

MEASUREMENTS OF  $B \rightarrow V\gamma$  DECAYS

A DISSERTATION  
SUBMITTED TO THE DEPARTMENT OF PHYSICS  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Aaron K. Yarritu  
August 2010

# Abstract

The standard model has been highly successful at describing current experimental data. However, extensions of the standard model predict particles that have masses at energy scales that are above the electroweak scale. The flavor-changing neutral current processes of the  $B$  meson are sensitive to the influences of these new physics contributions. These processes proceed through loop diagrams, thus allowing new physics to enter at the same order as the standard model. New physics may contribute to the enhancement or suppression of rate asymmetries or the decay rates of these processes.

The transition  $B \rightarrow V\gamma$  ( $V = K^*(892), \rho(770), \omega(782), \phi(1020)$ ) represents radiative decays of the  $B$  meson that proceed through penguin processes. Hadronic uncertainties limit the theoretical accuracy of the prediction of the branching fractions. However, uncertainties, both theoretical and experimental, are much reduced when considering quantities involving ratios of branching fractions, such as  $CP$  or isospin asymmetries. The most dominant exclusive radiative  $b \rightarrow s$  transition is  $B \rightarrow K^*\gamma$ . We present the best measurements of the branching fractions, direct  $CP$ , and isospin asymmetries of  $B \rightarrow K^*\gamma$ . The analogous  $b \rightarrow d$  transitions are  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$ , which are suppressed by a factor of  $|V_{td}/V_{ts}|^2 \sim 0.04$  relative to  $B \rightarrow K^*\gamma$ . A measurement of the branching fractions and isospin asymmetry of  $B^+ \rightarrow \rho^+\gamma$  and  $B^0 \rightarrow \rho^0\gamma$ , as well as a search for  $B \rightarrow \omega\gamma$ , are also given. These measurements are combined to calculate the ratio of CKM matrix elements  $|V_{td}/V_{ts}|$ , which corresponds to the length of one side of the unitary triangle. Finally, we present a search for the penguin annihilation process  $B \rightarrow \phi\gamma$ .

We use a sample of 383 million  $B\bar{B}$  events collected with the *BABAR* detector at the PEP-II asymmetric-energy  $B$  factory for the analysis of  $B \rightarrow K^*\gamma$ . We measure the branching fractions  $\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = (4.47 \pm 0.10 \pm 0.16) \times 10^{-5}$  and  $\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = (4.22 \pm 0.14 \pm 0.16) \times 10^{-5}$ . We measure the direct  $CP$  asymmetry to be  $-0.033 < \mathcal{A}_{CP}(B \rightarrow K^*\gamma) < 0.028$  and the isospin asymmetry to be  $0.017 < \Delta_{0-} < 0.116$ , where the limits are determined at the 90% C.L. and include both the statistical and systematic

uncertainties.

Using a sample of 347 million  $B\bar{B}$  events, we measure the branching fractions  $\mathcal{B}(B^+ \rightarrow \rho^+\gamma) = (1.10_{-0.33}^{+0.37} \pm 0.09) \times 10^{-6}$  and  $\mathcal{B}(B^0 \rightarrow \rho^0\gamma) = (0.79_{-0.20}^{+0.22} \pm 0.06) \times 10^{-6}$ , the isospin asymmetry  $\Delta = -0.35 \pm 0.27$ , and set a 90% C.L. upper limit  $\mathcal{B}(B \rightarrow \omega\gamma) < 0.78 \times 10^{-6}$ . We also measure the isospin-averaged branching fraction  $\mathcal{B}(B \rightarrow (\rho/\omega)\gamma) = (1.25_{-0.24}^{+0.25} \pm 0.09) \times 10^{-6}$ , from which we determine  $|V_{td}/V_{ts}| = 0.200_{-0.020}^{+0.021} \pm 0.015$ , where the first uncertainty is experimental and the second theoretical.

Finally, a sample of 124 million  $B\bar{B}$  events is used to set an upper limit of  $\mathcal{B}(B \rightarrow \phi\gamma) < 8.5 \times 10^{-7}$  at the 90% C.L.

# Acknowledgments

This thesis, like any, is the culmination of so many different people that it would be impossible to properly acknowledge everyone. The *BABAR* collaboration is comprised of numerous physicists who are spread around the world. There are many people who have not contributed to this measurement, but also helped me develop as a graduate student.

First, I would like to thank my parents, Mom and Dad, for providing me with the means and support for me to pursue my goals. Their love, support, and patience are unconditional. My advisor, David Leith, and Blair Ratcliff, have welcomed me in the group and provided guidance and helpful suggestions to point me in the right direction. Mark Convery and I have worked closely now for many years, and he has been invaluable for my education and learning experience. Jim Libby has also been responsible for me in this regard. Jaroslav Va'vra has guided me through the experimental phase of my thesis. Jochen Schwiening and Dave Aston have given their time to whatever computing question I have had, however small. Jose Benitez, my officemate, is always there to discuss my analysis, his analysis, or, sometimes more importantly, whatever concerns are happening in our lives. Thanks must also go out the rest of group B, Bill Dunwoodie, who is always ready to help the graduate students, and provide advice, Clive Field, Dave Muller, Jon Coleman, Kazuhito Suzuki, Veronique Ziegler, Adrienne Higashi. Lilian DePorcel and Giampiero Mancinelli have since left the group.

The Radiative Penguin AWG, and  $K^*\gamma$  group have contributed heavily to this thesis, both technically and with advice. The people involved are Kevin Flood, John Walsh, Jurgen Kroseberg, Gabriella Sciolla, Al Eisner, Gabriele Simi, Steve Sekula, Wouter Hulsbergen, Gavin Nesom, Jeff Berryhill, Tim Piatenko, Karsten Koeneke, Fang Fang, Ping Tang, Patrick Spradlin, Joe Tuggle, Joel Martinez, Debbie Bard, Kyle Knoepfel, Chris Schilling, Hojeong Kim, Jan Strube, and many others.

There are graduate students who have contributed to my life here at Stanford, who are Nicolas Berger, Tetiana Hryn'ova, Joerg Stelzer, Erkcan Özcan, Mark Tiller, Josh Thompson, Stephanie Majewski, and Wells Wulsin to name a few.

Thanks also go to my review committee members of Pat Burchat, Aaron Roodman, and JoAnne Hewett for taking time to read, and provide advice upon, this thesis.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgments</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Outline . . . . .	5
<b>2 PEP-II and the <i>BABAR</i> Detector</b>	<b>6</b>
2.1 The PEP-II Accelerator . . . . .	6
2.2 The <i>BABAR</i> Detector . . . . .	7
2.2.1 Silicon Vertex Tracker (SVT) . . . . .	9
2.2.2 Drift Chamber (DCH) . . . . .	12
2.2.3 Detector of Internally Reflected Cherenkov Light (DIRC) . . . . .	16
2.2.4 Electromagnetic Calorimeter (EMC) . . . . .	19
2.2.5 Instrumented Flux Return (IFR) . . . . .	22
2.2.6 Trigger . . . . .	24
<b>3 Photon Background in DIRC Fused Silica Bars</b>	<b>26</b>
3.1 Introduction . . . . .	26
3.2 Experimental Setup . . . . .	27
3.3 Acquired Data . . . . .	28
3.4 Supplemental Measurements . . . . .	32
3.4.1 Measurement of Reflectivity of EPOTEK-301-2 Glue . . . . .	32
3.4.2 Measurement of Scintillation . . . . .	32
3.5 Monte Carlo Simulation . . . . .	34
3.5.1 Delta Ray Simulation . . . . .	34
3.5.2 Reflectivity of Glue . . . . .	36

3.6	Conclusion . . . . .	37
<b>4</b>	<b>Theory</b>	<b>39</b>
4.1	Standard Model Lagrangian . . . . .	39
4.1.1	Fermion-Gauge Boson Interaction Lagrangian . . . . .	40
4.2	The CKM Matrix . . . . .	41
4.3	$B \rightarrow V\gamma$ Phenomenology . . . . .	43
4.3.1	Decay Dynamics . . . . .	44
4.3.2	Form Factors . . . . .	47
4.4	Measured Observables . . . . .	48
4.4.1	Branching Fractions . . . . .	48
4.4.2	CP Asymmetries . . . . .	49
4.4.3	Isospin Asymmetries . . . . .	52
4.4.4	$ V_{td}/V_{ts} $ . . . . .	53
<b>5</b>	<b>Event Selection, Optimization and Background Suppression</b>	<b>56</b>
5.1	Introduction . . . . .	56
5.1.1	Data . . . . .	57
5.1.2	Overview of Analysis Method . . . . .	61
5.2	Preliminary Event Selection . . . . .	62
5.2.1	Kinematic Variables . . . . .	64
5.3	Candidate Selection . . . . .	66
5.3.1	Particle Identification . . . . .	66
5.3.2	$K_S$ Selection . . . . .	67
5.3.3	$\pi^0$ Selection . . . . .	69
5.3.4	Vector Meson Selection . . . . .	70
5.3.5	High-Energy Photon Selection . . . . .	71
5.4	Continuum Background Suppression . . . . .	73
5.4.1	Input Variables . . . . .	74
5.4.2	Neural Network Output . . . . .	79
5.5	Cut Optimization . . . . .	85
5.5.1	Results of cut optimization . . . . .	88
<b>6</b>	<b>Measurement of <math>B \rightarrow K^*\gamma</math> Branching Fractions and Asymmetries</b>	<b>92</b>
6.1	B background . . . . .	92
6.2	Signal and Background Efficiencies . . . . .	96

6.3	Fit Parameterization . . . . .	97
6.3.1	PDF Line Shapes . . . . .	97
6.3.2	Correlation of Fit Observables . . . . .	99
6.3.3	Maximum Likelihood Fit . . . . .	105
6.4	Branching Fraction Fitting . . . . .	108
6.4.1	Toy Studies . . . . .	108
6.5	Branching Fraction Systematic Errors . . . . .	113
6.5.1	$B\bar{B}$ Production . . . . .	115
6.5.2	B Counting Systematic . . . . .	115
6.5.3	Tracking Efficiency . . . . .	115
6.5.4	Particle Identification Efficiency . . . . .	116
6.5.5	$K_S$ Efficiency . . . . .	116
6.5.6	Single Photon Efficiency . . . . .	116
6.5.7	$\pi^0$ Efficiency . . . . .	117
6.5.8	Photon Quality Selection . . . . .	117
6.5.9	$\pi^0/\eta$ Veto . . . . .	118
6.5.10	Neural Net systematic . . . . .	118
6.5.11	Fitting Systematic . . . . .	125
6.6	$CP$ Asymmetry . . . . .	128
6.6.1	$CP$ Asymmetry Systematic Errors . . . . .	128
6.7	Results from the Maximum Likelihood Fit . . . . .	131
6.7.1	Branching Fractions . . . . .	132
6.7.2	$CP$ Asymmetry . . . . .	139
6.8	Combined Results . . . . .	141
6.8.1	Covariance Matrix . . . . .	141
6.8.2	Branching Fractions . . . . .	141
6.8.3	Isospin Asymmetry . . . . .	143
6.8.4	$CP$ Asymmetry . . . . .	144
<b>7</b>	<b>Measurements of <math>B^+ \rightarrow \rho^+\gamma</math>, <math>B^0 \rightarrow \rho^0\gamma</math>, and <math>B \rightarrow \omega\gamma</math> Decays</b>	<b>145</b>
7.1	Fit Overview . . . . .	145
7.1.1	$B^0 \rightarrow \rho^0\gamma$ . . . . .	147
7.1.2	$B^+ \rightarrow \rho^+\gamma$ . . . . .	147
7.1.3	$B \rightarrow \omega\gamma$ . . . . .	150
7.2	Systematic Errors . . . . .	150

7.3	Results . . . . .	151
7.3.1	Branching Fractions . . . . .	151
7.3.2	Isospin Symmetry . . . . .	153
7.3.3	$ V_{td}/V_{ts} $ . . . . .	153
<b>8</b>	<b>Search for the Rare Decay <math>B \rightarrow \phi\gamma</math></b>	<b>156</b>
8.1	B background . . . . .	156
8.2	Fitting Procedure . . . . .	156
8.2.1	Monte Carlo Consistency . . . . .	158
8.2.2	Data Consistency and Background Estimation . . . . .	158
8.3	Systematic Errors . . . . .	159
8.4	Results . . . . .	160
<b>9</b>	<b>Conclusions</b>	<b>162</b>
9.1	$B \rightarrow K^*\gamma$ . . . . .	162
9.2	$B \rightarrow (\rho/\omega)\gamma$ . . . . .	164
9.3	$B \rightarrow \phi\gamma$ . . . . .	165
9.4	Future Prospects . . . . .	166
<b>A</b>	<b>OffPeak Data-Monte Carlo agreement</b>	<b>168</b>
A.1	$K^{*0} \rightarrow K^+\pi^-$ mode . . . . .	168
<b>B</b>	<b>Correlation of Fit Observables</b>	<b>173</b>
B.1	$K^{*0} \rightarrow K^+\pi^-$ mode . . . . .	173
<b>C</b>	<b>Neural Network Overview</b>	<b>177</b>
C.1	Neural Network Introduction . . . . .	177
C.1.1	Determination of Free Parameters . . . . .	177
C.1.2	Implementation of Neural Network . . . . .	178
<b>D</b>	<b>Signal Embedded Toy Studies</b>	<b>180</b>
D.1	$K^{*0} \rightarrow K^+\pi^-$ , $K^{*+} \rightarrow K_S\pi^+$ modes . . . . .	180
D.2	$K^{*0} \rightarrow K_S\pi^0$ , $K^{*+} \rightarrow K^+\pi^0$ modes . . . . .	182
<b>E</b>	<b>Parameter Comparison</b>	<b>186</b>
	<b>Bibliography</b>	<b>188</b>

# List of Tables

2.1	The specifications of the superlayer structure of the DCH. Listed is the number of cells per layer, the radius of the innermost sense wire, cell widths and wire angles. Widths and radii are specified at the center of the chamber.	13
4.1	Results of form factor calculations using lattice QCD, QCD sum rules, and light-cone QCD sum rules. For the lattice QCD entry, the first error is statistical, while the second error is systematic.	48
4.2	Current theoretical calculations of the $B \rightarrow K^*\gamma$ branching fraction	49
4.3	Previous measurements of the branching ratios and asymmetries for $B \rightarrow K^*\gamma$ . The last error on the isospin asymmetry for the <i>BABAR</i> measurement refers to the error on the production ratio of charged to neutral B events.	49
4.4	Theoretical predictions of the branching fractions for $B \rightarrow \rho\gamma$ and $B \rightarrow \omega\gamma$ .	50
4.5	Previous measurements of the branching ratios for $B \rightarrow \rho\gamma$ and $B \rightarrow \omega\gamma$ , where the limits are given by the 90% confidence level. The errors on the <i>Belle</i> measurements are statistical and systematic respectively, while the entry in paranthesis refers to the significance of the measurement.	51
4.6	Current theoretical calculations of isospin violation in $B \rightarrow K^*\gamma$ decays. $F^{B \rightarrow V}(q^2 = 0)$ is the $B \rightarrow K^*$ form factor whose estimates range from $0.23 \pm 0.06$ [27] to $0.38 \pm 0.06$ [76].	53
4.7	Previous measurements of $ V_{td}/V_{ts} $ . In the CDF entry, $\Delta m_d$ and $\Delta m_s$ denote the $B_d$ and $B_s$ mixing frequencies respectively. “Theor.” stands for the theoretical uncertainty.	55
5.1	Production cross-sections at a CM energy of 10.58 GeV [46]. The cross-section listed for $e^+e^-$ production includes the effect of limited detector acceptance.	57
5.2	Luminosity recorded over the history the <i>BABAR</i> detector.	58
5.3	Luminosity of dataset used by each analysis.	58

5.4	The isospin factors for the decay of the $K^*$ . . . . .	61
5.5	Summary of optimized selection criteria for $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$ . . .	89
5.6	Summary of optimized selection criteria for $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K_S\pi^0)$ . . .	89
5.7	Summary of optimized selection criteria for $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ . . .	90
5.8	Summary of optimized selection criteria for $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K_S\pi^+)$ . . .	90
5.9	Summary of optimized selection criteria for $B^0 \rightarrow \rho^0\gamma$ . . . . .	91
5.10	Summary of optimized selection criteria for $B^+ \rightarrow \rho^+\gamma$ . . . . .	91
5.11	Summary of optimized selection criteria for $B \rightarrow \omega\gamma$ . . . . .	91
6.1	Number of Monte Carlo events in the fit region scaled to $347 \text{ fb}^{-1}$ using the MC equivalent luminosities. The $B$ background is the sum of the three components described in section ???. The off-peak data is shown for comparison. . . . .	97
6.2	a) Signal MC simulation, b) Continuum MC simulation, and c) $B$ background MC simulation correlation coefficients for the $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$ mode. . . . .	99
6.3	a) Signal, b) Continuum, and c) $B$ background correlation coefficients for the $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K_S\pi^0)$ mode. . . . .	104
6.4	a) Signal, b) Continuum, and c) $B$ background correlation coefficients for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode. . . . .	104
6.5	a) Signal, b) Continuum, and c) $B$ background correlation coefficients for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K_S\pi^+)$ mode. . . . .	104
6.6	PDF parameters for the $K^{*0} \rightarrow K^+\pi^-$ and $K^{*+} \rightarrow K_S\pi^+$ modes. . . . .	106
6.7	PDF parameters for all $K^{*0} \rightarrow K_S\pi^0$ and $K^{*+} \rightarrow K^+\pi^0$ modes. . . . .	107
6.8	The pulls of the yields for each component for nominal fits (parameters behaving as in tables ?? and ??). All fits are signal embedded. . . . .	110
6.9	The major $B\bar{B}$ backgrounds of the $B \rightarrow K^*\gamma$ modes. Shown in parentheses is the percentage of the total $B\bar{B}$ background. . . . .	113
6.10	The pulls of the yields for each component for nominal fits (parameters behaving as in table ?? and ??). All fits are $B\bar{B}$ embedded. . . . .	113
6.11	The correction factor for signal efficiency for each decay mode. Included are the signal and $B\bar{B}$ embedded toy studies. . . . .	114

6.12	This table describes the factors needed to correct the Monte Carlo efficiency and their systematic errors. Shown in paranthesis is the section number that provides a description of the associated systematic. The final estimated efficiency is also given. . . . .	114
6.13	Signal efficiency and expected yields in the on-peak data set . . . . .	120
6.14	Number of events in the lower sideband and the signal region, the fitted signal yields, and the estimated background in the signal region in the on-peak dataset. . . . .	121
6.15	Signal yields in which the $b \rightarrow s\gamma$ yield has been varied by $\pm 20\%$ . The signal yield has been extracted with the nominal fit using the $B\bar{B}$ embedded toys. . . . .	126
6.16	Signal yields in which the $b \rightarrow s\gamma$ yield has been varied by $\pm 20\%$ . The signal yield has been extracted with the nominal fit using pure toys. However, the pure toys were generated by varying the $b \rightarrow s\gamma$ yield. . . . .	126
6.17	Values of $\mathcal{A}_{CP}$ measured in onpeak $m_{ES}$ sideband data. . . . .	128
6.18	Material model used for hadronic interaction asymmetry calculation . . . .	130
6.19	Component yields for the $K^{*0} \rightarrow K^+\pi^-$ mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure ??.	133
6.20	Component yields for the $K^{*0} \rightarrow K_S\pi^0$ mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure ??.	134
6.21	Component yields for the $K^{*+} \rightarrow K^+\pi^0$ mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure ??.	135
6.22	Component yields for the $K^{*+} \rightarrow K_S\pi^+$ mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure ??.	136
6.23	The signal yields from the fit to the on-peak data, signal efficiencies and total systematic uncertainty, isospin factors, and branching fractions. . . .	137
6.24	The fitted results of the $m_{K\pi}$ spectrum are compared with PDG [4] values for all four decay modes. . . . .	139
6.25	The individual $\mathcal{A}_{CP}$ values from the fit to the on-peak data. . . . .	140
6.26	The correlated systematic uncertainties of the four modes. ‘‘Common 5’’ refers to the five systematics that are common are common to all modes (B Counting, Photon Efficiency, Photon Quality, $\pi^0/\eta$ veto, and production uncertainty of $\Upsilon(4S) \rightarrow B\bar{B}$ ). The diagonal entries correspond to the total uncertainty for each mode. The matrix is symmetric, so all entries below the diagonal are omitted . . . . .	142

6.27	The calculated covariance matrix for all four modes. . . . .	142
6.28	The weights of the individual modes. . . . .	143
7.1	Number of Monte Carlo events in the fit region scaled to $316 \text{ fb}^{-1}$ using the MC luminosities for the $B \rightarrow (\rho/\omega)\gamma$ analysis. . . . .	146
7.2	$B\bar{B}$ background components used for the $B^0 \rightarrow \rho^0\gamma$ fit. . . . .	147
7.3	PDFs used for the $B^0 \rightarrow \rho^0\gamma$ fit. . . . .	148
7.4	$B\bar{B}$ background components used for the $B^+ \rightarrow \rho^+\gamma$ fit. . . . .	149
7.5	PDFs used for the $B^0 \rightarrow \rho^0\gamma$ fit. . . . .	149
7.6	PDFs used for the $B \rightarrow \omega\gamma$ fit. . . . .	150
7.7	Fractional systematic errors (in %) of the measured branching fractions. . . . .	151
7.8	The signal yield ( $n_{\text{sig}}$ ), significance ( $\Sigma$ ) in standard deviations including systematic errors, efficiency ( $\epsilon$ ), and branching fraction ( $\mathcal{B}$ ) for each mode. The errors on $n_{\text{sig}}$ are statistical only, while for the branching fraction the first error is statistical and the second systematic. . . . .	154
8.1	Origin of $B\bar{B}$ events that fall in the signal region. . . . .	157
8.2	This table shows the factors needed to correct the Monte Carlo efficiency and their systematic errors. . . . .	160
8.3	Confidence intervals for an expected background of 6 events . . . . .	161
9.1	Previous measurements of the branching ratios and asymmetries for $B \rightarrow K^*\gamma$ , as well as the measurements of this thesis . . . . .	162
9.2	Current theoretical calculations of the $B \rightarrow K^*\gamma$ branching fraction . . . . .	163
9.3	Previous measurements of the branching ratios for $B \rightarrow \rho\gamma$ and $B \rightarrow \omega\gamma$ , where the limits are given by the 90% confidence level, as well the measurements of this thesis. The errors on the <i>Belle</i> measurements are statistical and systematic, respectively, while the entry in paranthesis refers to the significance of the measurement. . . . .	163
9.4	Theoretical predictions of the branching fractions for $B \rightarrow \rho\gamma$ and $B \rightarrow \omega\gamma$ . 164	
9.5	Current theoretical calculations of isospin violation in $B \rightarrow K^*\gamma$ decays within the SM. $F^{B \rightarrow V}(q^2 = 0)$ is the $B \rightarrow K^*$ form factor whose estimates range from $0.23 \pm 0.06$ [27] to $0.38 \pm 0.06$ [76]. . . . .	164

9.6	The measurements of the branching ratios for $B \rightarrow \rho\gamma$ and $B \rightarrow \omega\gamma$ which were performed after the measurements described in this thesis. They are described in Ref. [77]. The entry in paranthesis refers to the significance of the measurement. . . . .	166
D.1	The pulls of the yields for each component for fits in which the signal MC has been eliminated for $m_{ES} < 5.27$ . All fits are signal embedded. . . . .	180
D.2	The pulls of the yields for Crystal Ball and Cruijff functions for the $K^{*0} \rightarrow K_S\pi^0$ mode. All fits are signal embedded. . . . .	184
D.3	The pulls of the yields for each component fits of less than 3 dimensions for the $K^{*+} \rightarrow K^+\pi^0$ mode, with the exception of line labeled “Cruijff”, which is in 3 dimensions. This line displays pulls of the yields for the Cruijff function. All fits are signal embedded. . . . .	184
E.1	A comparison of the results from the fit to OnPeak data with the full fit to Monte Carlo for the $K^{*0} \rightarrow K^+\pi^-$ and $K^{*+} \rightarrow K_S\pi^+$ modes. . . . .	186
E.2	A comparison of the results from the fit to OnPeak data with the full fit to Monte Carlo for the $K^{*0} \rightarrow K_S\pi^0$ and $K^{*+} \rightarrow K^+\pi^0$ modes. . . . .	187

# List of Figures

2.1	A rendering of the PEP-II accelerator and storage rings. Electrons and positrons are accelerated in bunches until they are extracted at the bypass lines. They eventually combine at the <i>BABAR</i> detector, located at IR-2. . . . .	6
2.2	An overview of the <i>BABAR</i> detector. Starting at the collision axis and moving radially outward, the sub-detectors are the Silicon Vertex Tracker (SVT), the Drift CHamber (DCH), the Detection of Internally Reflected Cherenkov Detector (DIRC) particle identification system, the ElectroMagnetic Calorimeter (EMC), and the Instrumented Flux Return (IFR). The solenoid provides a 1.5 T magnetic field that is necessary to perform momentum measurements of charged particles. . . . .	8
2.3	A schematic side view of the SVT. . . . .	10
2.4	A transverse view of the SVT. . . . .	10
2.5	A longitudinal view of the drift chamber with the units given in mm. The IP is offset by 370 mm from the chamber center to account for the asymmetric beams. . . . .	12
2.6	Schematic layout of the drift cells for the four innermost superlayers. Lines have been drawn between field wires to illustrate the cell boundaries. The numbers on the right side give the stereo angles (mrad) of the sense wires, while the 1 mm beryllium boundary is shown on the bottom. . . . .	14
2.7	An illustration of the DIRC particle identification system. The bar boxes house the fused silica bars which as serve radiators for the traversing particles. They are supported by the Central Support Tube (which also supports the DCH). The whole structure is supported by the Strong Support Tube. . . . .	17
2.8	A schematic of the DIRC bar. Cherenkov radiation emitted by the particle is channeled to the PMTs, which produce a signal to be read out by the front-end electronics. . . . .	18

2.9	A reconstructed dimuon event with PMT signals that have a measured time within the $\pm 300$ ns trigger window (left) and within 8 ns of the expected Cherenkov photon arrival time (right). . . . .	18
2.10	A longitudinal view of the top-half of the EMC with the dimensions given in mm. Displayed are the 56 axially-symmetric rings which compose the detector. The endcap is placed at the forward end to account for the asymmetric beam conditions. . . . .	20
2.11	A schematic of a crystal. The readout is mounted on the back end. . . . .	21
2.12	An overview of the IFR, in which the barrel and endcaps are indicated. All dimensions are given in millimeters. . . . .	23
2.13	Cross section of a planar RPC. . . . .	23
2.14	Cross section of an LST module. . . . .	24
3.1	The single photon angle distribution for muon tracks as reconstructed in the DIRC (dots) and the Monte Carlo prediction using the DIRC MC program (solid line). The signal is distributed around zero, while the background photons also lie outside the Cherenkov cone. The Monte Carlo only explains roughly 60% of the observed background. . . . .	27
3.2	Experimental setups used to study the photon background. The Cherenkov signal propagates to either (a) the mirror or (b) the photon trap. In the former case, a window of $\sim 36$ ns is available to study the early photon background activity before the Cherenkov signal returns; in the latter case a $\sim 70$ ns window is available . . . . .	29
3.3	Digital scope output from a single event for the mirror setup. The horizontal axis is time in terms of scope channels (1 scope channel = 0.4 ns) and the vertical axis is the amplitude. The large Cherenkov pulse arrives at channel 375, while the background pulses arrive earlier. . . . .	30
3.4	Digital scope output from a single event for the photon setup (a), and the output of the deconvolution algorithm (b). The horizontal axis is time in terms of scope channels (1 scope channel = 0.4 ns) and the vertical axis is the amplitude. The lack of the Cherenkov pulse is evident as a result of the photon trap. Also shown are the number and positions of the peaks using the peak finding algorithm (a), and the deconvolution algorithm (b). The two algorithms are consistent with each other. . . . .	31

3.5	Measured time distribution of photon pulses obtained from the peak finding algorithm using the mirror data. The Cherenkov signal arrives at channel 150. The hits before this time are all considered background. . . . .	31
3.6	Measured time distribution of photon pulses obtained from the peak finding algorithm using the photon trap data. The Cherenkov signal is clearly absent. Therefore, all hits are considered background. . . . .	32
3.7	The Fresnel curve is based on the measurement of the index of refraction. Also shown is the data from two different measurements of the reflectivity per bounce from the glue/silica interface. Finally, the curve with square references is used in the Monte-Carlo program that simulates the 4 m long bar cosmic ray experiment. . . . .	33
3.8	Time distribution of photon background pulses in the first 30 ns. The Monte Carlo only simulates delta rays and is represented by the smooth line, while the data (a combination of both the mirror and photon trap data) is given by the dots. Delta rays are the dominant component in the first background photon peak. . . . .	35
3.9	Time distribution of photon background pulses in the first 30 ns (a) and 0-70 ns (b), where the glue reflectivity is included in the simulation. The Monte Carlo is represented by the smooth line, while the data is given by the dots. For (a), the data is a combined of the mirror and photon trap data, while the data in (b) consists solely of the photon trap data. . . . .	37
3.10	Multiplicity distribution of background photons for the first 30 ns (a) and 0-70 ns (b). The Monte Carlo is represented by the smooth line, while the data is given by the dots. For (a), the data is a combined of the mirror and photon trap data, while the data in (b) consists solely of the photon trap data. The data and Monte agree well in terms of most probable number, which is three for (a) and five (b). . . . .	38
3.11	The single photon angle distribution for muon tracks as reconstructed in the DIRC for <i>BABAR</i> Monte Carlo. The effects of the glue are shown. Also shown for reference are the effects of no delta rays and no machine background (photon background coming from the beam interactions). . . .	38

4.1	The unitary triangles with the sides given in terms of expansion parameter $\lambda$ portrayed in the complex plane corresponding to a) Eq. ?? and b) Eq. ??. The openness of the former gives large $CP$ violation in decays such as $B \rightarrow J/\psi K_S$ and $B \rightarrow \pi\pi$ , while the latter gives small $CP$ violating effects in the decay of $B \rightarrow K^*\gamma$ . The angles given in (a) are determined by measuring time-dependent asymmetries. . . . .	43
4.2	Leading order short-distance Feynman digram for the $b \rightarrow (s, d)\gamma$ and transition . . . . .	44
4.3	One of the leading order Feynman diagrams contributing to the decay $B \rightarrow \phi\gamma$ in the Standard Model. . . . .	50
4.4	Experimental constraints on mSUGRA parameter space for the $B \rightarrow K^*\gamma$ isospin asymmetry (black), the $B \rightarrow X\gamma$ branching fraction (magenta). The green “Charged LSP” indicates the the cosmologically disfavored region if the lightest supersymmetric particle (LSP) is charged and R-parity is conserved. The grey “Excluded” region indicates constraints on the parameter space corresponding to direct searches for the Higgs and sparticles. The assumptions are $A_0 = 0$ , $\tan(\beta) = 30$ , and $\mu > 0$ . . . . .	54
5.1	Delivered and recorded luminosity as a function of time. . . . .	59
5.2	Number of tracks, $R_2$ , and CMS energy of highest energy photon before any cuts have been applied. . . . .	65
5.3	$m_{ES}$ and $\Delta E$ distributions after the preliminary cuts have been applied. . . . .	66
5.4	Pion efficiency and kaon misidentification rates of the $B \rightarrow \rho\gamma$ selector using the $D^*$ control sample. . . . .	68
5.5	Shown are the probability of the $\chi^2$ fit to the $K_S$ decay vertex, the $K_S$ invariant mass, and the $K_S$ flight significance. . . . .	69
5.6	The invariant mass of the $\pi^0$ for the $K^{*+} \rightarrow K^+\pi^0$ mode . . . . .	70
5.7	Shown are the $K^*$ invariant mass and the $\chi^2$ fit to the $K_S$ decay vertex. . . . .	71
5.8	Shown are the distance of the photon candidate to the nearest neutral bump, the distance to the nearest charged bump, and the output of the $\pi^0$ and $\eta$ veto likelihood functions. . . . .	73
5.9	The second moment ( $L_2^{bump}$ ) of the photon shower. . . . .	74
5.10	$\pi^0$ veto efficiencies for different veto methods. . . . .	75
5.11	$\eta$ veto efficiency for different veto methods. . . . .	76

5.12	$\cos \theta_T$ , $\cos \theta_B$ , and Legendre moments for signal (black dots) and continuum Monte Carlo (red line). The mode shown is $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$ and all cuts are applied except for the Neural Network cut. The plots are normalized to the same area. . . . .	80
5.13	$R'_2$ , $\Delta z$ significance, and subtagger output for signal (black dots) and continuum Monte Carlo (red line). The mode shown is $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$ and all cuts are applied except for the Neural Network cut. The plots are normalized to the same area. . . . .	81
5.14	a) $ \cos \theta_H $ , b) Net Flavor, and c) $\Delta z$ distributions for continuum background (black dots) and signal Monte Carlo (red line) for $B \rightarrow \phi\gamma$ . All cuts except the neural network cut are applied. . . . .	82
5.15	Every other forward (in the same hemisphere as the high energy photon) energy cone distribution for signal(red line) and continuum(black dots) Monte Carlo for $B \rightarrow \phi\gamma$ are shown. All cuts except the neural network cut are applied. . . . .	83
5.16	Every other backward (in the opposite hemisphere as the high energy photon) energy cone distribution for signal(red line) and continuum(black dots) Monte Carlo for $B \rightarrow \phi\gamma$ are shown. All cuts except the neural network cut are applied. . . . .	84
5.17	Neural Network output for signal and background . . . . .	85
5.18	Neural Network performance: Background Rejection vs. Selection efficiency	86
5.19	“Sensitivity”/signal efficiency as a function of neural net output for a slice of invariant phi mass. . . . .	89
6.1	The $B \rightarrow X_s^+\gamma$ generated hadronic mass spectrum for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode. The mass selection of $> 1.1 \text{ GeV}/c^2$ has been applied. In addition, all selection cuts have been applied. . . . .	93
6.2	$m_{ES}$ , $\Delta E$ , and $\cos \theta_H$ distributions of the inclusive $B \rightarrow X_s\gamma$ MC simulation for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode. All selection cuts have been applied, except for the $\Delta E$ and $\cos \theta_H$ distributions, in which case the respective selection has not been applied. . . . .	94

6.3	$m_{\text{ES}}$ , $\Delta E$ , and $\cos\theta_H$ distributions of the $B \rightarrow K^*\gamma$ MC simulation for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode, in which the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode has been removed. All selection cuts have been applied, except for the $\Delta E$ and $\cos\theta_H$ distributions, in which case the respective selection has not been applied. . . . .	95
6.4	$m_{\text{ES}}$ , $\Delta E$ , and $\cos\theta_H$ distributions of the $B$ background, not including $B \rightarrow K^*\gamma$ and $B \rightarrow X_s\gamma$ , for the $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ mode. All selection cuts have been applied, except for the $\Delta E$ and $\cos\theta_H$ distributions, in which case the respective selection has not been applied. . . . .	96
6.5	1 dimensional fits to the $K^{*0} \rightarrow K^+\pi^-$ mode. The dimensions are $\Delta E$ (top row), $m_{\text{ES}}$ (middle row), and $\cos\theta_H$ (bottom row). The components are signal (1 <sup>st</sup> column), continuum (2 <sup>nd</sup> column), and $B$ background (3 <sup>rd</sup> column). The helicity angle is determined by following the kaon daughter of the $K^*$ . . . . .	100
6.6	1 dimensional fits to the $K^{*0} \rightarrow K_S\pi^0$ mode. The dimensions are $\Delta E$ (top row), $m_{\text{ES}}$ (middle row), and $\cos\theta_H$ (bottom row). The components are signal (1 <sup>st</sup> column), continuum (2 <sup>nd</sup> column), and $B$ background (3 <sup>rd</sup> column). The helicity angle is determined by following the $K_S$ . . . . .	101
6.7	1 dimensional fits to the $K^{*+} \rightarrow K^+\pi^0$ mode. The dimensions are $\Delta E$ (top row), $m_{\text{ES}}$ (middle row), and $\cos\theta_H$ (bottom row). The components are signal (1 <sup>st</sup> column), continuum (2 <sup>nd</sup> column), and $B$ background (3 <sup>rd</sup> column). The helicity angle is determined by following the kaon. . . . .	102
6.8	1 dimensional fits to the $K^{*+} \rightarrow K_S\pi^+$ mode. The dimensions are $\Delta E$ (top row), $m_{\text{ES}}$ (middle row), and $\cos\theta_H$ (bottom row). The components are signal (1 <sup>st</sup> column), continuum (2 <sup>nd</sup> column), and $B$ background (3 <sup>rd</sup> column). The helicity angle is determined by following the pion. . . . .	103
6.9	$K^{*0} \rightarrow K^+\pi^-$ pure toy results . . . . .	109
6.10	$K^{*0} \rightarrow K^+\pi^-$ signal embedded toy results . . . . .	111
6.11	$K^{*0} \rightarrow K^+\pi^- B\bar{B}$ embedded toy results . . . . .	112
6.12	The likelihood fit on the $B \rightarrow D\pi$ control samples. . . . .	121
6.13	Comparison of neural net distributions for MC simulation and background subtracted on-peak data. . . . .	122
6.14	The comparison of the signal efficiency vs. cut value between Monte Carlo and background subtracted on-peak data. . . . .	123
6.15	The bin-by-bin $e_{\text{Data}}/e_{\text{MC}}$ . . . . .	124

6.16	$B\bar{B}$ distributions and fit parameters for the three dimensions of the fit for a) the nominal fit, b) when the $b \rightarrow s\gamma$ yield is decreased by 20%, and when c) the $b \rightarrow s\gamma$ yield is increased by 20%. The mode shown is $K^{*0} \rightarrow K^+\pi^-$ .	127
6.17	Total cross section difference for K-deuteron and $\pi$ -deuteron interactions from the PDG . . . . .	130
6.18	$p_{tab}$ distribution for K's and $\pi$ 's for truth-matched $K^{*0} \rightarrow K^+\pi^-$ signal decays. . . . .	131
6.19	$K^{*0} \rightarrow K^+\pi^-$ projection plots of the full fit to on-peak data . . . . .	133
6.20	$K^{*0} \rightarrow K_S\pi^0$ projection plots of the full fit to the on-peak data . . . . .	134
6.21	$K^{*+} \rightarrow K^+\pi^0$ projection plots of the full fit to the on-peak data . . . . .	135
6.22	$K^{*+} \rightarrow K_S\pi^+$ projection plots of the full fit to the on-peak data . . . . .	136
6.23	Negative maximum likelihood distributions of a signal embedded studies for a) $K^{*0} \rightarrow K^+\pi^-$ , b) $K^{*0} \rightarrow K_S\pi^0$ , c) $K^{*+} \rightarrow K^+\pi^0$ , d) $K^{*+} \rightarrow K_S\pi^+$ , obtained as described in the text. The negative maximum likelihoods from the on-peak data fit are marked by a red line. There are one thousand entries in each plot. . . . .	138
6.24	Relativistic P-wave Breit-Wigner line shape fit to the $K\pi$ invariant mass distribution of the sPlot of the on-peak dataset for the a) $K^{*0} \rightarrow K^+\pi^-$ , b) $K^{*0} \rightarrow K_S\pi^0$ , c) $K^{*+} \rightarrow K^+\pi^0$ , d) $K^{*+} \rightarrow K_S\pi^+$ modes. . . . .	140
7.1	Projections for the $\rho^0$ mode. . . . .	152
7.2	Projections for the $\rho^+$ mode. . . . .	153
7.3	Projections for the $\omega$ mode. . . . .	155
8.1	The four different regions used in estimating the number of background events in the signal region. . . . .	157
8.2	Argus fits to the $m_{ES}$ distributions in a) region 1, b) region 3, c) regions 1 and 3, d) regions 2 and 4. The data set used is continuum MC simulation. . . . .	158
8.3	The (a) $m_{ES}$ and (b) $\Delta E$ distributions of on-resonance data. Superimposed are (a) an Argus function and (b) a 1 <sup>st</sup> order polynomial that are obtained from fits to the data in regions 1 and 3. The different regions and fitting method are described in the text. . . . .	159
8.4	The 2D $\Delta E$ vs. $m_{ES}$ plane of the fitting region using real data. The box indicates the boundaries of the signal region, where 8 events lie. . . . .	161

9.1	The measurement of $B(B \rightarrow (\rho/\omega)\gamma)$ , when combined with $B(B \rightarrow K^*\gamma)$ , provides a measurement of the magnitude of one length of the unitary triangle $R_t$ (described in Section ??). This is outlined in green, and extends from the inner green circle to the outer green area. The measurement of $B_d/B_s$ mixing also gives a measurement of $R_t$ , and extends across the yellow shaded region. . . . .	165
A.1	Thrust angle, $\cos\theta_B$ , and Legendre moments for offpeak data(black dots) and continuum Monte Carlo (red line). The mode shown is $K^{*0} \rightarrow K^+\pi^-$ and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data. . . . .	169
A.2	$R'_2$ , $\Delta z$ significance, and subtagger output for offpeak data(black dots) and continuum Monte Carlo (red line). The mode shown is $K^{*0} \rightarrow K^+\pi^-$ and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data. . . . .	170
A.3	$R_2^{All}$ for offpeak data(black dots) and continuum Monte Carlo (red line). The mode shown is $K^{*0} \rightarrow K^+\pi^-$ and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data. . . . .	171
A.4	A comparison of offpeak data(black dots) and continuum Monte Carlo (red line) for all variables that were optimized. The mode shown is $K^{*0} \rightarrow K^+\pi^-$ and all cuts are applied to each plot except for the variable shown. The Monte Carlo is scaled to the data. . . . .	172
B.1	The left column shows the 2 dimensional scatter plots of the observables for signal Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot. . . . .	174
B.2	The left column shows the 2 dimensional scatter plots of the observables for continuum Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot. . . . .	175
B.3	The left column shows the 2 dimensional scatter plots of the observables for $B\bar{B}$ Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot. . . . .	176
D.1	Signal MC sidebands for the $m_{ES}$ distribution for all 4 modes . . . . .	181
D.2	$K^{*0} \rightarrow K^+\pi^-$ pion lab momentum and $\cos(\theta_H)$ distributions for truth-matched, $m_{ES} > 5.27$ (black) and untruthmatched, $m_{ES} < 5.27$ (red) . . .	182

D.3	Signal MC data and PDF fits for the (a) Cruijff and (b) Crystal Ball functions for the $K^{*0} \rightarrow K_S \pi^0$ mode. The left-hand side shows $m_{ES} < 5.27$ , while the right-hand side shows $m_{ES} > 5.27$ . We separate out the two regions for clarity. . . . .	183
D.4	Signal MC data and PDF fits for the (a) Cruijff and (b) Crystal Ball functions for the $K^{*+} \rightarrow K^+ \pi^0$ mode. The left-hand side shows $m_{ES} < 5.27$ , while the right-hand side shows $m_{ES} > 5.27$ . We separate out the two regions for clarity. . . . .	185



# Chapter 1

## Introduction

This thesis describes a measurement of the radiative decays  $B \rightarrow V\gamma$ , where  $V = K^*, \rho, \omega$ , or  $\phi$ . It also describes the theoretical tools and experimental apparatus that underlie these measurements. In addition, a study is presented that describes a new background source for the DIRC (Detection of Internally Reflected Cherenkov Detector). This chapter provides a general introduction.

### 1.1 Motivation

The past century has seen a vast extension in the library of scientific understanding. Einstein has forced physicists to reconsider how they perceive space and time, which then led him to reformulate gravity. The advent of quantum mechanics has led to a complete upheaval of the perception of matter. In place of point-like objects of the everyday world, there exists a formalism which speaks in terms of amplitudes and phases of a world which is unseen. The union of quantum mechanics and relativity has brought about an era in which matter has seemed to have lost its prominence. The advent of quantum field theory has led physicists to formulate theories by speaking of symmetry as the prime motivator. Indeed, it has become so central that when a symmetry is broken, understanding the mechanism often leads to new viewpoints. The culmination of the attempt to understand electromagnetism, weak, and strong interactions has become the standard model (SM) of particle physics.

