074 < \mathcal{A}(B \to (K^{*0}/K^{*+})\gamma) < 0.049.$$

8.5.2 Isospin-Asymmetry

The isospin-asymmetry defined in Eq. 41 can be expressed as,

$$\Delta_{0-}(B \to V\gamma) = \frac{I \times R^{+/0}\tau^+/\tau^0 - 1}{I \times R^{+/0}\tau^+/\tau^0 + 1}, \tag{76}$$

where $I = \mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B^+ \to K^{*+}\gamma)$, $R^{+/0} \equiv \Gamma(\Upsilon(4S) \to B^+B^-)/\Gamma(\Upsilon(4S) \to B^0\overline{B^0})$. Correlated experimental errors cancel in the ratio I. The measurement of the $R^{+/0}$ from *BABAR* [71] is consistent with 1.0. Similar to the treatment of the branching fractions in $B \to (K^{*0}/K^{*+})\gamma$ decays, we treat the 2.4% experimental error of $R^{+/0}$ as a systematic uncertainty. We use the same combination technique described in Appendix C to obtain

$$\Delta_{0-} = 0.050 \pm 0.045(\text{stat.}) \pm 0.028(\text{syst.}) \pm 0.024(R^{+/0}).$$

This restricts Δ_{0-} to the 90% confidence interval of

$$-0.046 < \Delta_{0-} < 0.146.$$

8.5.3 Summary of Asymmetries

The *CP*- and isospin-asymmetries of the $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays are plotted in Figure 46. Theoretical predictions based on the SM and other experimental results for these decays are also shown. The experimental results are in good agreement with each other, and with the precise SM predictions at a level of a few percent. The experimental results test the SM predictions at the few percent level.

As discussed previously, the Wilson coefficients summarize the short-distance physics. They behave like coupling constants. For radiative penguin B-decays, the dominant contribution is from the Wilson coefficient, C_7. New physics contributions may differ the SM predictions of these Wilson coefficients. In Ref. [42], the authors argue that in some new physics models, the Δ_{0-} is sensitive to the sign of C_7 and can be negative in some new physics
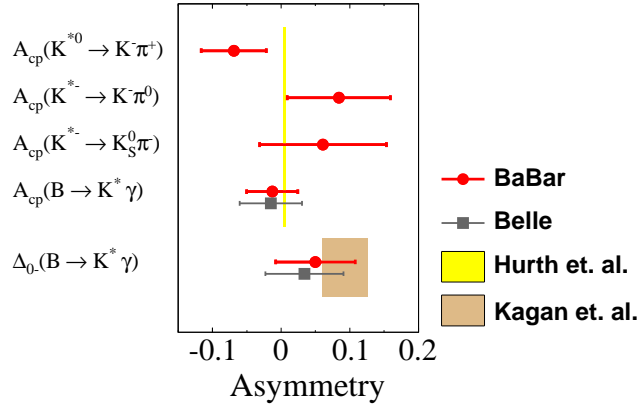
Figure 46: Summary of asymmetries for the $B \rightarrow (K^{*0}/K^{*+})\gamma$ decays. The *BABAR* result is compared to experimental result from Belle [31], and the theoretical predictions from Hurth *et al.* [41] and Kagan *et al.* [42].

models. Combined with Belle results, also shown in Figure 46, a positive Δ_{0-} is preferred, which is consistent with the SM prediction within error. However, the experimental error does not yet exclude the negative region completely.

9   Summary and Outlook

We have studied the radiative penguin decays using *BABAR* data. These results provide experimental tests of the SM predictions.

- The branching fractions of the $B \to (K^{*0}/K^{*+})\gamma$ decays are already much more precise than the theoretical predictions. These experimental results can give guidance for theoretical modeling of these decays; e.g., constrain the form factors of these decays.

- The *CP*-asymmetry and isospin asymmetry of the $B \to (K^{*0}/K^{*+})\gamma$ decays provide tests of the SM at a level of a few percent. They are in good agreement with the SM predictions, limiting the parameter space of new physics models [78].

- We have not yet found evidence for the $B \to (\rho/\omega)\gamma$ decays due to limited statistics. However, our results give the best upper limits at this moment. These results provide a competitive constraint on the $|V_{ts}|/|V_{ts}|$, using a ratio between the $B \to (K^{*0}/K^{*+})\gamma$ decays and the $B \to (\rho/\omega)\gamma$ decays. These results greatly enhance theoretical interest in radiative penguin decays.