The SM is a combination of two theories, quantum electroweak and quantum chromodynamics (QCD), that describes all known elementary particles and their interactions on a vast range of scales. The electroweak theory was established by 1979 in part

by experiments confirming the observation of neutral currents [1] and parity-violating asymmetries due to the interference of weak and electromagnetic electron scattering from nucleons using polarized electron beams performed at the Stanford Linear Accelerator Center (SLAC) [2]. QCD came to be widely accepted a few years later in part by electron-positron annihilation into three jets [3], and the observation of the logarithmic deviations from Bjorken scaling. In addition, the anomalously long lifetime of the  $J/\psi$  helped to confirm the asymptotic freedom of QCD.

Also, crucial to the development of the SM was the non-observation of flavor-changing neutral currents (FCNCs). The decays  $K \rightarrow \mu^+\mu^-$  and  $K \rightarrow \pi\nu\bar{\nu}$  are highly suppressed in the SM compared to the charged-current decay  $K \rightarrow \pi e\nu$ . For example, the branching ratios are given by [4]

$$\begin{aligned}\mathcal{B}(K_L \rightarrow \mu^+\mu^-) &= (6.84 \pm 0.11) \times 10^{-9}, \\ \mathcal{B}(K^+ \rightarrow \pi^+\nu\bar{\nu}) &= (1.5_{-0.9}^{+1.3}) \times 10^{-10}, \\ \mathcal{B}(K^+ \rightarrow \pi^0 e^- \nu_e) &= (5.05 \pm 0.06) \times 10^{-2}\end{aligned}\tag{1.1}$$

Eventually, the suppression of these decays came to be understood by introducing the charm quark. This kind of suppression is known as the Glashow-Iliopoulos-Maiani (GIM) mechanism. This prediction of a charm quark was confirmed by the discovery of the  $J/\psi$  particle. In addition, the  $b \rightarrow s, d$  transitions (along with  $B^0 - \bar{B}^0$  mixing) helped to establish the existence of a massive  $t$  quark.

It is now thirty years later and the SM has withstood an extensive range of tests, whose level of precision is such that  $\mathcal{O}(\alpha)$  radiative corrections are needed. The tests of QCD mostly lie in the high-energy range where QCD is perturbative. The strong coupling constant has been determined on energy scales ranging from 1.7 GeV to 189 GeV using a wide range of techniques [4]. These experiments are not only consistent among each other, but also clearly demonstrate the scale dependence of the coupling constant using the renormalization group equation and the independence of the coupling constant from quark flavor. Similarly, the electroweak gauge couplings of the quarks and leptons predicted by the symmetry of the SM have been verified to less than one percent by experiments at the Large Electron-Positron Collider (LEP) and the Stanford Linear Collider (SLC) [4, 5]. LEP experiments have explored much of the triple gauge coupling (TGC) parameter space, finding agreement with the SM within a few percent. These measurements verify that the TGC's conform to the gauge symmetries of the SM. Heavy flavor physics has been

explored in great depth by *BABAR* and *Belle*, helping to establish the magnitude of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements  $V_{cb}$  and  $V_{ub}$ , and the uncertainty on  $\sin 2\beta$  to a few percent. There are no significant deviations from the SM that cannot be accounted for by uncertainties related to experiment or theory.

Despite the phenomenological success of the SM, it remains to many just that, namely, a phenomenological model. It is widely believed that physics beyond the SM has to be present, mostly for reasons that pertain to theoretical aesthetics rather than experimental necessity. The SM is composed of a product of three distinct gauge groups with at least eighteen arbitrary parameters. These are the three gauge couplings, six quark and three charged lepton masses, three CKM angles with one CP violating phase, and two parameters to characterize the Higgs sector. In addition, there could be an additional strong CP violating parameter, which has been excluded by the non-existence of the electric dipole moment of the neutron. This is commonly referred to as the strong CP problem. The observations of neutrino mass brings in another seven parameters, which are three masses, three angles, and one phase. The results of the  $B$  factories have led to the conclusion that the  $CP$  violating phase in the SM is not enough to account for the baryon asymmetry that we observe today. While it is possible this asymmetry could be accounted for by a  $CP$  violating phase in the neutrino sector, new physics scenarios could also bring in other  $CP$  violating observables. The SM gives no explanation as to the origin of electroweak symmetry breaking, and, in addition, the origin of the left-handed currents in the weak sector. A more sophisticated theory is needed to provide theoretical depth.

The SM also gives no insight as to the hierarchy of masses, and, in fact, poses phenomenological problems related to all the masses. Radiative corrections to the Higgs mass have a quadratic dependence on the ultraviolet momentum cutoff, which is typically taken to be some very high energy scale at which new physics comes in to alter the high energy behavior of the SM. This is typically taken to be on the order of the Grand Unification Theory (GUT) scale ( $\sim 10^{16}$  GeV) or a little below the Planck mass ( $M_{pl} \sim 10^{19}$  GeV). Since electroweak data indicate that the Higgs mass is on the order of 100 GeV, huge cancellations must take place if the SM is taken as a fundamental theory. In fact, since all the masses in the SM are obtained through the coupling to the Higgs, this problem indirectly affects all of the fundamental particles. This condition is usually referred to as the gauge hierarchy problem. The Yukawa coupling of the top quark undergoes renormalization in which its own Yukawa corrections cancel against QCD corrections, thus creating a quasi-infrared fixed point [6] for the top quark mass. A wide range of initial

values for the Yukawa coupling will be attracted to this fixed point, which corresponds to the physical value of the top quark mass. Using the SM, one can calculate that the fixed point corresponds to  $m_t \sim 250 \text{ GeV}$ . However, in a theory with two Higgs doublets, such as the Minimal Supersymmetric Model (MSSM), the fixed point corresponds to  $m_t = 190 - 210 \text{ GeV} \sin \beta$  [7], where  $\tan \beta = v_2/v_1$  is the ratio of the two Higgs vacuum expectation values. This allows the top quark mass to lie closer to the experimentally determined value of  $m_t = 171.2 \pm 2.1$  [4].

These observations form a powerful subset of reasons why the SM is regarded as an effective theory, much like the Fermi interaction is an effective theory for the weak interactions. It is understood to provide a description of phenomenon up to some large energy scale, upon which a more sophisticated theory would become applicable. The nature of this theory is not known, although electroweak data are of sufficient precision as to provide constraints on new physics models. For example, certain classes of Technicolor models and a fourth generation of ordinary fermions are disfavored.

Direct detection of particles not in the SM is not possible at the  $B$  factories. This is relegated to the Large Hadron Collider, which will produce particles at TeV energies. However, the influences of new particles can be observed in rare radiative decays. These decays are mediated by loop diagrams at leading order, so that new particles that can be present in the loops can deviate SM predictions to a measurable degree. The new particles that may be present include, but are not limited to, the charged Higgs, the charginos, and other supersymmetric particles. Radiative decays also occur at rates proportional to  $G_F^2 \alpha$ , which is enhanced compared to non-radiative rare decays, which have rates proportional to  $G_F^2 \alpha^2$ . Precision measurements can not only test for SM predictions, but also restrict new physics parameter space.

The flavor-changing neutral currents  $b \rightarrow s\gamma$  and  $b \rightarrow d\gamma$  provide examples of decays which are sensitive to physics beyond the standard model, and are closely studied by the  $B$  factories. The former decay is CKM favored over the latter, thus making it more experimentally accessible. The inclusive calculation is more theoretically precise than the exclusive calculations, because they are not limited by our imperfect understanding of how the  $b$  and  $s, d$  quarks are bound into the initial and final state mesons, which is characterized by form factors. However, the inclusive measurement is more challenging than the exclusive counterparts, as a result of large backgrounds coming from the the lack of a hadronic mass selection and specific final states. The decay  $B \rightarrow K^*\gamma$  is an exclusive  $b \rightarrow s$  transition that is the most dominant radiative decay of the  $B$  meson. The corresponding decays  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$  are  $b \rightarrow d$  transitions, while the decay  $B \rightarrow \phi\gamma$

proceeds through a penguin annihilation process. Even though the branching fractions suffer from hadronic uncertainties, new physics can be extracted by considering quantities such as asymmetries and ratios of branching fractions. In this thesis, we measure the branching fractions, time-independent  $CP$ , and isospin asymmetries of  $B \rightarrow K^*\gamma$ . Also, a measurement of the branching fractions and isospin asymmetry of  $B^+ \rightarrow \rho^+\gamma$  and  $B^0 \rightarrow \rho^0\gamma$  is given, as well as an upper limit on the branching fraction of  $B \rightarrow \omega\gamma$ . We combine the measurements of  $B \rightarrow \rho\gamma$ ,  $B \rightarrow \omega\gamma$ , and  $B \rightarrow K^*\gamma$  to provide a measurement of the CKM element ratio  $|V_{td}/V_{ts}|$ . Finally, we also present a measurement of the branching fraction of  $B \rightarrow \phi\gamma$ . All of these decays are denoted by the notation  $B \rightarrow V\gamma$ , where  $V = K^*, \rho, \omega$ , or  $\phi$ .

## 1.2 Outline

The progression of this thesis will be to first explain the underlying theoretical concepts and experimental apparatus used to make the measurements related to  $B \rightarrow V\gamma$ , the analysis of which will be presented last. Thus, Chapter 2 gives an explanation of the elements in the *BABAR* detector. Chapter 3 provides a study of the DIRC fused silica bars. Chapter 4 formulates the classical SM Lagrangian, while also giving the theoretical tools needed for the phenomenology of  $B \rightarrow V\gamma$ . Chapter 5 will present the event selection, background suppression techniques, and selection optimization used for all the analyses. The next three chapters will discuss the techniques for the specific analyses of  $B \rightarrow K^*\gamma$ ,  $B \rightarrow (\rho/\omega)\gamma$ , and  $B \rightarrow \phi\gamma$ . Finally, Chapter 9 will conclude the thesis.

# Chapter 2

## PEP-II and the *BABAR* Detector

The primary goal of the *BABAR* experiment is to study  $CP$  violation through studies of  $B$  decays into neutral  $CP$  eigenstates. As such, the *BABAR* detector was optimized for this purpose. However, the detector is also suited to make precision measurements of the decays of bottom and charm mesons, and  $\tau$  leptons, as well as rare processes. The luminosity is provided by the PEP-II (Positron-Electron Project)  $B$  Factory, which was built specifically to produce copious amounts of  $B$  mesons for *BABAR*. This chapter will provide a description of the PEP-II accelerator, as well as the *BABAR* detector.

### 2.1 The PEP-II Accelerator

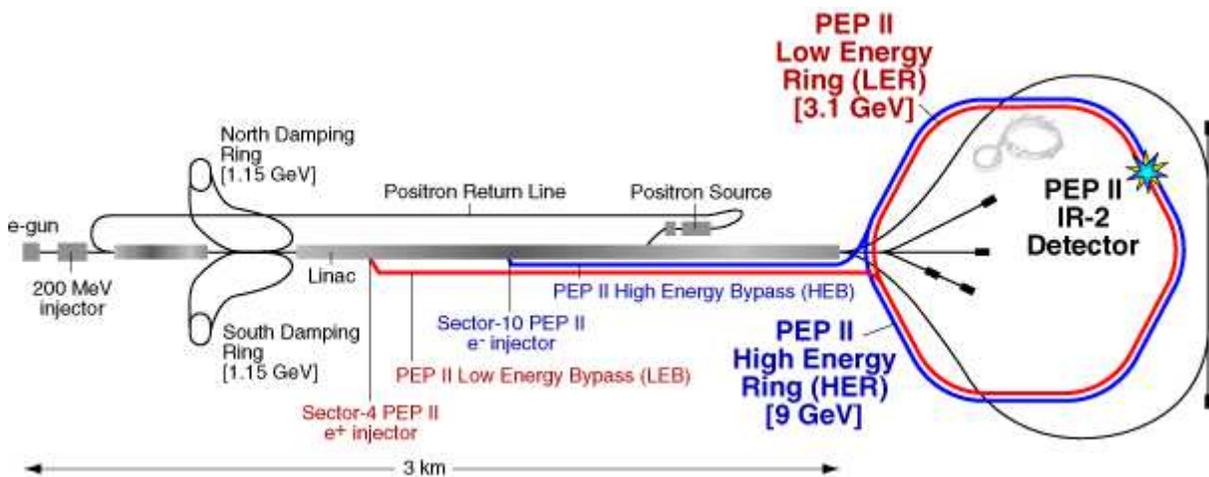


Figure 2.1: A rendering of the PEP-II accelerator and storage rings. Electrons and positrons are accelerated in bunches until they are extracted at the bypass lines. They eventually combine at the *BABAR* detector, located at IR-2.

A schematic of PEP-II [8] is shown in Fig. 2.1. Electrons are injected and accelerated in the linear accelerator (LINAC) until the High Energy Bypass (HEB) line, where they are extracted at their full energy of 9.0 GeV and transported to the High Energy Ring (HER). Some of the electrons are further conveyed until they collide with a tungsten target, which produces electron-positron pairs. The positrons are separated from these collisions and then returned back to the beginning of the LINAC and accelerated until 3.1 GeV, when they are extracted to the Low Energy Bypass (LEB) line. This line injects the positrons into the Low Energy Ring (LER).

The electrons and positrons are stored in the PEP-II rings with a storage time of approximately one hour. The circumference of the PEP-II rings is about 2.2 km, in which the LER sits above the HER. The LER has a nominal current of 2.14 A, which corresponds to roughly  $6 \times 10^{10}$  positrons per bunch, while the current for the HER is 0.99 A, or  $3 \times 10^{10}$  electrons per bunch. The bunch collision frequency is 238 MHz, which gives about a 4 ns time difference between bunches. They collide at the *BABAR* detector located at IR-2 in Fig. 2.1.

For the majority of running, the center-of mass (CM) energy of the beams is 10.58 GeV, which is the mass of the  $\Upsilon(4S)$  resonance. The asymmetry of the beam energies results in the  $\Upsilon(4S)$  being boosted in the lab frame with a Lorentz boost of  $\beta\gamma = 0.56$ . This allows for the distance between the two  $B$  vertices to be measured for time-dependent  $CP$  violation measurements.

As alluded to in the previous paragraph, not all of the data correspond to the  $\Upsilon(4S)$  resonance. Monte Carlo simulations of  $e^+e^- \rightarrow c\bar{c}, s\bar{s}, u\bar{u}, d\bar{d}$  events are less reliable than real data, and so approximately 12% of the data are taken about 40 MeV below this resonance to allow for studies of the non-resonant background. In addition,  $\sim 8\%$  of the data was taken at the  $\Upsilon(2S)$  and  $\Upsilon(3S)$  resonances.

PEP-II ran successfully from the fall of 1999 to the spring of 2008. It achieved a peak luminosity of about  $12.0 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ , surpassing its design of  $3.0 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ . The total integrated luminosity delivered to *BABAR* was  $557 \text{ fb}^{-1}$ .

## 2.2 The *BABAR* Detector

Key design requirements of the *BABAR* detector are

- large and uniform acceptance in CM frame. Thus, the detector is asymmetric in the boost direction;

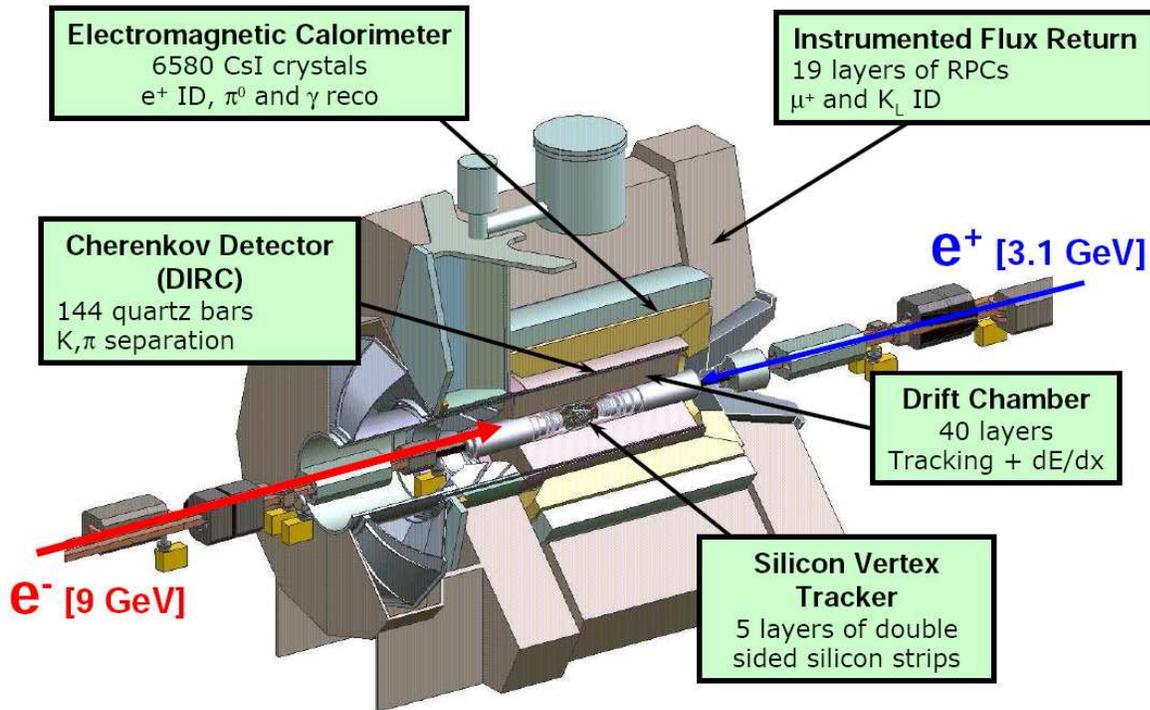


Figure 2.2: An overview of the *BABAR* detector. Starting at the collision axis and moving radially outward, the sub-detectors are the Silicon Vertex Tracker (SVT), the Drift Chamber (DCH), the Detection of Internally Reflected Cherenkov Detector (DIRC) particle identification system, the ElectroMagnetic Calorimeter (EMC), and the Instrumented Flux Return (IFR). The solenoid provides a 1.5 T magnetic field that is necessary to perform momentum measurements of charged particles.

- excellent reconstruction efficiency for charged particles down to 60 MeV/ $c$  and photons down to 20 MeV;
- very good charged particle momentum resolution, as well as energy and angular resolution for photons in the range 20 MeV to 4 GeV;
- very good vertex resolution for time-dependent  $CP$  measurements;
- excellent particle identification for leptons and hadrons. For this thesis,  $K/\pi$  separation is very important;
- a computing system that can control, process, and store large amounts of data;
- detector components that can operate under high background conditions.

Fig. 2.2 gives a schematic of the *BABAR* detector, which is composed of several sub-detectors that each fulfill one or more of the physics requirements listed above. The innermost detector is called the silicon vertex tracker (SVT), which is responsible for measuring decay vertices and track angles of charged particles near the interaction region. The momenta and angles of charged particles are measured with the drift chamber (DCH). Particle identification in the low momentum region using  $dE/dx$  is also done using this device, as well as the SVT. The DIRC performs particle identification of charged particles using Cherenkov light in the high momentum region. The detection of electromagnetic showers from photons and electrons is done using the electromagnetic calorimeter (EMC), which is an array of CsI crystals that is located just inside the solenoidal coil of the superconducting magnet. The identification of muons and neutral hadrons is done using the instrumented flux return (IFR), which has resistive plate chambers (RPC's) inserted into gaps in the steel flux return. However, beginning in 2004, due to poor performance, these RPC's were replaced by limited streamer tubes (LST's). All of these sub-detectors output data which is managed by a trigger and data acquisition system designed to maximize physics data acceptance, while minimizing deadtime and cost.

### 2.2.1 Silicon Vertex Tracker (SVT)

The SVT is responsible for measuring decay vertices and track trajectories near the interaction region. Since the measurement of time-dependent  $CP$  asymmetries relies heavily on knowledge of the  $B$  decay vertex, which is located near the interaction point (IP), the SVT is crucial for these kinds of measurements. The  $B$  meson has an average lifetime of  $\sim 250\mu m$  in the lab frame. For accurate determination of the  $CP$  asymmetry, a mean vertex resolution along the  $z$ -axis of at least  $80\mu m$  is necessary. Tracks with  $p_T < 120\text{ MeV}/c$  are not measured by the DCH, and so have to be reconstructed in the SVT. This occurs, for example, for the slow pions from  $D^*$  decays. The measurement of track angles provided by the SVT and DCH is also used as input to the DIRC, which relies on accurate knowledge of the point of entry into the quartz bars for particle identification. Finally, particle identification using energy loss along the track direction is performed by the SVT for low momentum tracks.

The SVT consists of five layers of double-sided silicon strip sensors, which are organized into 6, 6, 6, 16, and 18 modules respectively. The strips located on one side of the sensor run parallel to the beam and measure the azimuthal angle ( $\phi$  strips), while the strips on the other side run perpendicular and measure the  $z$  position ( $z$  strips). The innermost

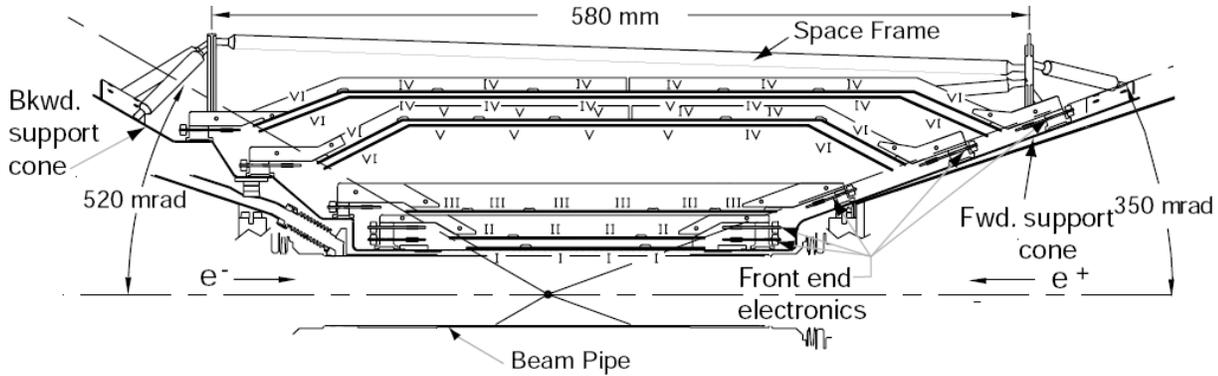


Figure 2.3: A schematic side view of the SVT.

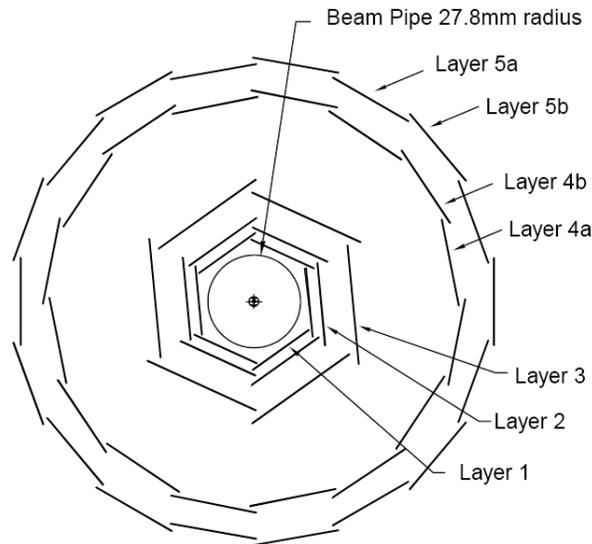


Figure 2.4: A transverse view of the SVT.

three layers are in barrel-type structures, and lie in the radial range  $\sim 3$  to  $\sim 6$  cm. They are used to make high-precision measurements close to the beam pipe. The outer two layers, which lie in the radial range  $\sim 12$  to  $\sim 15$  cm, are arch-shaped to minimize the amount of material required to cover the solid angle, and to also increase the angle of entry for particles that enter into the outer edge of the tracker. They are redundant to the three innermost layers, being mainly used for angle determination and linking tracks to the drift chamber. The cutaway side view of the SVT is shown in Fig. 2.3, while the end-on view is shown in Fig. 2.4. The SVT is restricted in the forward region by the B1 magnet, while the backward region is less constrained.

The five layers of the SVT are comprised of five different types of sensor shapes.

The smallest sensors are  $43 \times 42 \text{ mm}^2 (z \times \phi)$ , while the largest are  $68 \times 53 \text{ mm}^2$ . They are composed of n-type substrates with  $p^+$  and  $n^+$  strips on the two opposing sides. The depletion voltages are in the range 25-35 V, and the bias voltage is held at about 10 V above the depletion voltage. Charged particles which enter the substrate create electron-hole pairs which then move in the presence of the electric field. The strips then bring the signal to fanout circuits, which route the signal to the front-end electronics (FEE). The main component of the FEE is the ATOM (A Time-Over-Threshold Machine) chip, whose output is the time and time over threshold (ToT). Upon receipt of a L1 trigger, followed by a L1 accept command from the data acquisition system, this output is routed to the readout module (ROM). The reconstruction program then discards hits which fall more than 200ns from the event time (determined by the DCH), and groups the remaining hits according to position and time.

The movement of the SVT, both internally and as a whole, must be determined and corrected for, to correctly reconstruct the tracks. The first part of the correction procedure is the local alignment, which corrects for the relative motion of the 340 sensors. It uses dimuon events, cosmic rays, well-isolated high momentum tracks from hadronic events, and an optical survey to determine position information. Using all of this data, a  $\chi^2$  is formed for each sensor, and minimized with respect to the sensor's six local parameters. The local alignment of the sensors is expected to be relatively stable, and so is done infrequently, typically after magnet quenches or during detector access. In addition, the SVT, which is attached to the inside of the beryllium support tube, is not connected to the DCH to satisfy earthquake safety requirements, and so may move with respect to it due to factors such as thermal, mechanical, and magnetic stresses. The method to determine the relative positions of the two is called global alignment, which is the second part of the correction procedure. Tracks with a sufficient number of SVT and DCH hits are fit using SVT-only and DCH-only information. Then, the six global alignment parameters are determined by minimizing the difference between the track parameters obtained using the two different fits. The global alignment, performed about once every 2-3 hours, is done more frequently than local alignment.

The SVT performance is determined using various control samples, such as Bhabba and dimuon events. The track efficiency, defined as the ratio of the number of associated hits to the number of tracks crossing the active area of the module, is  $\sim 97\%$ . The hit resolution for both  $z$  and  $\phi$  varies from 15 to 40  $\mu\text{m}$  in the three innermost layers. In the two outer layers, the  $z$  resolution varies from 40 to 50  $\mu\text{m}$ , while the corresponding  $\phi$  resolution is  $\sim 15 \mu\text{m}$ .  $dE/dx$  information is also obtained using ToT information. The

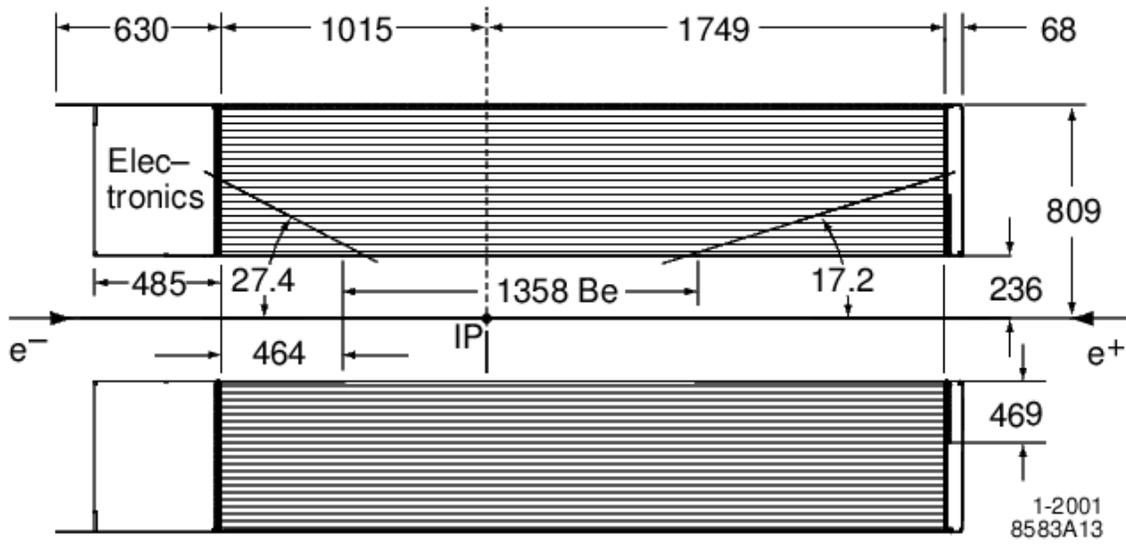


Figure 2.5: A longitudinal view of the drift chamber with the units given in mm. The IP is offset by 370 mm from the chamber center to account for the asymmetric beams.

number of samples used to obtain  $dE/dx$  is  $\sim 6$ , and the resolution is  $\sim 14\%$ . Up to  $500 \text{ MeV}/c$ , a  $2\sigma$  separation can be obtained between kaons and pions, and the same separation can be obtained between kaons and protons beyond  $1 \text{ GeV}/c$ .

### 2.2.2 Drift Chamber (DCH)

The drift chamber is responsible for measuring with high precision the momentum and angles of charged particles. The DCH is also crucial for reconstructing  $K_S$  decays that occur outside the SVT in the DCH volume. Thus, the DCH measures the longitudinal positions and momenta, as well as the transverse component. Simultaneously, the DCH provides particle identification for low momentum particles by measuring  $dE/dx$ . For example, it provides  $\pi/K$  separation up to  $700 \text{ MeV}/c$ . In the barrel region, the DIRC complements the DCH. However, in the extreme forward and backward regions, the DCH is the only particle identification device.

The DCH, whose side view is shown in Fig. 2.5, is about 3 m in length with an inner radius of  $\sim 23 \text{ cm}$ , and an outer radius of  $\sim 81 \text{ cm}$ . The inner cylinder is bounded by the support tube and is composed of a central 1-mm-thick beryllium tube with two 5-mm-thick aluminum extensions. It is kept thin in order to improve the track resolution for high momentum tracks, the matching of SVT to DCH tracks, and to reduce the background

from detector interactions. Track resolution is also improved by choosing a gas mixture that is 80% helium and 20% isobutane, which also serves to minimize multiple scattering. The material in the outer cylinder is minimized in order for the performance of the DIRC and EMC to be optimized. It is composed of two 1.6-mm-thick carbon fiber skins which enclose a 6-mm-thick hexagonal foam structure. Both are enclosed by aluminum foil on the inside and outside to provide RF shielding. The aluminum endplates are 24 mm thick. However, at the forward end, this thickness is reduced to 12 mm beyond a radius of 46.6 cm to minimize the material in front of the calorimeter endcap. For the same reason, the electronics and high-voltage supplies are mounted on the backward endplate.

There are a total of 28,768 wires in the DCH, which form 7,104 hexagonal drift cells arranged in 40 cylindrical layers. Each cell has one gold-coated tungsten-rhenium sense wire of 20 micron diameter, which is surrounded by 6 gold-coated aluminum field wires of 120 micron diameter. There are also aluminum guard and clearing wires which are of diameters 80 and 120 microns respectively. The 40 layers are clustered together in groups of 4 to form 10 superlayers. In order to obtain longitudinal position information, the wires in 6 of the superlayers are placed at small angles with respect to the z-axis. The stereo angles of the superlayers alternate between axial (A) and stereo (U,V) pairs, with the overall arrangement given by AUVAUVAUVA. This is shown in Fig. 2.6 for the first four superlayers. The complete specifications for the 10 superlayers are given in Table 2.1.

Table 2.1: The specifications of the superlayer structure of the DCH. Listed is the number of cells per layer, the radius of the innermost sense wire, cell widths and wire angles. Widths and radii are specified at the center of the chamber.

Superlayer number	Number of cells	Radius (mm)	Width (mm)	Angle (mrad)
1	96	260.4	17.0-19.4	0
2	112	312.4	17.5-19.5	45-50
3	128	363.4	17.8-19.6	-(53-57)
4	144	422.7	18.4-20.0	0
5	176	476.6	16.9-18.2	56-60
6	192	526.1	17.2-18.3	-(63-57)
7	208	585.4	17.7-18.8	0
8	224	636.7	17.8-18.8	65-69
9	240	688.0	18.0-18.9	-(72-76)
10	256	747.2	18.3-19.2	0

The field wires are held at ground potential, and the sense wires are at a voltage of

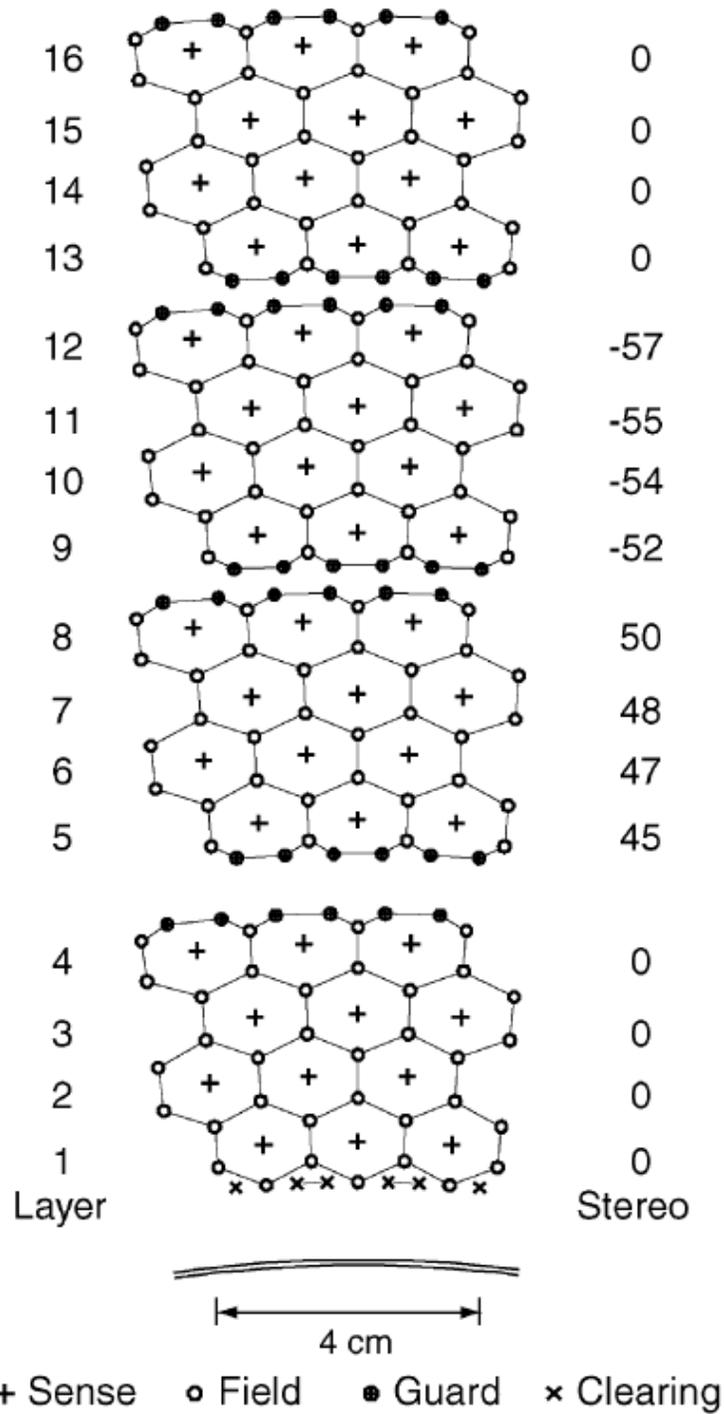


Figure 2.6: Schematic layout of the drift cells for the four innermost superlayers. Lines have been drawn between field wires to illustrate the cell boundaries. The numbers on the right side give the stereo angles (mrad) of the sense wires, while the 1 mm beryllium boundary is shown on the bottom.

1930 V (although there were brief periods during the first year of operation where it was held at 1900 V and 1960 V). The guard and clearing wires are at a voltage of 340 V and 825 V, respectively. Charged particles traversing the the DCH induce a trail of ionization. The electrons then drift in the gas towards the positively charge sense wires. As they near the wire, they start to accelerate and produce an “electron avalanche”. This avalanche ionizes the gas and the positively charged particles produced in the process causes an image charge on the sense wire. This charge then travels along the wire to be read out by the electronics. The gain is  $\sim 5 \times 10^4$  at the applied voltage. To maintain this gain at the boundary cells, guard wires are used. Clearing wires are used at the innermost boundary of layer 1 and the outermost boundary of layer 40 to collect charges created through photon conversions in the materials of the walls.

There are 16 sectors positioned symmetrically around the z-axis, each of which contains 3 front-end assemblies (FEAs). The sense wires are connected to the assemblies through service boards, which route the signal and HV distribution. In the FEA, a custom amplifier IC receives the input signal and produces a discriminator output signal for the drift time measurement and a shaped analog signal for the dE/dx measurement. This output is stored in the trigger latency buffer until an L1 accept signal comes and initiates the transfer to the readout buffer.

Dimuon events are used to determine the relationship between measured drift time and drift distance. The position resolution varies as a function of drift distance. It is lowest at 5 mm from the sense wire, where the resolution is  $\sim 0.1$  mm, and is highest near the cell edges (corresponding to a drift distance of  $\sim 10$  mm), where it is  $\sim 0.4$  mm. The dE/dx information is obtained through measurements of the total charge deposited in each drift cells. These measurements follow a Landau distribution, which if simply averaged do not yield the most-probable dE/dx value, and it is the most-probable value which exhibits a Bethe-Bloch dependence on momentum. Therefore, a truncated mean is calculated using the lowest 80% of the individual dE/dx measurements. After applying various corrections due to pressure, temperature, and cell geometry, the resolution obtained from a sample of Bhabba events is  $\sim 7.5\%$  of dE/dx for a typical track.

The tracking efficiency of the DCH, defined as the number of reconstructed tracks in the DCH to the number of tracks detected in the SVT, is  $\sim 98\%$ . The momentum resolution is determined from cosmic ray events. The upper and lower halves of the cosmic ray events are fit as two separate tracks, and the resolution is derived from the difference of the measured parameters for the two track halves. The transverse momentum resolution dependence  $p_T$  varies as  $\sim 0.5\%$  at  $\sim 0.5$  GeV/c to  $\sim 1\%$  at 5 GeV/c.

### 2.2.3 Detector of Internally Reflected Cherenkov Light (DIRC)

The function of the DIRC is to identify particles by reconstructing the Cherenkov angle from light produced as it passes through the DIRC bars. In the case of  $CP$  violation, the flavor of one of the  $B$  mesons must be determined in addition to fully reconstructing the other  $B$ . In this analysis,  $\pi/K$  separation is important. The momentum of the pions and kaons in the decay chain of  $B \rightarrow K^*\gamma$  is between 0.5 and 2 GeV/ $c$ , and so must be identified by a particle identification (PID) device other than the drift chamber, since  $dE/dx$  separation is only effective up to  $\sim 700$  MeV/ $c$ . The DIRC also has geometric requirements. It must be kept thin and uniform to minimize the energy resolution of the calorimeter. Also, in order to minimize the cost of the calorimeter, its radius must be kept to a minimum.

There are 144 DIRC bars composed of synthetic fused silica arranged evenly into 12 gas-tight bar boxes; within each barbox, the bars are optically isolated from each other by  $\sim 150$   $\mu\text{m}$  gap between each bar. Each bar is 17-mm-thick, 35-mm-wide, and 4.9-m-long, and is assembled from four 1.225 m pieces that are glued end-to-end. It is this glue interface that forms the basis of the study in chapter 3. Synthetic fused silica was chosen for the following reasons: its radiation hardness, long attenuation length, large index of refraction ( $n = 1.473$ ), and because within the wavelength acceptance of the DIRC (300 to 600 nm) it exhibits low chromatic dispersion. As illustrated in Fig. 2.7, the bar boxes are arranged in a 12-sided polygon around the beamline. The geometrical acceptance of the bars is  $\sim 94\%$  of the azimuthal angle and  $\sim 83\%$  of the cosine of the CM polar angle. The radial extent of the DIRC is 80 mm. Approximately 80% of the Cherenkov light is preserved through multiple bounces along the bar.

A mirror is placed at the forward end of each bar to reflect the photons back into the backward direction, where the photon detectors lie. To reduce the photon detection surface, a wedge is placed at the entrance of each bar, which reflects photons upward. It also reflects photons at large angles relative to the bar axis, which reduces the number of required photomultiplier tubes (PMTs). A common 10-mm-thick fused silica window is glued to the wedges contained in a barbox. This seals off the barbox from the water that is contained within the standoff box (SOB). This structure consists of a stainless steel cone and cylinder, and 12 sectors of PMTs. Each sector contains 896 PMTs. 6000 l of water was chosen to fill the standoff box because it's inexpensive, and has a refraction index close to that of fused silica. To reduce the magnetic field inside of the PMT sector, a bucking coil surrounds the standoff box as shown in Fig. 2.7.

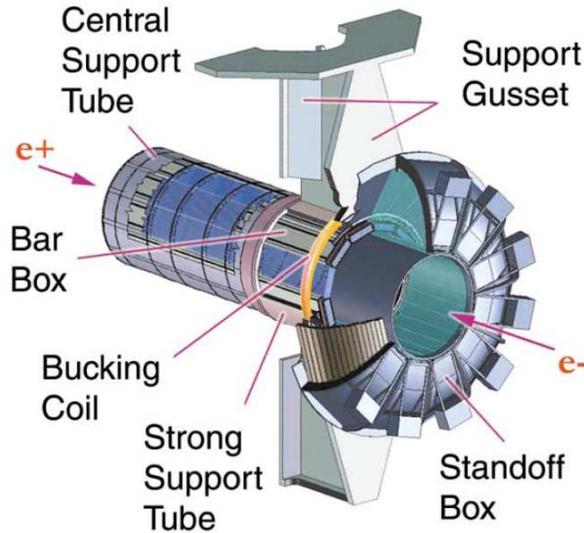


Figure 2.7: An illustration of the DIRC particle identification system. The bar boxes house the fused silica bars which serve as radiators for the traversing particles. They are supported by the Central Support Tube (which also supports the DCH). The whole structure is supported by the Strong Support Tube.

The 2.5 cm-diameter PMTs are attached to the inside of the standoff box. Each PMT has a hexagonal light catcher connected to the front of the photocathode, so that the surface area for light detection is  $\sim 90\%$  of the total area.

A schematic of a single DIRC bar, which also gives an illustration of the principle of the detector is shown in Fig. 2.8. Particles entering the fused silica bar with  $\beta = v/c \approx 1$  produce Cherenkov radiation which travels by total internal reflection down the bar. Eventually, the photons enter the standoff box and are detected by the PMTs. For a particle with  $\beta = 1$ , the expected number of photoelectrons is  $\sim 28$ . The front-end electronics (FEEs) are located outside the standoff box. There are 168 DIRC front-end boards (DFBs), each of which processes 64 PMT signals. The output of the electronics is the result of a time-to-digital converter (TDC), which waits for an L1 accept.

An unbinned maximum likelihood fit that uses both spatial and time information determines the Cherenkov angle of the track. To distinguish between signal and background, a selection based on the difference between the measured and expected photon arrival time is made, as illustrated in Fig. 2.9. This is calculated using the time-of-flight of the track, the propagation time of the photon in the DIRC bar, and the measured time of the candidate signal in the PMT. The output is a likelihood value for each of the five stable particles ( $e, \mu, \pi, K, p$ ).

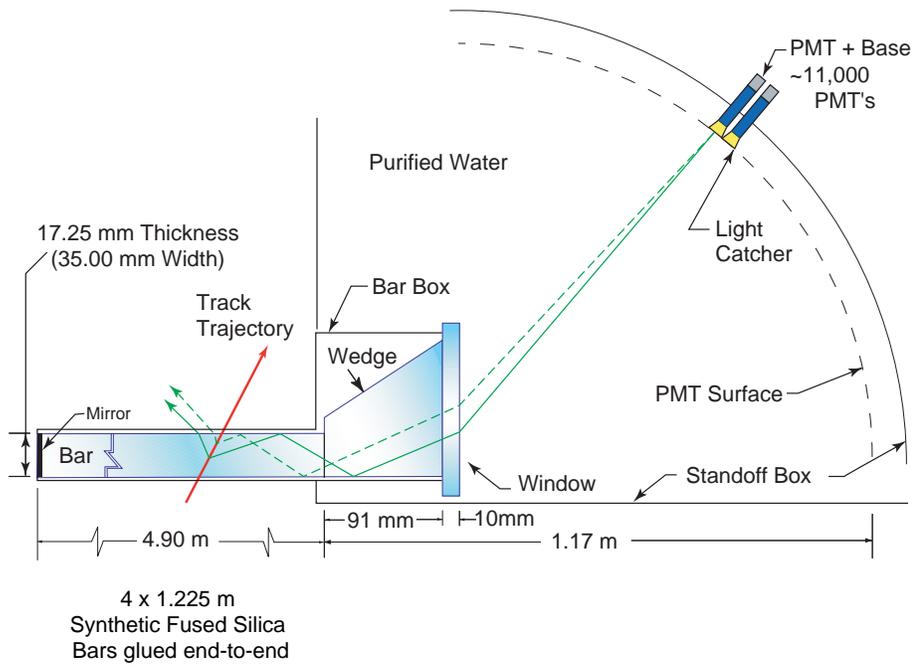


Figure 2.8: A schematic of the DIRC bar. Cherenkov radiation emitted by the particle is channeled to the PMTs, which produce a signal to be read out by the front-end electronics.

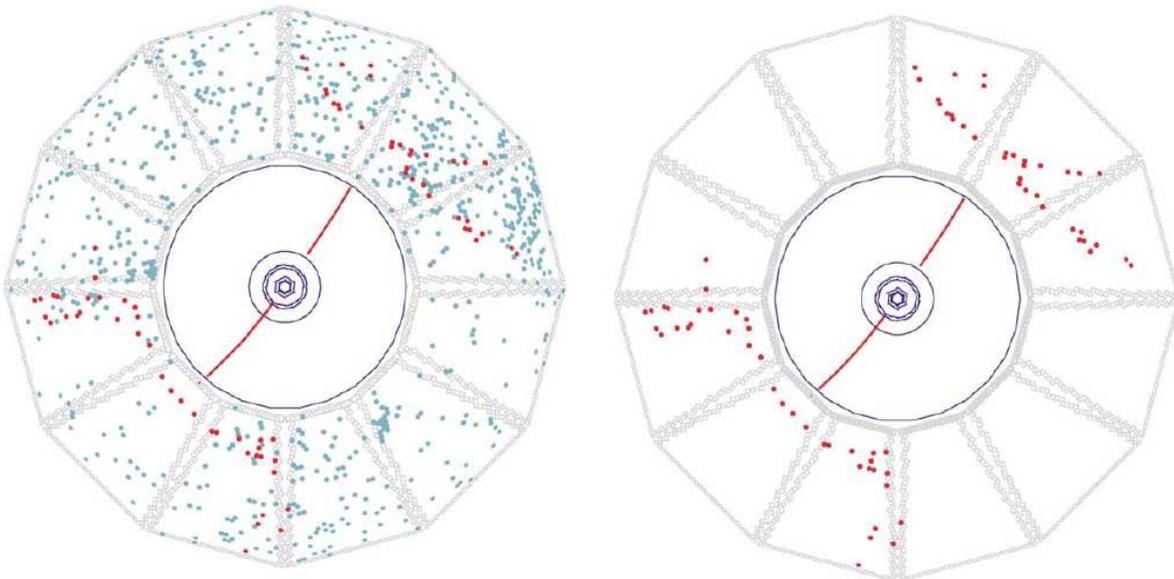


Figure 2.9: A reconstructed dimuon event with PMT signals that have a measured time within the  $\pm 300$  trigger window (left) and within 8 ns of the expected Cherenkov photon arrival time (right).

The single photon Cherenkov angle resolution is 9.8 mrad, which is determined from track uncertainties, chromatic dispersion, finite bar and PMT size, and geometric uncertainties of the DIRC (non-parallel sides of the bars, misalignment of the barbox-to-SOB, etc.). The measured time resolution is 1.7 ns, which is dominated by the intrinsic 1.5 ns time spread of the PMTs. At 3 GeV/ $c$ , a separation of  $4.2\sigma$  is achieved between kaons and pions. On average, the kaon selection efficiency is  $96 \pm 0.2\%$ , while the pion misidentification is  $2.1 \pm 0.1\%$ . The errors are statistical only.

### 2.2.4 Electromagnetic Calorimeter (EMC)

The electromagnetic calorimeter is used to measure electromagnetic showers in the energy range 20 MeV to 9 GeV. These showers are produced by particles, such as photons and electrons, which interact with the calorimeter. Electrons need to be identified to determine the flavor of the neutral  $B$  mesons in semileptonic decays. Neutral pions and etas are also detected by the calorimeter through their electromagnetic decays. In this analysis, the EMC is the only contributor to the detection of the high energy gamma in the decay  $B \rightarrow K^* \gamma$ .

The EMC consists of 6580 thallium-doped cesium iodide (CsI(Tl)) crystals that compose the barrel and endcap. 5760 crystals make up the barrel, which are arranged into 49 rings of 120 crystals each. The endcap contains 820 crystals that form 8 rings. Thallium-doped cesium iodide was chosen because it allows for good energy and angular resolution due to its high light yield (50,000  $\gamma$ /MeV) and small Moliere radius (3.9 cm). In addition, its short radiation length (1.85 cm) allows for good shower containment. The detector is illustrated in Fig. 2.10. Its solid-angle coverage in the c.m. system is 90%.

Each crystal has a tapered trapezoidal cross-section to account for the spreading of the electromagnetic shower. The area of the front face is  $\sim 4.7 \times 4.7 \text{ cm}^2$ , while the back face is  $\sim 6.1 \times 6.0 \text{ cm}^2$ . The crystal acts as a total-absorption scintillating instrument, and also totally internally reflects the majority of light that is incident upon its surface. Thus, the light is funneled to the back end of crystal, where the photodiodes lie. To aid in this process, the crystal is wrapped in white reflector material. Aluminum foil was placed on top of this to act as a Faraday shield, and further, the crystal was covered in mylar to provide electrical isolation from external support. A schematic of a crystal is given in Fig. 2.11.

Electrons or photons which enter the calorimeter generate an electromagnetic shower. To detect photons from this shower, two silicon PIN diodes of cross-sectional area  $2 \times 1 \text{ cm}^2$

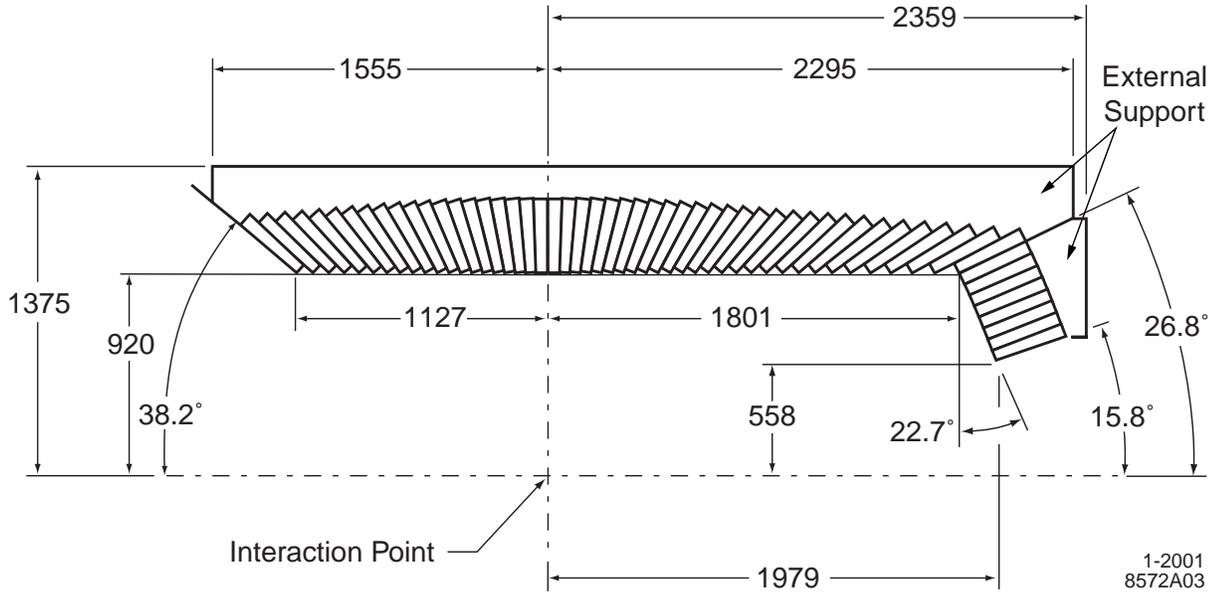


Figure 2.10: A longitudinal view of the top-half of the EMC with the dimensions given in mm. Displayed are the 56 axially-symmetric rings which compose the detector. The endcap is placed at the forward end to account for the asymmetric beam conditions.

are glued to a 1.2-mm-thick polystyrene substrate that is, in turn, glued to the back-end of the crystal. Photons that enter the depletion region of the photodiode create electron-hole pairs, which are swept away by the voltage difference of 50V. The quantum efficiency of the diode is 85% for the CsI(Tl) scintillation light. The signal is then amplified, and then extracted upon receipt of an L1 accept signal. It is translated into an amount of deposited energy, and corrected for energy losses in the back, rear, and between the crystals, as well as energy from the shower that is not associated with the cluster. A cluster is defined as an electromagnetic shower that is spread over one or more crystals. Reconstruction algorithms identify these clusters and discriminate them based on whether not they comes from a single energy maximum (called a bump), or from multiple energy maximums which make a merged cluster. In addition, these algorithms also determine whether or not a bump comes from a charged or neutral particle by attempting to associate a bump with a track. If such an association is possible, then the particle is assumed to be charged, otherwise, it is assumed to be neutral.

The energy resolution of the calorimeter is determined using control samples. With a radioactive source, it is measured at 6.13 MeV to be  $5.0 \pm 0.8\%$ , while from Bhabba scattering, the resolution is  $1.9 \pm 0.07\%$  at 7.5 GeV. The angular resolution is based an analysis of  $\pi^0$  and  $\eta$  decays, and is about 12 mrad at low energies ( $\sim 200$  MeV), and 3

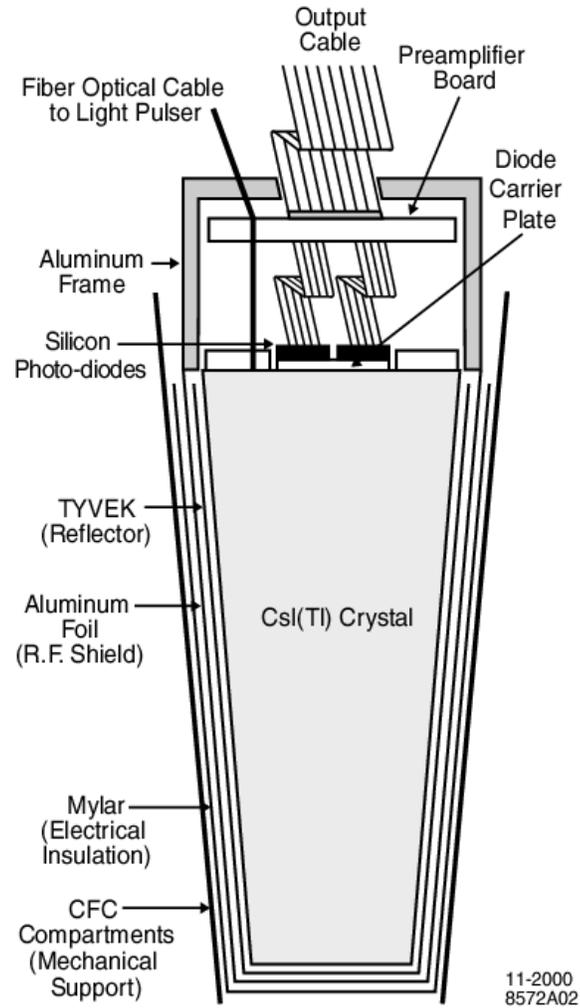


Figure 2.11: A schematic of a crystal. The readout is mounted on the back end.

mrad at high energies ( $\sim 3$  GeV). The EMC also separates electrons from hadrons. On average, the electron efficiency is 90% over the momentum range  $0.5 < p < 2$  GeV/ $c$ , while the pion misidentification rate is 0.2%.

### 2.2.5 Instrumented Flux Return (IFR)

The IFR serves two purposes. One is to provide a flux return for the magnetic field coming from the superconducting solenoid. The second is to identify muons and neutral hadrons, such as  $K_L$ 's and neutrons. Like electrons, muons play a role in identifying the flavor of a  $B$  in a semileptonic decay, as well as being used to reconstruct the  $J/\psi$ .  $K_L$ 's are detected by their hadronic showers in the iron, and are used, for example, for  $CP$  violation studies of  $B \rightarrow J/\psi K_L$ . Resistive plate chambers (RPCs) were initially used for particle detection in the IFR. However, after the first year of detector operation, the RPCs quickly degraded and were replaced by limited streamer tubes (LSTs) [9] in the central portion (barrel) of the detector.

There are a total of 806 RPC modules, which cover a total area of 2000 m<sup>2</sup>. 57 modules are in each of the six barrel sections, 108 modules are located in each of the two endcaps, and there are 32 cylindrical sections. The RPC's are placed in the gaps of the segmented steel, and comprise a total of 19 layers in the barrel, while the endcaps have 18 layers. An illustration is given in Fig. 2.12. Each planar RPC is comprised of two layers of 2-mm-thick bakelite sheets that are separated by a gap of 2 mm, which is filled with an argon-based gas mixture. Each external surface is coated with graphite, while the inner surfaces are treated with linseed oil. One layer is held at  $\sim 8$  kV, and the other is at ground. Streamers from ionizing particles are readout through capacitive strips, which are placed in orthogonal directions on either side the gap. A schematic of an RPC is given in Fig. 2.13. Signals are then digitized and passed to the ROMs, upon an L1 accept trigger.

The RPC's performance degraded from the beginning of the experiment, and the installation of the LST's began in 2004. It was completed in the fall of 2006. The inner 18 layers of the IFR were used for LST installation. However, to add in absorbing material, 6 of those layers contain brass, while the other 12 contain the LST tubes. A schematic of a tube is shown in Fig. 2.14. Each tube consists of about 7 or 8 cells, whose dimensions are 17mm wide, 15mm high, and 380 mm long. In the center of each resides a gold-plated anode wire, which is held in place by six wire holders placed inside the cell. The walls are coated internally with graphite paint and held at ground, while the wire is at 5.5 kV.

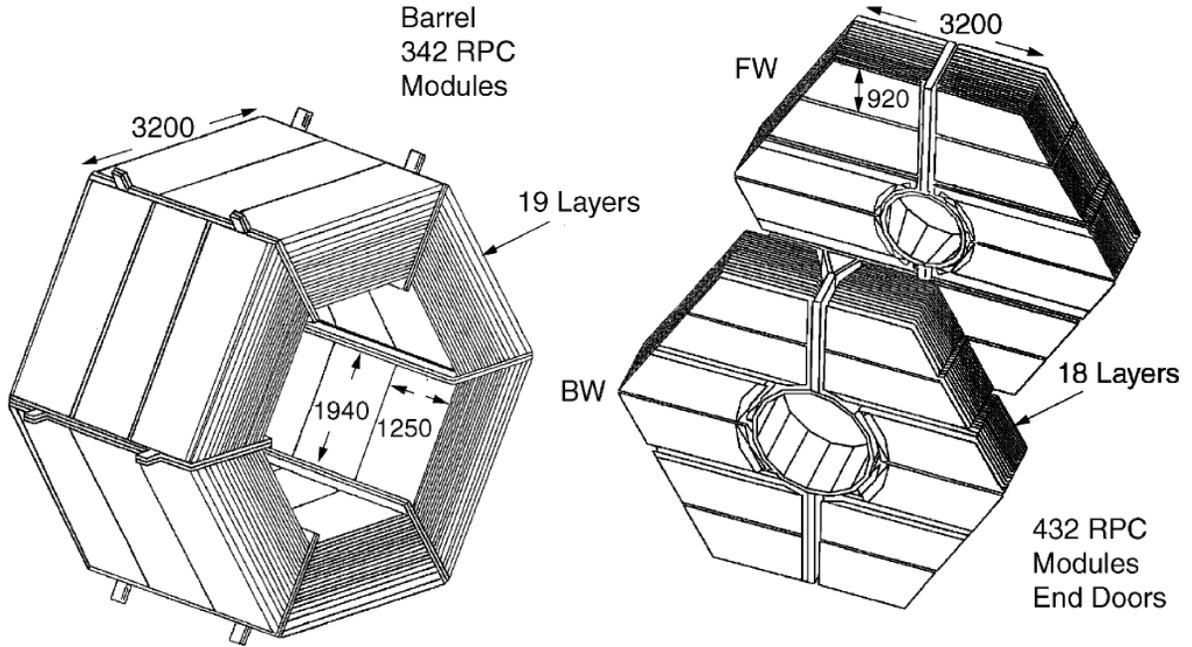


Figure 2.12: An overview of the IFR, in which the barrel and endcaps are indicated. All dimensions are given in millimeters.

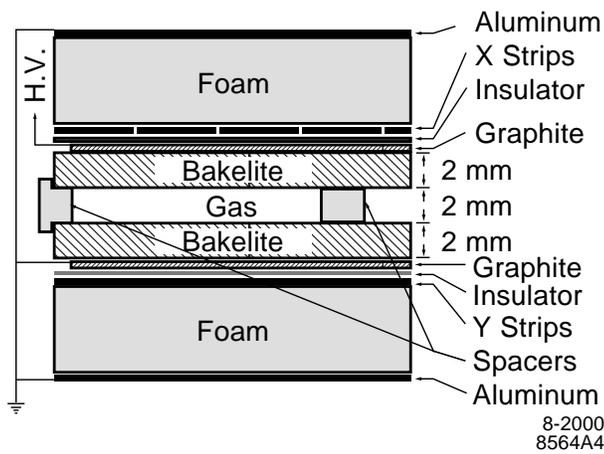


Figure 2.13: Cross section of a planar RPC.

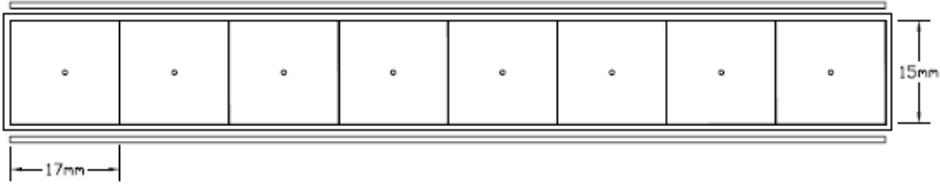


Figure 2.14: Cross section of an LST module.

The inside of each cell also contains a gas mixture which is argon, isobutane, and carbon dioxide in the ratio 3:8:89. A particle passing through the gas ionizes it and causes a streamer discharge, which can be readout from the wire, while also inducing a charge on a plane below the tube. Taken together with position of the layer, the full 3d information from the hit can be determined.

From 2000 to 2005 the muon efficiency and pion rejection rates steadily declined. For example, for a sample of high energy muons ( $2 \text{ GeV}/c < p < 4 \text{ GeV}/c$ ), and a pion rejection rate of 96%, the muon efficiency decreased from 88% to 76% using a neural net based muon selection algorithm. The installation of the LST's increased the muon efficiency to 92% for the same pion rejection rate.

### 2.2.6 Trigger

The trigger system is designed to select events of interest ( $e^+e^- \rightarrow q\bar{q}, e^+e^- \rightarrow l^+l^-$ ), while at the same time efficiently rejecting background events. The trigger efficiency for  $B\bar{B}$  events must be above 99%, while the efficiency for continuum events is required to be at least 95%. The beam backgrounds must held to a minimum, as the total event rate must be under 120 Hz. The trigger is implemented in two stages. The first stage, L1, is in hardware, while the second, L3, is based in software.

The L1 trigger uses three detectors as input, which are the drift chamber trigger (DCT), the electromagnetic trigger (EMT), and the instrumented flux return trigger (IFT). The former two satisfy all of the trigger requirements independently with high efficiency, and so are largely redundant, while the IFT is used mainly for diagnostic purposes. All three L1 triggers generate trigger primitives, which contain information about the position and energy of a particle. These primitives are passed to the global trigger (GLT) every 134 ns, which processes them and sends them to the Fast Control and Timing System (FCTS). This system can mask or prescale any trigger. If one remains, then an L1 accept signal is sent out for the event to be readout.

The input to the DCT consist of one bit for each of the 7104 DCH cells, where the bit contains time information from the sense wire associated with that cell. This information is passed to 24 Track Segment Finder (TSF) modules, which find track segments in an adjacent set of cells that span all four layers of a superlayer. The Binary Link Tracker (BLT) then receives these track segments and forms complete tracks. In addition, eight transverse momentum discriminators (PTDs), determine if the track segments have a  $p_t$  greater than some minimum value. The output of the DCT is a set of trigger primitives which categorize of the output of the BLT and PTDs into short tracks (tracks reaching DCH superlayer 5), long tracks (tracks reaching DCH superlayer 10), and high  $p_t$  tracks ( $p_t > 800 \text{ MeV}/c$ ).

The EMT treats the calorimeter as if it divided into 280 towers,  $7 \times 40$  ( $\theta \times \phi$ ). The barrel has 240 towers, each of which contains 24 crystals in a  $8 \times 3$  ( $\theta \times \phi$ ) array. The endcap contains 40 towers, each of which contains 19-22 crystals. All crystal energies in each tower above 20 MeV are sent to the EMT. 10 Trigger Processor Boards (TPBs) determine the energies in the 40  $\phi$ -sectors and different ranges in  $\theta$ . The output of the EMT is five trigger primitives which categorize the crystal energies using different minimum thresholds, and locations in the EMC.

The function of the L3 trigger software is to reconstruct and classify events using the output of the L1 triggers, as well as the complete event data. The trigger is executed in three phases. The first phase defines the L3 input lines. In the second, classification tests are performed which produce pass-fail output flags. These tests are comprised of two algorithms, which are a track finding algorithm for the DCH and a clustering algorithm for the EMC. The DCH algorithm determines the five track helix parameters with those tracks with  $p_t > 250 \text{ MeV}/c$ , while the EMC algorithm identifies energy clusters with requirements such that minimum ionizing particles can be found. In the last phase, L3 output lines are formed.

The efficiency of the two trigger levels are derived from Monte Carlo simulation. The combined L1 and L3 triggers achieve an efficiency of above 99.9% for  $B\bar{B}$  events, and above 92% for continuum events.

# Chapter 3

## Photon Background in DIRC Fused Silica Bars

### 3.1 Introduction

As mentioned in Section 2.2.3, the DIRC is responsible for particle identification beyond a momentum of  $\sim 700 \text{ MeV}/c$ . When a particle passes through the DIRC, it emits photons at an angle with respect to the track that is dependent on its velocity and the index of refraction of quartz. Figure 3.1 shows this measured angle subtracted from the expected angle per reconstructed photon for  $e^+e^- \rightarrow \mu^+\mu^-$  events for both data and Monte Carlo (MC) simulation. Approximately only 60 % of the observed background is explained by Monte Carlo.

Understanding the photon background has consequences for particle identification. The DIRC uses an unbinned extended maximum likelihood fit that incorporates both space and time measurements, as well as the number of photons. The likelihood per track is given by

$$L \propto \mu^{N_\gamma^{assoc}} \frac{e^{-\mu}}{N_\gamma^{assoc}!} \prod_{i=1}^{N_\gamma^{assoc}} \left[ (1-r)S_i(\theta_C, \delta t) + rB_i \right] \quad (3.1)$$

where  $r$  is the background fraction,  $N_\gamma^{assoc}$  is the number of associated photon hits,  $\mu = N_\gamma^{exp} + rN_\gamma^{assoc}$  is the sum of the expected number of signal photons  $N_\gamma^{exp}$  and background photons.  $S_i(\theta_C, \delta t)$  is the signal estimator that is a function of the measured Cherenkov angle  $\theta_C$  and the difference between measured and expected photon propagation time  $\delta t$ , and  $B_i$  is the flat background contribution. Due to poor understanding

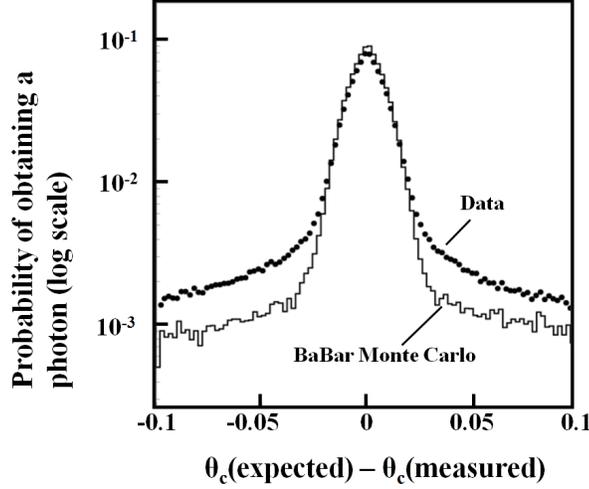


Figure 3.1: The single photon angle distribution for muon tracks as reconstructed in the DIRC (dots) and the Monte Carlo prediction using the DIRC MC program (solid line). The signal is distributed around zero, while the background photons also lie outside the Cherenkov cone. The Monte Carlo only explains roughly 60% of the observed background.

of the background, the mean number of background photons  $rN_{\gamma}^{assoc}$  must be obtained through the fit. However, this number can be understood more completely by knowing the expected number of background photons.

The goal of this analysis, therefore, is to achieve a greater understanding of the origin of the photons which are not understood by the Monte Carlo. This is done by performing an experiment, described in Section 3.2, in which cosmic rays are used to generate Cherenkov radiation. A detailed Monte Carlo simulation (independent of the Babar DIRC Monte Carlo) is then used to simulate the experiment. The conclusion of this study is that there are two major mechanisms that contribute to the background. They are the Cherenkov photons generated by delta rays originating from the primary charged particles (cosmic rays) and reflections from the EPOTEK-301-2 glue [10] and fused silica interfaces.

## 3.2 Experimental Setup

The experimental setups are shown in Fig. 3.2. Three standard DIRC bars with dimensions  $17 \text{ mm} \times 31 \text{ mm} \times 1225 \text{ mm}$  are glued together with EPOTEK-301-2 glue, where the glue joint has a thickness of  $25 \text{ }\mu\text{m}$ . A photomultiplier tube (PMT) type Quantacon XP2020 Philips [11] is attached directly to each bar with the EPOTEK-301-2 glue. On this side of the bar, the surfaces are composed of fused silica, EPOTEK-301-2

glue, Borosilicate glass, and the bi-alkali photocathode. In setup of Fig. 3.2a, the bar end opposite to the photomultiplier is equipped with a mirror, which is air-coupled with a spring applying pressure against the bar. The charged particles enter the bar with an angle of  $56.5^\circ$  with respect to the bar normal and  $\sim 50$ -cm from the left end of the bar in Fig. 3.1, pointing away from the phototube. The particle trajectory is defined by two entrance scintillation counters limiting the angular acceptance to  $\pm 5^\circ$  with respect to the mean of  $56.5^\circ$ . The 12-in-thick lead shielding in front of the exit counter provides a selection of the minimum muon energy of 0.4 GeV. The track angle ensures that all internally reflected Cherenkov light first travels through the bar away from the PMT and reflects from the mirror before it arrives in the phototube (or leaves the bar because the internal reflection condition is not fulfilled). This leaves a time window of about 36 ns for collecting light from the bar before the Cherenkov signal. In setup Fig. 3.1b, the Cherenkov photons are efficiently absorbed by the photon trap so that a study of the background photons which would normally arrive in coincidence with the Cherenkov peak is possible, essentially extending the observation time interval to 70 ns. The photon trap is an aluminum box coupled to the bar with a pipe surrounding the bar end. Both devices are filled with a fluid which matches the refraction index of fused silica and their walls are covered with photon-absorbing cloth.

### 3.3 Acquired Data

The PMT signal was amplified  $\times 10$  with a LeCroy fast amplifier and the output recorded with a HP digital scope read out by a MAC ICC computer with CAMAC-based GPIB interface. Figs. 3.3 and 3.4(a) show the raw waveforms for a single event with the mirror and the photon trap, respectively. In Fig. 3.3 one clearly observes the Cherenkov signal arriving at channel number 375. The earlier activity in front of that channel can be attributed to background photons. In Fig. 3.4(a) the Cherenkov signal is absorbed efficiently by the photon trap. The remaining pulses slowly diminish toward the end of the 70 ns window. A peak finding algorithm was applied offline: the waveform is differentiated channel by channel and a peak is localized if the waveform starts dropping for at least five consecutive channels. The result was checked with a deconvolution algorithm, which takes the single-photon pulse shape into account. The deconvolution algorithm assumes a standard PMT amplifier pulse shape of the form  $t \exp(-t/\tau)$ , where the shaping constant  $\tau$  is assumed to be 20ns. The raw scope waveform is converted to the deconvoluted one

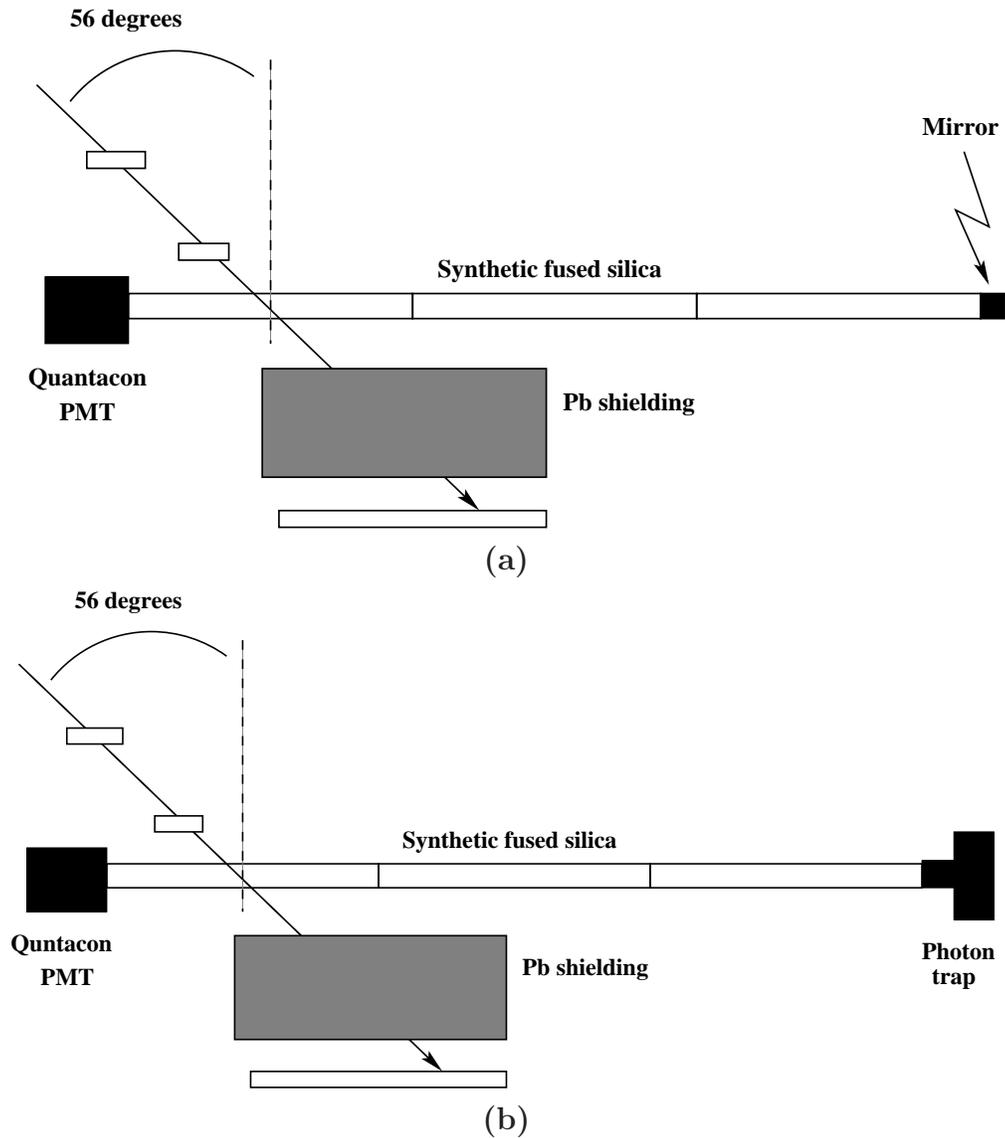


Figure 3.2: Experimental setups used to study the photon background. The Cherenkov signal propagates to either (a) the mirror or (b) the photon trap. In the former case, a window of  $\sim 36$  ns is available to study the early photon background activity before the Cherenkov signal returns; in the latter case a  $\sim 70$  ns window is available

using the equation

$$d_i = r_i - 2 \exp^{-T/\tau} r_{i-1} + \exp^{-2T/\tau} r_{i-2}, \quad (3.2)$$

where  $T$  is the scope sampling time (0.4ns),  $r_i$  is the  $i$ -th PMT output sample, and  $d_i$  is the  $i$ -th deconvoluted sample. The deconvoluted waveform is then subject to a threshold cut to eliminate unwanted noise pulses. The output is shown in Fig. 3.4(b). The two algorithms produced similar results for the number of peaks and peak positions.

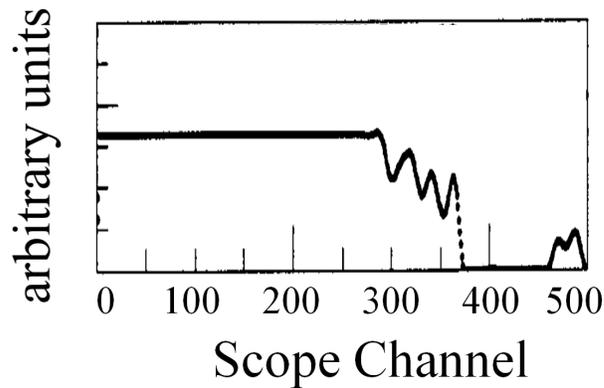


Figure 3.3: Digital scope output from a single event for the mirror setup. The horizontal axis is time in terms of scope channels (1 scope channel = 0.4 ns) and the vertical axis is the amplitude. The large Cherenkov pulse arrives at channel 375, while the background pulses arrive earlier.

Fig. 3.5 shows the result of the peak finding algorithm for the mirror setup. The first peak near channel 115 is due to delta-ray electrons (electrons which have been liberated from an atom by the passing muon) traveling fast enough to produce Cherenkov photons by themselves in somewhat randomized directions. As will be discussed later, the shoulder at channel 130 is explained by the reflection of photons at the first glue plane. The Cherenkov pulse arrives at channel 150. There is a drop-off of photoelectrons after this channel because the counting algorithm is overwhelmed by the number of photoelectrons. The background features show in Fig. 3.5 are also evident in Fig. 3.6, which shows the result of the peak finding algorithm for the photon trap setup. However, in this case, the Cherenkov pulse is absent. In addition, the application of a LeCroy TDC allowed one to make a leading edge “single hit” distribution. This was consistent with a leading edge “single hit” algorithm applied to the recorded pulse shapes.

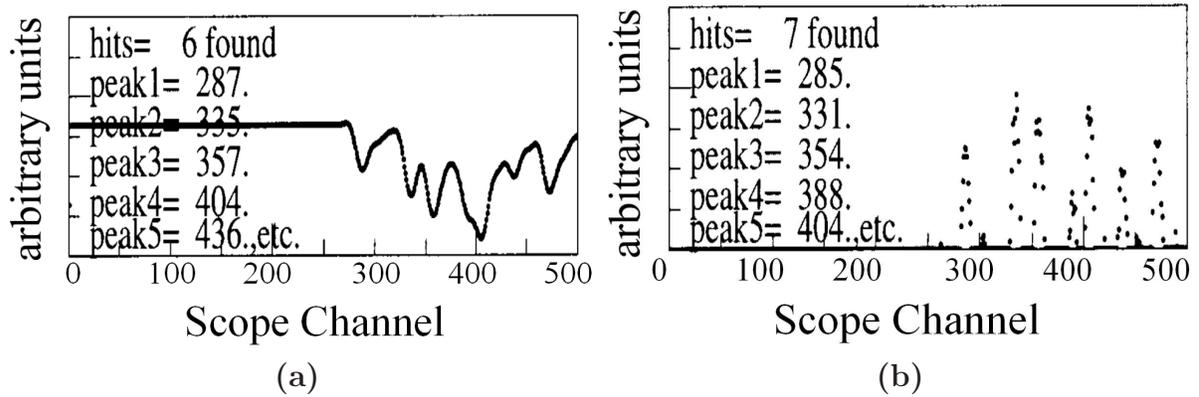


Figure 3.4: Digital scope output from a single event for the photon setup (a), and the output of the deconvolution algorithm (b). The horizontal axis is time in terms of scope channels (1 scope channel = 0.4 ns) and the vertical axis is the amplitude. The lack of the Cherenkov pulse is evident as a result of the photon trap. Also shown are the number and positions of the peaks using the peak finding algorithm (a), and the deconvolution algorithm (b). The two algorithms are consistent with each other.

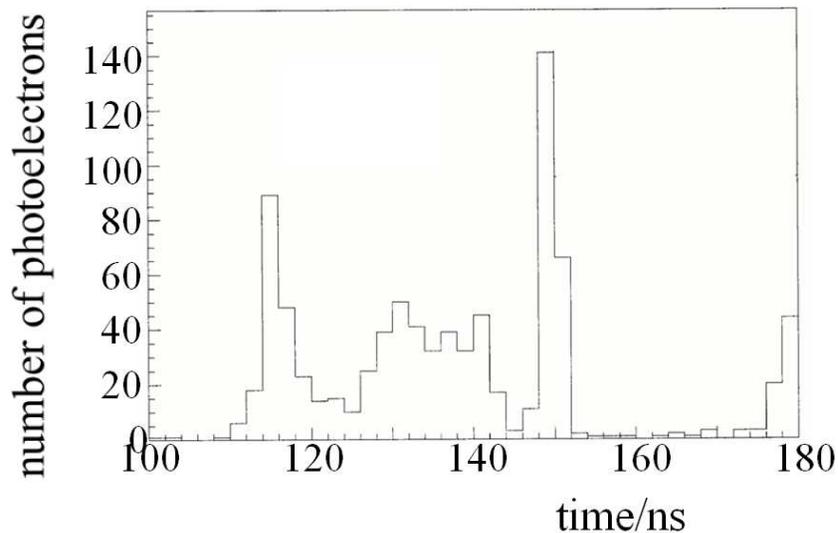


Figure 3.5: Measured time distribution of photon pulses obtained from the peak finding algorithm using the mirror data. The Cherenkov signal arrives at channel 150. The hits before this time are all considered background.

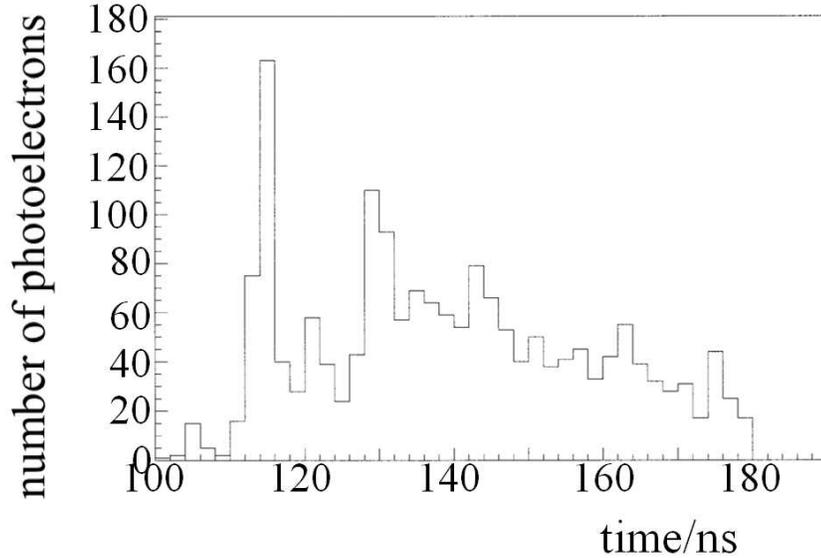


Figure 3.6: Measured time distribution of photon pulses obtained from the peak finding algorithm using the photon trap data. The Cherenkov signal is clearly absent. Therefore, all hits are considered background.

## 3.4 Supplemental Measurements

### 3.4.1 Measurement of Reflectivity of EPOTEK-301-2 Glue

The index of refraction of the EPOTEK-301-2 glue as a function of wavelength was measured in a separate experiment using four different wavelengths [12]. Based on the results, the Fresnel reflectivity at a particular wavelength is calculated for the Transverse Electric (TE) mode. The reflectivity was also directly measured with a 442 nm laser [12]. The reflectivity as a function of the angle of incidence to the surface quartz-glass for the different methods is shown in Fig. 3.7. The measured reflectivity was significantly higher than what was predicted by the Fresnel theory. For our Monte Carlo simulation, a second order polynomial curve was adjusted to describe the distinct features of the arrival-time spectrum in the data. This yielded the curve in Fig. 3.7 which is consistent with the measured values of the reflectivity.

### 3.4.2 Measurement of Scintillation

The scintillation rate of quartz was measured with a Fe source which primarily emits 5.9 keV X-rays. Its emission of energetic gamma rays, which can create Compton electrons with enough energy to produce their own Cherenkov photons, is negligible. The only light

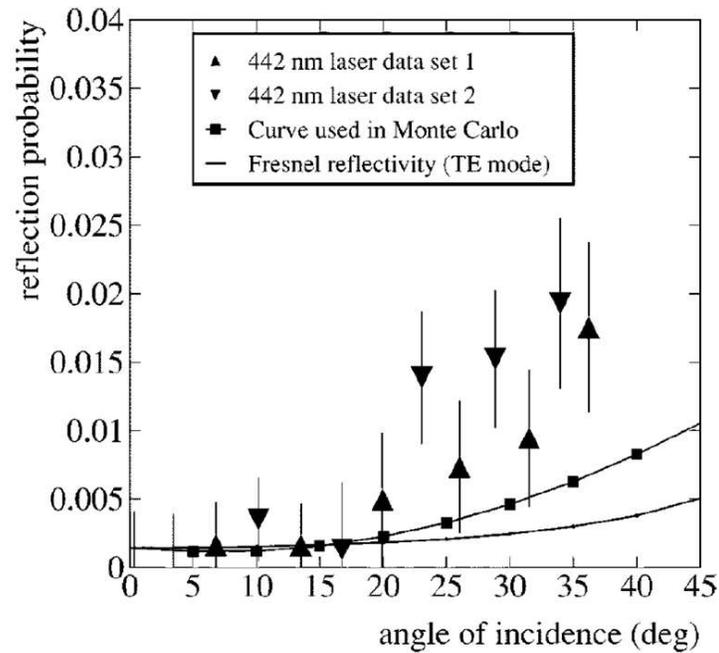


Figure 3.7: The Fresnel curve is based on the measurement of the index of refraction. Also shown is the data from two different measurements of the reflectivity per bounce from the glue/silica interface. Finally, the curve with square references is used in the Monte-Carlo program that simulates the 4 m long bar cosmic ray experiment.

production mechanism is scintillation light by the photoelectric effect: the X-rays have enough energy to kick off a bound electron, which can then travel to nearby atoms and excite them, thus producing scintillation photons. The source emits X-rays at a rate of  $3.6 \times 10^7$  counts/min. When it was placed directly on the bar, a rate of  $(1.96 \pm 0.3) \times 10^3$  counts/min was observed in the PMT next to it. According to the Monte Carlo simulation, the PMT acceptance was 5%. Therefore, the probability for a single 5.9 keV X-ray to produce a scintillation photon is  $(1.96 \times 10^3 / 3.6 \times 10^7) / 0.05 \sim 1.1 \times 10^{-3}$ . Assuming that all of the energy deposited by the traversing muon is used to create scintillation photons (the minimum ionizing muon deposits 13.5 MeV) there would be about  $(1.1 \times 10^{-3} / .0059) \times 13.5 \sim 0.12$  detected scintillation photons per muon. This number is negligible compared to the  $\approx 5$  observed background photons in a time window of 0-70 ns.

## 3.5 Monte Carlo Simulation

Both experimental setups described in Section 3.2 were simulated based on the Monte Carlo program of a single DIRC bar geometry [13]. This generates a photon spectrum according to the physics of Cherenkov radiation and traces photons from their point of generation through the bar taking into account wavelength dependent bulk-material attenuation, bar surface scattering, and mirror reflection. Cherenkov light is generated for all charged particles above their Cherenkov threshold. Delta-ray generation from the primary particle track is simulated using the FLUKA package [14] and the optical properties of the glue.