By summer of 2006, *BABAR* expects to collect a total data sample of more than 500 fb^{-1}. With this *BABAR* data set and assuming $\mathcal{B}(B^+ \to \rho^+\gamma) = 1.0 \times 10^{-6}$, a better than 5σ significant measurement of the $b \to d\gamma$ decay is expected, as shown in Figure 47, where the projection of the relative errors of the branching fractions are shown for individual decay modes and the combined result. Figure 47 also shows the projection of the relative error on the $|V_{td}|/|V_{ts}|$ using this analysis. It is very likely that we will determine $|V_{td}|/|V_{ts}|$ with less than a 15% error within two years. As shown, the theoretical error in the form factor ratio is the largest theoretical uncertainty. Theoretical improvement is needed. Improved measurements on the *CP*- and isospin-asymmetries of these decays can provide very stringent
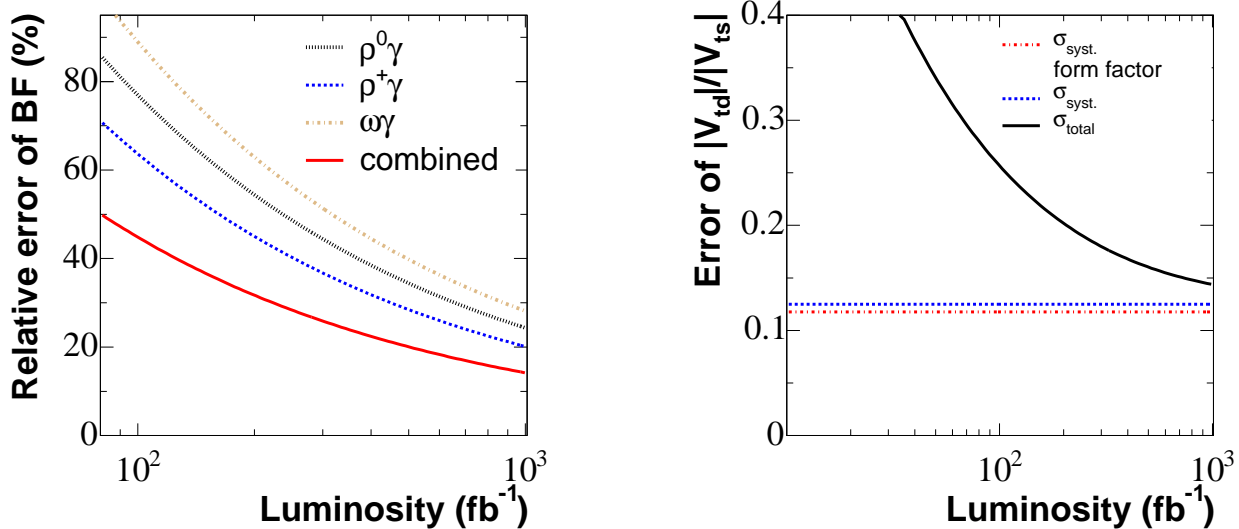
Figure 47: Projection of relative error on the $B \to (\rho/\omega)\gamma$ branching fractions (left) and the $|V_{td}|/|V_{ts}|$ (right) using this analysis. Projection of relative error on $|V_{td}|/|V_{ts}|$ using this analysis. (see text for details)

tests of the SM, hopefully providing hints of physics beyond the SM.

A   The Systematic Errors of Neural Networks

For the analysises discussed in this dissertation, the extensive usage of the neural network (NN) has brought excellent improvement on the background rejection. The improvement mainly results from the possibility of using of additional background suppression variables and the account for the correlations among background suppression variables. About 30-50% more background rejection is obtained in the analysis of the $B \to (K^{*0}/K^{*+})\gamma$ decays than in the previous analysis of the same decays [77], where cuts on background suppression variables are directly applied. The detailed discussion of these benifits can be refered in Section 5, and a general discussion of the neural network can be found here [63]. The training of the NNs is performed using the Stuttgart Neural Network Simulator (SNNS) [65] over Monte Carlo data samples. The NN distributions are then validated using control samples obtained from the *BABAR* data, then we sign the systematic errors accordingly. The detailed procedures are described as follows.

A.1   Introduction

The NN distributions of the $B \to V\gamma$ signal decays are checked using $B \to D\pi$ control samples. For the $B \to D\pi$ control samples, the background suppression variables are calculated in the same way as for the $B \to V\gamma$ decays, by treating the bachelor pion in the $B \to D\pi$ decay as the signal photon, as shown in Figure 48. Since most of the background suppression variables are related to the rest of the event[38], similar distributions on these variables are obtained for the $B \to V\gamma$ decays and the $B \to D\pi$ decays. Therefore the $B \to D\pi$ decays are a very good control sample of the NNs used in the $B \to V\gamma$ decays.

Table 10 summarizes the on-resonance data of the $B \to D\pi$ control samples used to validate the NNs in the $B \to V\gamma$ decays. The corresponding signal MCs are also used. The

---

[38]The exception is for kinematic variables on the signal side.

control samples are selected in such a way that the charge of the B mesons is the same for the $B \to V\gamma$ decay and the corresponding control sample. The NNs used in $B \to V\gamma$ decays are applied to the corresponding $B \to D\pi$ control sample. The NN outputs of data and MC are compared to each other to check for any data-MC difference, and systematic errors are assigned accordingly. Notice that we use different NN configurations for each decay mode. In the following, we identify the NN by the $B \to V\gamma$ decay mode, e.g. the $B^0 \to \rho^0\gamma$ NN. The control sample used to validate the NN is implied, as shown in Table 10.

A.2 Signal Selection

Identical reconstruction schemes are followed to obtain the $B \to D\pi$ candidates on both the on-resonance data and the signal MCs. Event-selection criteria are applied to suppress the continuum background and the B-background. These selection criteria include
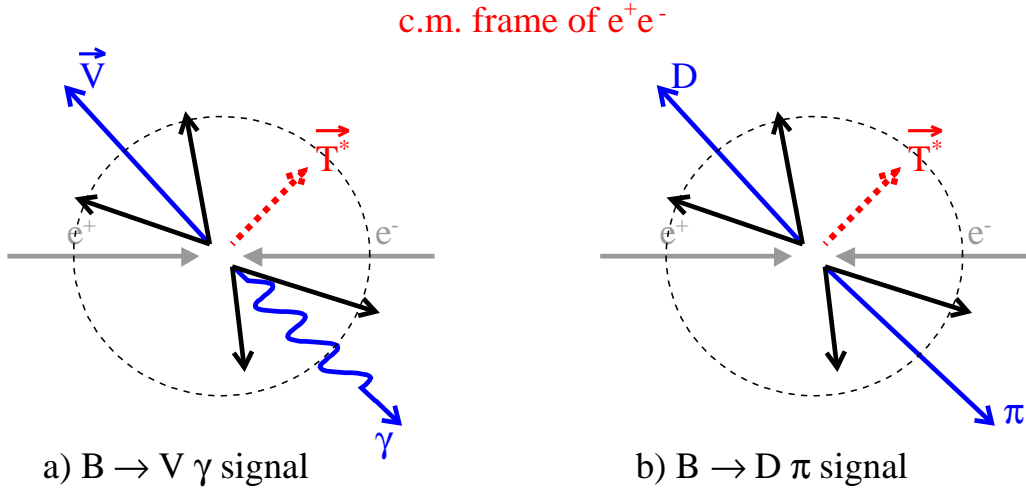