### 3.5.1 Delta Ray Simulation

The simulated light yield originating from delta rays only is shown in Fig. 3.8 together with the data from our mirror setup. The data is the combined mirror and photon trap data, since they agree in their characteristics within the first 30 ns. The Monte Carlo curve shows the absolute number of photons for the corresponding number of data tracks. It is evident that delta rays describe the time behavior of the earliest background photons in the 30 ns window before the signal while the overall spectrum is not well reproduced. It is interesting to note that scintillation photons emitted in random direction along the particle track arrive at a similar time spectrum only that their rate is strongly suppressed compared to delta rays.

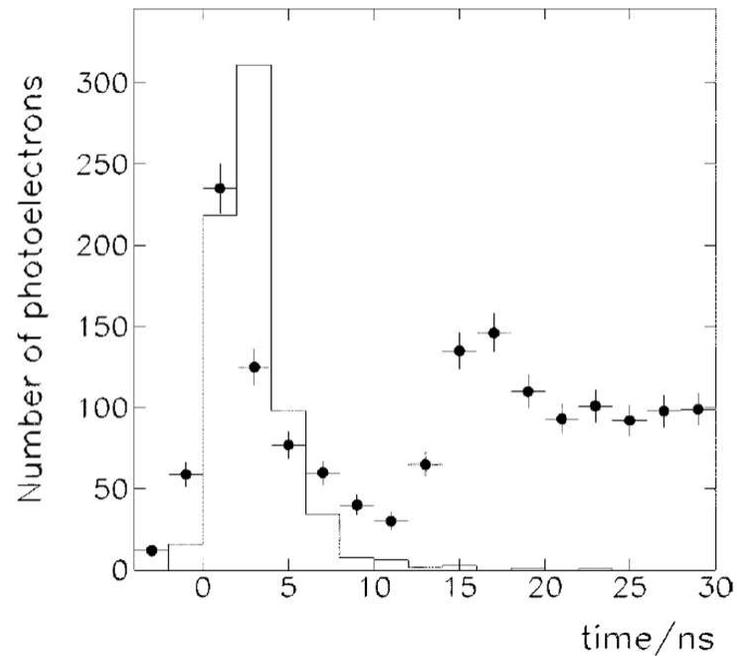


Figure 3.8: Time distribution of photon background pulses in the first 30 ns. The Monte Carlo only simulates delta rays and is represented by the smooth line, while the data (a combination of both the mirror and photon trap data) is given by the dots. Delta rays are the dominant component in the first background photon peak.

### 3.5.2 Reflectivity of Glue

It is obvious from Fig. 3.8 that delta rays do not explain the background features satisfactorily. The distance to the first glue joint between the bars corresponds roughly to the 15 ns, where the shoulder shows up in the data. Therefore, to more accurately describe the data, the reflection of photons was included from the glue/fused silica interfaces which so far was neglected in the DIRC simulations. The modified spectrum is shown in Fig. 3.9(a) together with both the mirror and photon trap data. The Monte Carlo describes the first photon arrival peak, due to delta rays, well. The ratio between the second glue reflection peak and the first delta-ray peak was tuned using the reflectivity per angle of incidence (Fig. 3.7). The analysis of the reflection was done as follows. First, the measured refraction index of the glue [12] and that of fused silica were used and the TE and Transverse Magnetic (TM) reflections were calculated using the Fresnel theory. In this case, the second peak in Fig. 3.9(a) was underestimated by a factor of five. Second, the refraction index of the glue was tuned to obtain the best agreement with the 4-m-long bar data. Again, the TE and TM modes were calculated appropriately, according to Fresnel theory. However, following this procedure, unphysical values of the refraction index of the EPOTEK-301-2 glue were obtained. Third, the data from the direct measurement of the reflectivity of the glue/fused silica interface for the TE mode at 442 nm was fit to. The results of this fit were used to model the reflection. In this case, the Monte Carlo overestimated the size of the second peak in Fig. 3.9(a). Finally, the fit was tuned to the direct measurement of the reflectivity to achieve agreement with the data in Fig. 3.9(a). The result of this tuning is shown in Fig. 3.7. The tuned curve explains the basic features of the measured data, although the time between the first and second peaks in the Monte Carlo is  $\sim 1 - 2$  ns greater than the time in the data. In summary, the reflection at the glue/fused silica interface may not follow the simple Fresnel theory. In fact, the deviation from this theory may point to a more complex situation at the interface between bars.

Fig. 3.9(b) shows the photon-trap data extending the time interval without a Cherenkov-light signal to 70 ns together with the Monte Carlo generated with the tuned reflectivity curve. The basic features of the distribution are reproduced well. It shows that nearly all background photons are collected within the 70 ns.

Using the reflectivity-tuned Monte Carlo the most probable number of background photons in the first 30 ns is determined to be three, which is shown in Fig. 3.10(a). For the full 70-ns interval, the number of background photons is five (Fig. 3.10(b)). About 96 photoelectrons from the Cherenkov signal is expected at the dip angle of  $56.5^\circ$ , so that

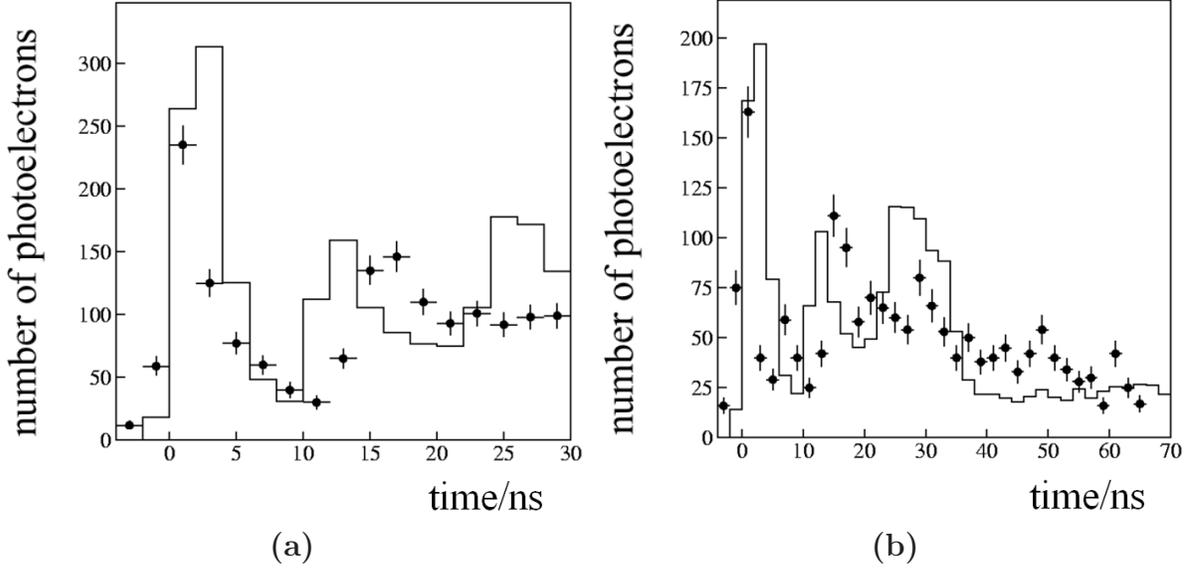


Figure 3.9: Time distribution of photon background pulses in the first 30 ns (a) and 0-70 ns (b), where the glue reflectivity is included in the simulation. The Monte Carlo is represented by the smooth line, while the data is given by the dots. For (a), the data is a combined of the mirror and photon trap data, while the data in (b) consists solely of the photon trap data.

the most probable number represents approximately 5% of the proper DIRC Cherenkov signal. However, the distribution has a long tail caused by the delta-ray contribution as shown in Fig. 3.10(a) and (b). Since the real data have a finite pulse shaping time, there is a natural upper limit on the number of pulses one can measure and the tail is suppressed.

## 3.6 Conclusion

It has been shown that the two major contributions to the photon background in DIRC fused silica bars in the BABAR experiment are as follows: a) Cherenkov photons generated by delta-ray electrons and b) reflection of all photons from the silica/glue interfaces. The reflection is caused by the difference in the refraction index of the glue and fused silica. The light yield due to scintillation is negligible. The results of the glue reflection when put into the *BABAR* Monte Carlo for  $e^+e^- \rightarrow \mu^+\mu^-$  events are shown in Fig. 3.11. The effect of the glue is that it reflects photons coming from the main Cherenkov peak, which causes them to come outside of the timing selection for the DIRC ( $\delta t = 8\text{ns}$ ). Thus, it causes the Cherenkov angle distribution to be somewhat less populated than without the glue.

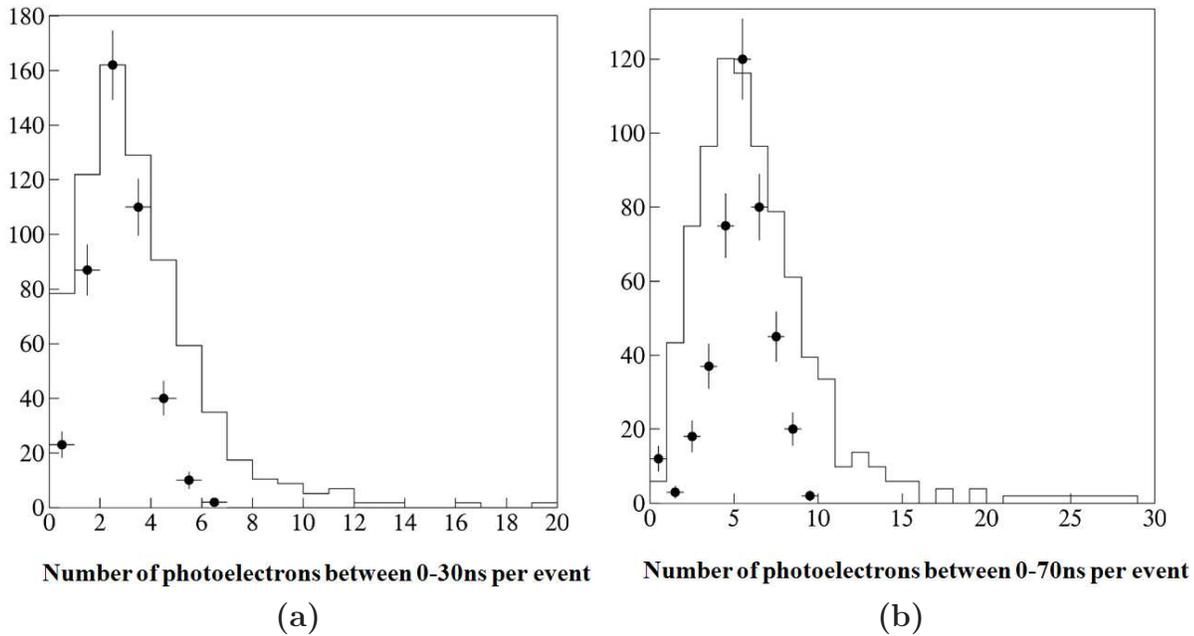


Figure 3.10: Multiplicity distribution of background photons for the first 30 ns (a) and 0-70 ns (b). The Monte Carlo is represented by the smooth line, while the data is given by the dots. For (a), the data is a combined of the mirror and photon trap data, while the data in (b) consists solely of the photon trap data. The data and Monte agree well in terms of most probable number, which is three for (a) and five (b).

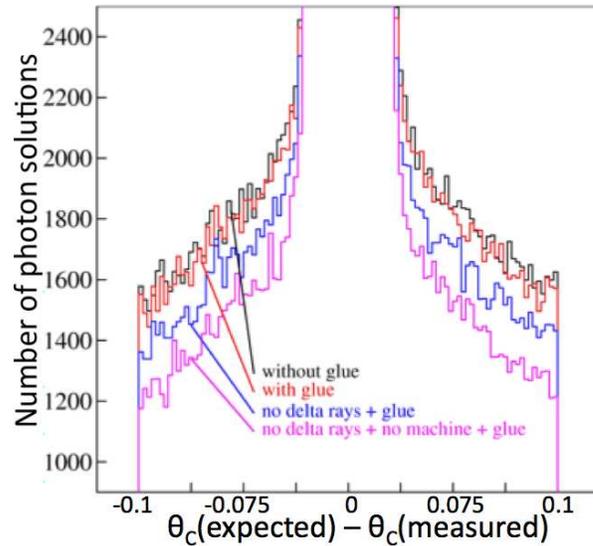


Figure 3.11: The single photon angle distribution for muon tracks as reconstructed in the DIRC for *BABAR* Monte Carlo. The effects of the glue are shown. Also shown for reference are the effects of no delta rays and no machine background (photon background coming from the beam interactions).

# Chapter 4

## Theory

This chapter provides an overview of the classical SM Lagrangian relevant to  $B \rightarrow V\gamma$ , along with discussing the theoretical calculations for the rates and asymmetries for  $B \rightarrow V\gamma$ .

### 4.1 Standard Model Lagrangian

The SM is a renormalizable quantum field theory that makes predictions in terms of parameters that must be experimentally obtained, such as the masses and couplings. It describes the electromagnetic, weak, and strong interactions based on the principles of gauge symmetries, Lorentz invariance, and renormalizability. The gauge group is a Lie group, according to which the fields transform according to a unitary finite-dimensional representation. The theory of the strong interactions is described by the non-abelian gauge group  $SU(3)_C$ , where the  $C$  represents the color charge, under which only the quarks transform non-trivially. The electromagnetic and weak interactions are described by the group  $SU(2)_L \times U(1)_Y$ , where the  $L$  indicates that the group acts only on the left-handed components of the field, and  $Y$  represents weak hypercharge. The quantum number associated with group  $SU(2)$  is called weak isospin. The product of all three groups,  $SU(3)_C \times SU(2)_L \times U(1)_Y$ , forms the gauge group of the SM under which all fields transform. To endow mass, this group is spontaneously broken to  $SU(3)_C \times U(1)_Q$ , where  $Q$  is the electric charge, by the vacuum expectation value of the Higgs field. This provides the SM a way to give mass while at the same time maintaining renormalizability, and provides the last ingredient to give the SM Lagrangian its final form.

### 4.1.1 Fermion-Gauge Boson Interaction Lagrangian

The contributions to penguin processes are described by the following Lagrangian, which, in terms of mass eigenstates, gives the fermion ( $\nu, e, u, d$ ) interactions with the gauge bosons. The bosons are denoted by  $Z_\mu, A_\mu, G_\mu^a$ , and  $W_\mu^\pm$ , where the  $\mu$  and  $a$  denote the vector and color index respectively. It is given by

$$\begin{aligned}
\mathcal{L}_f = & \sum_{f=\nu,e,u,d} \sum_{j=1}^3 \left[ \bar{f}^j i \not{\partial} f^j + \frac{e}{s_W c_W} (\tau_f^3 \bar{f}_L^j \gamma^\mu f_L^j - s_w^2 Q_f \bar{f}^j \gamma^\mu f^j) Z_\mu - e Q_f \bar{f}^j \gamma^\mu f^j A_\mu \right] \\
& + \sum_{f=u,d} \sum_{j=1}^3 g_s \bar{f}^j \gamma^\mu T^a f^j G_\mu^a + \sum_{i,j=1}^3 \frac{e}{\sqrt{2} s_W} \left( V_{ij} \bar{u}_L^i \gamma^\mu d_L^j W_\mu^+ + V_{ij}^\dagger \bar{d}_L^i \gamma^\mu u_L^j W_\mu^- \right) \\
& + \sum_{j=1}^3 \frac{e}{\sqrt{2} s_W} \left( \bar{\nu}_L^j \gamma^\mu e_L^j W_\mu^+ + \bar{e}_L^j \gamma^\mu \nu_L^j W_\mu^- \right), \tag{4.1}
\end{aligned}$$

where  $j$  denotes the family index,  $s_W = \sin \theta_W$  and  $c_W = \cos \theta_W$ , where  $\theta_W$  is the weak mixing angle,  $\tau_f^3$  is the third component of weak-isospin,  $Q_f$  is the charge of the fermion, and  $V$  is the CKM matrix, which is unitary by construction. Since the rotation matrices that transform the gauge eigenstates into the mass eigenstates have dropped out of the terms involving the fermion fields and the neutral gauge bosons, there are no FCNCs at tree level. FCNCs therefore occur at the loop level, and even then are further suppressed because of the GIM mechanism. This is the observation that, owing to the unitarity of the CKM matrix, the amplitude of FCNCs would vanish in the limit of equal quark masses of the intermediate states. Since the process  $b \rightarrow s\gamma$  is a FCNC, the GIM mechanism applies to it.

The  $V - A$  structure of the weak interactions provides the mechanism to violate the discrete symmetries of parity  $P$  and charge conjugation  $C$ . Under the transformation  $P$ ,  $C$ , and time reversal  $T$ , a spinor  $\psi$  and vector field  $V_\mu$  transform as

$$\begin{aligned}
P\psi(t, \vec{x})P^{-1} &= \gamma^0 \psi(t, -\vec{x}), & PV_\mu(t, \vec{x})P^{-1} &= V^\mu(t, -\vec{x}), \\
C\psi(t, \vec{x})C^{-1} &= i\gamma^2 \gamma^0 \bar{\psi}^T(t, \vec{x}), & CV_\mu(t, \vec{x})C^{-1} &= -V_\mu(t, \vec{x}), \\
T\psi(t, \vec{x})T^{-1} &= \gamma^1 \gamma^2 \psi(-t, \vec{x}), & TV_\mu(t, \vec{x})T^{-1} &= V^\mu(-t, \vec{x}).
\end{aligned} \tag{4.2}$$

Using Eq. 4.2, one can show that the terms in Eq. 4.1 that couple the up- and down-type quarks to the  $W$  bosons transform under  $P$  and  $C$  as

$$\begin{aligned}
P \left( V_{ij} \bar{u}_L^i \gamma^\mu d_L^j W_\mu^+ + V_{ij}^\dagger \bar{d}_L^i \gamma^\mu u_L^j W_\mu^- \right) P^{-1} &= V_{ij} \bar{u}_R^i \gamma^\mu d_R^j W_\mu^+ + V_{ij}^\dagger \bar{d}_R^i \gamma^\mu u_R^j W_\mu^-, \\
C \left( V_{ij} \bar{u}_L^i \gamma^\mu d_L^j W_\mu^+ + V_{ij}^\dagger \bar{d}_L^i \gamma^\mu u_L^j W_\mu^- \right) C^{-1} &= V_{ij} \bar{d}_R^j \gamma^\mu u_R^i W_\mu^- + V_{ij}^\dagger \bar{u}_R^j \gamma^\mu d_R^i W_\mu^+.
\end{aligned} \tag{4.3}$$

Essentially, the operators  $C$  and  $P$  change the chirality of the current that is coupled to the  $W$  boson. Since the right-handed components of the up- and down-type quarks do not couple to  $W^\pm$ ,  $C$  and  $P$  are said to be violated maximally. Combining the operations in Eq. 4.3, the transformation under  $CP$  is

$$CP \left( V_{ij} \bar{u}_L^i \gamma^\mu d_L^j W_\mu^+ + V_{ij}^\dagger \bar{d}_L^i \gamma^\mu u_L^j W_\mu^- \right) CP^{-1} = V_{ji}^\dagger \bar{u}_L^i \gamma^\mu d_L^j W_\mu^+ + V_{ji} \bar{d}_L^i \gamma^\mu u_L^j W_\mu^-. \tag{4.4}$$

Therefore, the condition for  $CP$  violation in the SM becomes

$$V \neq V^* \tag{4.5}$$

so that the reality of the CKM matrix implies  $CP$  conservation. The operation  $CPT$  is a symmetry of the SM by the  $CPT$  theorem, which states that  $CPT$  is a symmetry of any Lorentz invariant local quantum field theory with a Hermitian Hamiltonian.

## 4.2 The CKM Matrix

The CKM matrix is conventionally written as

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{4.6}$$

where,  $V_{ud}$ , for example, provides the coupling between the  $u$  and  $d$  quark. The matrix can be expressed in terms of four parameters, three angles and one phase, which manifestly shows its unitarity [15],

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{4.7}$$

where  $c_{ij} = \cos \theta_{ij}$  and  $s_{ij} = \sin \theta_{ij}$ . The three generations of the SM is a necessary

condition for the CKM matrix to have an imaginary  $CP$ -violating phase, since a  $2 \times 2$  matrix can be expressed in terms of solely real parameters. However, it is not sufficient, as it can be shown that the phase can be removed if two quarks of the same charge had equal masses, or if the value of any of the three mixing angles were 0 or  $\pi/2$  [16]. Additional phases can be brought in with the addition of another quark generation. Practically, the Wolfenstein parameterization [17] is used, which expands each element in powers of the sine of the Cabibbo angle  $\theta_c$ ,  $\lambda = \sin \theta_c = |V_{us}| = 0.22$ . Up to  $\mathcal{O}(\lambda^4)$ , the expansion is

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} - (\frac{1}{8} + \frac{A^2}{2})\lambda^4 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 + A\lambda^4(\frac{1}{2} - \rho - i\eta) & 1 - A^2\frac{\lambda^4}{2} \end{pmatrix}, \quad (4.8)$$

where the relationship between (4.7) and (4.8) is given by

$$s_{12} = \lambda, \quad s_{23} = A\lambda^2, \quad s_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta). \quad (4.9)$$

Eq. 4.8 shows explicitly that the CKM matrix is nearly diagonal and symmetric, a pattern which is not theoretically understood. The unitarity of the matrix implies relationships among its elements. In particular, there are twelve conditions, two of which are

$$V_{cd}V_{cb}^* + V_{td}V_{tb}^* + V_{ud}V_{ub}^* = 0, \quad (4.10)$$

$$V_{tb}V_{ts}^* + V_{ub}V_{us}^* + V_{cb}V_{cs}^* = 0. \quad (4.11)$$

However, new physics which brings in additional generations could make these conditions untrue. Each of these equations may be written in terms of the parameters  $\rho$ ,  $A$ ,  $\lambda$ , and  $\eta$ , which leads to Fig. 4.1. The understanding and verification of Fig. 4.1(a) is the primary motivation for the  $B$ -factories. Fig. 4.1(b) relates to the decay  $B \rightarrow K^*\gamma$ . This may be understood by considering the leading order short-distance diagram of  $B \rightarrow K^*\gamma$ , which is shown in Fig. 4.2. The intermediary particles are given by  $u$ ,  $c$ , and  $t$ , which together with the quarks  $b$  and  $s$  form the sides of Fig. 4.1(b). The spear-shaped nature of Fig. 4.1(b) leads to small  $CP$ -violating effects, which may be contrasted with Fig. 4.1(a). There is also a unitary triangle that leads to a somewhat larger time-independent  $CP$  violating asymmetry in  $B \rightarrow \rho\gamma$ . However, the statistics of the dataset of this thesis is such that this asymmetry is not measurable. This is also true of the time-dependent  $CP$  asymmetry in  $B^0 \rightarrow \rho^0\gamma$ , for which a measurement has been only obtained by the

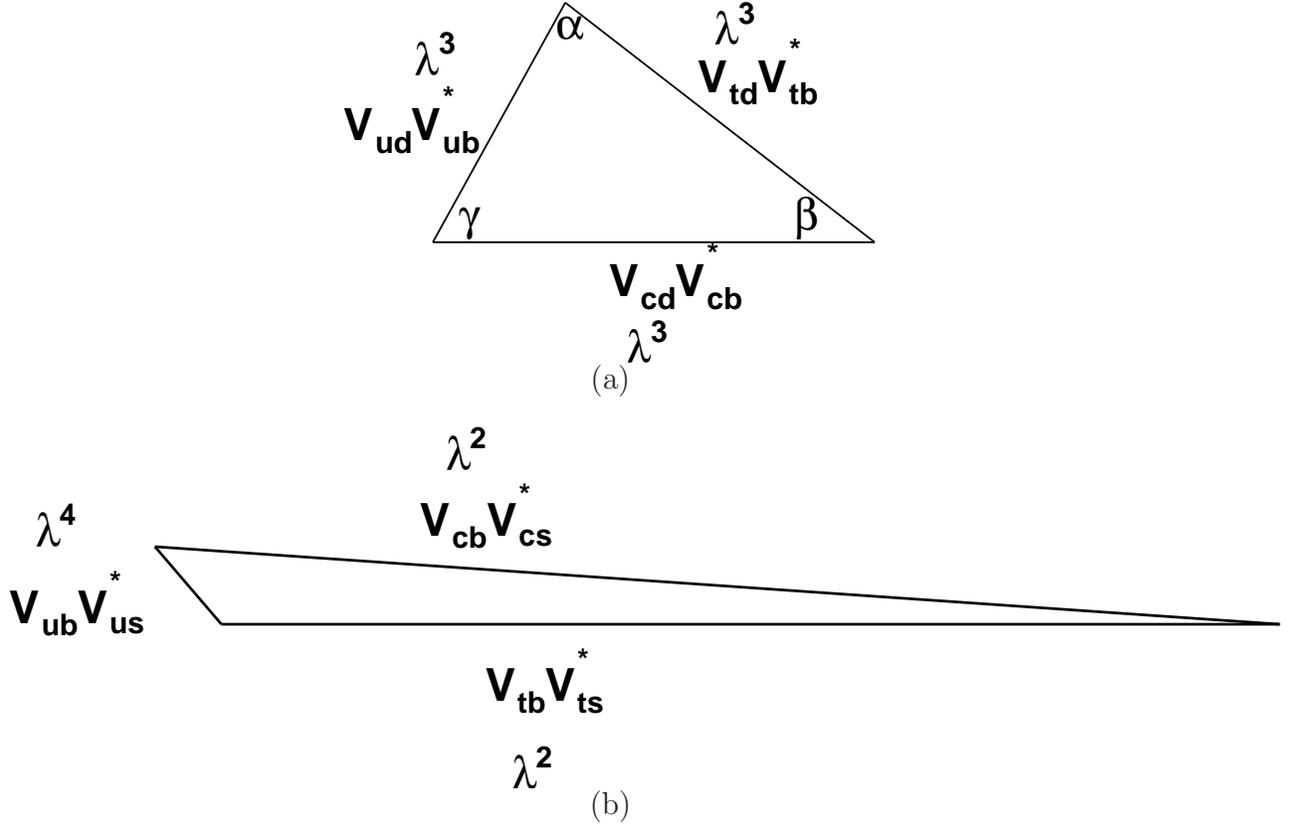


Figure 4.1: The unitary triangles with the sides given in terms of expansion parameter  $\lambda$  portrayed in the complex plane corresponding to a) Eq. 4.10 and b) Eq. 4.11. The openness of the former gives large  $CP$  violation in decays such as  $B \rightarrow J/\psi K_S$  and  $B \rightarrow \pi\pi$ , while the latter gives small  $CP$  violating effects in the decay of  $B \rightarrow K^*\gamma$ . The angles given in (a) are determined by measuring time-dependent asymmetries.

*Belle* experiment [18].  $CP$  asymmetries in  $B \rightarrow K^*\gamma$ , including the time-dependent  $CP$  asymmetry in  $B \rightarrow K^*\gamma(K^{*0} \rightarrow K_S\pi^0)$ , will be discussed further in Section 4.4.2.

### 4.3 $B \rightarrow V\gamma$ Phenomenology

In this thesis, the decays  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-, K_S\pi^0)$ ,  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0, K_S\pi^+)$ ,  $B^0 \rightarrow \rho^0\gamma(\rho^0 \rightarrow \pi^+\pi^-)$ ,  $B^+ \rightarrow \rho^+\gamma(\rho^+ \rightarrow \pi^+\pi^0)$ ,  $B \rightarrow \omega\gamma(\omega \rightarrow \pi^+\pi^-\pi^0)$ , and  $B \rightarrow \phi\gamma(\phi \rightarrow K^+K^-)$  are measured. This section provides a discussion of the kinematic amplitudes involved in these transitions, as well as the line shape used to model the vector ( $K^*, \rho, \phi$ ) resonance.

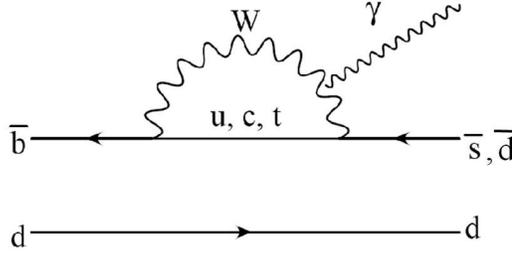


Figure 4.2: Leading order short-distance Feynman digram for the  $b \rightarrow (s, d)\gamma$  transition

### 4.3.1 Decay Dynamics

#### 4.3.1.1 Helicity Amplitudes

The amplitude of a resonance of spin  $J$  decaying into two particles,

$$J \rightarrow 1 + 2, \quad (4.12)$$

can be written in terms of helicity states of the particles involved. Let  $M$  denote the  $z$ -component of the spin of resonance  $J$  of mass  $m$  along an arbitrarily defined  $z$ -axis. In the rest frame of the resonance  $J$ , the spherical angle  $\Omega = (\theta, \phi)$  is defined by the direction of the momentum of particle 1. With the helicities of particles 1 and 2 given by  $\lambda_1, \lambda_2$ , the amplitude is given as [19]

$$A = \sqrt{\frac{2J+1}{4\pi}} F_{\lambda_1 \lambda_2}^J D_{M\lambda}^{J*}(\phi, \theta, 0), \quad (4.13)$$

where  $\lambda = \lambda_1 - \lambda_2$  is the spin projection of the resonance  $J$  along the direction of the momentum of particle 1,  $F_{\lambda_1 \lambda_2}^J$  is the helicity amplitude, and  $D_{M\lambda}^{J*}(\phi, \theta, 0)$  are the  $D$ -functions.

The chiral nature of the weak interactions can be used to simplify the possible helicities of the final states for amplitudes that are evaluated in the SM for  $B \rightarrow K^* \gamma$ ,  $B \rightarrow \rho \gamma$ , and  $B \rightarrow \omega \gamma$ . The  $b \rightarrow q \gamma$  ( $q = (s, d)$ ) amplitude can be described by the effective Hamiltonian

$$\mathcal{H}_{eff} = \frac{G_F e}{16\sqrt{2}\pi^2} m_b F_{\mu\nu} [\bar{q}_L \sigma^{\mu\nu} b_R + \bar{q}_R \sigma^{\mu\nu} b_L] \quad (4.14)$$

where the CKM matrix elements and color indices have been ignored. The chiral and helicity eigenstates for a spin 1/2 fermion can be related by

$$\begin{aligned}
|\lambda_{-}\rangle &= |-\rangle + \frac{m}{2E}|+\rangle + \mathcal{O}\left(\frac{m^2}{E^2}\right) \\
|\lambda_{+}\rangle &= |+\rangle + \frac{m}{2E}|-\rangle + \mathcal{O}\left(\frac{m^2}{E^2}\right)
\end{aligned} \tag{4.15}$$

where  $|\lambda_{-}\rangle, |\lambda_{+}\rangle$  represent the left and right helicity eigenstates, and  $|-\rangle, |+\rangle$  represent the left and right chiral eigenstates. The expansion is in powers of  $m/E$ , where  $m$  is mass of the particle, and  $E$  is the total energy. To leading order in  $m/E$ , the first term in Eq. 4.14 has the helicity structure  $b_R \rightarrow q_L\gamma_L$ , while the second term mediates the opposite transition. However, the  $W$  boson only couples to left-handed external chiral states, so the  $b_R \rightarrow q_L\gamma_L$  transition brings in a factor of  $m_b$  and the  $b_L \rightarrow q_R\gamma_R$  transition brings in a factor of  $m_q$  to the amplitude upon considering Eq. 4.15. So, the emission of a right-handed polarized photon is suppressed roughly by a factor of  $m_q/m_b$  in a  $b$  decay. In the case of a  $\bar{b}$  decay, the left-handed polarized photon is suppressed by the same factor. This suppression has been experimentally probed by measuring the time-dependent  $CP$  asymmetry of  $B \rightarrow K^*\gamma(K^{*0} \rightarrow K_S\pi^0)$ , as will be discussed in Section 4.4.2.2.

### 4.3.1.2 Angular Distributions

The amplitude for the total decay chain  $B \rightarrow V\gamma(V \rightarrow P_1P_2)$  can be written as

$$A(M, \lambda_\gamma) \sim \sum_{\lambda_V} D_{M, \lambda_\gamma - \lambda_V}^{*s(B)}(\phi, \theta, 0) A_{\lambda_\gamma \lambda_V} \times D_{-\lambda_V, \lambda_{P_1} - \lambda_{P_2}}^{*s(V)}(\phi', \theta', 0) B_{\lambda_{P_1} \lambda_{P_2}}. \tag{4.16}$$

Here, the  $z$ -axis is the arbitrarily defined spin-quantization axis for  $M$ , while the  $z'$ -axis is the direction of  $\gamma$  in the rest frame of the  $B$ , so that the spin projection of  $V$  along  $z'$  is  $M_1 = -\lambda_V$ . Inserting the known values of the spins, squaring the amplitude, noting that  $d_{10}^1(\theta') = -\sin\theta'/\sqrt{2}$ , and neglecting rotationally invariant terms, the angular distribution is

$$\frac{d\Gamma}{d\cos\theta'} \sim \sin^2\theta'. \tag{4.17}$$

This distribution will be measured in this thesis.

For  $B \rightarrow \omega\gamma(\omega \rightarrow \pi^+\pi^-\pi^0)$ , the decay chain considered is  $\omega \rightarrow \text{di-pion}(\pi_1\pi_2)\pi_3, (\pi_1\pi_2) \rightarrow \pi_1\pi_2$ . The amplitude can be written as

$$A(\lambda_V) \sim \sum_{\lambda_{\pi_1\pi_2}} D_{-\lambda_V, \lambda_{\pi_1\pi_2} - \lambda_{\pi_3}}^{*s(\omega)}(\phi, \theta, 0) A_{\lambda_{\pi_1\pi_2} \lambda_{\pi_3}} \times D_{\lambda_{\pi_1\pi_2}, \lambda_{\pi_1} - \lambda_{\pi_2}}^{*s(\pi_1\pi_2)}(\phi', \theta', 0) B_{\lambda_{P_1} \lambda_{P_2}}, \quad (4.18)$$

where the  $z$ -axis is in the direction of the  $\gamma$  in the rest frame of the  $B$ , and the  $z'$  axis is in the direction of the  $\pi_1\pi_2$  in the rest frame of the  $\omega$ . Due to parity invariance, the  $\pi_1\pi_2$  system cannot have helicity zero. After computation, and permuting the pion charge over  $\pi_{1,2,3}$ , the angular distribution is given by

$$\frac{d\Gamma}{d \cos \theta_n d \cos \theta_{\pi^+\pi^-}} \sim |A'(m_{\pi^+\pi^-}^2, m_{\pi^+\pi^0}^2)|^2 \sin^2 \theta_{\pi^-\pi^+} \sin^2 \theta_n \quad (4.19)$$

where  $\theta_n$  is the angle between the decay plane normal and the  $B$  in the  $\omega$  rest frame,  $\theta_{\pi^+\pi^-}$  is the angle between the  $\pi^+$  and  $\pi^0$  in the  $\pi^+\pi^-$  rest frame. The  $\theta_{\pi^+\pi^-}$  angle is called the Dalitz angle, while  $\theta_n$  is the helicity angle. We have combined the coefficients  $A_{\lambda_{\pi_1\pi_2} \lambda_{\pi_3}}$ ,  $B_{\lambda_{P_1} \lambda_{P_2}}$  into a single coefficient  $A'$ . The coefficient  $A'$  contains information about possible resonances, which in this case is given by the  $\rho$  resonance. Therefore, the mass dependence of the coefficient is given by the  $\rho$  Breit-Wigner distribution. The two-pion invariant mass and the Dalitz angle are not independent. Therefore, the  $A'$  coefficient can cause deviations from the  $\sin^2 \theta$  distribution. However, it is a feature of the  $\omega$  resonance that this does not occur because the limited phase space available in the  $\omega$  decay causes the two-pion invariant mass to occur far down the  $\rho$  Breit-Wigner distribution. Therefore,  $A' \approx \text{constant}$ , which causes the Dalitz angle to have a  $\sin^2 \theta$  distribution. Both the Dalitz angle and the helicity angle of the  $\omega$  will be measured in this thesis.

#### 4.3.1.3 Parameterization of the $V$ resonance

In the rest frame of the  $V$ , where  $V = (K^*, \rho, \phi)$ , the daughters are emitted with relativistic velocities, so that it is appropriate to use a relativistic Breit-Wigner distribution to characterize the  $V$  resonance. In addition, an energy dependent width is utilized that provides a high-side tail. Let  $m$  be the reconstructed  $V$  mass, and  $m_0$  be the nominal  $V$  mass. Then, the distribution is given by

$$f_{BW}(m) = C_{BW} \frac{p \left( \frac{p}{p_0} \right)^{2l} (B^l(p, p_0))^2}{(m^2 - m_0^2)^2 + (m_0 \Gamma_{rel})^2} \quad (4.20)$$

where

$$\Gamma_{rel} = \Gamma \left( \frac{p}{p_0} \right)^{2l+1} \left( \frac{m_0}{m} \right) (B^l(p, p_0))^2. \quad (4.21)$$

Here,  $p$  is momentum of one of the daughters in the rest frame of the  $V$  evaluated as a function of  $m$ ,  $p_0$  is  $p$  evaluated at  $m_0$ ,  $l$  is the relative orbital angular momentum of the two daughters, and  $\Gamma$  is the nominal width. The factor  $(p/p_0)((m_0/m)$  accounts for the variations of two-body phase space of the width, while the term  $(p/p_0)^{2L}$  is a centrifugal barrier factor.  $B^l(p, p_0)$  is the Blatt-Weisskopf barrier functions [20]

$$B^1(p, p_0) = \frac{F_1(p)}{F_1(p_0)}, \quad (4.22)$$

where

$$F_1(p) = \frac{1}{\sqrt{(pR)^2 + 1}}, \quad (4.23)$$

is the Blatt-Weisskopf barrier function for a spin 1 resonance. The constant  $R$  is  $5 \text{ GeV}^{-1}$ , which is approximately  $1fm$ , the range of the strong interactions. The formulation used here is chosen such that  $F_1(1/R) = 1$ . These functions arise because of the need of the spin of the resonance to be translated into angular momentum of the decay products. Since this is a strong decay, which has a very limited range, decay products of low momentum will be unable to generate enough angular momentum to conserve the spin. These functions account for this effect. Although Eq. 4.20 is valid for any resonance that has a 2-body decay, the statistics of the samples are such that a distribution can be obtained only for the  $K^*$  resonance. Therefore, in this thesis, this distribution will be used to validate that the signal events of the data sample are populated by the  $K^*$  resonance.

### 4.3.2 Form Factors

The  $B \rightarrow V$  ( $V = \rho, \omega, K^*$ ) form factor is defined by

$$\langle V(p_V, \eta) | i q^\nu \bar{s} \sigma_{\mu\nu} b | \bar{B}(p_B) \rangle = -i F^{B \rightarrow V}(q^2) \epsilon_{\mu\alpha\beta\rho} \eta^\alpha p^\beta q^\rho \quad (4.24)$$

where  $p = p_V + p_B$ ,  $q = p_B - p_V$ ,  $\eta$  is the polarization vector of  $V$ , and  $q^2 = 0$ . It can be evaluated by using a variety of different approaches [21, 22, 23], including lattice QCD, QCD sum rules, and light-cone QCD sum rules. A summary is shown in Table 4.1.

Table 4.1: Results of form factor calculations using lattice QCD, QCD sum rules, and light-cone QCD sum rules. For the lattice QCD entry, the first error is statistical, while the second error is systematic.

	Lattice QCD [21]	QCD sum rules [22]	Light-cone QCD sum rules [23]
$F^{B \rightarrow K^*}(0)$	$0.24 \pm 0.03 \pm_{-0.01}^{+0.04}$	$0.35 \pm 0.05$	$0.31 \pm 0.04$
$F^{B \rightarrow \rho}(0)$			$0.27 \pm 0.04$
$F^{B \rightarrow \omega}(0)$			$0.25 \pm 0.04$

## 4.4 Measured Observables

### 4.4.1 Branching Fractions

The branching fractions can be evaluated in a variety of methods, which include heavy quark effective theory (HQET), perturbative QCD (pQCD), and soft-collinear effective theory (SCET). They all involve factorizing the matrix elements into hard and soft dynamics. The theoretical calculations for the branching fractions for  $B \rightarrow K^* \gamma$  are shown in Table 4.2, while the experimental measurements are shown in Table 4.3. Due to hadronic uncertainties (represented by the forms factors listed in Table 4.1), the theoretical uncertainties are much larger than the experimental ones. To leading order in HQET, the  $\mathcal{B}(B \rightarrow K^* \gamma) \propto |F^{B \rightarrow K^*}(0)|^2$  and  $\mathcal{B}(B \rightarrow \rho \gamma) \propto |F^{B \rightarrow \rho}(0)|^2$  [26]. As shown in Table 4.2, the central value of the theoretically predicted branching fractions have come down as a function of time. This is partially due to the fact that the form factor  $F^{B \rightarrow K^*}$  has decreased with time. Ref. [76], which was published in 1998, uses QCD sum rules to give  $F^{B \rightarrow K^*} = 0.38 \pm 0.06$ , while the most recent prediction gives  $F^{B \rightarrow K^*} = 0.31 \pm 0.04$  [23]. Another factor leading to the different predictions of Ref. [24] through Ref. [28] is the number of diagrams computed increased. However, because the large uncertainties on the predictions, the measured branching fractions of this thesis are consistent with all of the predictions. The theoretical predictions for  $B \rightarrow \rho \gamma$  and  $B \rightarrow \omega \gamma$  are given in Table 4.4, while the measurements are given in Table 4.5.

Prior to the measurement described in this thesis, the mode  $B^0 \rightarrow \rho^0 \gamma$  has only been observed by the *Belle* collaboration, while the modes  $B^+ \rightarrow \rho^+ \gamma$  and  $B \rightarrow \omega \gamma$  were undetected. The pure annihilation radiative decay  $B \rightarrow \phi \gamma$  is further suppressed in the Standard Model relative to  $B \rightarrow K^* \gamma$  and  $B \rightarrow \rho \gamma$ . The theoretical predictions for  $B \rightarrow \phi \gamma$  are  $\sim 10^{-11}$  [35, 36], and a leading order Feynman diagram is given in Fig. 4.3. A previous search has bounded the branching fraction to be  $\mathcal{B}(B \rightarrow \phi \gamma) < 3.3 \times 10^{-6}$  [29].

Table 4.2: Current theoretical calculations of the  $B \rightarrow K^*\gamma$  branching fraction

$\mathcal{B}(B \rightarrow K^*\gamma)$ (NLO)	Reference
$(7.25 \pm 2.6) \times 10^{-5}$	Ali and Parkhomenko [24]
$(7.9_{-3.0}^{+3.5}) \times 10^{-5}$	Beneke, Feldmann and Seidel [25]
$(7.09_{-2.27}^{+2.47}) \times 10^{-5}$	Bosch and Buchalla [26]
$(5.8 \pm 2.9) \times 10^{-5}$ ( $B^0 \rightarrow K^{*0}\gamma$ ) $(6.0 \pm 3.0) \times 10^{-5}$ ( $B^+ \rightarrow K^{*+}\gamma$ )	Matsumori, Sanda, and Keum [27]
$(4.3 \pm 1.4) \times 10^{-5}$ ( $B^0 \rightarrow K^{*0}\gamma$ ) $(4.6 \pm 1.4) \times 10^{-5}$ ( $B^+ \rightarrow K^{*+}\gamma$ )	Ali, Pecjak, and Greub [28]

Table 4.3: Previous measurements of the branching ratios and asymmetries for  $B \rightarrow K^*\gamma$ . The last error on the isospin asymmetry for the *BABAR* measurement refers to the error on the production ratio of charged to neutral B events.