Figure 48: The event shape of the $B \to V\gamma$ signal a) and the $B \to D\pi$ signal b) in the c.m. frame of the e^+e^--system. After singling out the reconstructed signal, indicated by labeled solid lines, the rest of event (ROE), indicated by the unlabeled solid lines, are grouped together. The thrust axis of ROE, $\vec{T}^*$, is indicated by the dashed arrow.

charged particle identification, requirements on the charged tracks, etc. After these cuts, the B-background in the signal region is estimated to be between 1-2% of the events in the signal region, which is defined to be a region of $5.27 < \mathrm{m}_{ES} < 5.29$ GeV/c^2 and $|\Delta E^*| < 0.05$ GeV. A small fraction of continuum background still presents in the signal region for the on-resonance data, and we estimated the yield of this background using a 1-dimensional unbinned likelihood fit over the m_{ES}. The likelihood includes two components, a Gaussian component and an ARGUS component and uses all the events in the fit region, which is defined to be a region of $5.2 < \mathrm{m}_{ES} < 5.29$ GeV/c^2 and $|\Delta E^*| < 0.05$ GeV. Some examples of the fits are shown in Figure 49. The yields of the signal and the continuum background in the signal region are then obtained.

A.3 NN Distributions

The events in the signal region are chosen to make a comparison of the NN distributions between the data and the MC. For this purpose, 100-bin histograms are used. For the on-resonance data, the remaining continuum background is subtracted using the events in the lower sideband, which is defined to be a region of $5.2 < \mathrm{m}_{ES} < 5.27$ GeV/c^2 and $|\Delta E^*| <$

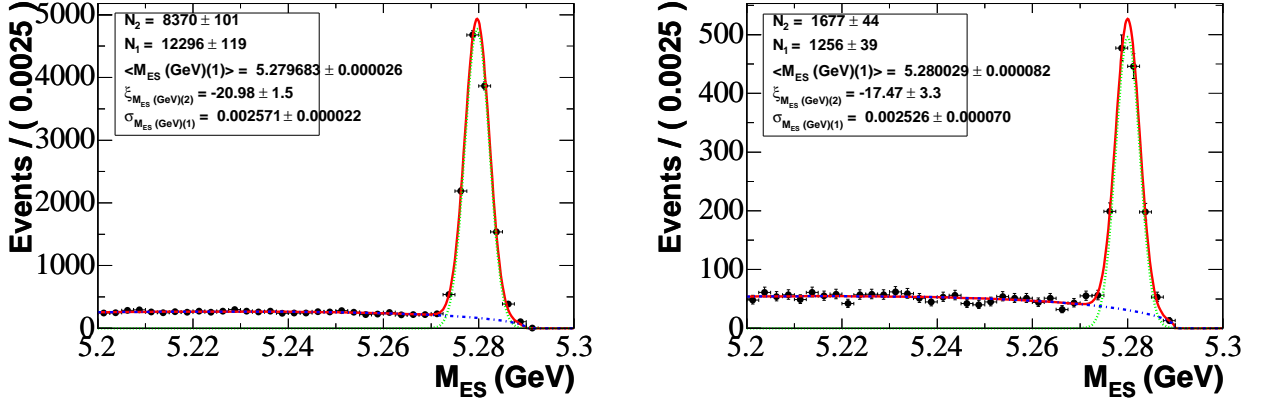| Modes	Control sample	$\mathcal{B}(\times 10^{-6})$	Luminosity (fb^{-1})
$K^{*0} \to K^+\pi^-$ $K^{*0} \to K^0_s\pi^0$	$B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$	273 ± 41	81.9
$B^0 \to \rho^0\gamma$ $B^0 \to \omega\gamma$	$B^0 \to D^-\pi^+, D^- \to K^0_s\pi^-$	42 ± 6	167.0
$K^{*+} \to K^+\pi^0$ $K^{*+} \to K^0_s\pi^+$	$B^+ \to D^0\pi^+, D^0 \to K^+\pi^-$	201 ± 20	81.9
$B^+ \to \rho^+\gamma$	$B^+ \to D^0\pi^+, D^0 \to K^+\pi^-$	201 ± 20	167

Table 10: Control samples used in the NN validation.

Figure 49: Left: the likelihood fit on the $B^- \to D^0\pi^-$ ($D^- \to K^+\pi^+\pi^-$) on-peak data. Right: the likelihood fit on the $B^0 \to D^-\pi^+$ ($D^- \to K_S\pi^-$) on-peak data. 167 fb^{-1} of on-resonance data are used here. The Gaussian-like solid curve is the signal. The dashed curve is the total background. The solid curve overlaid with the data points is the total fit.

0.05 GeV, where there are no signal events. The i-th bin of the signal NN distribution for the on-resonance data can be expressed as,

$$h_{total}(i) - N_{bkg}/N_{lower} \cdot h_{lower}(i), \tag{77}$$

where h_{total} and h_{lower} are the histogram distributions for all events in the signal region and the lower sideband. N_{bkg} and N_{lower} are the estimated number of background events in the signal region and the total events in the lower sideband. Figure 50 shows the comparison of the NN distributions between the background-subtracted data and the MC for each decay mode. Very good agreement between the data and MC is seen.

A.4 Efficiency Comparison

With the same notation used in Eq. 77 for the on-resonance data, the signal efficiency with a NN cut value at x can be expressed as:

$$\left(\int_x^{1.0} h_{total} - \frac{N_{bkg}}{N_{lower}} \cdot \int_x^{1.0} h_{lower} \right) / N_{sig}, \tag{78}$$
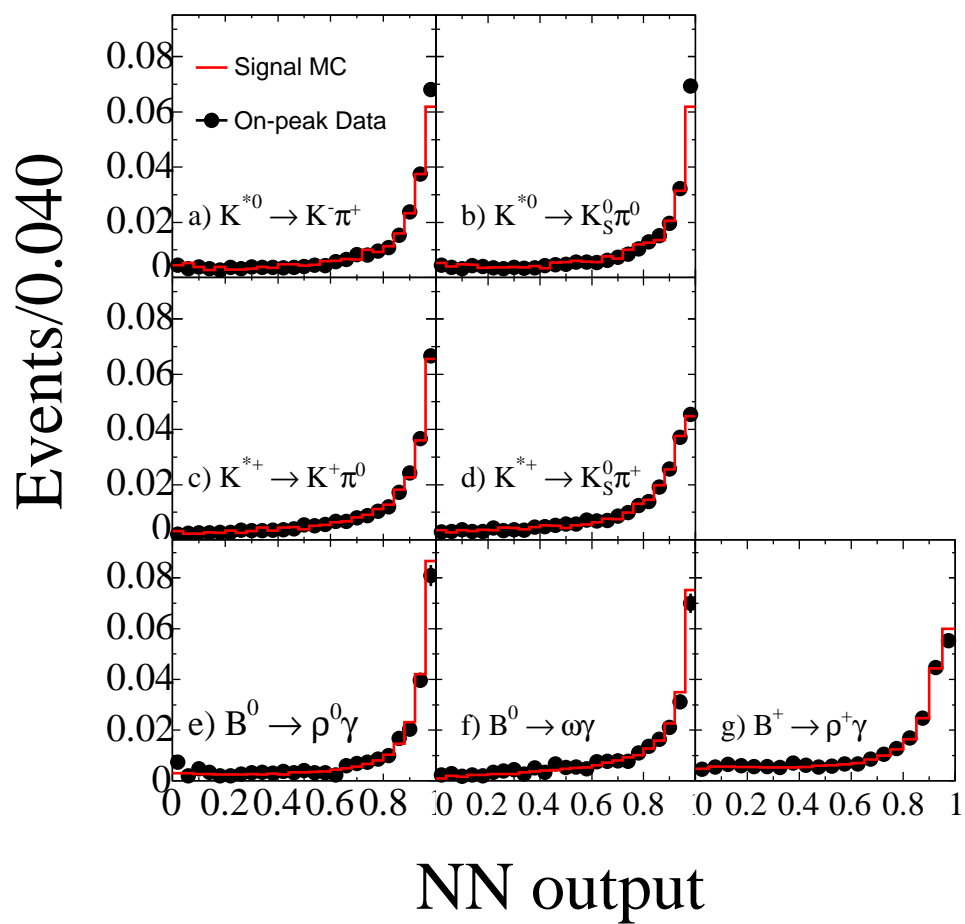
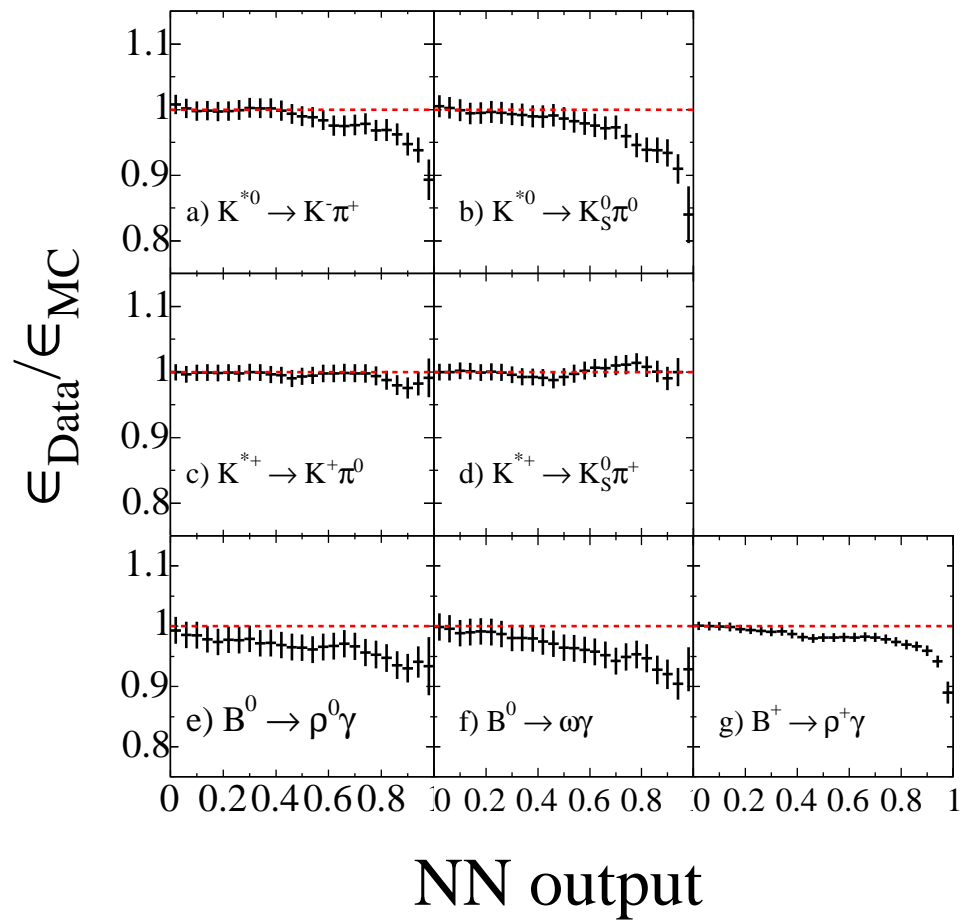Figure 50: Comparisons of the NN distributions between the data and MC.