	CLEOII [29] 9.2 fb <sup>-1</sup>	<i>BABAR</i> [30] 81.9 fb <sup>-1</sup>	<i>Belle</i> [31] 78.0 fb <sup>-1</sup>
$B^0 \rightarrow K^{*0}\gamma$ ( $\times 10^{-5}$ )	$4.55_{-0.68}^{+0.72} \pm 0.34$	$3.92 \pm 0.20 \pm 0.24$	$4.01 \pm 0.21 \pm 0.17$
$B^+ \rightarrow K^{*+}\gamma$ ( $\times 10^{-5}$ )	$3.76_{-0.83}^{+0.89} \pm 0.28$	$3.87 \pm 0.28 \pm 0.26$	$4.25 \pm 0.31 \pm 0.24$
$\mathcal{A}_{CP}$	$+0.08 \pm 0.13 \pm 0.03$	$-0.013 \pm 0.036 \pm 0.010$	$-0.015 \pm 0.044 \pm 0.012$
Isospin asymmetry	N/A	$+0.050 \pm 0.045 \pm 0.028 \pm 0.024$ (measured parameter: $\Delta_{0-}$ )	$+0.012 \pm 0.044 \pm 0.026$ (measured parameter: $\Delta_{0+}$ )

## 4.4.2 CP Asymmetries

The  $CP$  asymmetry is defined as

$$\mathcal{A}_{CP} = \frac{\Gamma(\bar{B} \rightarrow \bar{K}^*\gamma) - \Gamma(B \rightarrow K^*\gamma)}{\Gamma(\bar{B} \rightarrow \bar{K}^*\gamma) + \Gamma(B \rightarrow K^*\gamma)}, \quad (4.25)$$

In this section, the direct  $CP$  asymmetry from the decays  $B \rightarrow K^*\gamma(K^{*0} \rightarrow K^+\pi^-)$ ,  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0, K^{*+} \rightarrow K_S\pi^+)$  and the time-dependent  $CP$  asymmetry involving the decay  $B \rightarrow K^*\gamma(K^{*0} \rightarrow K_S\pi^0)$  will be discussed briefly.

### 4.4.2.1 Time-independent

The total amplitude of  $B \rightarrow K^*\gamma$  and its conjugate can be written as

Table 4.4: Theoretical predictions of the branching fractions for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$ .

Mode	Branching fraction ( $\times 10^{-6}$ )		
	Ref. [32]	Ref. [26]	Ref. [23]
$B^+ \rightarrow \rho^+\gamma$	$1.37 \pm 0.28$	$1.58^{+0.53}_{-0.46}$	$1.16 \pm 0.26$
$B^0 \rightarrow \rho^0\gamma$	$0.65 \pm 0.12$		$0.55 \pm 0.13$
$B \rightarrow \omega\gamma$	$0.53 \pm 0.12$		$0.44 \pm 0.10$

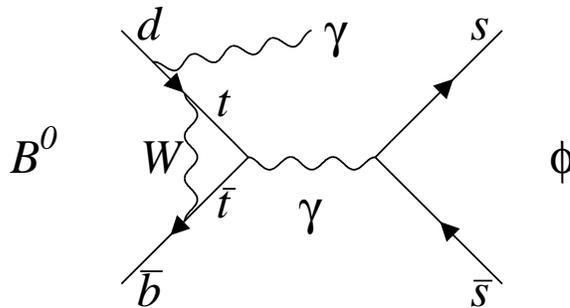
Figure 4.3: One of the leading order Feynman diagrams contributing to the decay  $B \rightarrow \phi\gamma$  in the Standard Model.

Table 4.5: Previous measurements of the branching ratios for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$ , where the limits are given by the 90% confidence level. The errors on the *Belle* measurements are statistical and systematic respectively, while the entry in paranthesis refers to the significance of the measurement.

	CLEOII [29] 9.2 fb <sup>-1</sup>	BABAR [33] 191 fb <sup>-1</sup>	<i>Belle</i> [34] 350.9 fb <sup>-1</sup>
$B^+ \rightarrow \rho^+\gamma(\times 10^{-6})$	< 13	< 1.8	$0.55_{-0.36-0.08}^{+0.42+0.36}$ (1.6 $\sigma$ )
$B^0 \rightarrow \rho^0\gamma(\times 10^{-6})$	< 17	< 0.4	$1.25_{-0.33-0.06}^{+0.37+0.07}$ (5.2 $\sigma$ )
$B \rightarrow \omega\gamma(\times 10^{-6})$	< 92	< 1.0	$0.56_{-0.27-0.10}^{+0.34+0.05}$ (2.3 $\sigma$ )
$B \rightarrow (\rho/\omega)\gamma(\times 10^{-6})$		< 1.2	$1.32_{-0.31-0.09}^{+0.34+0.10}$ (5.1 $\sigma$ )

$$A(\bar{B} \rightarrow \bar{K}^*\gamma) = \lambda_t^* A_t e^{i\delta_t} + \lambda_c^* A_c e^{i\delta_c} + \lambda_u^* A_u e^{i\delta_u}, \quad (4.26)$$

$$A(B \rightarrow K^*\gamma) = \lambda_t A_t e^{i\delta_t} + \lambda_c A_c e^{i\delta_c} + \lambda_u A_u e^{i\delta_u}, \quad (4.27)$$

where  $\lambda_t = V_{ts}^* V_{tb}$ ,  $\lambda_c = V_{cs}^* V_{cb}$ ,  $\lambda_u = V_{us}^* V_{ub}$ , the weak ( $\lambda_t, \lambda_c, \lambda_u$ ) and strong ( $\delta_t, \delta_c, \delta_u$ ) phases have been separated out, and  $A_t, A_c, A_u$  are the real contributions to the amplitude associated with the corresponding CKM elements. The strong phases arise from  $\mathcal{O}(\alpha_s)$  contributions which also involve the light quarks [37]. Using Eqs. 4.26 and 4.27, the  $CP$  asymmetry can be expressed as

$$\mathcal{A}_{CP} = \frac{N}{D}, \quad (4.28)$$

where

$$N = A_t A_c \sin(\delta_t - \delta_c) \text{Im}(\lambda_t \lambda_c^*) + A_c A_u \sin(\delta_c - \delta_u) \text{Im}(\lambda_c \lambda_u^*) + A_u A_t \sin(\delta_u - \delta_t) \text{Im}(\lambda_u \lambda_t^*), \quad (4.29)$$

$$D = (A_t^2 |\lambda_t|^2 + A_c^2 |\lambda_c|^2 + A_u^2 |\lambda_u|^2)/2 + A_t A_c \cos(\delta_t - \delta_u) \text{Re}(\lambda_t \lambda_c^*) + A_c A_u \cos(\delta_c - \delta_u) \text{Re}(\lambda_c \lambda_u^*) + A_u A_t \cos(\delta_u - \delta_t) \text{Re}(\lambda_u \lambda_t^*). \quad (4.30)$$

Thus, there must be at least two amplitudes with different weak and strong phases for a non-vanishing  $CP$  asymmetry. Eqs. 4.29 and 4.30 show that the asymmetry is determined

not only by the phases, but also by the relative magnitudes of the amplitudes. Since the dominant amplitude is given by  $A_t$ , and  $\text{Im}(\lambda_t\lambda_c^*)$ ,  $\text{Im}(\lambda_c\lambda_u^*)$ , and  $\text{Im}(\lambda_u\lambda_t^*)$  are  $\sim \lambda^6$ , the asymmetry is highly suppressed. The SM gives  $\mathcal{A}_{CP} \sim -0.5\%$  [26, 27]. The  $CP$  asymmetry of inclusive  $b \rightarrow s\gamma$  also has a magnitude that is less than 1%. However, new physics could enhance this to be 10-50% [38].

#### 4.4.2.2 Time-dependent

As mentioned in Section 4.3.1, the photon emitted in a radiative  $b(\bar{b})$  decay is predominantly left(right)-handed polarized.  $CP$  asymmetries in  $B$  decays which involve the interference with and without mixing require that both  $B$  and  $\bar{B}$  decay to a common state, which means interference will only take place with states with a common photon helicity. Therefore, new physics models which are left-right symmetric can be tested by measuring the time-dependent  $CP$  asymmetry of  $B \rightarrow K^*\gamma(K^{*0} \rightarrow K_S\pi^0)$ . The SM prediction is given by [39]

$$\mathcal{A}_{CP}(B \rightarrow K^*\gamma(K^{*0} \rightarrow K_S\pi^0)) \approx (2m_s/m_b) \sin(2\beta) \sin(\Delta mt), \quad (4.31)$$

which is very small. However, in left-right symmetric models, the asymmetry can be 50% or larger [39]. This asymmetry has been measured, and is consistent with SM expectations [40, 41]. It will not be measured in this thesis.

### 4.4.3 Isospin Asymmetries

#### 4.4.3.1 $B \rightarrow K^*\gamma$

The isospin asymmetry is defined as

$$\Delta_{0-} = \frac{\Gamma(\bar{B} \rightarrow \bar{K}^*\gamma) - \Gamma(B^- \rightarrow K^{*-}\gamma)}{\Gamma(\bar{B} \rightarrow \bar{K}^*\gamma) + \Gamma(B^- \rightarrow K^{*-}\gamma)}. \quad (4.32)$$

In this thesis, we assume  $CP$  asymmetry when making the measurement. The SM predictions for the isospin asymmetry are given in Table 4.6. Since the theoretical hadronic and correlated experimental systematic uncertainties partially cancel when taking the ratio, the isospin asymmetry provides a good sensitivity to new physics. For example, one new physics model is the minimal supergravity model (mSUGRA), where the number of parameters is reduced to five, which are  $m_0$  (common mass of scalar particles at the

supersymmetric grand unification scale),  $m_{1/2}$  (universal gaugino mass),  $A_0$  (universal trilinear SUSY breaking parameter),  $u$  (the sign of the Higgs mixing parameter), and  $\tan(\beta)$  (the ratio of the two Higgs vacuum expectation values). When considering this model, Ref. [43] indicates that the isospin asymmetry could provide more severe restrictions on parameter space than the ones coming from the branching fraction of  $B \rightarrow X\gamma$ . This is displayed in Fig. 4.4 for  $A_0 = 0$ ,  $\tan(\beta) = 30$ , and  $\mu > 0$ .

Table 4.6: Current theoretical calculations of isospin violation in  $B \rightarrow K^*\gamma$  decays.  $F^{B \rightarrow V}(q^2 = 0)$  is the  $B \rightarrow K^*$  form factor whose estimates range from  $0.23 \pm 0.06$  [27] to  $0.38 \pm 0.06$  [76].

Isospin violation	Reference
$(+8.0_{-3.2}^{+2.1})\% \times 0.3/T_1^{B \rightarrow K^*}$ ( $\Delta_{0-}$ )	Kagan and Neubert [42]
$(+2.7 \pm 0.8)\%$ ( $\Delta_{0+}$ )	Matsumori, Sanda, and Keum [27]

#### 4.4.3.2 $B \rightarrow \rho\gamma$

The isospin-violating ratio in  $B \rightarrow \rho\gamma$  is defined as

$$\Delta = \frac{\Gamma(B^+ \rightarrow \rho^+\gamma)}{2\Gamma(B^0 \rightarrow \rho^0\gamma)} - 1. \quad (4.33)$$

According to isospin, we expect  $\Gamma(B^+ \rightarrow \rho^+\gamma) = 2\Gamma(B^0 \rightarrow \rho^0\gamma)$ . This ratio is dependent on the CKM angle  $\alpha$  [32], so, in principle, a measurement of this quantity could yield an independent determination of this angle. Within the SM, the isospin asymmetry is dominated by weak annihilation contributions. The SM expectation using QCD factorization is  $\Delta = (2.9 \pm 2.1)\%$  [32]. However, new physics could alter this value. In this thesis, since the branching fractions of  $B^+ \rightarrow \rho^+\gamma$  and  $B^0 \rightarrow \rho^0\gamma$  are measured, an estimate of this quantity is given.

#### 4.4.4 $|V_{td}/V_{ts}|$

When the unitary triangle in Fig. 4.1(a) is rescaled by dividing each side by  $V_{cd}V_{cb}^*$ , the length of the side opposite the angle  $\gamma$  ( $R_t$ ) is given by

$$R_t \equiv \sqrt{(1 - \bar{\rho})^2 + \bar{\eta}^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{ts}} \right|, \quad (4.34)$$

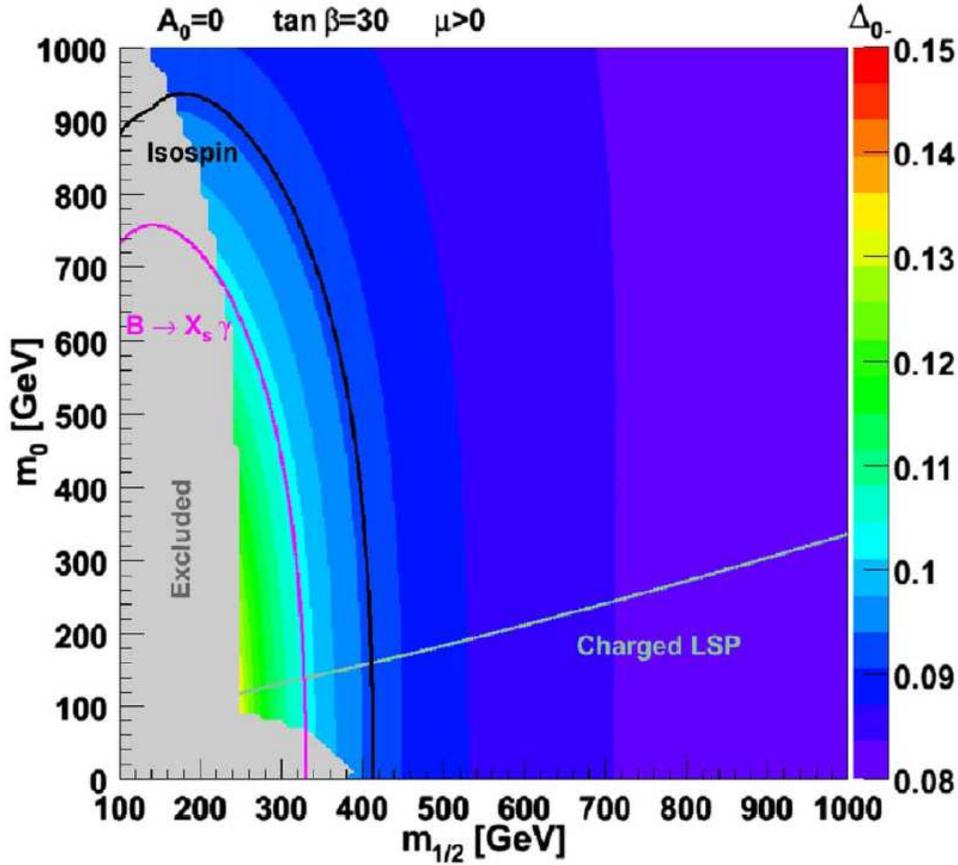


Figure 4.4: Experimental constraints on mSUGRA parameter space for the  $B \rightarrow K^* \gamma$  isospin asymmetry (black), the  $B \rightarrow X \gamma$  branching fraction (magenta). The green “Charged LSP” indicates the the cosmologically disfavored region if the lightest supersymmetric particle (LSP) is charged and R-parity is conserved. The grey “Excluded” region indicates constraints on the parameter space corresponding to direct searches for the Higgs and sparticles. The assumptions are  $A_0 = 0$ ,  $\tan(\beta) = 30$ , and  $\mu > 0$ .

where  $\bar{\rho} = \rho(1 - \lambda^2/2)$  and  $\bar{\eta} = \eta(1 - \lambda^2/2)$ . Therefore, a measurement of the CKM ratio  $|V_{td}/V_{ts}|$  determines one of the sides of the unitary triangle. To gain precision with the limited statistics in  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$  for CKM phenomenology, we define the isospin averaged branching fraction

$$\mathcal{B}[B \rightarrow (\rho/\omega)\gamma] \equiv \frac{1}{2} \left\{ \mathcal{B}(B^+ \rightarrow \rho^+\gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} [\mathcal{B}(B^0 \rightarrow \rho^0\gamma) + \mathcal{B}(B \rightarrow \omega\gamma)] \right\}. \quad (4.35)$$

Non-perturbative QCD uncertainties partially cancel when taking the ratio

$$\frac{\mathcal{B}[B \rightarrow (\rho/\omega)\gamma]}{\mathcal{B}[B \rightarrow K^*\gamma]} = S \left| \frac{V_{td}}{V_{ts}} \right|^2 \left( \frac{1 - m_\rho^2/M_B^2}{1 - m_{K^*}^2/M_B^2} \right)^3 \zeta^2 [1 + \Delta R], \quad (4.36)$$

where  $S$  is 1 for  $\rho^+$  and  $\frac{1}{2}$  for  $\rho^0$  and  $\omega$ ,  $m$  is the particle mass,  $\zeta = 0.85 \pm 0.07$  is the ratio of form factors for the decays  $B \rightarrow (\rho/\omega)\gamma$  and  $B \rightarrow K^*\gamma$ , and  $\Delta R = 0.1 \pm 0.1$  accounts for different dynamics, including weak annihilation contributions. The latter two quantities are obtained from Ref. [44]. In this thesis, a measurement of this ratio will be used to extract the quantity  $|V_{td}/V_{ts}|$ , which can also be obtained from the ratio of  $B_d$  and  $B_s$  mixing frequencies. A summary of previous measurements of this quantity is given in Table 4.7. Physics beyond the SM could affect differently  $B \rightarrow (\rho/\omega)\gamma$  and  $B_d/B_s$  mixing, and hence create inconsistencies between the results obtained from the two methods.

Table 4.7: Previous measurements of  $|V_{td}/V_{ts}|$ . In the CDF entry,  $\Delta m_d$  and  $\Delta m_s$  denote the  $B_d$  and  $B_s$  mixing frequencies respectively. ‘‘Theor.’’ stands for the theoretical uncertainty.

Collaboration	$ V_{td}/V_{ts} $
CDF [45]	$0.2060 \pm 0.0007 (\Delta m_s)_{-0.0060}^{+0.0081} (\Delta m_d + \text{theor.})$
Belle [34]	$0.199_{-0.025}^{+0.026} (\text{exp})_{-0.015}^{+0.018} (\text{theor})$

# Chapter 5

## Event Selection, Continuum Background Suppression, and Selection Optimization

In the previous chapter, the theoretical implications of the  $B \rightarrow V\gamma$  observables were discussed. In this chapter, the discussion will turn to the measurements of these quantities. After a brief introduction in Section 5.1, the event and candidate selection for all of the modes used in this study is discussed in Sections 5.2 and 5.3, along with a description of the primary kinematic variables used to distinguish signal from background. Then, Section 5.4 will describe how the continuum background is dealt with. The final Section 5.5 discusses how the selection criteria were optimized. The next three chapters will discuss the fitting procedure and the results for  $B \rightarrow K^*\gamma$  and  $B \rightarrow \phi\gamma$ . In the case of  $B \rightarrow (\rho/\omega)\gamma$ , only the results will be presented.

### 5.1 Introduction

As mentioned in Section 4.4.1, there are previous measurements made by *BABAR* of  $B \rightarrow K^*\gamma$  [30] and  $B \rightarrow (\rho/\omega)\gamma$  [33]. The current analyses builds on and improves upon those analyses. The  $B \rightarrow \phi\gamma$  mode is similar to the former two modes, and much of the analysis methods for  $B \rightarrow \phi\gamma$  is borrowed from them. Eight different decay modes are considered:

- $B^0 \rightarrow K^{*0}\gamma$  ( $K^{*0} \rightarrow K^+\pi^-$ ,  $K^{*0} \rightarrow K_S\pi^0$ )
- $B^+ \rightarrow K^{*+}\gamma$  ( $K^{*+} \rightarrow K^+\pi^0$ ,  $K^{*+} \rightarrow K_S\pi^+$ )

- $B^0 \rightarrow \rho^0 \gamma$  ( $\rho^0 \rightarrow \pi^+ \pi^-$ )
- $B^+ \rightarrow \rho^+ \gamma$  ( $\rho^+ \rightarrow \pi^+ \pi^0$ )
- $B \rightarrow \omega \gamma$  ( $\omega \rightarrow \pi^+ \pi^- \pi^0$ )
- $B \rightarrow \phi \gamma$  ( $\phi \rightarrow K^+ K^-$ ).

The reconstruction of these modes proceeds in three different analyses: the measurement of  $B \rightarrow K^* \gamma$  constitutes one analysis, the measurement of  $B^+ \rightarrow \rho^+ \gamma$  and  $B \rightarrow \omega \gamma$  comprises another, and the final analysis is the search for  $B \rightarrow \phi \gamma$ .

The cross-sections for  $e^+e^-$  collisions to produce quarks and leptons at a CM energy of 10.58 GeV (the  $\Upsilon(4S)$  mass) are given in Table 5.1. The production of light quark pairs ( $c\bar{c}, s\bar{s}, u\bar{u}, d\bar{d}$ ), which are known as continuum events, is about three times the production of  $e^+e^- \rightarrow b\bar{b}$ . Given this fact, and since jets from continuum events contain high energy gammas which can simulate the high energy gamma coming from the signal decay, continuum processes form the majority of the background in these analyses. Background that results from  $B$  decays also plays a role in the analyses. The importance of the  $B$  background depends on the particular analysis, and will be discussed further below.

Table 5.1: Production cross-sections at a CM energy of 10.58 GeV [46]. The cross-section listed for  $e^+e^-$  production includes the effect of limited detector acceptance.

$e^+e^- \rightarrow$	Cross-section(nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$\tau^+\tau^-$	0.89
$\mu^+\mu^-$	1.16
$e^+e^-$	$\sim 40$

### 5.1.1 Data

The *BABAR* experiment and the PEP-II accelerator ran from October of 1999 to April of 2008, with data begin broken into seven run periods. The recorded luminosity is given in Table 5.2. The ‘‘Onpeak’’ data is recorded at the  $\Upsilon(4S)$  mass, while the ‘‘Offpeak’’

data is taken at 200 MeV below the  $\Upsilon(4S)$  resonance. The first six runs were taken at the  $\Upsilon(4S)$  mass, and the precision on the number of  $B\bar{B}$  pairs corresponding to each run is 1.1%. Run 7 was used to explore other  $\Upsilon$  resonances, such as the  $\Upsilon(2S)$  and the  $\Upsilon(3S)$ . Table 5.3 lists the luminosity of each dataset that was used by each analysis. The  $B \rightarrow \phi\gamma$  analysis used Runs 1-4, the  $B \rightarrow (\rho/\omega)\gamma$  analysis used Runs 1-4 and a substantial fraction of Run 5 (this was done in order to make the International Conference of High-Energy Physics (ICHEP) summer deadline), while the  $B \rightarrow K^*\gamma$  analysis used Runs 1-5. Fig. 5.1 shows the luminosity delivered by PEP-II and recorded by *BABAR* during the course of the entire experiment.

Table 5.2: Luminosity recorded over the history the *BABAR* detector.

Run	Date Range	Onpeak ( $\text{fb}^{-1}$ )	Offpeak ( $\text{fb}^{-1}$ )	$B\bar{B}$ pairs ( $\times 10^6$ )
1	Oct 1999-Oct 2000	20.43	2.62	22.43
2	Feb 2001-June 2002	61.15	6.92	67.46
3	Dec 2002-June 2003	32.31	2.47	35.61
4	Sep 2003-July 2004	100.31	10.12	110.49
5	Apr 2005-Aug 2006	133.27	14.49	147.19
6	Jan 2007-Sep 2007	78.78	7.88	86.88
7	Dec 2007-Apr 2008	Other $\Upsilon$ resonances		

Table 5.3: Luminosity of dataset used by each analysis.

Analysis	Date Range	Onpeak ( $\text{fb}^{-1}$ )	Offpeak ( $\text{fb}^{-1}$ )	$B\bar{B}$ pairs ( $\times 10^6$ )
$B \rightarrow \phi\gamma$	Oct 1999-July 2004	214.2	22.13	235.99
$B \rightarrow (\rho/\omega)\gamma$	Oct 1999-June 2006	316.29	27.75	347.50
$B \rightarrow K^*\gamma$	Oct 1999-Aug 2006	347.46	36.62	383.18

### 5.1.1.1 Simulated Data

This thesis makes extensive use of simulated events, known as Monte Carlo (MC) simulation samples. These events are generated according to models that simulate a particular process (for example  $e^+e^- \rightarrow B^0\bar{B}^0$ ). Each unstable particle generated in the simulation is either forced to decay to a particular mode (for example,  $B \rightarrow K^*\gamma(K^{*0} \rightarrow$

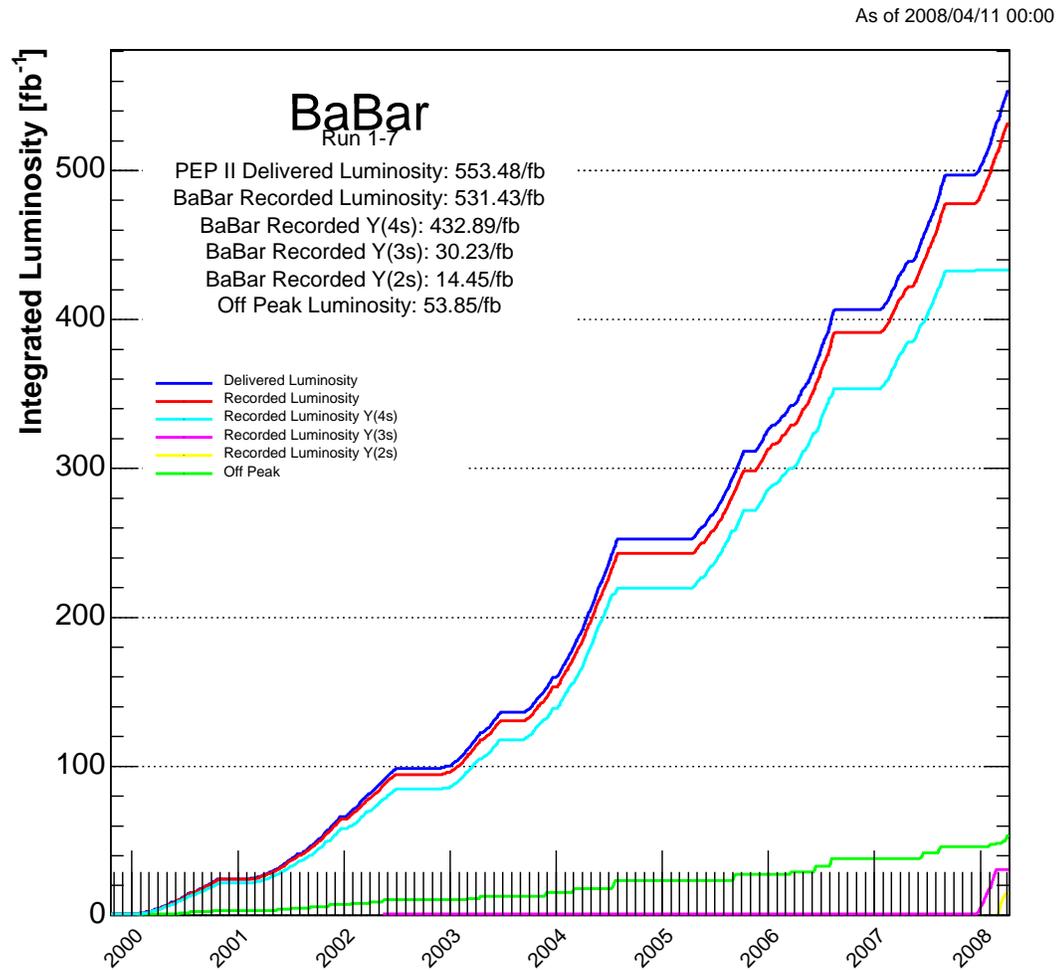


Figure 5.1: Delivered and recorded luminosity as a function of time.

$K^+\pi^-$ ) or to one of many possible modes with a probability (usually taken from the Particle Data Group (PDG)) of decaying to a particular mode. The generated particles traverse a detailed model of the detector. Version 4 of the **GEANT** software package [47] is used to simulate the interactions of the particles with the detector material and the responses of the detector subsystems. Real data obtained by reading out the detector at random beam crossings are merged with simulated events in order to add realistic beam backgrounds to the simulation.

The MC simulation samples are used for three phases in each analysis: cut optimization, the development of the neural network, and to construct the fit models. We use MC simulation samples that reproduce signal and continuum events in all three phases. To understand the background coming from  $B$  decays, we use samples of  $e^+e^- \rightarrow B\bar{B}$  MC simulation, where both  $B$ 's are allowed to decay to any one of their modes (“generic  $B\bar{B}$ ”), and also MC simulation where one  $B$  is constrained to a particular mode, and the other  $B$  is allowed to decay to any mode (“exclusive  $B\bar{B}$ ”). This MC simulation is used in the construction of the fit models. The size of each MC simulation sample varies widely. However, in general, compared to the onpeak data, the size of the continuum MC simulation is approximately the same, the “generic  $B\bar{B}$ ” is about three times, and the “exclusive  $B\bar{B}$ ” sample is  $\approx 100$  or greater. The offpeak data is used to validate the continuum MC simulation. In the case of  $B \rightarrow \phi\gamma$ , the onpeak data is used to validate the fit procedure and extract the signal yield, while in other two analyses it is used only for the latter.

Finally, each generated particle that traverses the detector may be reconstructed or not. The reconstructed properties, e.g., four-momentum, may be different from its generated ones. In addition to having each reconstructed particle be recorded by the event, each event also has a list of generated particles for the event. An attempt is made to map each reconstructed particle to its generated counterpart. If such an association is possible, then the particle is termed “truth-matched”. We will make use of this in the cut optimization, when we will only use candidates in the signal MC simulation that can be truth-matched. This is because we want our cut selection to project out true signal events. In addition, we also make use of truth-matching when we study biases that occur in the  $B \rightarrow K^*\gamma$  fit model (appendix D).

### 5.1.2 Overview of Analysis Method

All of the intermediate vector mesons in the signal modes decay strongly. The branching fractions of the  $K^*$  are determined using isospin symmetry and are given by the square of the Clebsch-Gordon coefficients (Table 5.4). When the  $K^*$  meson decays to a  $K_S$ , the  $K_S$  is reconstructed from the  $K_S \rightarrow \pi^+\pi^-$  decay only, where we assume a branching fraction of 69.2% [48]. The branching fractions of the  $\rho$ ,  $\omega$ , and  $\phi$  used in this thesis are  $\mathcal{B}(\rho^0 \rightarrow \pi^+\pi^-) \sim 100\%$ ,  $\mathcal{B}(\rho^+ \rightarrow \pi^+\pi^0) \sim 100\%$ ,  $\mathcal{B}(\omega \rightarrow \pi^+\pi^-\pi^0) = (89.1 \pm 0.7)\%$  [48], and  $\mathcal{B}(\phi \rightarrow K^+K^-) = (49.1 \pm 0.6)\%$  [49].

Table 5.4: The isospin factors for the decay of the  $K^*$ .

Decay Mode	Isospin Factor
$K^{*0} \rightarrow K^+\pi^-$	0.667
$K^{*0} \rightarrow K_S\pi^0$	0.333
$K^{*+} \rightarrow K^+\pi^0$	0.333
$K^{*+} \rightarrow K_S\pi^+$	0.667

The  $\phi(1020)$  has a width of  $\sim 4$  MeV and the  $\omega(782)$  has a width of  $\sim 8$  MeV; therefore, a tight cut on the invariant mass of the resonance is a powerful handle to reject combinatorial background. In contrast, the  $K^*(892)$  has a width of  $\sim 50$  MeV and the  $\rho(770)$  has a width of  $\sim 150$  MeV, which provides less stringent background rejection.

In signal decays, due to the two-body kinematics, the low mass of the particles in the final state, and the mass of the  $B$  meson, the  $\gamma$  and the meson are produced with very high momentum in the  $B$  rest frame. Requiring only the presence of the high energy photon allows us to reduce  $\approx 85\%$  of the combinatorial background because the energy spectrum of the photons decreases exponentially with energy. After applying the remaining selection criteria, the leftover continuum background is dealt with by a neural network, the description of which is in appendix C. The  $B$  background, except for the analysis of  $B \rightarrow \phi\gamma$ , for which the  $B$  background is negligible, is dealt with by performing a maximum likelihood fit.

The general outline of the analysis procedure is as follows:

- *Preliminary Event Selection:* Because of the large data sample collected by the BABAR detector, it is impractical for the analysts to deal with a dataset that contains every event. Therefore, preliminary event selections are designed to vastly reduce the size of the dataset while maintaining a high signal efficiency. In addition, they

greatly reduce the amount of disk space required to store and CPU time required to process the events.

- *Cut optimization:* Selection criteria, such as the invariant mass of the vector resonance and a number of variables based on photon quality, are optimized by maximizing the statistical signal significance, which is defined as  $S/\sqrt{S+B}$ , where  $S$  and  $B$  are the rates for signal and background, respectively. The cuts are optimized for each mode individually.
- *Continuum background suppression:* A number of variables are combined together into a neural network (NN) and are used to suppress the continuum background.
- *Signal yield extraction:* We use an extended unbinned maximum likelihood fit in order to extract the signal yield for the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analyses. The number of dimensions and components varies depending on the sensitivity needed for the analysis. For the  $B \rightarrow \phi\gamma$  analysis, we count of the number of events in a defined signal region, and compare it to the expected number of background events obtained from a fit. This is done because we do not expect a signal.

## 5.2 Preliminary Event Selection

Generic  $B\bar{B}$  decays are characterized by high multiplicity and isotropic event topology in the center-of-mass system (CMS) resulting from the cascade of decays originating from the  $B$  meson, which is slowly moving in this frame. In contrast, background electromagnetic processes are characterized by low multiplicity (typically two tracks), and continuum events have a more collimated jet structure from the hadronic interactions of the fast moving quark-antiquark pair. We, therefore, require at least three tracks that individually satisfy the following requirements:

- The distance of closest approach to the beamspot in the x-y plane to be  $|d_{xy}| < 1.5$  cm and along the z-axis to be  $|d_z| < 10$  cm, which rejects fake tracks and background tracks not originating from the beamspot.
- At least 12 hits in the DCH to ensure good track quality.
- The laboratory momentum must be  $|p| < 10$  GeV/c. A momentum greater than this is not consistent with the beam energies.

- The trajectory of each track is fit. The fit is required to converge.
- The transverse momentum must be  $p_T > 100 \text{ MeV}/c$ , which reduces beam related background.

This set of track selection criteria is known as the `GoodTracksLoose`, and every charged track in all three analyses is required to satisfy them, except for the ones coming from a  $K_S$ . Figure 5.2(a) shows the track multiplicity for the signal MC simulation and offpeak data. The majority of the events containing two tracks are from QED processes, such as Bhabba scattering. The event topology is characterized as the ratio of the 2nd to the 0th Fox-Wolfram moments,  $R_2$ , which is shown in Fig. 5.2(b) after the requirement for the number of tracks. For a completely isotropic event,  $R_2 = 0$ , while for a completely collimated event,  $R_2 = 1$ . We require  $R_2 < 0.9$ . The peak at one for the offpeak data is from leftover electromagnetic processes.

Figure 5.2(c) displays the CMS energy distribution of the highest energy photon (a photon is identified as a “bump”, which is described in Section 2.2.4, in the EMC) in the event after the requirement of the number of tracks and  $R_2$ . In the CMS, the  $B$  meson is moving slowly, so that this photon is almost mono-energetic with  $E_{CMS}^* \approx 2.5 \text{ GeV}$ . We require  $1.5 < E_{CMS}^* < 3.5 \text{ GeV}$ .

High energy  $\pi^0$ 's are more likely to have decay products with a small opening angle. The two photons can then have two local maxima in the EMC that are sufficiently near each other such that the reconstruction algorithm cannot distinguish them. Such a situation is called a merged  $\pi^0$ , and tends to have a wider lateral moment than isolated photons. Hadronic interactions with the detector also have the same effect. The lateral moment is defined as

$$\text{LAT} = \frac{\sum_{i=2}^{N_{\text{crystals}}} E_i \Delta x_i}{\sum_{i=0}^{N_{\text{crystals}}} E_i \Delta x_i}, \quad (5.1)$$

where  $N_{\text{crystals}}$  is the number of crystals in an EMC cluster, each with energy  $E_i$  and ordered from the most energetic to the least. For  $i \geq 2$ ,  $\Delta x_i$  is the distance from the cluster local maximum to crystal  $i$ , while  $\Delta x_{0,1} = 5 \text{ cm}$ . An EMC bump with  $N_{\text{crystals}} \leq 2$  has a LAT of zero by definition. We require  $\text{LAT} < 0.8$ . In addition, the each  $\gamma$  from the  $\pi^0$  is required to have  $E_\gamma > 30 \text{ MeV}$  in the laboratory frame.

There are also a number of other selection criteria designated as preliminary. These include the invariant mass of the light vector meson ( $0.5 < m_{K^*} < 1.3 \text{ GeV}/c^2$ ,  $0.5 < m_{\rho^0} < 1.2 \text{ GeV}/c^2$ ,  $0.5 < m_{\rho^+} < 1.3 \text{ GeV}/c^2$ ,  $0.732 < m_\omega < 0.832 \text{ GeV}/c^2$ ,  $0.7 < m_\phi < 1.5 \text{ GeV}/c^2$ )

and the pseudoscalars which decay inside the detector ( $0.115 < m_{\pi^0} < 0.150 \text{ GeV}/c^2$ ,  $0.47 < m_{K_S} < 0.522 \text{ GeV}/c^2$ ). These requirements will be optimized and described further in the Section 5.3.

The preliminary criteria also include requirements on not only the properties of the  $\pi^0$ , but also its daughters. Both  $\gamma$  daughters are subject to the same requirements as the highest energy photon, with the exception of the selection on the center-of-mass energy. In addition, the  $\pi^0$  candidate is required to have an energy greater than 200 MeV in the laboratory frame, and is subject to a mass constrained fit in which the both  $\gamma$  daughters are required to originate from the primary vertex. This is done to recalculate the momentum of the daughters. The primary vertex is estimated per event by using selected charged tracks in the event, where each track is required to have an impact parameter in the x-y plane, calculated at the nominal interaction point, of less than 1 mm.

Finally, the last two variables used in the preliminary event selection are described in Section 5.2.1.

### 5.2.1 Kinematic Variables

Two powerful variables that are used to discriminate between signal and background are  $m_{ES}$  and  $\Delta E$ . For perfectly measured particles, they are completely anti-correlated. However, experimental resolution allows us to treat them as uncorrelated quantities.

The final state particles are used to reconstruct the vector meson, which is then combined with the  $\gamma$  candidate to form a  $B$  candidate. For each  $B$  candidate,  $\Delta E$  is defined in the CMS as:

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

where the “\*” indicates a CMS quantity,  $E_{\gamma}^*$  is the energy of the gamma,  $E_V^*$  is the energy of the vector meson, and  $E_{\text{beam}}^*$  is the beam energy (5.29 GeV in the CMS for on-resonance data). With the expectation that each  $B$  meson carries half the beam energy in the CMS this quantity measures how close the energy of the  $B$ -candidate is to the expected energy determined from the beam, which is more precisely known than any of the detector-reconstructed kinematic quantities in this analysis. Thus, for signal events  $\Delta E$  is centered around zero; we require  $|\Delta E| < 0.6 \text{ GeV}$  for the preliminary event selection. In the present case, the resolution of  $\Delta E$  is dominated by the high energy  $\gamma$ . Fig. 5.3(b) shows the comparison between  $K^{*0} \rightarrow K^+\pi^-$  signal and continuum MC simulation for

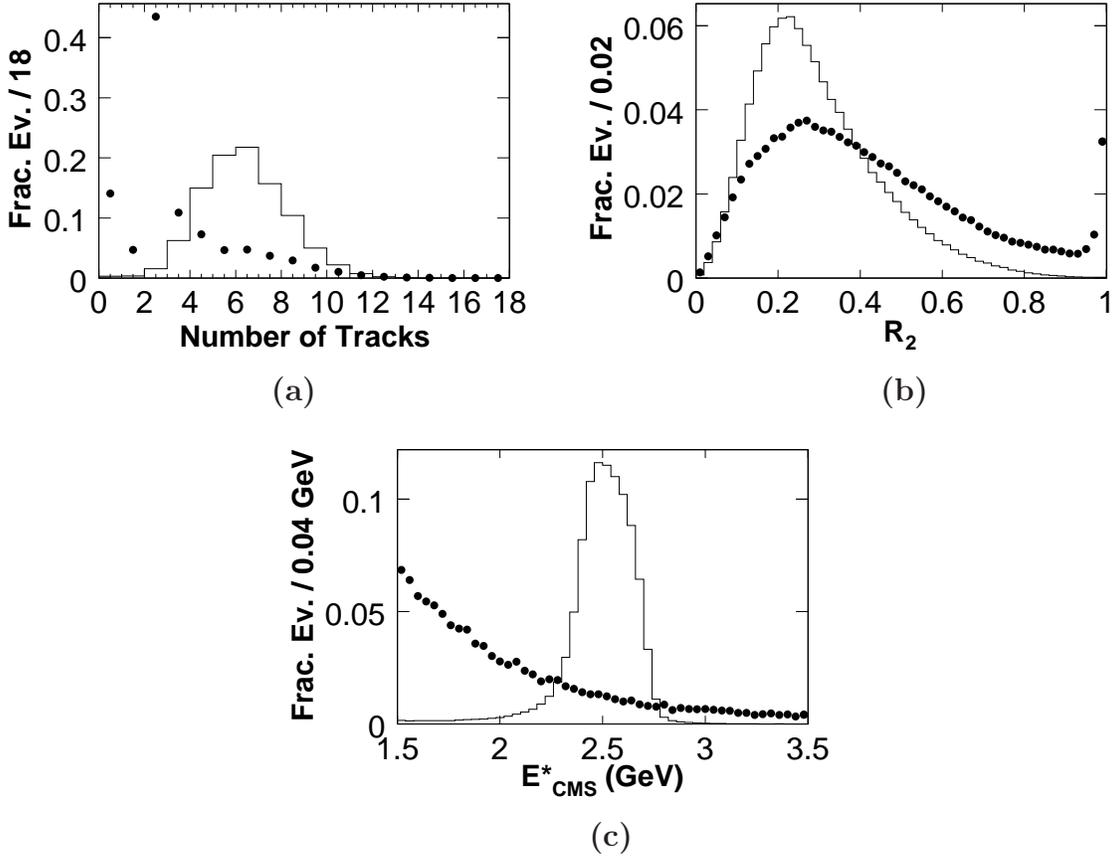


Figure 5.2: The distributions of the number of tracks (a),  $R_2$  (b), and the CMS energy of the highest energy photon (c) of the event. The solid line refers to  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  signal MC simulation, and the black dots represent the offpeak data. The distributions have been normalized to unity.

the  $\Delta E$  distribution. The pronounced tail of the  $\Delta E$  distributions at negative values arise because of the energy leakage in the calorimeter, as well as  $\pi^0, \eta$  decays where one photon has been assumed to be the signal photon. In the case of the signal MC simulation, this is mis-reconstructed signal.

The energy-substituted mass  $m_{\text{ES}}$  of the  $B$  meson candidate is reconstructed by using the initial state energy (the beam energy) and the momentum of the final state particles. It is defined by

$$m_{\text{ES}} = \sqrt{E_{\text{beam}}^{*2} - (\vec{p}_\gamma^* + \vec{p}_V^*)^2},$$

where  $\vec{p}_\gamma^*$  is the three momentum of the  $\gamma$  candidate and the three momentum of the

vector meson candidate is represented by  $\vec{p}_V^*$ . For signal events,  $m_{ES}$  is centered at the  $B$  meson mass ( $5.28 \text{ GeV}/c^2$  [4]). The  $m_{ES}$  distribution for  $K^{*0} \rightarrow K^+\pi^-$  signal and continuum MC simulation is shown in Fig. 5.3(a). We require  $5.1 < m_{ES} < 5.5 \text{ GeV}/c^2$ .

Furthermore, each analysis is “blind”, which means that the contents of the  $m_{ES}$  and  $\Delta E$  region which contain the signal events (defined by  $5.27 < m_{ES} < 5.29 \text{ GeV}/c^2$  and  $-0.2 < \Delta E < 0.1 \text{ GeV}$ ) is kept hidden from the analyst. This is to prevent a biased analysis.