Figure 51: Comparisons of the NN cut efficiencies between the data and the MC.

where N_{sig} is the total signal yield in the signal region. Figure 51 shows the comparisons of this efficiency between the data and the MC for each NN. In general, very good agreement is seen. Small efficiency discrepancies are seen when a cut is applied around one. This difference is understood to come from the small B-background in the signal region. With a tighter cut on the NN distributions, this B-background plays a more important role in the background subtraction.

A.5 Results

Based on the results in Figure 51, systematic errors of the NN are assigned as follows.

A.5.1 Cut Efficiency

For every decay mode, a loose cut is applied to the NN distribution to reject a large fraction of the continuum background. The cut value of each mode is shown in Table 3 and Table 4, typically around 0.6. As shown in Figure 51, the efficiency difference between the data and the MC around this cut value is almost negligible, so we decide not to correct the signal efficiency. To be conservative, the larger of the relative error of the efficiency ratio and the deviation of the efficiency from 1.0 is taken as the systematic error of the NN cut. These errors are relative to the signal efficiency and 3.0%, 3.5%, 2.7%, 2.8%, 4.6%, 4.6%, and 1.8% for $K^{*0} \rightarrow K^+\pi^-$, $K^{*0} \rightarrow K_s^0\pi^0$, $K^{*+} \rightarrow K^+\pi^0$, $K^{*+} \rightarrow K_s^0\pi^+$, $B^0 \rightarrow \rho^0\gamma$, $B^0 \rightarrow \omega\gamma$, and $B^+ \rightarrow \rho^+\gamma$, respectively. These errors are dominated by the statistics of the control samples.

A.5.2 NN Distributions in the Fits

For $B \rightarrow (\rho/\omega)\gamma$ decays, the shape of the neural net distribution is extracted from Monte Carlo and fixed in the likelihood fit. We scale the signal neural net shape by the data/Monte Carlo ratio in the control sample. Then, we fit the unblinded data set with the modified neural net shape. The shift of the signal yield is about 2.5%. The data over

Monte Carlo discrepancy in the control sample is understood to be an effect of the remaining B-background on the level of a few percent. Furthermore, the bias on the signal yield due to the NN shape is also correlated with the uncertainty caused when applying a loose cut. Therefore, we decided to ignore the effect from the signal neural net shape. A crosscheck is performed by using the $B \to (K^{*0}/K^{*+})\gamma$ control samples. We extracted the signal neural net shape from the $B \to (K^{*0}/K^{*+})\gamma$ control sample (see Section 5.6.3) and fit this shape to the pure toy Monte Carlos used in Section 5.6. The average bias on the signal yield is 0.3 events in $B^0 \to \rho^0\gamma$ decay mode, while we expected 17.1 events. This further constrains the effect of NN shape on the signal yield to be very tiny.

B   Individual Functions to Compose the PDFs

Table 11 summarizes all the function used to compose the component PDFs, $\mathcal{P}(\vec{x_j}; \vec{\alpha_i})$, for the $B \to (K^{*0}/K^{*+})\gamma$ decays and $B \to (\rho/\omega)\gamma$ decays.

Var.	Components	$K^{*0} \to K^+\pi^-$	$K^{*0} \to K_s^0\pi^0$	$K^{*+} \to K^+\pi^0$	$K^{*+} \to K_s^0\pi^+$	$B^0 \to \rho^0\gamma$	$B^0 \to \omega\gamma$	$B^+ \to \rho^+\gamma$
m_{ES}	Sig.	GAUSS	GAUSS	CB	GAUSS	CB	CB	CB
	$K^*\gamma$ bkg.	-	-	-	-	CB	-	CB
	$\rho\pi^0(\eta)$ bkg.	-	-	-	-	CB	CB	CB
	cont.bkg.	ARGUS	ARGUS	ARGUS	ARGUS	ARGUS	ARGUS	ARGUS
	comb.B.bkg.	NOVO	NOVO	NOVO	NOVO	KEYS	-	KEYS
ΔE^*	Sig.	CB	CB	CB	CB	CB	CB	CB
	$K^*\gamma$ bkg.	-	-	-	-	CB	-	CB
	$\rho\pi^0(\eta)$ bkg.	-	-	-	-	CB	CB	CB
	cont.bkg.	1st Poly.	1st Poly.	1st Poly.	2nd Poly.	2nd Poly.	2nd Poly.	2nd Poly.
	comb.B.bkg.	EXP	GAUSS	GAUSS	GAUSS	KEYS	-	KEYS
$\mathcal{N}$	Sig.	-	-	-	-	CB	KEYS	CB
	$K^*\gamma$ bkg.	-	-	-	-	sig. NN	-	sig. NN
	$\rho\pi^0(\eta)$ bkg.	-	-	-	-	sig. NN	sig. NN	sig. NN
	cont.bkg.	-	-	-	-	1D HIST	1D HIST	1D HIST
	comb.B.bkg.	-	-	-	-	sig. NN	-	sig. NN
$\mathcal{F}$	Sig.	-	-	-	-	KEYS	KEYS	KEYS
	$K^*\gamma$ bkg.	-	-	-	-	KEYS	-	KEYS
	$\rho\pi^0(\eta)$ bkg.	-	-	-	-	KEYS	KEYS	KEYS
	cont.bkg.	-	-	-	-	KEYS	KEYS	KEYS
	comb.B.bkg.	-	-	-	-	KEYS	-	KEYS

Table 11: Summary of individual PDFs used to compose the component PDFs. Here $\mathcal{N}$ is the NN output variable, $\mathcal{F}$ is the Fisher variable, "GAUSS" donates the Gaussian function, "CB" donates the Crystal-Ball function [52], "1D HIST" donates a 1-dimensional histogram PDF, "NOVO" donates the Novosibirsk function [68], "EXP" donates the exponential function, "i^{th} Poly" donates i-th Polynomial function, "ARGUS" donates the ARGUS function [53], and "KEYS" PDF is described in Ref. [69].

C   The Combining Method in $B \to (K^{*0}/K^{*+})\gamma$ Analysis

A χ^2 method, described in *Probability and Statistics in Particle Physics* [81], is used for combining two measurements with partially correlated systematic errors. The method described in the book is for general use, but a special case of it, where all measurements are of a single quantity, is applied here.

The method starts with a vector of n measurements,

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}. \tag{79}$$

The fitted parameters are kept in a vector called $\vec{\theta}$, which may have many components, but in our case has only one since we are fitting for a single branching fraction. The "predictions", $\vec{f}$ for the $\vec{y}$s are related to the fitted parameters by the A matrix

$$\vec{f} = A\vec{\theta}. \tag{80}$$

In our case, the A matrix is a column vector of 1's.

Errors and correlations of measurements are described by a symmetric covariance matrix:

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdot & \cdot & \sigma_{1n}^2 \\ \sigma_{12}^2 & \sigma_2^2 & \cdot & \cdot & \sigma_{2n}^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{1n}^2 & \sigma_{2n}^2 & \cdot & \cdot & \sigma_n^2 \end{pmatrix}, \tag{81}$$

where σ_i is the total (statistical and systematic errors added in quadrature) error of a measurement in mode i, and σ_{ij} is the part of the systematic error that is correlated between modes i and j.

A χ^2 is then defined as

$$\chi^2 \equiv (\vec{y} - A\theta)^T V^{-1} (\vec{y} - A\theta), \tag{82}$$

where θ is the combined branching fraction to be calculated. If we minimize this χ^2 with respect to θ, we find

$$\theta = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y}. \tag{83}$$

The vector that is dotted with the $\vec{y}$ is the "weights" that are assigned to each measurement. We define the weight vector, $\bar{w}$ as: $\bar{w} = (A^T V^{-1} A)^{-1} A^T V^{-1}$. The total error on the combined branching ratio is given by:

$$\sigma_{\bar{y}}^2 = (A^T V^{-1} A)^{-1}. \tag{84}$$

In order to define off-diagonal elements of the V matrix, we decide which systematic error are common between modes. The common errors, such as number of B candidates, the efficiencies of the charged tracks, neutral candidates, and K_s^0's, for each mode-pair are added in quadrature to get the corresponding off-diagonal element. One somewhat tricky point is what to do for a common systematic that contributes different amounts to two modes. For example, the tracking systematic contributes 1.6% to the $K^{*0} \to K^+\pi^-$ mode and only 0.8% to the $K^{*+} \to K^+\pi^0$ mode. The correct procedure is to take $\sqrt{0.8 \times 1.6}\%$ as the correlated error. Following these calculations, the final covariance matrix is shown in Eq. 85,

$$
\begin{array}{cccc}
 & K^{*0} \to K^+\pi^- & K^{*0} \to K_s^0\pi^0 & K^{*+} \to K^+\pi^0 & K^{*+} \to K_s^0\pi^+
\end{array}
$$

$$
\begin{array}{c}
K^{*0} \to K^+\pi^- \\
K^{*0} \to K_s^0\pi^0 \\
K^{*+} \to K^+\pi^0 \\
K^{*+} \to K_s^0\pi^+
\end{array}
\left(
\begin{array}{cccc}
8.71 \times 10^{-12} & 4.21 \times 10^{-12} & 5.28 \times 10^{-12} & 3.67 \times 10^{-12} \\
. & 12.37 \times 10^{-11} & 1.35 \times 10^{-11} & 5.07 \times 10^{-12} \\
. & . & 4.09 \times 10^{-11} & 4.73 \times 10^{-12} \\
. & . & . & 1.65 \times 10^{-11}
\end{array}
\right). \tag{85}
$$

D Systematic Error of the Combined $B \to (\rho/\omega)\gamma$ Branching Fraction

The sources of systematic uncertainties of the combined "effective signal yield" is being categorized as:

1. The effect from the individual fit systematic errors, summarized in Table 7.

2. The effect from the uncertainties of the signal efficiencies, shown in Table 6.

For the individual fit systematic errors, we vary each individual signal yield by one σ separately and take the bias on the combined effective signal yield as the systematic uncertainty. However, for the systematic uncertainties from the signal efficiencies, the situation is complicated by the fact that some of the systematic uncertainties, shown in Table 6, are correlated among modes. We explicitly break the systematic uncertainties on the signal efficiencies into two categories:

- The uncertainties correlated among decay modes, which include the errors from B counting, tracking efficiency, PID systematic, π/γ efficiency, π^0/η veto, and photon distance cut.

- The uncertainties uncorrelated among decay modes, which include the neural net systematic uncertainties and the "MC statistics/fit bias/B-background modeling".

For the correlated uncertainties of the signal efficiencies, we vary the signal efficiency on the same side for each decay mode simultaneously; the shift on the signal yield is taken as the systematic uncertainty due to the correlated part of signal efficiencies. For the uncorrelated part, we vary the signal efficiencies separately; the bias on the combined effective signal yield is taken as the corresponding systematic uncertainty due to the uncorrelated uncertainties of the efficiencies.