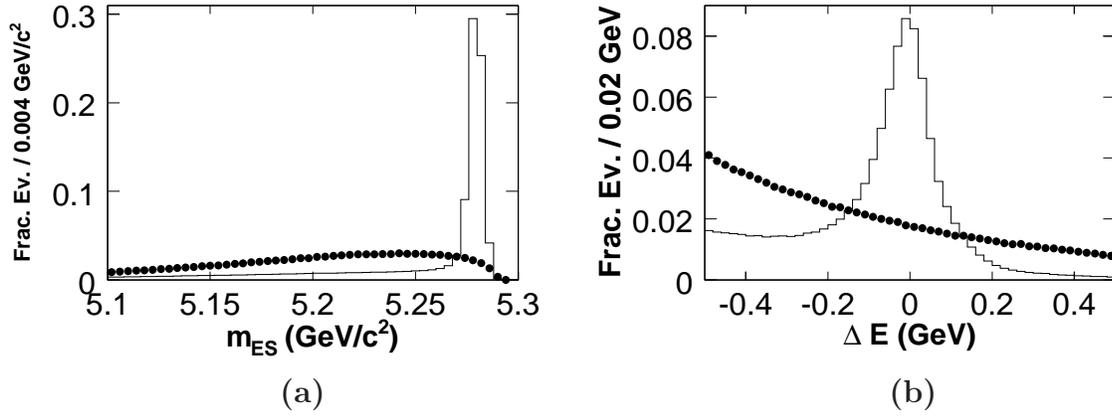


Figure 5.3: The  $m_{ES}$  (a) and  $\Delta E$  (b) distributions after the preliminary cuts have been applied. The solid line refers to the  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  signal MC simulation, while the black dots represent the continuum MC simulation. The distributions have been normalized to unity.

## 5.3 Candidate Selection

This section describes the selections related to individual particles. Some of these selections are optimized as described in Section 5.5, and will be stated explicitly.

### 5.3.1 Particle Identification

Kaons and pions are identified by using  $dE/dx$  information from the SVT and DCH, as well as using the Cherenkov angle measurement and the number of Cherenkov photons from the DIRC. These measurements are used to form a likelihood  $L_i$ , which is the product of the individual likelihoods for the SVT, DCH, and DIRC, for a particle type  $i$  ( $i = K, \pi, p$ ). The SVT and DCH are applicable approximately in the momentum range  $|p| < 0.7 \text{ GeV}/c$ , while the DIRC is valid for  $|p| > 0.6 \text{ GeV}/c$ . Electrons are identified

using a product of likelihoods  $L'_i (i = e, \pi, K, p)$  using the deposited energy and the lateral shapes of showers in the EMC, as well as information from the DCH and DIRC. Muons are chosen using a cut-based selector which uses primarily IFR information.

For the  $B \rightarrow K^*\gamma$  analysis, the kaons are selected with the kaon **Tight** selection, which is designed to keep the pion misidentification rate below 5%. This selection requires  $L_K/(L_K+L_\pi) > 0.9$  and  $L_K/(L_K+L_p) > 0.2$ , and, if the momentum of the kaon candidate is above 0.4 GeV/c,  $L'_e/(L'_e + L'_\pi + L'_K + L'_p) < 0.95$ . This latter requirement is known as the **Tight** veto for electrons. An equivalent selection is made for the kaons from the  $B \rightarrow \phi\gamma$  analysis. The pions of the  $B \rightarrow K^*\gamma$  analysis are required to fail the kaon **Tight** selector. This is done because the pion efficiency is greater with this condition than with a selection based on a pion selector, due to the pion selector's requirement of the particle being in the DIRC acceptance. By requiring the pion candidate to fail the kaon **Tight** selector, we accept pions that fall outside of the DIRC acceptance, as well as inside. For the  $B \rightarrow (\rho/\omega)\gamma$  analysis, in which the  $B \rightarrow K^*\gamma$  background is important, the choice of selector was optimized in a previous analysis [33]. Two pion selectors are utilized: **VeryTight**, which requires that  $L_K/(L_K + L_\pi) < 0.2$ ,  $L_p/(L_p + L_\pi) < 0.5$ , and **Tight**, which requires  $L_K/(L_K + L_\pi) < 0.5$  and  $L_p/(L_p + L_\pi) < 0.98$ . The **VeryTight** selector also requires the veto of a muon selector in which the  $\pi, e$ , and  $K$  misidentification rates are below 2.5%, and the muon efficiency is  $\approx 70\%$ . This is known as the **VeryTight** selector for muons. Additionally, both selectors require the electron **Tight** veto. For the  $B \rightarrow \rho\gamma$  decay modes, a selection based on the pion **VeryTight** criteria is made. In addition, a selection of 0.0001 is made on the minimal significance criteria of the Poisson probability of the number of photons seen in the DIRC for a particular candidate being different from the expected number for the assigned pion hypothesis. Using a  $D^* \rightarrow D^0\pi^+(D^0 \rightarrow K^+\pi^-)$  control sample, Fig 5.4 shows the pion efficiencies and kaon mis-identification rates for the  $B \rightarrow \rho\gamma$  selector. For  $B \rightarrow \omega\gamma$ , the  $B \rightarrow K^*\gamma$  background is not quite as significant, so the pion **Tight** criteria is used with no photon consistency requirement.

### 5.3.2 $K_S$ Selection

As mentioned previously, the  $K_S$  is reconstructed through the mode  $K_S \rightarrow \pi^+\pi^-$ . The track parameters and covariance matrix of the two pions are used to form a  $\chi^2$ , which is minimized to extract the vertex and the 4-momentum of the  $K_S$ . In addition, various constraints are used in the  $\chi^2$  fit. These include a vertex constraint, which requires that the production vertex of each pion is equal to the decay vertex of the  $K_S$ , the kinematic

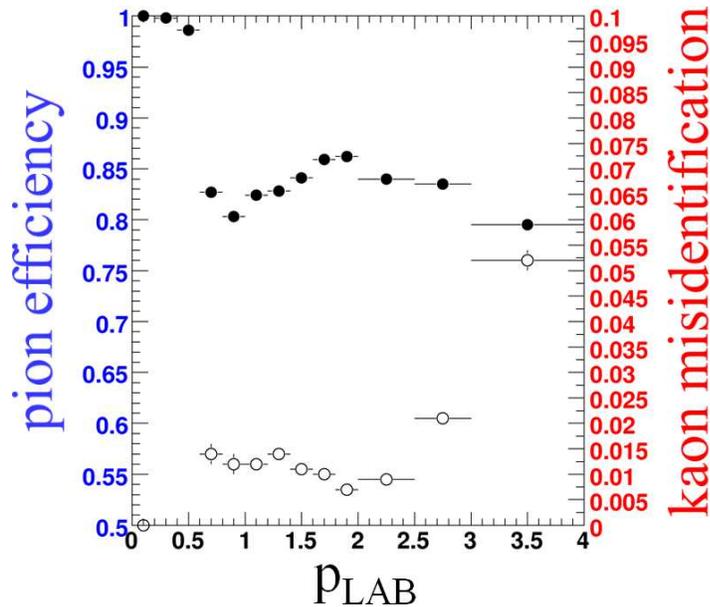


Figure 5.4: Pion efficiency and kaon misidentification rates of the  $B \rightarrow \rho\gamma$  selector using the  $D^*$  control sample. The filled circles represent the  $\pi^\pm$  efficiency and use the left-hand scale, while the open circles represent the  $K^\pm$  mis-identification rates and use the right-hand scale.

constraint, which requires the the 4-momentum of the  $K_S$  is equal to the sum of the two pions, and a mass constraint, which assures the invariant mass of the  $K_S$  is equal to its mass hypothesis. Fig. 5.5(a) shows the  $\chi^2$  probability distribution for signal and continuum MC simulation for the  $K^{*+} \rightarrow K_S\pi^+$  mode. We require the probability of  $\chi^2 > 0.001$ , which essentially eliminates badly reconstructed candidates. In addition, we make a requirement on the invariant mass of the two pions. This is optimized, and for the  $K^{*0} \rightarrow K_S\pi^0$  mode is  $0.486 < m_{\pi^+\pi^-} < 0.521 \text{ GeV}/c^2$ , while the  $K^{*+} \rightarrow K_S\pi^+$  mode is  $0.483 < m_{\pi^+\pi^-} < 0.521 \text{ GeV}/c^2$ . Fig. 5.5(b) displays this after the  $\chi^2$  probability requirement. Finally, the  $K_S$  decay length and error is extracted from a  $\chi^2$  fit to the entire decay chain [50]. The constraints of the fit are the vertex and kinematic constraints, as well as a beamspot constraint, which forces the  $\Upsilon(4S)$  candidate to originate from the beamspot, the beam-energy constraint, which requires the energy of the  $\Upsilon(4S)$  candidate to be equal to the energy of the beam, and the lifetime constraint, which sets the sum of the lifetimes of the 2  $B$  candidates to be equal to twice the nominal value of the lifetime of the  $B$  meson. The latter constraint is applied as a Gaussian, in which the width of the Gaussian is equal to the root-mean-square of the lifetime distribution. The flight significance is defined as the reconstructed decay length of the  $K_S$  divided by its

uncertainty, and is shown after the probability of  $\chi^2$  selection and invariant mass selection in Fig. 5.5(c). Combinatorial background will come primarily from the event vertex, while there is a decay length associated with real  $K_S$ 's. Therefore, the latter will have a larger flight significance than the former. This is also optimized, and for the  $K^{*0} \rightarrow K_S\pi^0$  mode is required to be greater than 9.3, while the  $K^{*+} \rightarrow K_S\pi^+$  selection is greater than 10.27.

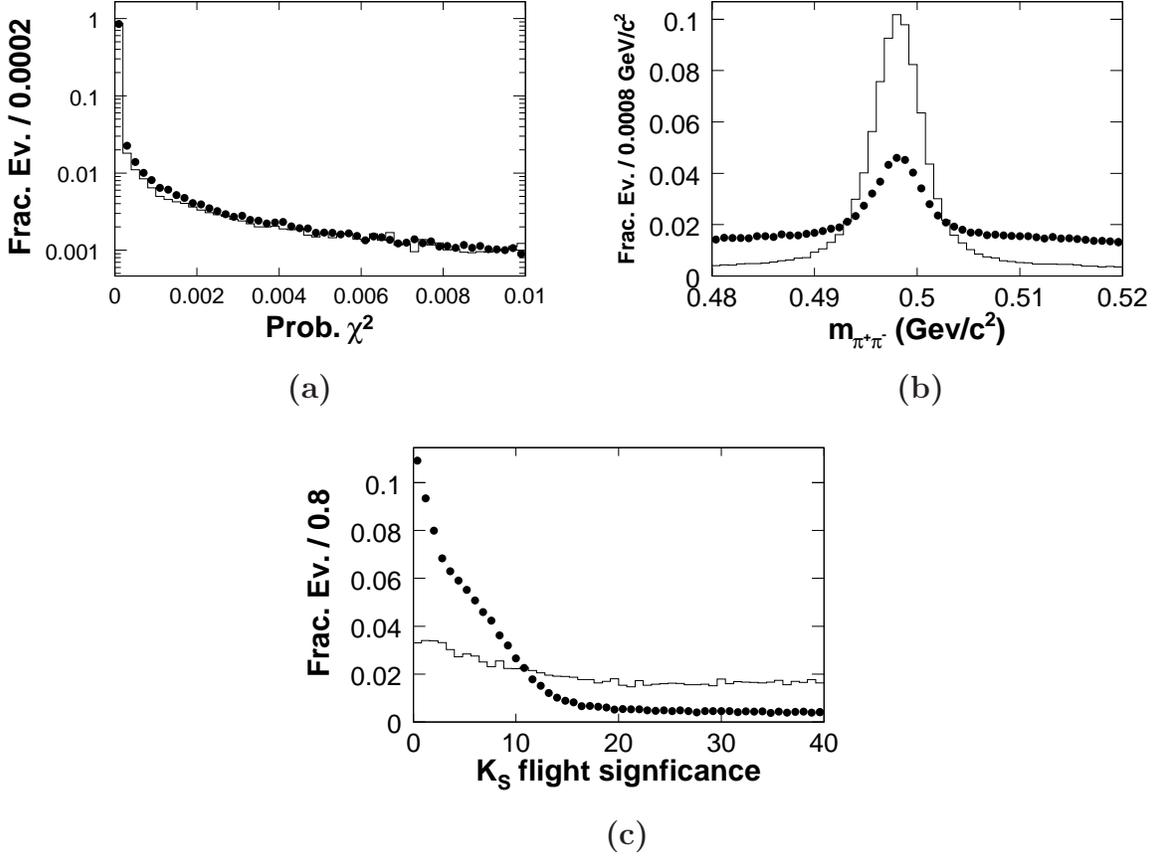


Figure 5.5: The probability of the  $\chi^2$  fit to the  $K_S$  decay vertex (a), the  $K_S$  invariant mass (b), and the  $K_S$  flight significance (c) are displayed for signal MC simulation (solid line) and continuum MC simulation (black dots). The distributions have been normalized to unity and the mode shown is  $K^{*+} \rightarrow K_S\pi^+$ .

### 5.3.3 $\pi^0$ Selection

Fig. 5.6 shows the invariant mass of the  $\pi^0$  for the  $K^{*+} \rightarrow K^+\pi^0$  mode. As can be seen, additional discrimination between signal and continuum is small, and the optimized selection for the  $K^{*0} \rightarrow K_S\pi^0$  mode is  $0.112 < m_{\gamma\gamma} < 0.150 \text{ GeV}/c^2$ , while for the  $K^{*+} \rightarrow K^+\pi^0$  mode it is  $0.114 < m_{\gamma\gamma} < 0.15 \text{ GeV}/c^2$ . For the  $B^+ \rightarrow \rho^+\gamma$ , the selection is

$0.117 < m_{\gamma\gamma} < 0.148 \text{ GeV}/c^2$ , and  $0.122 < m_{\gamma\gamma} < 0.150 \text{ GeV}/c^2$  for  $B \rightarrow \omega\gamma$ . In addition, for the  $B \rightarrow (\rho/\omega)\gamma$  analysis, we consider the cosine of the opening angle the  $\gamma$  daughters of the  $\pi^0$ :  $\cos\theta_{\gamma\gamma}$ . Continuum events are marked by high-energy  $\pi^0$ 's, which will have a small opening angle, while softer  $\pi^0$ 's will have a larger opening angle.

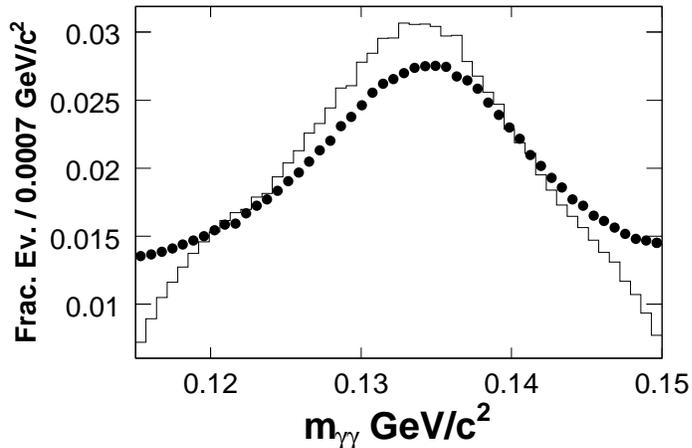


Figure 5.6: The invariant mass of the  $\pi^0$  for signal MC simulation (solid line) and continuum MC simulation (black dots) for the  $K^{*+} \rightarrow K^+\pi^0$  mode.

### 5.3.4 Vector Meson Selection

For the cases in which the vector meson decays to two charged tracks, a vertex for the meson is calculated using a  $\chi^2$  fit, in which the vertex and kinematic constraints are used. For the  $K^{*0} \rightarrow K^+\pi^-$  and the  $\phi \rightarrow K^+K^-$  modes, the selection on the  $\chi^2$  probability is fixed to be greater than 0.01. For the  $\rho^0 \rightarrow \pi^+\pi^-$  and  $\omega \rightarrow \pi^+\pi^-\pi^0$  modes, the  $\chi^2$  probability selection is optimized, with the result that  $\text{Prob}(\chi_{\rho^0}^2) > 0.008$  and  $\text{Prob}(\chi_{\omega}^2) > 0.0$ . For the  $\omega \rightarrow \pi^+\pi^-\pi^0$  mode, there is an optimized  $K_S$  veto in which the distance from the vertex of the  $\omega$  candidate and the beamspot is required to be less than 0.189 cm. For all modes, an optimized selection is made on the invariant mass of the vector meson. The selections for the  $B \rightarrow K^*\gamma$  analysis are  $0.781 < m_{K^+\pi^-} < 1.086 \text{ GeV}/c^2$ ,  $0.820 < m_{K_S\pi^0} < 1.000 \text{ GeV}/c^2$ ,  $0.786 < m_{K^+\pi^0} < 1.000 \text{ GeV}/c^2$ , and  $0.792 < m_{K_S\pi^+} < 1.03 \text{ GeV}/c^2$ . The optimized selections for the  $B \rightarrow (\rho/\omega)\gamma$  analysis give  $0.633 < m_{\pi^+\pi^-} < 0.967 \text{ GeV}/c^2$ ,  $0.636 < m_{\pi^+\pi^0} < 0.932 \text{ GeV}/c^2$ , and  $0.764 < m_{\pi^+\pi^-\pi^0} < 0.795 \text{ GeV}/c^2$ , while the invariant mass of the  $\phi$  in the  $B \rightarrow \phi\gamma$  analysis is  $1.011 < m_{K^+K^-} < 1.029 \text{ GeV}/c^2$ . For the  $K^{*0} \rightarrow K^+\pi^-$  mode, the invariant mass and the probability of  $\chi^2$  of the fit is shown

in Fig. 5.7. The separation of the signal and continuum MC simulation is evident in Fig. 5.7(b).

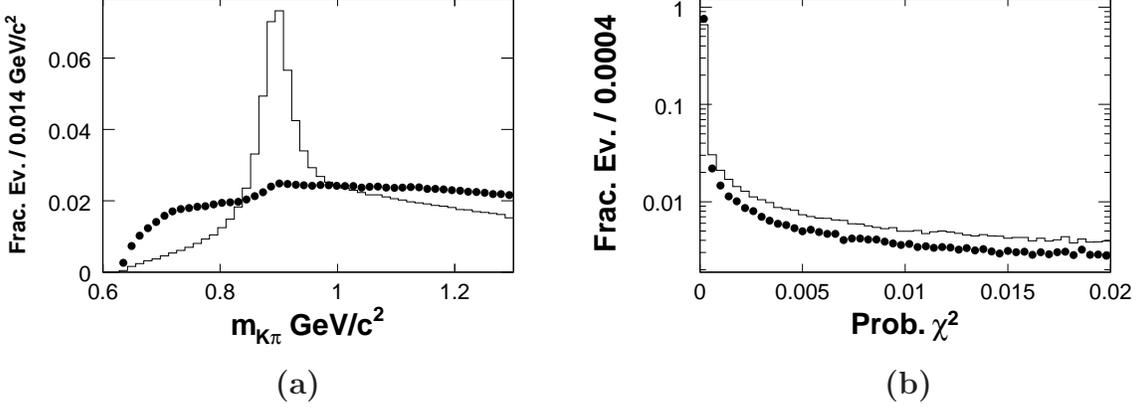


Figure 5.7: The  $K^*$  invariant mass (a) and the probability of the  $\chi^2$  fit to the  $K^*$  decay vertex (b) are displayed for signal MC simulation (solid line) and continuum MC simulation (black dots). The distributions have been normalized to unity and the mode shown is  $K^{*0} \rightarrow K^+\pi^-$ .

### 5.3.5 High-Energy Photon Selection

As mentioned in Section 5.2, a photon is identified as a local energy maximum in the EMC, which is not associated with any charged track. To further ensure that the photon and a charged track do not overlap, as well as eliminating EMC clusters at the edge of the EMC, we restrict the range of  $\cos\theta$  to be  $[-0.74, 0.93]$ , where  $\theta$  is the polar angle with respect to the  $z$ -axis, which points in the direction of the HER beam. The tracking coverage outside this range is dominated by the DCH and SVT. In addition, the photon must not contain any crystals that have been identified as noisy or dead by the online monitoring. However, sometimes the monitoring may miss noisy channels. In this case, the number of crystals of the EMC bump will be one or a few. Therefore, we require the number of crystals that the photon shower traverses to be greater than 4. Hadronic neutral objects which interact with the EMC can have multiple bumps near each other. We require the EMC bump associated with the high-energy photon candidate to be 25 cm from the closest charged or neutral EMC bump. These distributions are shown in Figs. 5.8(a) and (b). Hadronic objects and merged  $\pi^0$ 's also have a wider shower shape, as mentioned in Section 5.2. We define the second moment of the bump, which characterizes the lateral profile of energy distribution across the crystals, to be

$$L_2^{bump} = \sum_{i=0}^{N_{\text{crystals}}} \frac{E_i [(\theta_i - \theta_c)^2 + (\phi_i - \phi_c)^2]}{\sum_i E_i}, \quad (5.2)$$

where  $\theta_c, \phi_c$  are the angular coordinates of the centroid maxima, and  $E_i, \theta_i, \phi_i$  are the energy and coordinates of the  $i$ th crystal in the bump. This is optimized, and we require  $L_2^{bump} < 0.002$ . This distribution is shown in Fig. 5.9. The differences of hadronic to electromagnetic showers are also characterized by the variable `s9s25`, which is the ratio of the sums of the energies of the central 9 crystals to the central 25 crystals surrounding the centroid. This variable is used only for the  $B \rightarrow (\rho/\omega)\gamma$  analysis, and has smaller values for hadronic showers than for electromagnetic ones.

Finally, we veto the  $\pi^0 \rightarrow \gamma\gamma$  and  $\eta \rightarrow \gamma\gamma$  decays by forming a likelihood ratio for the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analyses. Photons from  $\pi^0$  and  $\eta$  decays are a major source of background, as the branching fractions of the decays of these mesons into two photons are 99% and 39% respectively. To suppress these backgrounds, we associate the photon candidate ( $\gamma_1$ ) with all other EMC bumps in the event which are not matched to a track ( $\gamma_2$ ), and form  $\pi^0 / \eta$  likelihood functions. For the  $B \rightarrow (\rho/\omega)\gamma$  analysis, we also associate each  $\gamma_1$  with photons that have converted to electron-positron pairs. These converted photons are required to satisfy the following: the invariant mass  $m(e^+e^-)$  smaller than 30 MeV, have a distance of closest approach (DOCA) to the beamspot in the x-y plane  $< 0.5$  cm, and have a DOCA to the beamspot along the z-axis of  $< 1.0$  cm. We form the invariant mass  $m(\gamma_1 e^+ e^-)$ . The output of the likelihood functions is a number between 0 (signal) and 1 ( $\pi^0$  or  $\eta$ ), and we veto the event if either the likelihood functions or  $m(\gamma_1 e^+ e^-)$  is consistent with that of a  $\pi^0$  or  $\eta$ . The  $\pi^0$  and  $\eta$  likelihood functions are defined as

$$\mathcal{LR}(\theta) = \frac{p(M(\gamma_1\gamma_2), E_{\gamma_2}|\theta)}{p(M(\gamma_1\gamma_2), E_{\gamma_2}|\rho\gamma) + p(M(\gamma_1\gamma_2), E_{\gamma_2}|\theta)},$$

where  $\theta$  is either  $\pi^0$  or  $\eta$ , and  $p$  is probability density function (PDF) in terms of the invariant mass of a pair of photons,  $M(\gamma_1\gamma_2)$ , and the energy of  $\gamma_2$  in the lab frame,  $E_{\gamma_2}$ . We determine the PDF's using signal and continuum MC simulation. Figs. 5.8(c) and (d) show the output of the  $\pi^0$  and  $\eta$  likelihood functions.

For the  $B \rightarrow \phi\gamma$  analysis, we form a mass veto only. The photon candidate ( $\gamma_1$ ) is combined with candidates ( $\gamma_2$ ) in the event. To avoid beam background contamination, we select  $\gamma_2$  candidates with an energy greater than 50 MeV for the  $\pi^0$  veto and greater than 250 MeV for the  $\eta$  veto. Then, the  $\gamma\gamma$  invariant mass of the remaining candidates

is formed. If any combination has an invariant mass  $\pm 0.02$  GeV from the  $\pi^0$  mass or  $\pm 0.04$  GeV from the  $\eta$  mass, then it is excluded from the reconstruction. The performance of the vetoes is given in Figures 5.10 and 5.11 for  $B^0 \rightarrow \rho^0 \gamma$  signal MC simulation.

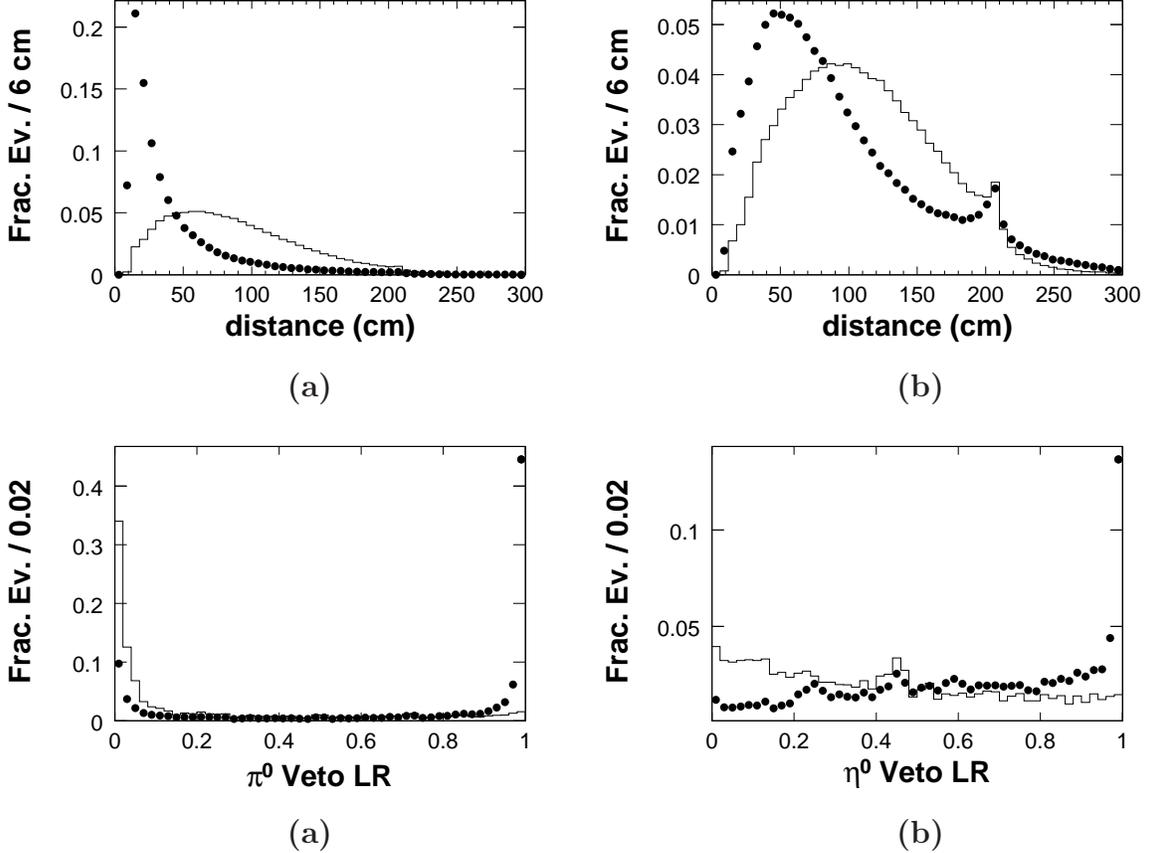


Figure 5.8: The distance of the photon candidate to the nearest neutral bump (a), the distance to the nearest charged bump (b), and the output of the  $\pi^0$  (c) and  $\eta$  (d) veto likelihood functions are displayed for signal MC simulation (solid line) and continuum MC simulation (black dots). The distributions have been normalized to unity and the mode shown is  $K^{*0} \rightarrow K^+ \pi^-$ .

## 5.4 Continuum Background Suppression

The dominant background in all of the analyses comes from the continuum, which mostly occurs in back-to-back high energy jets. The background is mostly from events that contain  $\pi^0$ 's and  $\eta$ 's. Continuum events have a much different topology than signal  $B$  events, which are isotropic due to the small 3-momentum of the  $B$  meson in the  $\Upsilon(4S)$  rest frame. Therefore, we consider variables that exploit these differences, which are

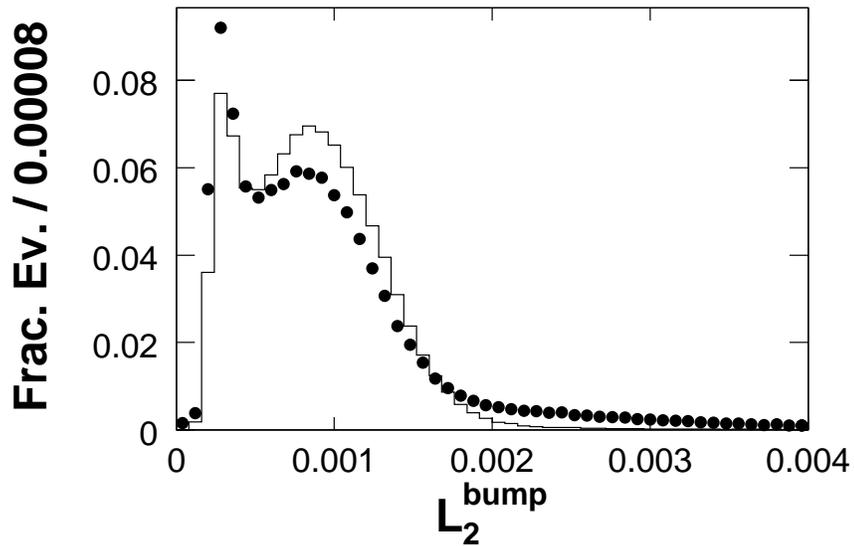


Figure 5.9: The second moment ( $L_2^{bump}$ ) of the photon shower is displayed for signal MC simulation (solid line) and continuum MC simulation (black dots). The distribution has been normalized to unity and the mode shown is  $K^{*0} \rightarrow K^+\pi^-$ .

called event-shape variables. In addition, continuum events proceed through the strong interactions, which do not allow for flavor changing currents, while  $B$  decays proceed through the weak interactions, which do allow for these kinds of currents. We also consider variables to distinguish between these two types of situations. Each analysis has a unique set of variables as inputs to the neural network, which are described in the following sections.

### 5.4.1 Input Variables

The inputs to the neural networks for all three analyses will be first described. During this discussion, the differences in the three analyses will be mentioned. Following this discussion is a summary of inputs to each analysis.

The inputs are the following:

- $\cos \theta_T$

The thrust angle  $\theta_T$  is the angle between the high-energy photon and the thrust axis of the rest of the event (ROE) computed in the CM frame. The thrust axis of the ROE is defined as the axis which maximizes the sum of the particle momenta along

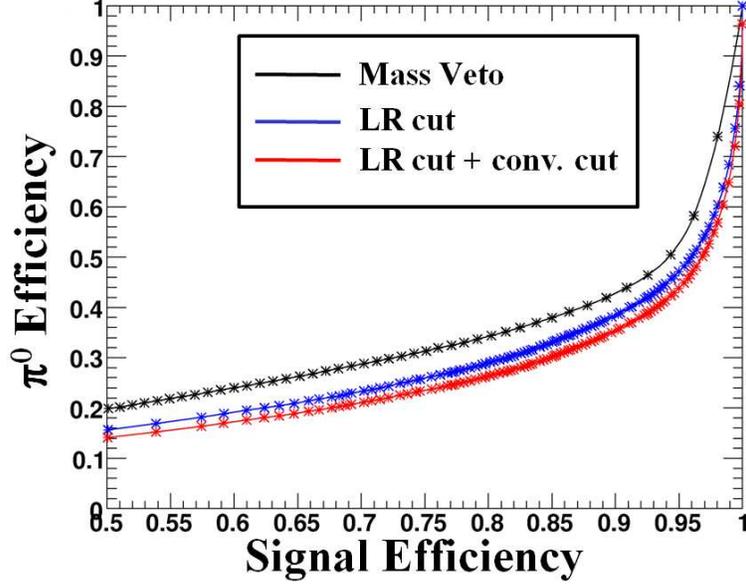


Figure 5.10:  $\pi^0$  efficiency vs. signal efficiency for different veto methods. The black curve is for the veto using only  $M(\gamma_1\gamma_2)$ , the blue curve is for the likelihood method, and the red curve is for the likelihood veto plus converted photons.

this axis, where the particles are not associated with the signal  $B$  candidate. Since  $B$  decays are isotropic in the CM frame, the distribution of  $\cos\theta_T$  is thus expected to be flat. This is not true for continuum events, where the high-energy photon originates usually from a high energy  $\pi^0$  or  $\eta$  decaying in one of these jets. Thus, the direction of the high energy photon is almost identical with the direction of one of the jets. For continuum events,  $\cos\theta_T$  is thus expected to be close to  $-1$  or  $1$ .

- $\cos\theta_B$

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

- Legendre moments and Energy Cones

We further take advantage of the differences in energy distribution between continuum and signal events by considering two different, but equivalent, ways to represent the same information. The two sets of variables are called Legendre moments and

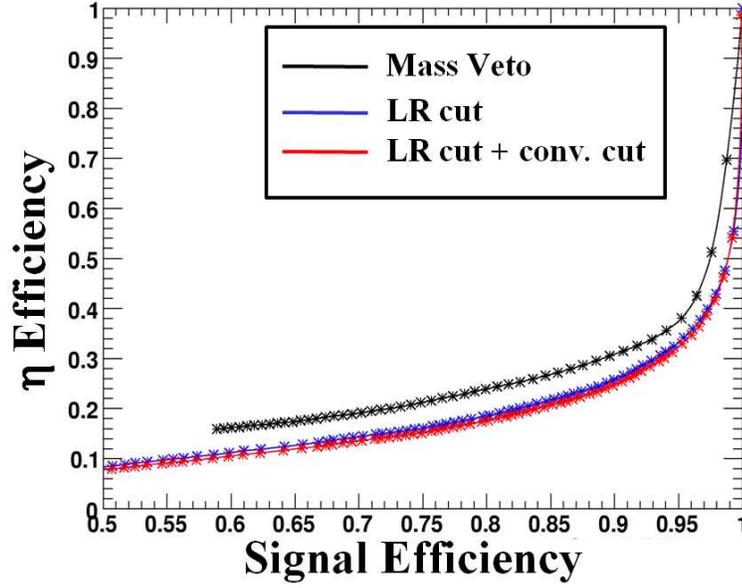


Figure 5.11:  $\eta$  efficiency vs. signal efficiency for different veto methods. The black curve is for the veto using only  $M(\gamma_1\gamma_2)$ , the blue curve is for the likelihood method, and the red curve is for the likelihood veto plus converted photons.

energy cones. The Legendre moments are defined as follows:

$$L_i = \sum_{j \text{ in ROE}} |\vec{P}_j| \times |\cos \theta_j|^i$$

$\theta_j$  is the angle between the  $j$ th particle and a specific event axis, both in the CM frame. Two event axes are used in this analysis: the direction of high-energy photon and the direction of the thrust-axis  $\vec{T}^*$  of the ROE.  $\vec{P}_j$  is the momentum vector of the  $j$ th particle. In addition, only two Legendre moments are considered,  $L_1$  and  $L_2$ , and both are normalized to  $L_0$ . In continuum events,  $|\cos \theta_j|$  has values that are biased towards unity, while in signal events this is not true. Therefore, the Legendre moments have greater values for continuum events than for signal events. An alternate way to represent the same information is to consider the energy distribution of the tracks and neutral candidates in the ROE, and form energy cones. The energy is binned into cones of 10-degree increments in the CMS about the photon, for a total of 18 energy cones. Continuum events have energy mainly distributed in a jet, whose center is the photon candidate. Therefore, these events have particles with higher energies at small angles. Signal events are more isotropically distributed, and do not have this behavior. The Legendre moments are

used in the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analyses, while the energy cones are used in the  $B \rightarrow \phi\gamma$  analysis. They are equivalent, and the reason for the differences in the analyses is simply historical. The Legendre moments are considered an improvement because they encode equivalent information more simply.

- $R'_2$

$R'_2$  is  $R_2$  calculated in a frame in the high-energy photon has been subtracted from the total event. In this frame, an ISR event behaves like a two-jet continuum event, thus this variable tends to have larger values for ISR events.

- $\Delta z$  significance

Using a global decay chain fit, described in Section 5.3.2, we obtain a value for the distance between the vertices of the  $B$  candidates ( $\Delta z$ ), as well as the resulting error ( $\sigma_z$ ). The  $\Delta z$  significance is defined as  $\Delta z/\sigma_z$ . Due to the lifetime of the  $B$ , signal  $B$  events have a larger significance than continuum events, which originate from the primary vertex.

- Flavor Identification of  $B^0/\bar{B}^0$

Continuum events proceed strongly, so should have minimal net flavor production. In contrast,  $B^0/\bar{B}^0$  events decay weakly, which allows for flavor-changing currents. We consider two different ways to take advantage of this information. One is neural network based, while the other simply counts the net flavor production in the ROE. The neural network based identification is described first, which is followed by the description of counting the net flavor production.

*BABAR* utilizes several neural networks in order to distinguish  $B^0$  from  $\bar{B}^0$  mesons. They are described in detail in Ref. [51]. Each neural network combines variables together, and the output is a number between  $-1$  and  $1$ , where  $-1$  indicates a  $\bar{B}^0$  and  $1$  indicates a  $B^0$ . The number  $0$  corresponds to the case where no flavor was identified. We do not distinguish between  $B^0$  and  $\bar{B}^0$ , but instead utilize the fact that continuum events will have a neural network output near  $0$ , while  $B\bar{B}$  events will have an output near  $\pm 1$ . Each of the four algorithms displayed in Fig. 5.13 corresponds to a different neural network. The “Electron Tag” and “Muon Tag” neural networks use the charge of the electron or muon, respectively, to identify the flavor of the  $B^0$  meson. The “Slow-Pion Tag” and the “Kaon Slow-Pion Tag” use the slow pion from  $D^*$  decays, and the correlations between the kaon and slow

pion from  $D^*$  decays, respectively, for flavor identification. Specifically, this set of variables was used for the  $B \rightarrow K^*\gamma$  analysis. For the  $B \rightarrow (\rho/\omega)\gamma$  analysis, in addition to using the output the neural networks, we use the inputs of those neural networks. This is in order to increase the sensitivity of those inputs for continuum rejection, rather than flavor identification. The neural network for the  $B \rightarrow (\rho/\omega)\gamma$  analysis was not done by the author, and so the details of the inputs, of which there are 18, will not be given here. The details can be found in Ref. [52].

The  $B \rightarrow \phi\gamma$  analysis uses the net flavor production of the ROE. This is defined by the following quantity

$$N_{\mathcal{F}} = |N_e| + |N_\mu| + |N_K| + |N_{Sl.\pi}| + N_{K_S^0} \quad (5.3)$$

where

1. Net kaon number:  $N_K =$  Number of  $K^+$  minus the number of  $K^-$ .
2. Net electron number:  $N_e =$  Number of  $e^+$  minus the number of  $e^-$ .
3. Net muon number:  $N_\mu =$  Number of  $\mu^+$  minus the number of  $\mu^-$ .
4. Net slow  $\pi$  number:  $N_{Sl.\pi} =$  Number of slow  $\pi^+$  minus the number of slow  $\pi^-$ .
5.  $K_S^0$  number:  $N_{K_S^0} =$  Number of  $K_S^0$ .

The selection of the various components of  $N_{\mathcal{F}}$  proceeds in the following way:

1. Kaon: **Tight** selection.
2. Electron: **Tight** selection (the conjugate of electron **Tight** veto) with  $p_{CMS} > 0.5 \text{ GeV}/c$ .
3. Muon: **Tight** selection (similar to the **VeryTight** selection, except that the  $\pi, e,$  and  $K$  misidentification rates are below 3.0%, and the muon efficiency is  $\approx 74\%$ ) with  $p_{CMS} > 1.0 \text{ GeV}/c$ .
4. Slow  $\pi$ :  $p_{CMS} < 0.250 \text{ GeV}/c$  with  $\cos\theta_{Thrust/p_\pi} > 0.8$ ,  $d_0 < 0.5 \text{ cm}$  where  $\theta_{Thrust/p_\pi}$  is the angle between the thrust axis of the event and the three momentum of the pion and  $d_0$  is the distance of closest approach to the primary vertex.
5.  $K_S^0$ : Successfully vertexed candidate with  $M_{K_S^0} = [0.480, 0.516] \text{ GeV}/c^2$ , at least 1 mm displacement from the  $\phi \rightarrow K^+K^-$  vertex and  $\cos\theta_{Disp/p_{K_S^0}} > 0.98$  where

$\theta_{Disp/p_{K_S^0}}$  is the angle between the difference of the  $K_S^0$  and the  $\phi$  vertexes and the  $\phi$  lab three momentum.

The candidates are selected in the order given *i.e.* Kaons are classified first, electron candidates classified second from the remaining tracks, muons classified third, etc. In all cases, the tracks do not overlap the reconstructed  $B \rightarrow \phi\gamma$  decay.

The reason for the differences in the analyses is, like the case of the energy distribution variables, historical. The neural network based set of variables is an improvement over the net flavor production because it considers more information in a more efficient way.

- $R_2^{All}$

This was described in Section 5.2.

- $\cos\theta_H$

The helicity angle  $\theta_H$  is the angle between the momentum of one of the daughters of the vector meson and the inverse of the momentum of the reconstructed  $B$  candidate in the rest frame of the vector meson. As shown in Section 4.3.1, for signal events,  $\cos\theta_H$  has a distribution of  $\sin^2\theta$ . Continuum events have a flat distribution in this variable.

The  $B \rightarrow K^*\gamma$  analysis used the variables  $\cos\theta_T$ ,  $\cos\theta_B$ , Legendre moments,  $R_2'$ ,  $\Delta z$  significance, flavor identification, and  $R_2^{All}$ . A comparison between signal and continuum MC simulation for each input for the  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  mode can be found in Figs. 5.12 and 5.13, while a comparison between offpeak data and continuum MC simulation can be found in Appendix A. All four modes of the  $B \rightarrow K^*\gamma$  analysis use identical input variables. The  $B \rightarrow (\rho/\omega)\gamma$  analysis used the same input variables as the  $B \rightarrow K^*\gamma$  analysis, with the exception of the difference described for the flavor identification variables. Finally, the  $B \rightarrow \phi\gamma$  analysis used  $\cos\theta_T$ ,  $\cos\theta_B$ ,  $\cos\theta_H$ ,  $R_2'$ ,  $\Delta z$ , energy cones, and net flavor in the ROE. A comparison between signal and continuum MC simulation for the inputs that have not already been shown for  $B \rightarrow K^*\gamma$  can be found in Figs. 5.14 through 5.16.

### 5.4.2 Neural Network Output

The above variables are combined into a neural network, which is developed using the ‘‘Stuttgart Neural Network Simulator’’ (SNNS) [53]. We divide the MC simulation sample

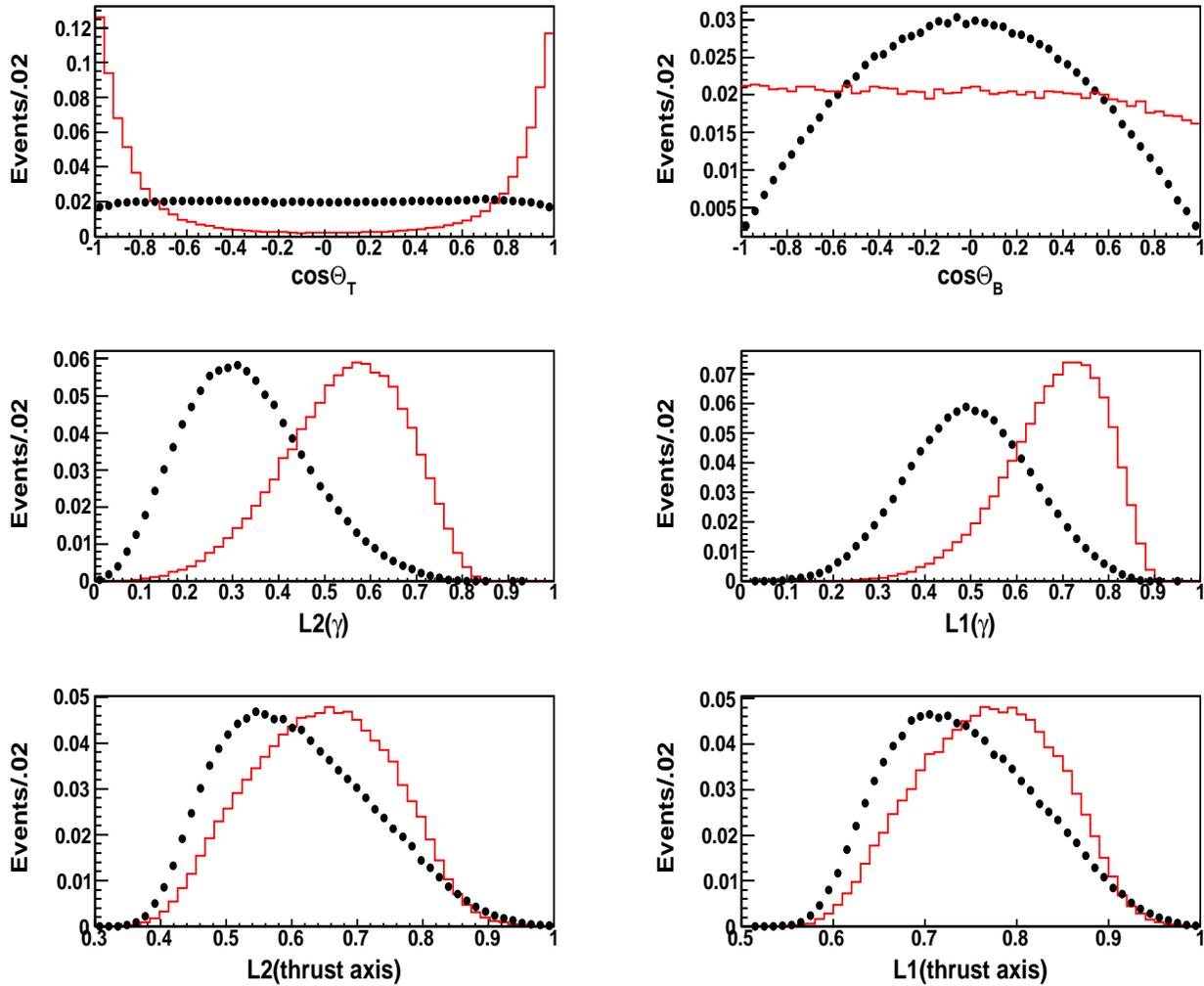


Figure 5.12:  $\cos\theta_T$ ,  $\cos\theta_B$ , and Legendre moments for signal (black dots) and continuum Monte Carlo (red line). The mode shown is  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  and all cuts are applied except for the Neural Network cut. The plots are normalized to the same area.