Description	$B^0 \to \rho^0\gamma$ mode	$B^+ \to \rho^+\gamma$ mode	$B^0 \to \omega\gamma$ mode	error (events)
Corr. eff. var.	0.151-0.165	0.123-0.141	0.080-0.092	$^{+15.6}_{-14.3}$
Uncorr. eff. var. $\{$	0.140-0.176	-	-	$^{+17.2}_{-16.7}$
	-	0.121-0.143	-	$^{+4.6}_{-5.5}$
	-	-	0.080-0.092	$^{+4.0}_{-4.6}$
signal bias $\{$	$^{+1.7}_{-1.6}$	-	-	$^{+10.9}_{-9.8}$
	-	$^{+2.2}_{-2.2}$	-	$^{+12.9}_{-12.6}$
	-	-	$^{+1.3}_{-1.9}$	$^{+27.5}_{-34.6}$
Total				$^{+40.2}_{-44.5}$

Table 12: Summary of all the systematic uncertainties for the combined effective signal yield.

Following the approach suggested above, we obtained the systematic uncertainty for each case and the results are summarized in Table 12. The total systematic uncertainty for the combined "effective yield" is $^{+40.2}_{-44.5}$.

References

[1] C.N. Yang, *Elementary Particles* (Princeton University Press, Princeton, N.J., 1961).

[2] L.M. Lederman, Am. J. Phys. **38**, 129 (1970).

[3] S. Weinberg, *The Quantum Theory of Fields* (Press Syndicate of the University of Cambridge, New York, 1995).

[4] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[5] A. Salam, in *Elementary Particle Theory*, edited by N. Svartholm (Almquist and Wiksells, Stockholm, 1969).

[6] S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[7] P. Ramond, *Field Theory: A Modern Primer* (Addison-Wesley Publishing Company, 1990).

[8] D. Boutigny *et al.* [*BABAR* Collaboration], SLAC-R-504, 1998.

[9] M. Chaichian and N.F. Nelipa, *Introduction to Gauge Field Theories* (Springer-Verlag, New York, 1984).

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

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

[12] G. Savard *et al.*, Phys. Rev. Lett. **74**, 1521 (1995).

[13] G.J. Gilman, K. Kleinknecht, and B. Renk, Phys. Rev. D **66**, 113 (2002).

[14] R.N. Mohapatra, *Unification and Supersymmetry* (Springer-Verlag, New York, 2003).

[15] M. Drees, hep-ph/9611409.

[16] I.J.R. Aitchison, hep-ph/0505105.

[17] Y. Okada, M. Yamaguchi, and T. Yanagida, Prog. Theor. Phys. **85**, 1 (1991).

[18] R. Barate *et al.* [ALEPH Collaboration], Phys. Lett. B **440**, 419 (1998).

[19] V.D. Barger, M.S. Berger, and R.J.N. Phillips, Phys. Rev. Lett. **70**, 1368 (1993).

[20] R. Flores, K.A. Olive, and D. Thomas, Phys. Lett. B **245**, 509 (1990).

[21] J.H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).

[22] B. Aubert *et al.* [*BABAR* Collaboration], Nucl. Instr. Meth. A **479**, 1 (2002).

[23] A. Abashian *et al.* [Belle Collaboration], Nucl. Instr. Meth. A **479**, 117 (2002).

[24] B. Aubert *et al.* [*BABAR* Collaboration], hep-ex/0408075.

[25] H.D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973).

[26] K.G. Wilson, Phys. Rev. **179**, 1499 (1969).

[27] S.W. Bosch and G. Buchalla, Nucl. Phys. B **621**, 459 (2002).

[28] G. Buchalla, A.J. Buras, and M.E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996).

[29] P. Ball and V.M. Braun, Phys. Rev. D **58**, 094016 (1998).

[30] T.E. Coan *et al.* [CLEO Collaboration], Phys. Rev. Lett. **84**, 5283 (2000).

[31] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0239, 2003.

[32] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **88**, 101805 (2002).

[33] H. Cheng and C.K. Chua, hep-ph/0401141.

[34] D. Becirevic, hep-ph/0211340.

[35] A. Ali, E. Lunghi, and A.Y. Parkhomenko, hep-ph/0405075.

[36] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **92**, 111801 (2002).

[37] S.W. Bosch and G. Buchalla, hep-ph/0408231.

[38] S. Eidelman *et al.*, Phys. Lett. B **592**, 1 (2004).

[39] Heavy Flavor Average Group, hep-ex/0412073.

[40] A. Ali and A.Y. Parkhomenko, Eur. Phys. Jour. C **23**, 89 (2002).

[41] T. Hurth, E. Lunghi, and W. Porod, hep-ex/0312260.

[42] A.L. Kagan and M. Neubert, Phys. Lett. B **539**, 227 (2002).

[43] A.J. Buras *et al.*, Nucl. Phys. B **347**, 491 (1990).

[44] H. Wittig, hep-ph/0310329.

[45] S. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977).

[46] L. Lista *et al.* [*BABAR* Collaboration], Nucl. Phys. Proc. Suppl. **66**, 541 (1998).

[47] P. Billoir, Nucl. Instr. Meth. A **225**, 225 (1984).

[48] D.N. Brown, E.A. Charles, and D.A. Roberts, *the BABAR Track Fitting Algorithm*, Proceedings of the International Conference on Computing in High Energy and Nuclear Physics, Padova, Italy, edited by M. Mazzucato and M. Michelotto (INFN, Padua, Italy, 2001).

[49] T. Sjöstrand, PYTHIA 5.7 and JETSET 7.4 Physics and Manual, CERN-TH 7112, 1993.

[50] A. Ryd *et al.*, *BABAR* Analysis Document (BAD) 522 (unpublished).

[51] D.N. Brown *et al.* [*BABAR* Collaboration], hep-ex/0305085.

[52] T. Skwarnicki [Crystall Ball Collaboration], DESY F31-86-02.

[53] H. Albrecht *et al.* [ARGUS Collaboration], Phys. Lett. B **185**, 218 (1987).

[54] D.S. Best, C.P. Jessop, and H. Tanaka, *BABAR* Analysis Document (BAD) 201 (unpublished).

[55] G. Mancinelli and S.M. Spanier, *BABAR* Analysis Document (BAD) 116 (unpublished).

[56] Vertexing and Composition Tools Group, *BABAR* Analysis Document (BAD) 102 (unpublished).

[57] B. Aubert *et al.*, *BABAR* Analysis Document (BAD) 720 (unpublished). Presented at 39th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions, La Thuile, Italy, 2004.

[58] B. Aubert *et al.*, *BABAR* Analysis Document (BAD) 700 (submitted to Phys. Rev. Lett.).

[59] B. Aubert *et al.*, *BABAR* Analysis Document (BAD) 705 (unpublished). Presented at 39th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions, La Thuile, Italy, 2004.

[60] J. Dragic *et al.* [Belle Collaboration], hep-ex/0405068.

[61] R. Aleksan*et al.*, *BABAR* Analysis Document (BAD) 705 (unpublished).

[62] J. Beringer, *BABAR* Analysis Document (BAD) 317 (unpublished).

[63] P.C. Bhat, *Multivariate Methods in High Energy Physics: The Neural Network Revolution* (World Scientific, 2000).

[64] M.R. Convery *et al.*, *BABAR* Analysis Document (BAD) 354 (unpublished).

[65] A. Zell *et al.*, SNNS (Stuttgart Neural Network Simulator) User Manual, Version 4.2 (unpublished).

[66] R.A. Fisher, Annals. Eugenics **7**, 179 (1936).

[67] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification* (Wiley, New York, 2001).

[68] D.S. Best and C.P. Jessop, *BABAR* Analysis Document (BAD) 204 (unpublished).

[69] K.S. Cranmer, hep-ex/0011057.

[70] C. Hearty, *BABAR* Analysis Document (BAD) 134 (unpublished).

[71] B. Aubert *et al.* [*BABAR* Collaboration], hep-ex/0401028.

[72] M. Bona *et al.*, *BABAR* Analysis Document (BAD) 324 (unpublished).

[73] T. Allmendinger *et al.*, *BABAR* Analysis Document (BAD) 867 (unpublished).

[74] A. Anjomshoaa *et al.*, *BABAR* Analysis Document (BAD) 20 (unpublished).

[75] A.V. Gritsan, *BABAR* Analysis Document (BAD) 677 (unpublished).

[76] M.T. Allen *et al.*, *BABAR* Analysis Document (BAD) 870 (unpublished).

[77] T. Colberg *et al.*, *BABAR* Analysis Document (BAD) 33 (unpublished).

[78] Z. Xiao and C. Zhuang, hep-ph/0310097.

[79] K. Abe *et al.* [Belle Collaboration], hep-ex/0408138.

[80] B. Grinstein and D. Pirjol, Phys. Rev. D **62**, 093002 (2000).

[81] A.G. Frodesen, O. Skjeggestad, and H. Tøfte, *Probability and Statistics in Particle Physics* (Universitetsforlaget, Norway, 1979).

Constraining $|V_{td}|/|V_{ts}|$ using Radiative Penguin $B \to V(K^*/\rho/\omega)\gamma$ Decays

Ping Tan

Under the supervision of Assistant Professor Sridhara Dasu

At the University of Wisconsin — Madison

Exclusive radiative penguin B decays, $B \to (K^{*0}/K^{*+})\gamma$ and $B \to (\rho/\omega)\gamma$, are flavor-changing neutral-current (FCNC) processes. Studies of these decays are of special interest in testing Standard Model (SM) predictions and searching for other beyond-the-SM FCNC interactions. Using 89×10^6 $B\overline{B}$ pairs from $BABAR$, we measure the branching fraction ($\mathcal{B}$), CP-asymmetry ($\mathcal{A}$), and isospin asymmetry (Δ_{0-}) of $B \to (K^{*0}/K^{*+})\gamma$ as follows:

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = 3.92 \pm 0.20(stat.) \pm 0.24(syst.),$$
$$\mathcal{B}(B^+ \to K^{*+}\gamma) = 3.87 \pm 0.28(stat.) \pm 0.26(syst.),$$
$$\mathcal{A}(B \to K^*\gamma) = -0.013 \pm 0.36(stat.) \pm 0.10(syst.),$$
$$\Delta_{0-}(B \to K^*\gamma) = 0.050 \pm 0.045(stat.) \pm 0.028(syst.) \pm 0.024(R^{+/0}).$$

The 90% confidence intervals for the CP-asymmetry and the isospin-asymmetry in the $B \to K^*\gamma$ decay are given as:

$$-0.074 < \mathcal{A}(B \to K^*\gamma) < 0.049,$$
$$-0.046 < \Delta_{0-}(B \to K^*\gamma) < 0.146.$$

We also search for $B \to (\rho/\omega)\gamma$ decays using 211×10^6 $B\overline{B}$ pairs from $BABAR$. No evidence for these decays is found. We set the upper limits at 90% confidence level for these decays:

$$\mathcal{B}(B^0 \to \rho^0\gamma) < 0.4 \times 10^{-6},$$
$$\mathcal{B}(B^+ \to \rho^+\gamma) < 1.8 \times 10^{-6},$$
$$\mathcal{B}(B^0 \to \omega\gamma) < 1.0 \times 10^{-6},$$
$$\overline{\mathcal{B}}(B \to (\rho/\omega)\gamma) < 1.2 \times 10^{-6}.$$

These results are in good agreement with the SM predictions. The branching fractions of these decays are then used to constrain the ratio $|V_{td}|/|V_{ts}|$.

Sridhara Dasu (Adviser)