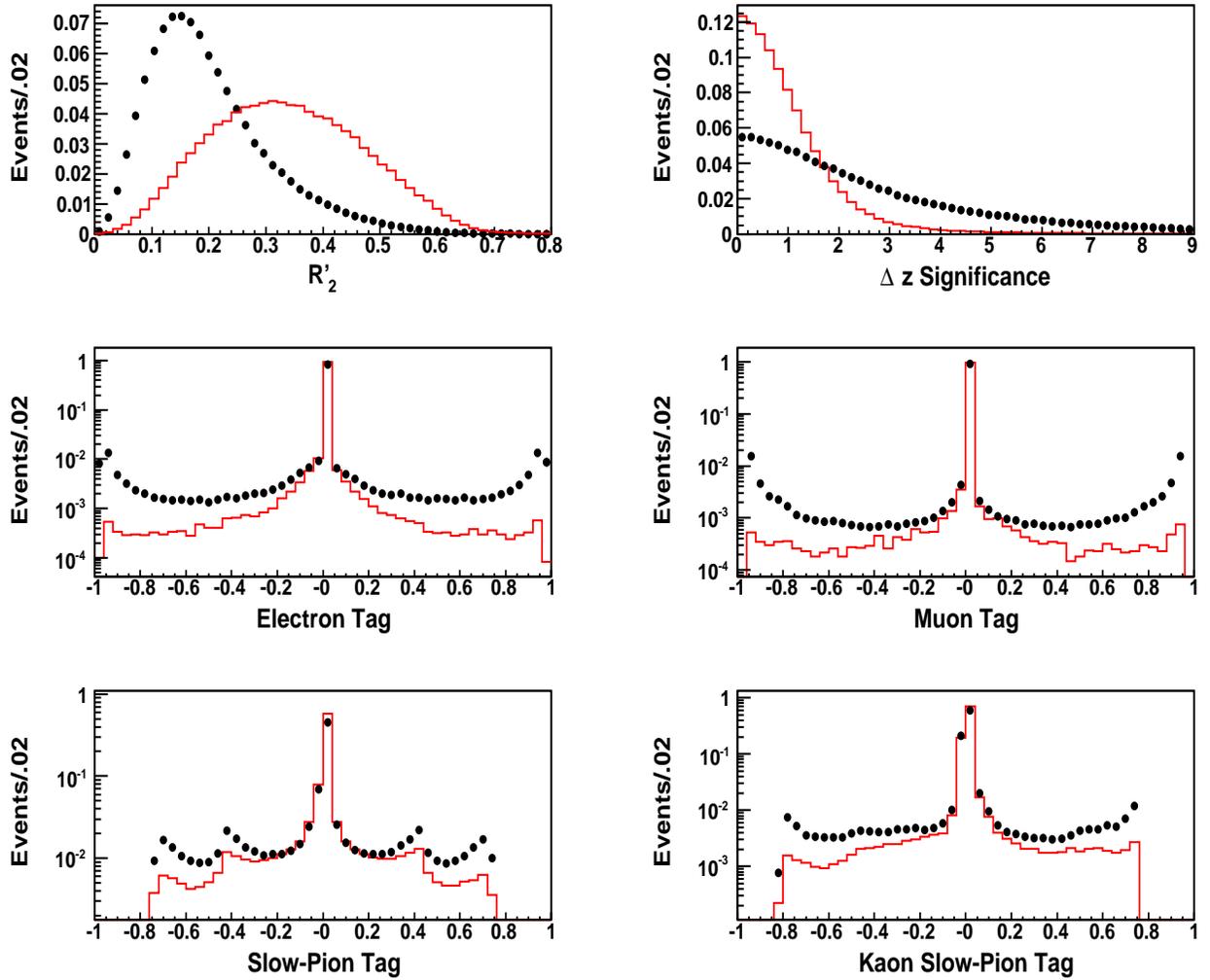


Figure 5.13:  $R'_2$ ,  $\Delta z$  significance, and subtagger output for signal (black dots) and continuum Monte Carlo (red line). The mode shown is  $B^0 \rightarrow K^{*0} \gamma (K^{*0} \rightarrow K^+ \pi^-)$  and all cuts are applied except for the Neural Network cut. The plots are normalized to the same area.

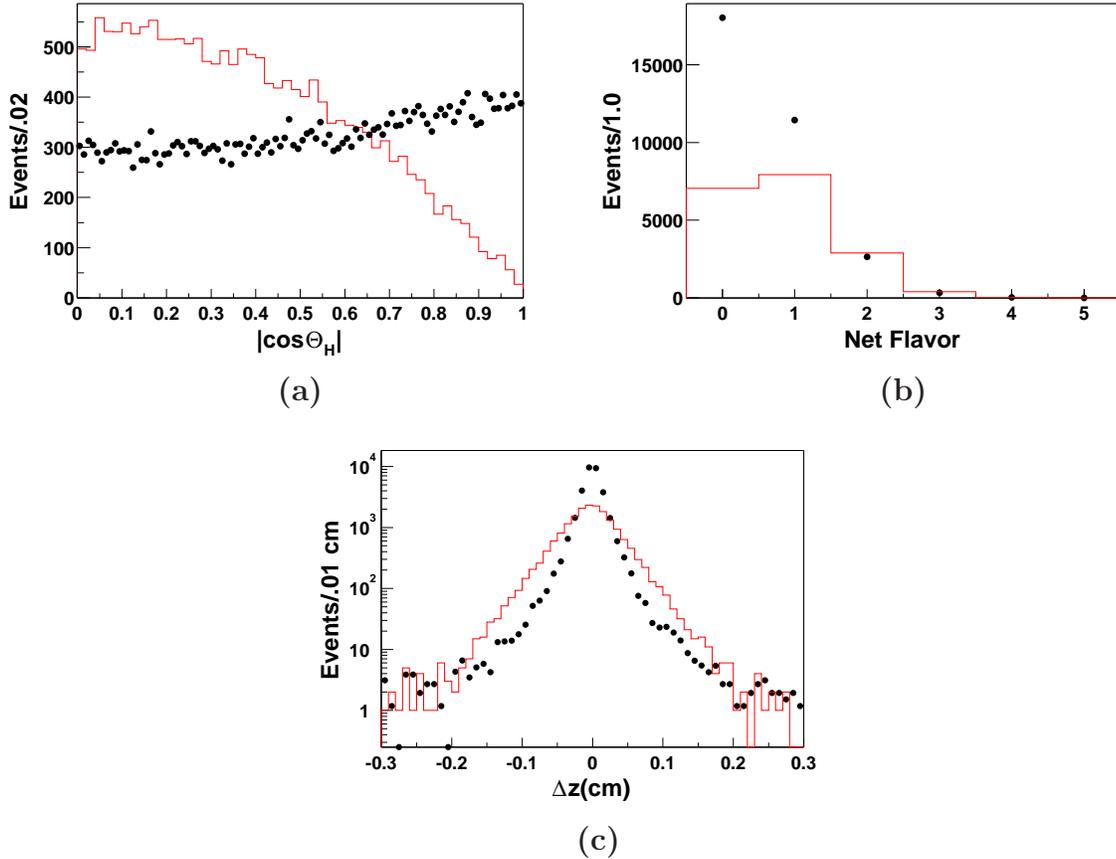


Figure 5.14: a)  $|\cos\theta_H|$ , b) Net Flavor, and c)  $\Delta z$  distributions for continuum background (black dots) and signal Monte Carlo (red line) for  $B \rightarrow \phi\gamma$ . All cuts except the neural network cut are applied.

into two parts, one is used to train the neural network to optimize it for signal/background separation, and the other is a validation sample, which is used to validate that the neural network behaves appropriately. The two samples are statistically independent of each other. For the background sample, we use  $uds$ ,  $c\bar{c}$ , and  $\tau^+\tau^-$  MC simulation, while the signal mode that pertains to the particular analysis is used for the signal sample. Details of the training of the neural network can be found in appendix C.

Figure 5.17 shows the neural network output of the validation sample for the  $B \rightarrow K^*\gamma$  analysis. The large separation of continuum background and signal events is evident. Figure 5.18 shows the background rejection vs. selection efficiency curves using the variables from the previous analysis [30] and for the  $B \rightarrow K^*\gamma$  analysis. It should be noted that the “old” curve is not taken directly from the previous analysis, but instead is created from the current Monte Carlo dataset using the variables from the previous analysis. For three

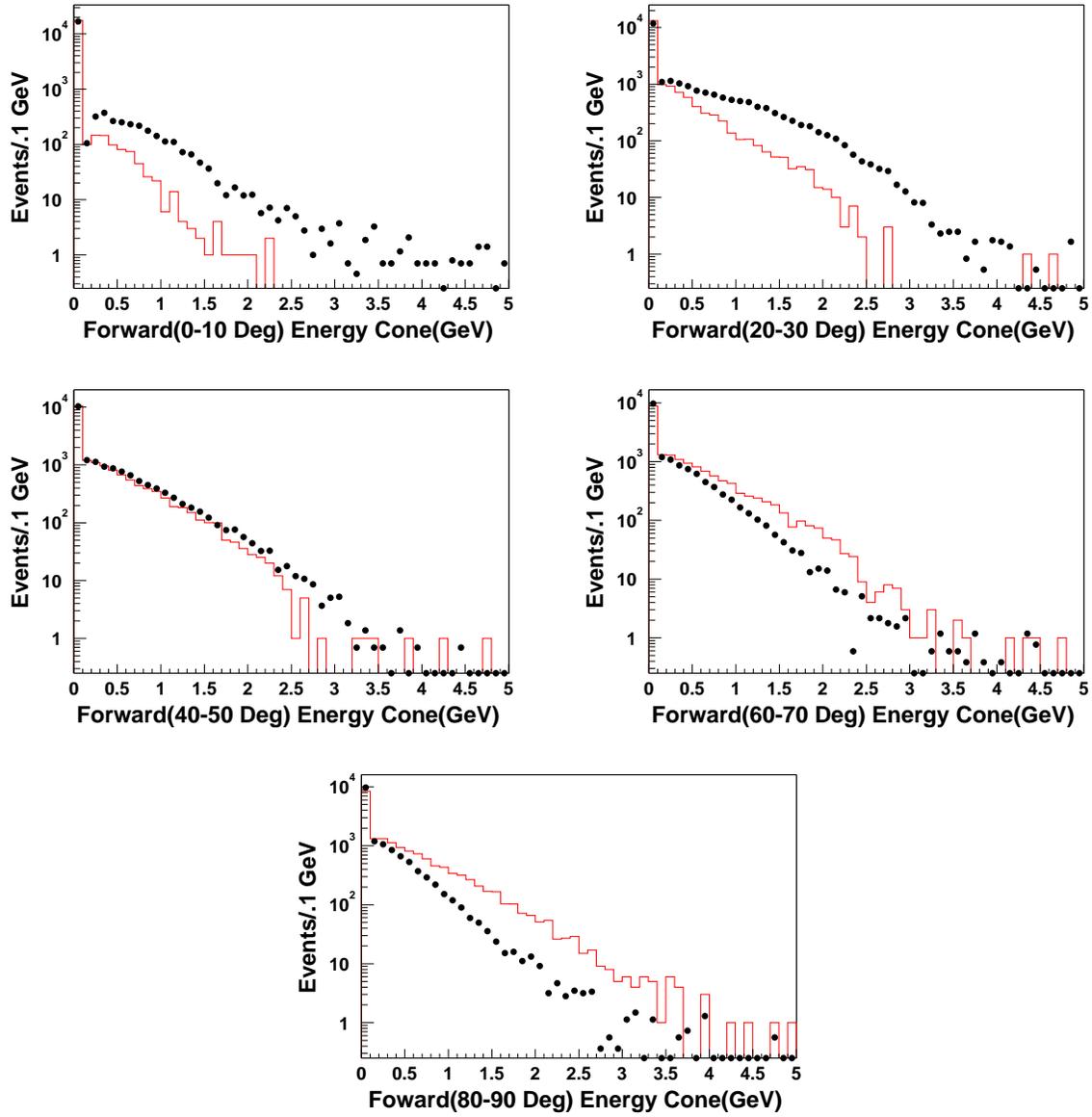


Figure 5.15: Every other forward (in the same hemisphere as the high energy photon) energy cone distribution for signal(red line) and continuum(black dots) Monte Carlo for  $B \rightarrow \phi\gamma$  are shown. All cuts except the neural network cut are applied.

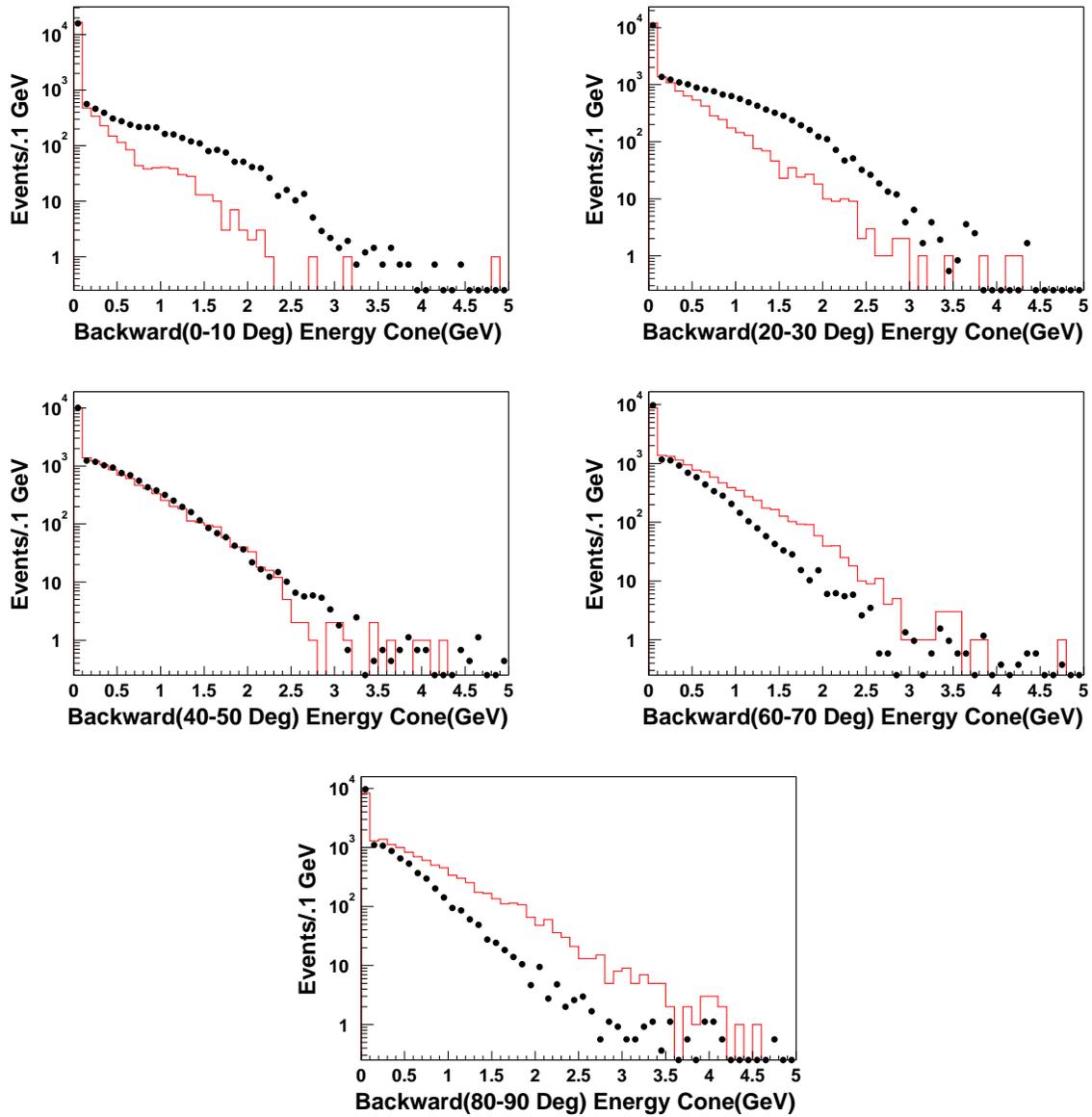


Figure 5.16: Every other backward (in the opposite hemisphere as the high energy photon) energy cone distribution for signal (red line) and continuum (black dots) Monte Carlo for  $B \rightarrow \phi\gamma$  are shown. All cuts except the neural network cut are applied.

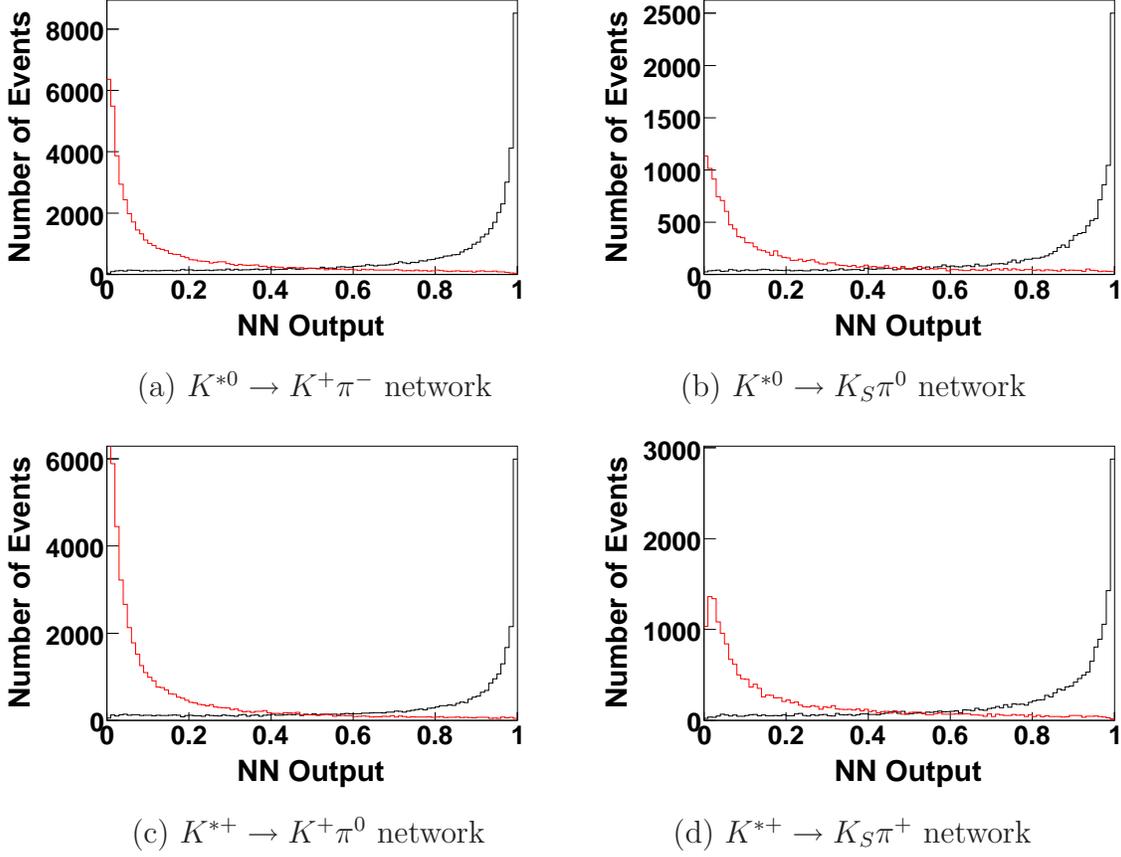


Figure 5.17: The neural network output of the validation sample for the background (red) and signal (black) for the a)  $K^{*0} \rightarrow K^+\pi^-$ , b)  $K^{*0} \rightarrow K_S\pi^0$ , c)  $K^{*+} \rightarrow K^+\pi^0$ , and d)  $K^{*+} \rightarrow K_S\pi^+$  networks.

of the neural networks, there is about a 40-50% improvement, while the  $K^{*0} \rightarrow K_S\pi^0$  mode remains the same. The neural network output is optimized.

## 5.5 Cut Optimization

Some of the cuts described in Section 5.3 have been explicitly described as being optimized, while others are held fixed. In this section, we describe how we deal with the optimization procedure. To optimize the cuts, we use “truth-matched” (described in Section 5.1.1.1) signal and continuum MC simulation samples.

Once all the fixed cuts have been applied to the MC simulation samples, we further constrain the events to lie in the region  $5.2 < m_{ES} < 5.3 \text{ GeV}/c^2$ ,  $-0.3 < \Delta E < 0.3 \text{ GeV}$  for the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analysis, and  $5.1 < m_{ES} < 5.29 \text{ GeV}/c^2$ ,  $-0.3 <$

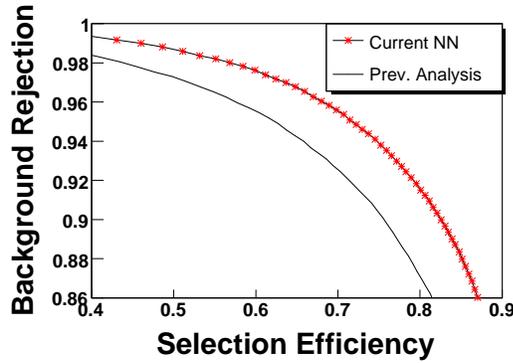
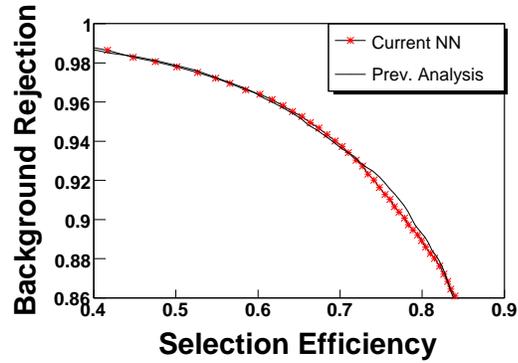
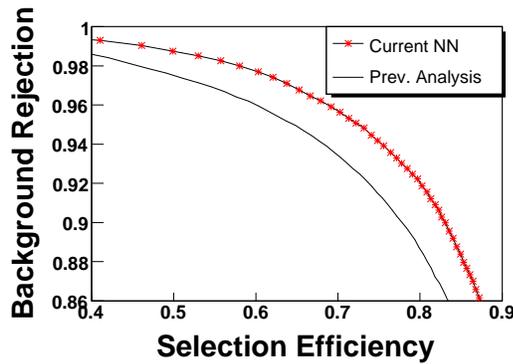
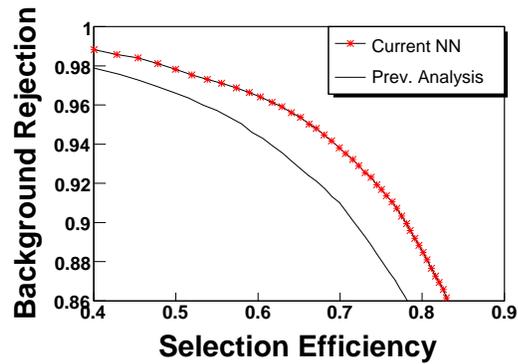
(a)  $K^{*0} \rightarrow K^+\pi^-$  network(b)  $K^{*0} \rightarrow K_S\pi^0$  network(c)  $K^{*+} \rightarrow K^+\pi^0$  network(d)  $K^{*+} \rightarrow K_S\pi^+$  network

Figure 5.18: The background rejection vs. selection efficiency for the a)  $K^{*0} \rightarrow K^+\pi^-$ , b)  $K^{*0} \rightarrow K_S\pi^0$ , c)  $K^{*+} \rightarrow K^+\pi^0$ , and d)  $K^{*+} \rightarrow K_S\pi^+$  networks for the validation sample. The solid line is the curve made using the variables from the previous analysis [30], while the starred curve uses the variables described in this section. All cuts except the neural net cut are applied.

$\Delta E < 0.3 \text{ GeV}$  for the  $B \rightarrow \phi\gamma$  analysis. These events are then fed as input to the StatPatternRecognition (SPR) package [54], which utilizes an algorithm [55] to find the best region (defined by an optimized figure of merit) that is composed of all the selection variables that are optimized. Here, the figure of merit is the statistical signal significance  $S/\sqrt{S+B}$ . The exact set of variables used by the algorithm depends on the mode considered. The procedure described below only applies to the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analyses. The  $B \rightarrow \phi\gamma$  analysis uses a different procedure, which will be described at the end of this section.

The algorithm searches the region in two stages. In the beginning, the first stage considers the entire parameter space and divides each dimension into two regions, in which the point of separation corresponding to each dimension is found by iterating through all possible separations maximizing the figure of merit. Once an optimal separation is found, the region with the maximal figure of merit is considered and each dimension is again divided into two. The same process is performed and iterated upon. The rate at which the regions are shrunk is controlled by the maximum number of events that can be removed at each step. This number of events is given by the user, and is optimized to give the maximum figure of merit. In the  $B \rightarrow K^*\gamma$  analysis, the optimal number of events (given as a percentage of the total) is found to be 30% for the  $K^{*0} \rightarrow K^+\pi^-$  mode, 70% for  $K^{*0} \rightarrow K_S\pi^0$  mode, 65% for the  $K^{*+} \rightarrow K^+\pi^0$  mode, and 20% for the  $K^{*+} \rightarrow K_S\pi^+$  mode. For the  $B \rightarrow (\rho/\omega)\gamma$  analysis, optimization gave 30% for the  $\rho^0 \rightarrow \pi^+\pi^-$  mode, and 40% for the  $\rho^+ \rightarrow \pi^+\pi^0$  and  $\omega \rightarrow \pi^+\pi^-\pi^0$  modes. The regions are gradually reduced in size until the figure of merit can no longer be improved, at which time the process is terminated. The second stage gradually increases the size of the region being considered to optimize the figure of merit. Finally, the selected region is discarded and the entire process begins again. This is because a more optimal region may be found in a region that had not previously been considered. In order to qualitatively assess the performance of the optimization procedure, plots of the relevant distributions in signal and continuum MC simulation samples are compared, and have been shown throughout this section.

To increase the statistics of our final optimization sample, we use all the candidates that pass fixed cuts and fall within the  $m_{\text{ES}}-\Delta E$  region described above. However, the quantity of interest is the signal significance contained within the signal region ( $5.27 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$ ,  $-0.2 < \Delta E < 0.1 \text{ GeV}$ ). Therefore, we rescale the distributions by the ratio of the number of events found in the signal region to the number in the larger  $m_{\text{ES}}-\Delta E$  region. All of the MC simulation samples are also independently scaled to  $347 \text{ fb}^{-1}$ .

We include  $\cos\theta_H$ , which is used in the fit, as one of the variables that is optimized in the procedure. This variable provides good signal/background separation, which is taken advantage of in the fitting procedure itself. Therefore, we choose to include this variable in the optimization to obtain the best possible cuts for the rest of the variables, while simulating the separation provided by  $\cos\theta_H$  in the fit. The limits for  $\cos\theta_H$  obtained are then relaxed for the candidates that are used in the fit. For the  $B \rightarrow (\rho/\omega)\gamma$  analysis, we use the neural network output as one of the dimensions of the fit. Therefore, for this analysis only, we relax the limit on this variable as well.

For the  $B \rightarrow \phi\gamma$  analysis, only two variables were optimized, the neural network output and the invariant mass of the  $\phi$ , by using an iterative procedure which minimizes the “sensitivity”/signal efficiency as a function of these two variables. The sensitivity is defined in Ref. [56] as the average upper limit of the number of signal events that would be obtained by an ensemble of experiments with the expected background and no true signal. In order to determine this number, the number of background events, which is a function of the neural network and phi mass selection, in the signal region must be determined. This is done as follows.

In bins of the neural net output and phi mass, we project the continuum background events onto the fit region. The  $m_{\text{ES}}$  projection is then fitted with an Argus shape, while the  $\Delta E$  distribution is fitted with a 1<sup>st</sup> order polynomial. Both distributions are then integrated in the signal region ( $5.27 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$ ,  $-0.2 < \Delta E < 0.1 \text{ GeV}$ ) and the fit region. The ratio of these integrals multiplied by the number of events in the fit region yields the number of background events in the signal region for that neural network cut. This number is then converted to the sensitivity, which, when divided by the signal efficiency, yields the expected upper limit. This is shown in Fig. 5.19 as a function of the neural network output (for a slice of phi mass).

### 5.5.1 Results of cut optimization

Tables 5.5 through 5.8 summarize the final results of cut optimization for all four modes from the  $B \rightarrow K^*\gamma$  analysis. The  $K^*$  mass cut was optimized separately for the  $K^{*0} \rightarrow K^+\pi^-$  mode because the results from SPR were not optimal. The results for the  $B \rightarrow (\rho/\omega)\gamma$  analysis are summarized in Tables 5.9 through 5.11. For the  $B \rightarrow \phi\gamma$  analysis, the results of the procedure described above are  $1.011 < m_{K^+K^-} < 1.029 \text{ GeV}/c^2$  and a neural network output  $> 0.7$ .

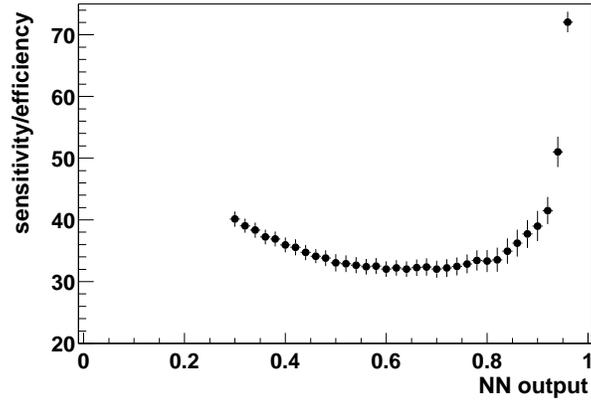


Figure 5.19: “Sensitivity”/signal efficiency as a function of neural net output for a slice of invariant phi mass.

Table 5.5: Summary of optimized selection criteria for  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  (mass given in  $\text{GeV}/c^2$ ).

Variable	Lower limit	Upper limit
Neural Network output	0.692	1
$m_{K\pi}$	0.781	1.086
$\cos\theta_H$	-0.813	0.893
$\eta$ veto	0	0.970
$\pi^0$ veto	0	0.874
$\gamma$ 2 <sup>nd</sup> moment	0	0.002

Table 5.6: Summary of optimized selection criteria for  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K_S\pi^0)$  (mass given in  $\text{GeV}/c^2$ ).

Variable	Lower limit	Upper limit
Neural Network output	0.767	1
$m_{K_S\pi^0}$	0.820	1.00
$\cos\theta_H$	-0.932	0.582
$\eta$ veto	0	0.974
$\pi^0$ veto	0	0.892
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$m_{\pi^0}$	0.112	0.151
$m_{K_S}$	0.486	0.521
$K_S$ flight significance	9.3	none

Table 5.7: Summary of optimized selection criteria for  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  (mass given in GeV/c<sup>2</sup>).

Variable	Lower limit	Upper limit
Neural Network output	0.824	1
$m_{K\pi^0}$	0.786	1.00
$\cos\theta_H$	-0.949	0.718
$\eta$ veto	0	0.987
$\pi^0$ veto	0	0.918
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$m_{\pi^0}$	0.114	0.151

Table 5.8: Summary of optimized selection criteria for  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K_S\pi^+)$  (mass given in GeV/c<sup>2</sup>).

Variable	Lower limit	Upper limit
Neural Network output	0.710	1
$m_{K_S\pi}$	0.792	1.03
$\cos\theta_H$	-0.846	0.90
$\eta$ veto	0	0.977
$\pi^0$ veto	0	0.906
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$m_{K_S}$	0.483	0.521
$K_S$ flight significance	10.27	none

Table 5.9: Summary of optimized selection criteria for  $B^0 \rightarrow \rho^0 \gamma$  (mass given in  $\text{GeV}/c^2$ )

Variable	Lower limit	Upper limit
Neural Network output	0.978	1
$m_{\pi\pi}$	0.633	0.957
$\text{Prob}(\chi^2_{\rho^0})$	0.008	1
$\cos \theta_{hel}$	-0.807	0.786
$\eta$ veto	0	0.918
$\pi^0$ veto	0	0.737
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$\gamma \frac{s_9}{s_{25}}$	0.938	1

Table 5.10: Summary of optimized selection criteria for  $B^+ \rightarrow \rho^+ \gamma$  (mass given in  $\text{GeV}/c^2$ )

Variable	Lower limit	Upper limit
Neural Network output	0.939	1
$m_{\pi\pi^0}$	0.636	0.932
$\cos \theta_{hel}$	-0.733	0.660
$\eta$ veto	0	0.932
$\pi^0$ veto	0	0.944
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$\gamma \frac{s_9}{s_{25}}$	0.944	0.996
$\cos \theta_{\gamma\gamma}$	0.789	0.997
$m_{\pi^0}$	0.117	0.148

Table 5.11: Summary of optimized selection criteria for  $B \rightarrow \omega \gamma$  (mass given in  $\text{GeV}/c^2$ ), distance in cm)

Variable	Lower limit	Upper limit
Neural Network output	0.776	1
$m_{\pi\pi\pi^0}$	0.764	0.795
$\text{Prob}(\chi^2_{\omega})$	0.000	1
$\cos \theta_{hel}$	-0.717	0.838
$\cos \theta_{Dalitz}$	-0.843	1
Flight distance	0	0.189
$\eta$ veto	0	0.939
$\pi^0$ veto	0	0.767
$\gamma$ 2 <sup>nd</sup> moment	0	0.002
$\gamma \frac{s_9}{s_{25}}$	0.945	0.990
$\cos \theta_{\gamma\gamma}$	0.413	1
$m_{\pi^0}$	0.122	1

# Chapter 6

## Measurement of Branching Fractions, $CP$ , and Isospin Asymmetries of $B \rightarrow K^* \gamma$

The previous chapter dealt with the event and candidate selection, and also described how the continuum background is suppressed with the application of the neural network. Briefly, it was mentioned in Section 5.1 that  $B$  decays also contribute to the background. In addition, the whole discussion was done in the context of the  $B \rightarrow V\gamma$  decay modes that are studied in this thesis. This chapter will focus on the  $B \rightarrow K^*\gamma$  analysis. It will detail the treatment of the  $B$  background (Section 6.1), and describe the fitting procedure to extract the branching fractions and the time-independent  $CP$  asymmetry, which is defined in Eq. 6.10.

### 6.1 B background

Approximately 30% of the background comes from  $B$  decays, which is categorized into three types. The dominant component, about 70% of the  $B$  background, comes from  $B \rightarrow X_s \gamma$  decays, where  $X_s$  is a state that has an invariant mass above the  $K^*$ , and can be either resonant or non-resonant. However, we model this component from inclusive  $B \rightarrow X_s \gamma$  MC simulation, where the only resonance is the  $K^*$ . We treat the effect of not including resonances as a systematic error, which will be described in Section 6.5.11.1. Figure 6.1 shows the hadronic mass spectrum from the inclusive MC simulation, where we have excluded the  $K^*$  resonance by placing a hadronic mass cut of  $> 1.1 \text{ GeV}/c^2$

on the generated mass. Effectively, the s-quark fragments into kaon and pions whose invariant mass is above  $1.1 \text{ GeV}/c^2$ . Most of the background that is misreconstructed comes from instances where there are three particles in the hadronic final state, and one of the particles has not been reconstructed. As shown in Figure 6.2, this component peaks in  $m_{\text{ES}}$  at the *B* meson mass, but, due to the missing particle, peaks at negative  $\Delta E$ . The  $\cos\theta_H$  distribution peaks at one. This is the result of a soft pion in the lab frame being used in the  $K^* \rightarrow K\pi$  decay. When this is done, the kaon has a high momentum in the lab frame, and the momentum vector nearly points in the same direction as  $K^*$ . As a consequence, in the  $K^*$  rest frame, the *B* and *K* momentum vectors nearly point in opposite directions.

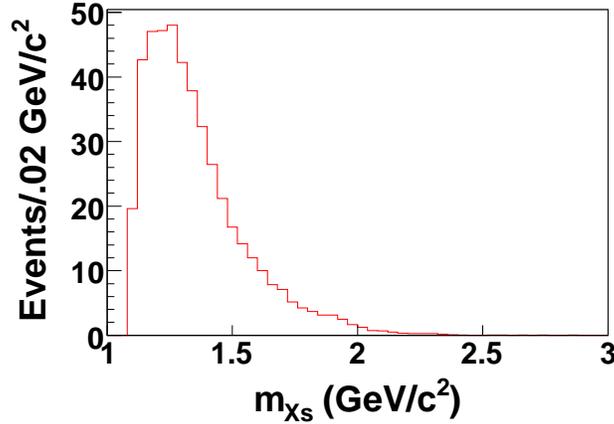


Figure 6.1: The  $B \rightarrow X_s^+ \gamma$  generated hadronic mass spectrum for the  $B^+ \rightarrow K^{*+} \gamma (K^{*+} \rightarrow K^+ \pi^0)$  mode. The mass selection of  $> 1.1 \text{ GeV}/c^2$  has been applied. In addition, all selection cuts have been applied.

The second component of the *B* background comes from  $B \rightarrow K^* \gamma$  decays other than the signal mode. For example, in the  $B^0 \rightarrow K^{*0} \gamma (K^{*0} \rightarrow K_S \pi^0)$  mode, about 15% of the *B* background comes from the  $B^+ \rightarrow K^{*+} \gamma (K^{*+} \rightarrow K_S \pi^+)$  mode, where the  $\pi^0$  is taken from the other side of the event. In general (as an average for all four  $B \rightarrow K^* \gamma$  decays), this component comprises approximately 10% of the *B* background. It is modeled by a  $B \rightarrow K^* \gamma$  MC simulation sample in which the  $K^*$  decays to any allowed mode, but the signal mode has been excluded from the sample. As displayed in Figure 6.3, the component peaks in  $m_{\text{ES}}$  at the *B* meson mass, but has a  $\Delta E$  distribution that has a negative peak because the other side of the event typically contains softer particles.

The last component ( $\approx 5\%$  of the *B* background) is made of the remaining *B* decays. It is comprised mostly of  $B \rightarrow K^* \eta$  and  $B \rightarrow K^* \pi^0$ , where one of the photons coming

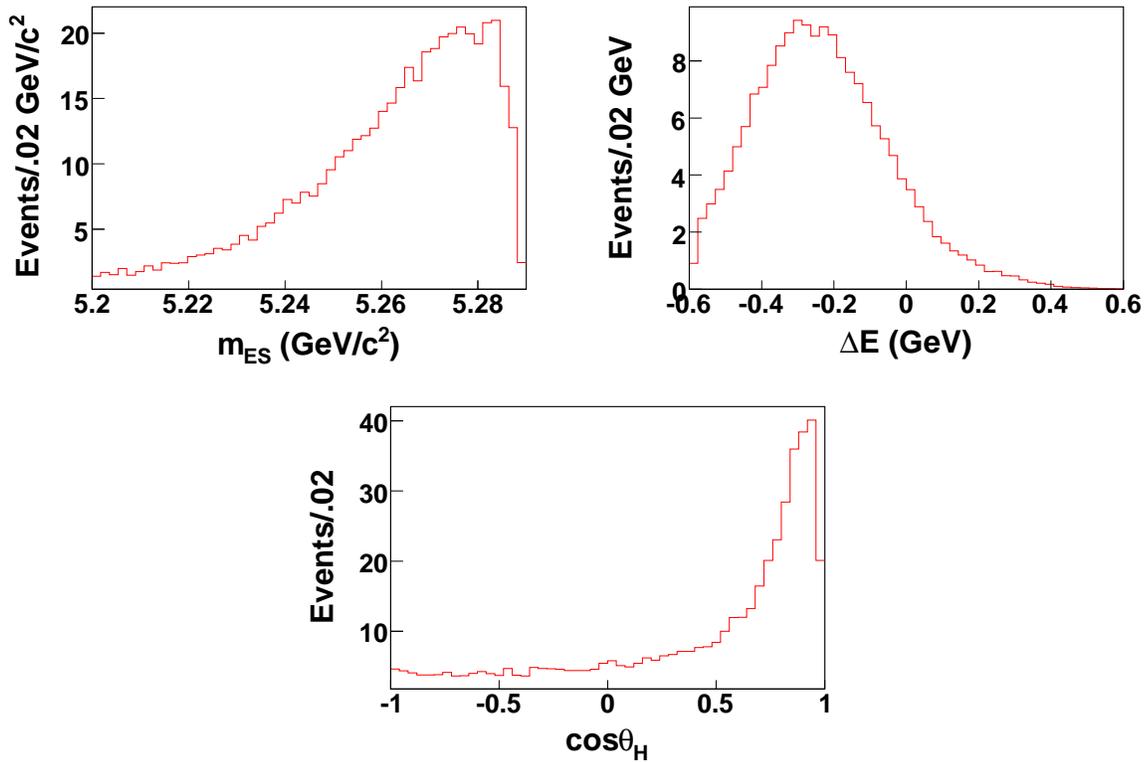


Figure 6.2:  $m_{ES}$ ,  $\Delta E$ , and  $\cos \theta_H$  distributions of the inclusive  $B \rightarrow X_s \gamma$  MC simulation for the  $B^+ \rightarrow K^{*+} \gamma (K^{*+} \rightarrow K^+ \pi^0)$  mode. All selection cuts have been applied, except for the  $\Delta E$  and  $\cos \theta_H$  distributions, in which case the respective selection has not been applied.

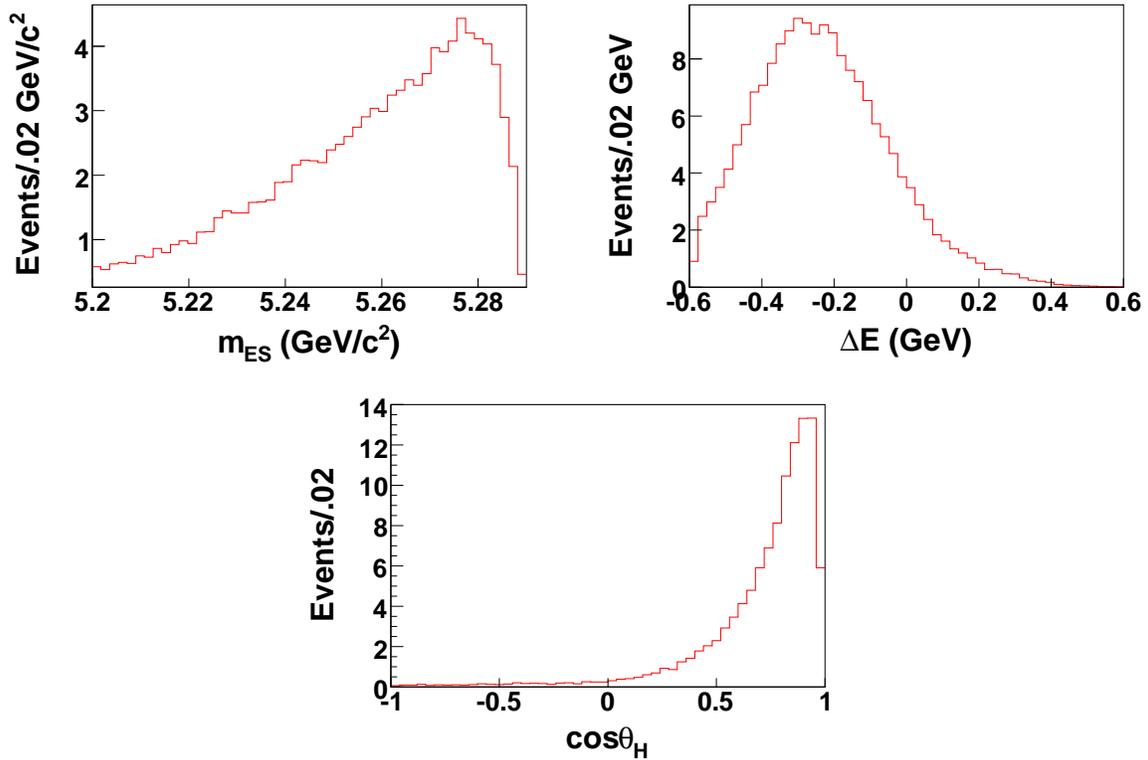


Figure 6.3:  $m_{ES}$ ,  $\Delta E$ , and  $\cos\theta_H$  distributions of the  $B \rightarrow K^*\gamma$  MC simulation for the  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  mode, in which the  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  mode has been removed. All selection cuts have been applied, except for the  $\Delta E$  and  $\cos\theta_H$  distributions, in which case the respective selection has not been applied.

from the  $\eta, \pi^0$  decay has been missed in the reconstruction. The MC simulation sample that describes this component is made of  $B$  decays in which the  $B$  is allowed to decay generically, and the  $B \rightarrow K^*\gamma$  and  $B \rightarrow X_s\gamma$  decays have been removed. The distributions are shown in figure 6.4. The  $\cos\theta_H$  distribution has a  $\cos^2\theta_H$ -like dependence because  $d\Gamma/d\cos\theta_H \sim \cos^2\theta_H$  for  $B \rightarrow K^*P$ , where  $P = \eta, \pi^0$ .

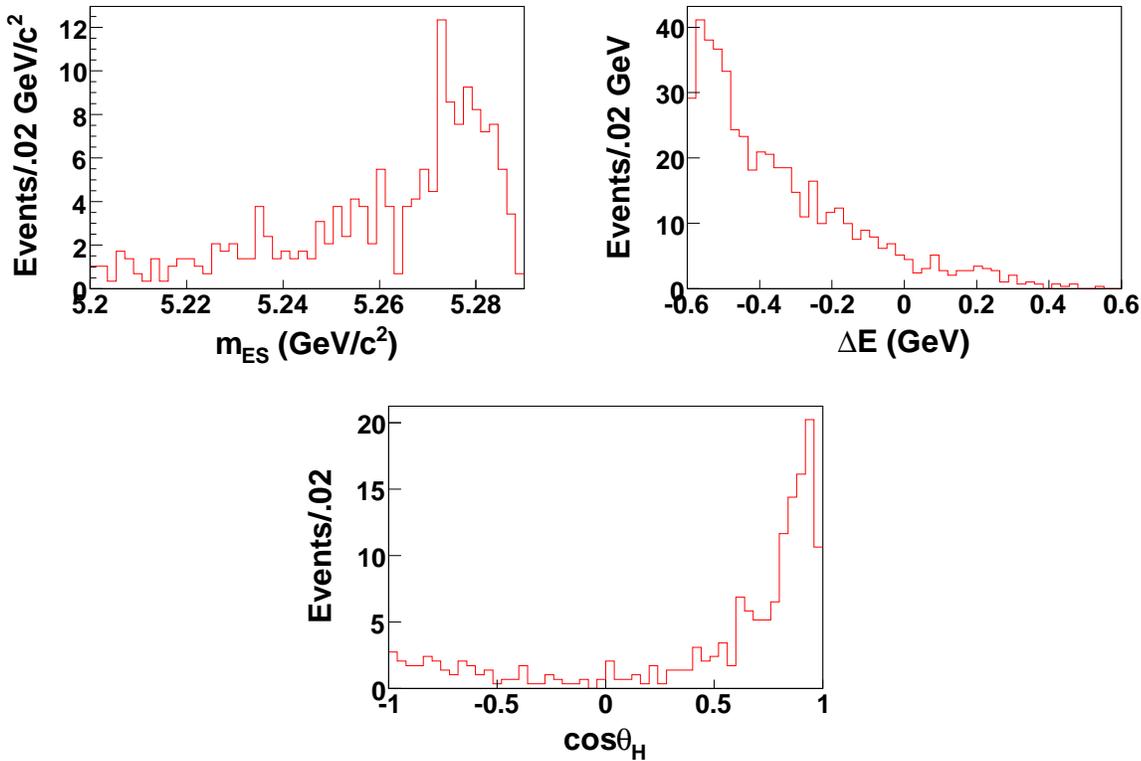


Figure 6.4:  $m_{ES}$ ,  $\Delta E$ , and  $\cos\theta_H$  distributions of the  $B$  background, not including  $B \rightarrow K^*\gamma$  and  $B \rightarrow X_s\gamma$ , for the  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  mode. All selection cuts have been applied, except for the  $\Delta E$  and  $\cos\theta_H$  distributions, in which case the respective selection has not been applied.

## 6.2 Signal and Background Efficiencies

After applying all of the selection criteria to each of the MC simulation samples, we further restrict the candidates to lie in the region  $5.22 < m_{ES} < 5.29 \text{ GeV}/c^2, |\Delta E| < 0.3 \text{ GeV}, |\cos\theta_H| < 0.75$ . This region is called the fit region, and will be used as the domain in which to perform the likelihood fit. In each event, this region can also have multiple  $B$  candidates. There are approximately 1.1 candidates per event in signal Monte

Carlo. We use only the  $B$  candidate that has the  $K^*$  candidate whose mass is closest to the nominal mass [48]. Table 6.1 shows the expected number of events for the signal, continuum, and  $B\bar{B}$  categories in  $347 \text{ fb}^{-1}$ . The signal efficiencies, as determined using MC simulation signal events, are  $22.9 \pm 0.05\%$ ,  $14.5 \pm 0.05\%$ ,  $16.7 \pm 0.07\%$ , and  $21.5 \pm 0.08\%$  in the fit region for the  $K^{*0} \rightarrow K^+\pi^-$ ,  $K^{*0} \rightarrow K_S\pi^0$ ,  $K^{*+} \rightarrow K^+\pi^0$ , and  $K^{*+} \rightarrow K_S\pi^+$  modes respectively.

Table 6.1: Number of Monte Carlo events in the fit region scaled to  $347 \text{ fb}^{-1}$  using the MC equivalent luminosities. The  $B$  background is the sum of the three components described in section 6.1. The off-peak data is shown for comparison.

	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$	$K^{*+} \rightarrow K^+\pi^0$	$K^{*+} \rightarrow K_S\pi^+$
Signal	$2327 \pm 5$	$254 \pm 1$	$848 \pm 4$	$758 \pm 3$
Continuum + $\tau^+\tau^-$	$2893 \pm 45$	$505 \pm 20$	$1493 \pm 32$	$151 \pm 33$
$B\bar{B}$	$609 \pm 6$	$108 \pm 2$	$379 \pm 5$	$287 \pm 4$
Off-Peak Data	$3313 \pm 202$	$688 \pm 88$	$1600 \pm 14$	$1525 \pm 14$

## 6.3 Fit Parameterization

We extract the signal yield by using a three dimensional maximum likelihood fit, in which we use  $m_{\text{ES}}$ ,  $\Delta E$ , and  $\cos\theta_H$  as the dimensions. We also consider three components: signal, continuum, and  $B\bar{B}$ . All of the  $B$  background components discussed in Section 6.1 are considered as a sum in the  $B\bar{B}$  component. Only the candidates that lie in the fit region and pass the multiple candidate selection are considered, as discussed in Section 6.2.

### 6.3.1 PDF Line Shapes

The first step toward parameterizing the fit model is determining the PDF line shapes by performing one dimensional fits to the Monte Carlo sample that represents each component. The signal  $m_{\text{ES}}$  distribution for the  $K^{*0} \rightarrow K_S\pi^0$  and  $K^{*+} \rightarrow K^+\pi^0$  modes is described by a Crystal Ball function [57]. This is defined as

$$f_{CB}(x) = C_{CB} \cdot \begin{cases} e^{-\frac{(x-\langle x \rangle)^2}{2\sigma^2}} & \text{for } x > \langle x \rangle - \alpha\sigma \\ \frac{\left(\frac{n}{\alpha}\right)^n \cdot e^{-\frac{\alpha^2}{2}}}{\left(\frac{\langle x \rangle - x}{\sigma} + \frac{n}{\alpha} - \alpha\right)^n} & \text{for } x \leq \langle x \rangle - \alpha\sigma \end{cases} \quad (6.1)$$

where  $\langle x \rangle$  is the mean of the core gaussian,  $\sigma$  is the width, and  $\alpha$  and  $n$  are two tail parameters. For the  $K^{*0} \rightarrow K^+ \pi^-$  and  $K^{*+} \rightarrow K_S \pi^+$  modes, the signal  $m_{\text{ES}}$  distribution is parameterized as a Cruiff function. The Cruiff function is defined as a piece-wise function

$$f_{Cruiff}(x) = C_{Cruiff} \cdot \exp\left(-\frac{-(x-\mu)^2}{\sigma_{L,R}^2 + \alpha_{L,R}(x-\mu)^2}\right) \quad (6.2)$$

defined to the left (L) and right (R) of  $\mu$ , which is the peak position of the distribution. Here,  $\sigma_{L,R}$  and  $\alpha_{L,R}$  are the widths and measures of the tails, respectively, to the left and right of the peak. We constrain  $\sigma_L = \sigma_R$  in the Cruiff function, because the signal distribution is approximately Gaussian in the core. The motivation to define two different parameterizations (one for the  $K^{*0} \rightarrow K_S \pi^0$  and  $K^{*+} \rightarrow K^+ \pi^0$  modes, and the other for the  $K^{*0} \rightarrow K^+ \pi^-$  and  $K^{*+} \rightarrow K_S \pi^+$  modes) of the signal  $m_{\text{ES}}$  distribution is related to the evaluation of the systematic errors, which is described in appendix D. The Cruiff function also describes the signal  $\Delta E$  distribution for each mode, but with different values for the parameters. The  $\cos \theta_H$  distribution for the signal component is modeled by a 2<sup>nd</sup> order polynomial.

For the continuum hypothesis, the  $m_{\text{ES}}$  PDF is parameterized by an ARGUS function [58]. The functional form is

$$f_{Argus}(x) = C_{Argus} \cdot \frac{x}{E_{BEAM}} \cdot \sqrt{1 - \frac{x^2}{E_{BEAM}^2}} \cdot e^{-\xi \left(1 - \frac{x^2}{E_{BEAM}^2}\right)} \quad (6.3)$$

where  $\xi$  as the Argus parameter and  $E_{BEAM}$  is the argus endpoint. Here,  $E_{BEAM} = \sqrt{s}/2$ , where  $\sqrt{s}$  is the center of mass energy. The continuum  $\Delta E$  and  $\cos \theta_H$  shapes are modeled by a first- or second-order polynomial.

Various functional forms are used to describe the  $B\bar{B}$  background. In particular, the  $m_{\text{ES}}$  distribution is parameterized by a Novosibirsk function, which is defined as

$$f_{Novosibirsk}(x) = C_{Novosibirsk} \cdot \exp\left(-0.5 \frac{\ln^2[1 + \Lambda\tau \cdot (x - \mu)]}{\tau^2} + \tau^2\right) \quad (6.4)$$

where  $\Lambda = \sinh(\tau\sqrt{\ln 4})/(\sigma\tau\sqrt{\ln 4})$ , the peak position is  $\mu$ , the width is  $\sigma$ , and  $\tau$  is the tail parameter. The  $\Delta E$  distribution is described by a Crystal Ball function, while the  $\cos\theta_H$  distribution is modeled by a 4<sup>th</sup> order polynomial.

The full fit range is utilized in all nine fits for each separate mode, except for the signal  $m_{\text{ES}}$  fit. For the well measured  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes, the  $m_{\text{ES}}$  fit range is narrowed to  $5.27 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$ , while for the  $K^{*0} \rightarrow K_S\pi^0$  and  $K^{*+} \rightarrow K^+\pi^0$  modes, its range is set to  $5.26 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$ , due to the  $\pi^0$  tails. This restriction occurs because the vast majority of the signal distribution is in the narrower range. The one dimensional fits for each fit observable and component are shown in Figures 6.5 through 6.8. The parameters of the distributions are displayed on the plots. For the  $K^{*0} \rightarrow K^+\pi^-$ ,  $K^{*0} \rightarrow K_S\pi^0$ , and  $K^{*+} \rightarrow K^+\pi^0$  modes, the helicity angle is determined by following the kaon daughter of the  $K^*$ , while for the  $K^{*+} \rightarrow K_S\pi^+$  mode it is determined by the following the pion daughter.

### 6.3.2 Correlation of Fit Observables

We also parameterize the degree of correlation between the fit observables. The correlation coefficients between each dimension for each component are listed in Tables 6.2 through 6.5, and the profile plots for the  $K^{*0} \rightarrow K^+\pi^-$  mode are given in appendix B. As shown for the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes,  $m_{\text{ES}}$  and  $\Delta E$  are uncorrelated in signal MC simulation. However, for the two  $\pi^0$  modes for the signal MC simulation, there exists a slight correlation between  $m_{\text{ES}}$  and  $\Delta E$ . This is because there is energy loss in the calorimeter, which creates tails on the distributions on both  $m_{\text{ES}}$  and  $\Delta E$  in the negative direction.

Table 6.2: a) Signal MC simulation, b) Continuum MC simulation, and c)  $B$  background MC simulation correlation coefficients for the  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow K^+\pi^-)$  mode.

	$m_{\text{ES}}$	$\Delta E$
$\Delta E$	0.0183	
$\cos\theta_H$	-0.0657	-0.0015

a) Signal MC simulation

	$m_{\text{ES}}$	$\Delta E$
$\Delta E$	-0.0266	
$\cos\theta_H$	-0.0104	-0.0211

b) Continuum MC simulation

	$m_{\text{ES}}$	$\Delta E$
$\Delta E$	-0.1014	
$\cos\theta_H$	-0.1305	0.2776

c)  $B$  background MC simulation

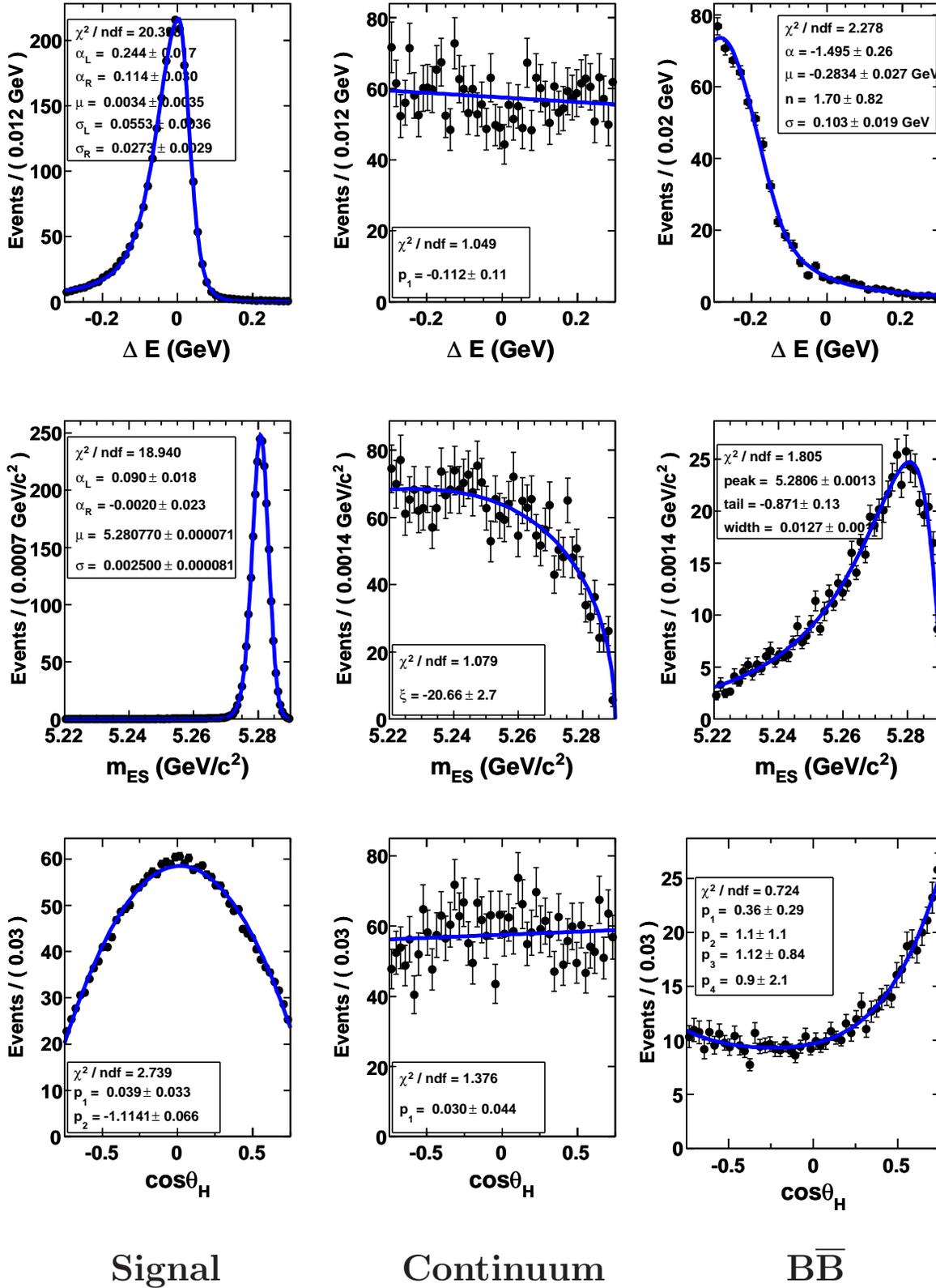


Figure 6.5: 1 dimensional fits to the  $K^{*0} \rightarrow K^+ \pi^-$  mode. The dimensions are  $\Delta E$  (top row),  $m_{ES}$  (middle row), and  $\cos\theta_H$  (bottom row). The components are signal (1<sup>st</sup> column), continuum (2<sup>nd</sup> column), and  $B\bar{B}$  background (3<sup>rd</sup> column). The helicity angle is determined by following the kaon daughter of the  $K^*$ .



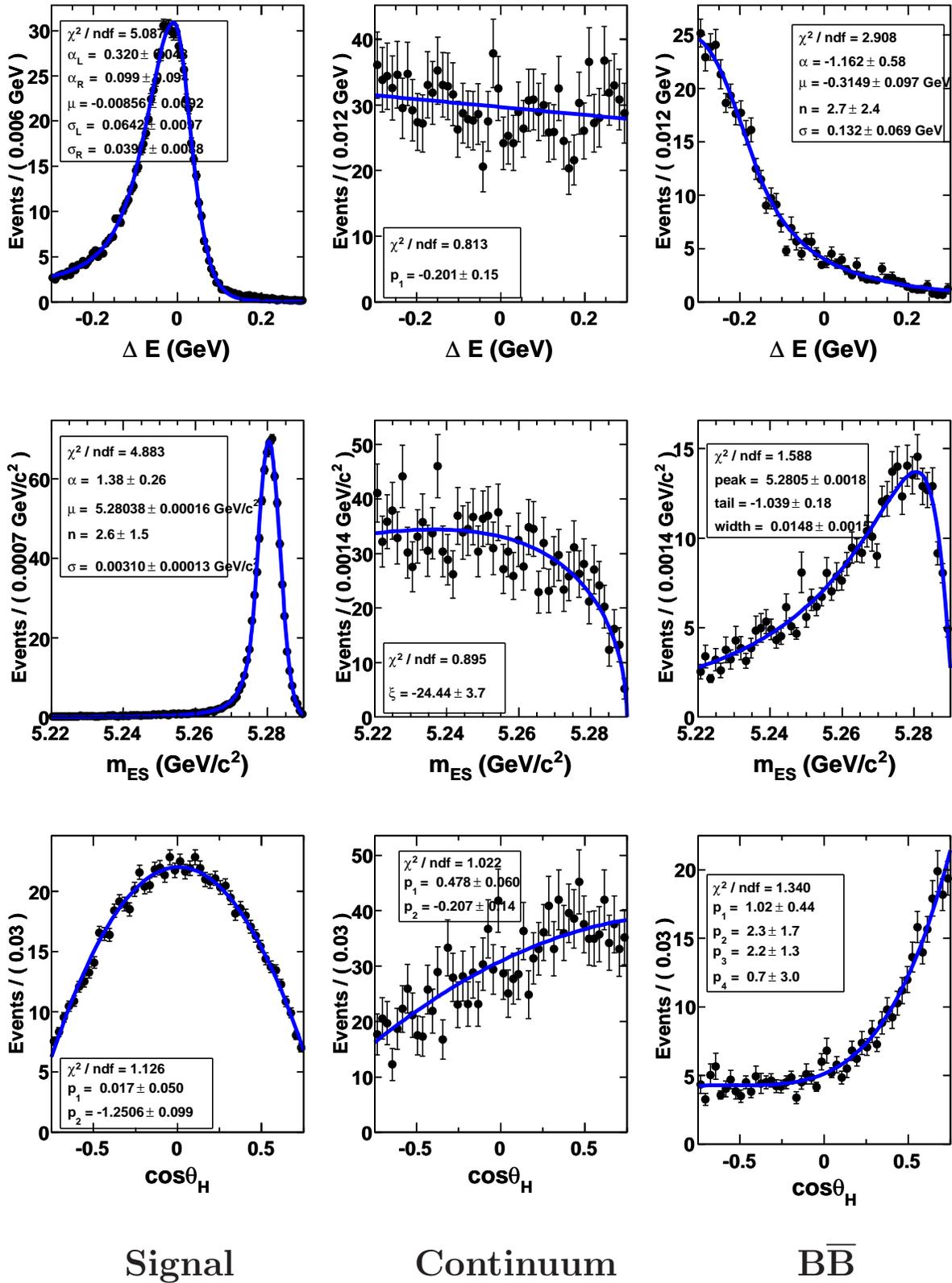


Figure 6.7: 1 dimensional fits to the  $K^{*+} \rightarrow K^+ \pi^0$  mode. The dimensions are  $\Delta E$  (top row),  $m_{ES}$  (middle row), and  $\cos\theta_H$  (bottom row). The components are signal (1<sup>st</sup> column), continuum (2<sup>nd</sup> column), and  $B$  background (3<sup>rd</sup> column). The helicity angle is determined by following the kaon.

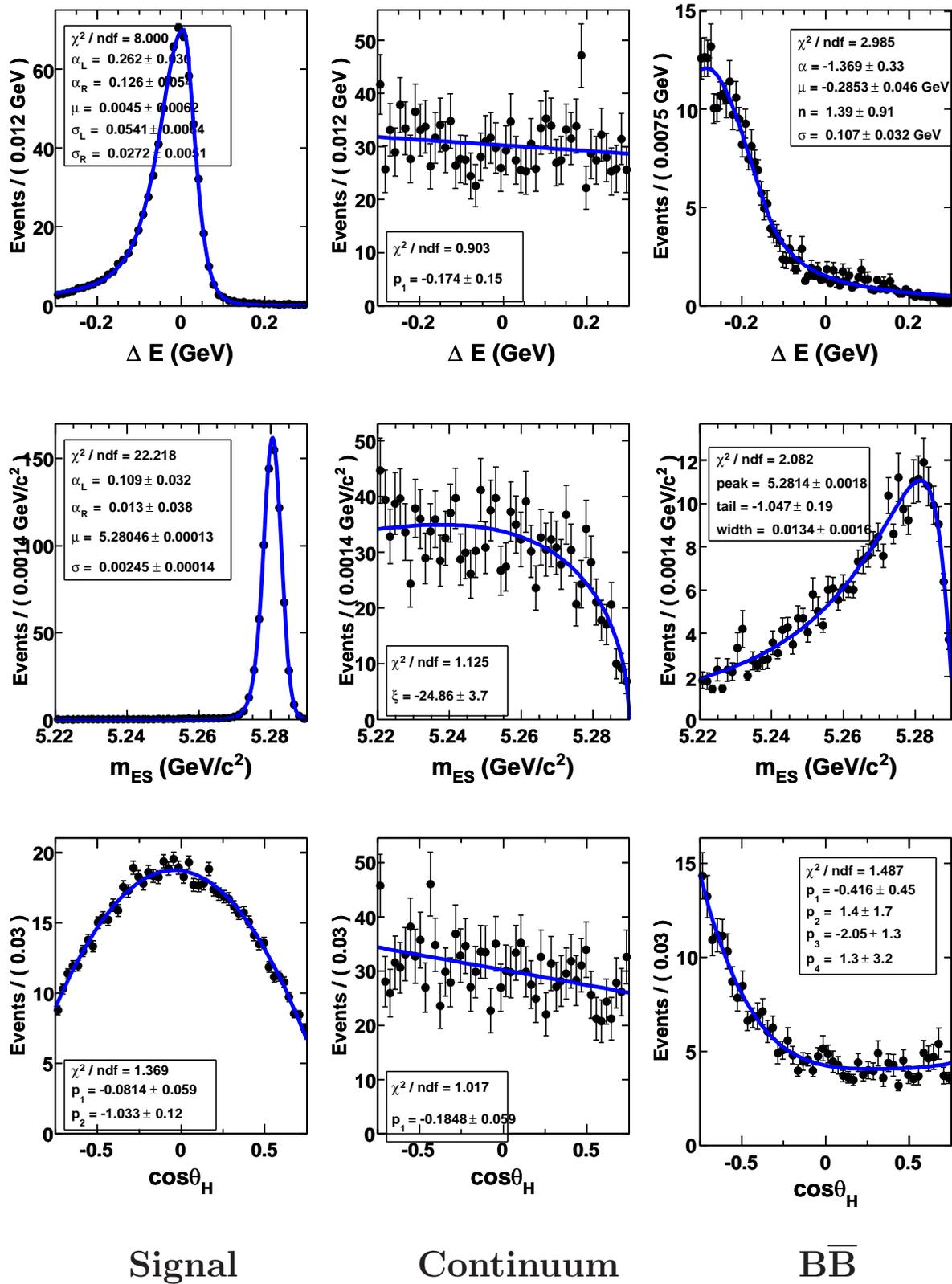


Figure 6.8: 1 dimensional fits to the  $K^{*+} \rightarrow K_S\pi^+$  mode. The dimensions are  $\Delta E$  (top row),  $m_{ES}$  (middle row), and  $\cos\theta_H$  (bottom row). The components are signal (1<sup>st</sup> column), continuum (2<sup>nd</sup> column), and  $B$  background (3<sup>rd</sup> column). The helicity angle is determined by following the pion.

Table 6.3: a) Signal, b) Continuum, and c)  $B$  background correlation coefficients for the  $B^0 \rightarrow K^{*0} \gamma (K^{*0} \rightarrow K_S \pi^0)$  mode.

	$m_{ES}$	$\Delta E$
$\Delta E$	0.1382	
$\cos \theta_H$	-0.1129	-0.0057

a) Signal MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	-0.0149	
$\cos \theta_H$	-0.0186	-0.0645

b) Continuum MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	0.0200	
$\cos \theta_H$	-0.1315	0.1607

c)  $B$  background MC simulationTable 6.4: a) Signal, b) Continuum, and c)  $B$  background correlation coefficients for the  $B^+ \rightarrow K^{*+} \gamma (K^{*+} \rightarrow K^+ \pi^0)$  mode.

	$m_{ES}$	$\Delta E$
$\Delta E$	0.1419	
$\cos \theta_H$	-0.0945	-0.0047

a) Signal MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	-0.0506	
$\cos \theta_H$	0.0072	0.0150

b) Continuum MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	-0.0076	
$\cos \theta_H$	-0.1046	0.2452

c)  $B$  background MC simulationTable 6.5: a) Signal, b) Continuum, and c)  $B$  background correlation coefficients for the  $B^+ \rightarrow K^{*+} \gamma (K^{*+} \rightarrow K_S \pi^+)$  mode.

	$m_{ES}$	$\Delta E$
$\Delta E$	0.0403	
$\cos \theta_H$	0.0974	-0.0036

a) Signal MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	-0.0070	
$\cos \theta_H$	-0.0272	0.0205

b) Continuum MC simulation

	$m_{ES}$	$\Delta E$
$\Delta E$	-0.0525	
$\cos \theta_H$	0.1392	-0.1494

c)  $B$  background MC simulation

### 6.3.3 Maximum Likelihood Fit

The largest linear correlation coefficients are in the  $B$  background sample, in which the largest is 0.27. We ignore the correlations in the maximum likelihood fit by using PDFs that are products of one dimensional PDFs. We construct a separate fit for each mode, and study the effects of correlations using embedded toys. The unbinned extended likelihood function to extract the signal yield for the branching fraction measurements is defined as

$$\mathcal{L} = \exp\left(-\sum_{i=1}^M n_i\right) \cdot \prod_{j=1}^N \left[ \sum_{i=1}^M n_i \mathcal{P}(\vec{x}_j; \vec{\alpha}_i) \right] \quad (6.5)$$

where  $N$  is the number of events,  $M = 3$  is the number of hypotheses (signal, continuum, and  $B\bar{B}$ ), and  $n_i$  represents the yield of a particular hypothesis.  $\mathcal{P}_i$  is the product of one-dimensional PDFs over the three dimensions  $\vec{x}$  ( $m_{ES}$ ,  $\Delta E$ ,  $\cos\theta_H$ ), and  $\vec{\alpha}$  represents the fit parameters. The inclusion of the  $\cos\theta_H$  observable suppresses the  $B\bar{B}$  background. As shown in Figs. 6.2, 6.3, and 6.4, the  $\cos\theta_H$  distribution is not consistent with  $\sin^2\theta$ .

The parameterization for  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes are shown in Table 6.6, while the  $K^{*0} \rightarrow K_S\pi^0$  and  $K^{*+} \rightarrow K^+\pi^0$  modes are shown in Table 6.7. When the parameter is labeled as floating, it is determined by performing a fit to the full three-dimensional MC simulation sample, while the fixed designation means that it is determined by the one-dimensional fits, and fixed in the full fit. For the branching fraction fits, the yields are also floated.

#### 6.3.3.1 Fit for the $CP$ asymmetry

For the  $\mathcal{A}_{CP}$  fit, we divide the dataset, excluding the  $K^{*0} \rightarrow K_S\pi^0$  mode, into two subsamples, which correspond to  $B \rightarrow \bar{K}^*\gamma$  and  $\bar{B} \rightarrow K^*\gamma$  decays. Each of the decay modes can be identified as a  $K^*$  or a  $\bar{K}^*$  by the charge of the final states. We then perform a simultaneous fit to the two subsamples, in which all shape parameters are assumed to be flavor independent, and therefore shared. Each mode has a separate fit. The fraction of the yield of each component is extracted from the fit, and converted to  $\mathcal{A}_{CP}$  using the equation

$$\mathcal{A}_{CP} = \frac{n_{\bar{B}} - n_B}{n_{\bar{B}} + n_B} = 1 - 2f_B, \quad (6.6)$$

where  $n_{\bar{B}}$  is number of  $\bar{B}$ 's in the sample,  $n_B$  is the number of  $B$ 's, and  $f_B$  is the fraction

Table 6.6: PDF parameters for the  $K^{*0} \rightarrow K^+ \pi^-$  and  $K^{*+} \rightarrow K_S \pi^+$  modes.

Data Set	PDF	Parameter	float/fix
Signal MC	Cruijff( $m_{\text{ES}}$ )	$\langle m_{\text{ES}} \rangle$ (GeV/ $c^2$ )	float
		$\sigma_{Lm_{\text{ES}}}$ (GeV/ $c^2$ )	float
		$\sigma_{Rm_{\text{ES}}}$ (GeV/ $c^2$ )	fix to $\sigma_{Lm_{\text{ES}}}$
		$\alpha_{Lm_{\text{ES}}}$	fix
		$\alpha_{Rm_{\text{ES}}}$	fix
	Cruijff( $\Delta E$ )	$\langle \Delta E \rangle$ (GeV)	float
		$\sigma_{L\Delta E}$ (GeV)	float
		$\sigma_{R\Delta E}$ (GeV)	float
		$\alpha_{L\Delta E}$	fix
		$\alpha_{R\Delta E}$	fix
	$2^{\text{nd}}$ -order polynomial( $\cos(\theta_H)$ )		float
Continuum MC	Argus( $m_{\text{ES}}$ )	$\xi_{m_{\text{ES}}}$	float
		$E_{\text{BEAM}}$ (GeV)	fix
	$1^{\text{st}}$ -order polynomial( $\Delta E$ )	$P_1$	float
	$1^{\text{st}}$ -order polynomial( $\cos(\theta_H)$ )	$P_1$	float
$BB$ background MC	Novosibirsk( $m_{\text{ES}}$ )	$\langle m_{\text{ES}} \rangle$ (GeV/ $c^2$ )	fix
		$\sigma_{m_{\text{ES}}}$ (GeV/ $c^2$ )	fix
		$\tau_{Bbkg}$	fix
	Crystal Ball( $\Delta E$ )	$\langle \Delta E \rangle$ (GeV)	fix
		$\sigma_{\Delta E}$ (GeV)	fix
		$\alpha_{\Delta E}$	fix
		$n_{\Delta E}$	fix
		$4^{\text{th}}$ -order polynomial( $\cos(\theta_H)$ )	$P_1 - P_4$

Table 6.7: PDF parameters for all  $K^{*0} \rightarrow K_S\pi^0$  and  $K^{*+} \rightarrow K^+\pi^0$  modes.

Data Set	PDF	Parameter	float/fix	
Signal MC	Crystal Ball( $m_{\text{ES}}$ )	$\langle m_{\text{ES}} \rangle$ (GeV/c <sup>2</sup> )	float	
		$\sigma_{m_{\text{ES}}}$ (GeV/c <sup>2</sup> )	float	
		$\alpha_{m_{\text{ES}}}$	fix	
		$n_{m_{\text{ES}}}$	fix	
	Cruiff( $\Delta E$ )	$\langle \Delta E \rangle$ (GeV)	float	
		$\sigma_{L\Delta E}$ (GeV)	float	
		$\sigma_{R\Delta E}$ (GeV)	float	
		$\alpha_{L\Delta E}$	fix	
		$\alpha_{R\Delta E}$	fix	
		$2^{\text{nd}}$ -order polynomial( $\cos(\theta_H)$ )		float
Continuum MC	Argus( $m_{\text{ES}}$ )	$\xi_{m_{\text{ES}}}$	float	
		$E_{\text{BEAM}}$ (GeV)	fix	
	$K^{*0} \rightarrow K_S\pi^0$ $K^{*+} \rightarrow K^+\pi^0$	$1^{\text{st}}$ -order polynomial( $\Delta E$ )	$P_1$	float
		$1^{\text{st}}$ -order polynomial( $\cos(\theta_H)$ )	$P_1$	float
		$2^{\text{nd}}$ -order polynomial( $\cos(\theta_H)$ )	$P_1, P_2$	float
$BB$ background MC	Novosibirsk( $m_{\text{ES}}$ )	$\langle m_{\text{ES}} \rangle$ (GeV/c <sup>2</sup> )	fix	
		$\sigma_{m_{\text{ES}}}$ (GeV/c <sup>2</sup> )	fix	
		$\tau_{Bbkg}$	fix	
	Crystal Ball( $\Delta E$ )	$\langle \Delta E \rangle$ (GeV)	fix	
		$\sigma_{\Delta E}$ (GeV)	fix	
		$\alpha_{\Delta E}$	fix	
		$n_{\Delta E}$	fix	
		$4^{\text{th}}$ -order polynomial( $\cos(\theta_H)$ )	$P_1 - P_4$	fix

of the  $B$ 's in the sample. In all other respects, the  $\mathcal{A}_{CP}$  fit is identical to the fit for the branching fractions.

## 6.4 Branching Fraction Fitting

The previous section constructed each PDF line shape and the result was applicable to both the branching fraction and  $CP$  asymmetry fits. This section will describe the branching fraction fits, the toy studies performed, and the evaluation of the systematic errors.

### 6.4.1 Toy Studies

This section describes the toy studies done to validate the fits for the branching fractions. Two types of toy studies are performed, pure and embedded toys. The pure toys are defined as a series of fits performed in which the events are generated according to the PDFs defined in Section 6.3.1. These events are then fit with the nominal fit. The embedded toys studies are also a series of fits. However, instead of all the components being generated from the PDFs, a component, chosen as either signal (Signal Embedded) or  $B\bar{B}$  ( $B$  Embedded), is from the MC simulation sample.

#### 6.4.1.1 Pure Toys

The pure toy studies primarily test if there is any bias in the fitting procedure, such as problems with the fitting code. For each mode, there were 1000 pure toy experiments performed, with a convergence rate of approximately 99%. Figure 6.9 shows the pull and parameter distributions for the 3 component yields for the  $K^{*0} \rightarrow K^+ \pi^-$  mode. The pull is defined as

$$\text{Pull} = \frac{n_{\text{extracted}} - n_{\text{generated}}}{\sigma}, \quad (6.7)$$

where  $n_{\text{extracted}}$  is the number of extracted events,  $n_{\text{generated}}$  is the number of generated events, and  $\sigma$  is the error on the extracted yield. In the pure toys, we find no significant bias in the signal yield.

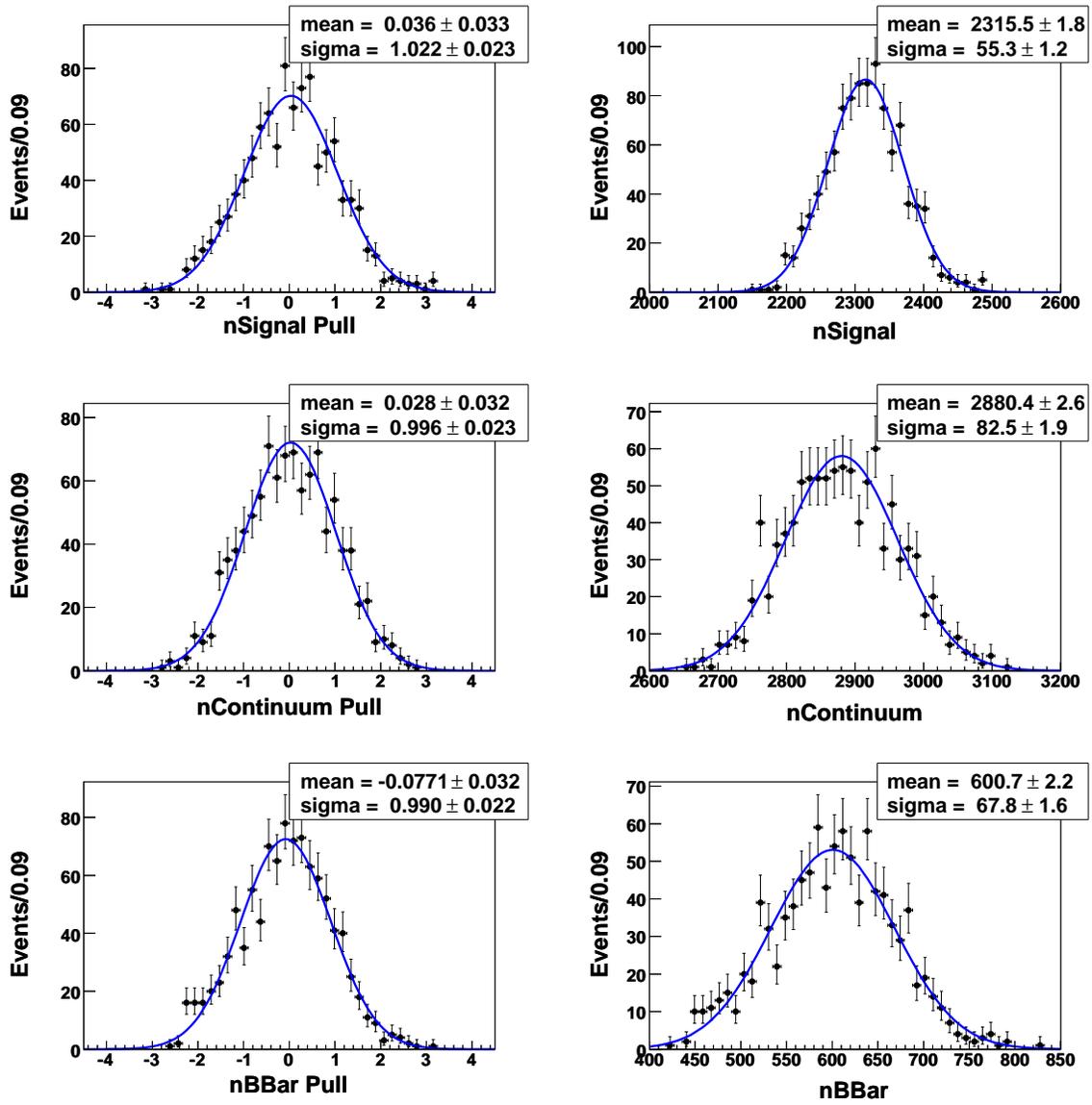


Figure 6.9:  $K^{*0} \rightarrow K^+\pi^-$  pure toy results for 1000 experiments for the 3 different component yields.

### 6.4.1.2 Signal Embedded Toys

The embedded toy studies characterize the deviation of the fit model from the actual distribution of events. The deviations can be caused by correlations among the observables as shown in 6.3.2, or by the PDF shape not modeling the one dimensional distribution well. In this section, we embed signal Monte Carlo events only. We randomly select events from the signal MC simulation sample. The number of events drawn is Poissonian distributed around the expected number in  $347 \text{ fb}^{-1}$  (Table 6.1). In addition, we make sure the samples are statistically independent. Some bias on the signal yield is present (Table 6.8). The cause of the bias is investigated in appendix D. Figure 6.10 shows the pull and yield distributions from this study for the  $K^{*0} \rightarrow K^+ \pi^-$  mode. Table 6.11 summarizes the correction factors, defined as the average extracted to the expected signal yield, to the signal efficiencies associated with the fit bias. The systematic error is determined as half of the fit bias.

Table 6.8: The pulls of the yields for each component for nominal fits (parameters behaving as in tables 6.6 and 6.7). All fits are signal embedded.

	$B^0 \rightarrow K^{*0} \gamma$		$B^+ \rightarrow K^{*+} \gamma$	
	$K^{*0} \rightarrow K^+ \pi^-$	$K^{*0} \rightarrow K_S \pi^0$	$K^{*+} \rightarrow K^+ \pi^0$	$K^{*+} \rightarrow K_S \pi^+$
Nominal Pull				
Sig. Pull	$-0.814 \pm 0.12$	$-0.630 \pm 0.062$	$-0.751 \pm 0.15$	$-0.697 \pm 0.12$
Cont. Pull	$0.56 \pm 0.11$	$0.092 \pm 0.058$	$0.12 \pm 0.12$	$0.24 \pm 0.11$
$B\bar{B}$	$-0.076 \pm 0.098$	$0.206 \pm 0.061$	$0.22 \pm 0.13$	$-0.10 \pm 0.10$

### 6.4.1.3 $B\bar{B}$ Embedded Toys

In addition to the signal embedded toys, we perform independent B background embedded toy experiments. The major  $B\bar{B}$  backgrounds are listed in Table 6.9. These backgrounds are weighted by luminosity and embedded into the background.

On average, approximately 90% of the total  $B\bar{B}$  background is embedded per mode. The results are shown in Table 6.10. The bias on the signal yield is almost negligible. The correction factors resulting from all the embedded toy studies are shown in Table 6.11. The systematic error is determined in Section 6.5.11.1.

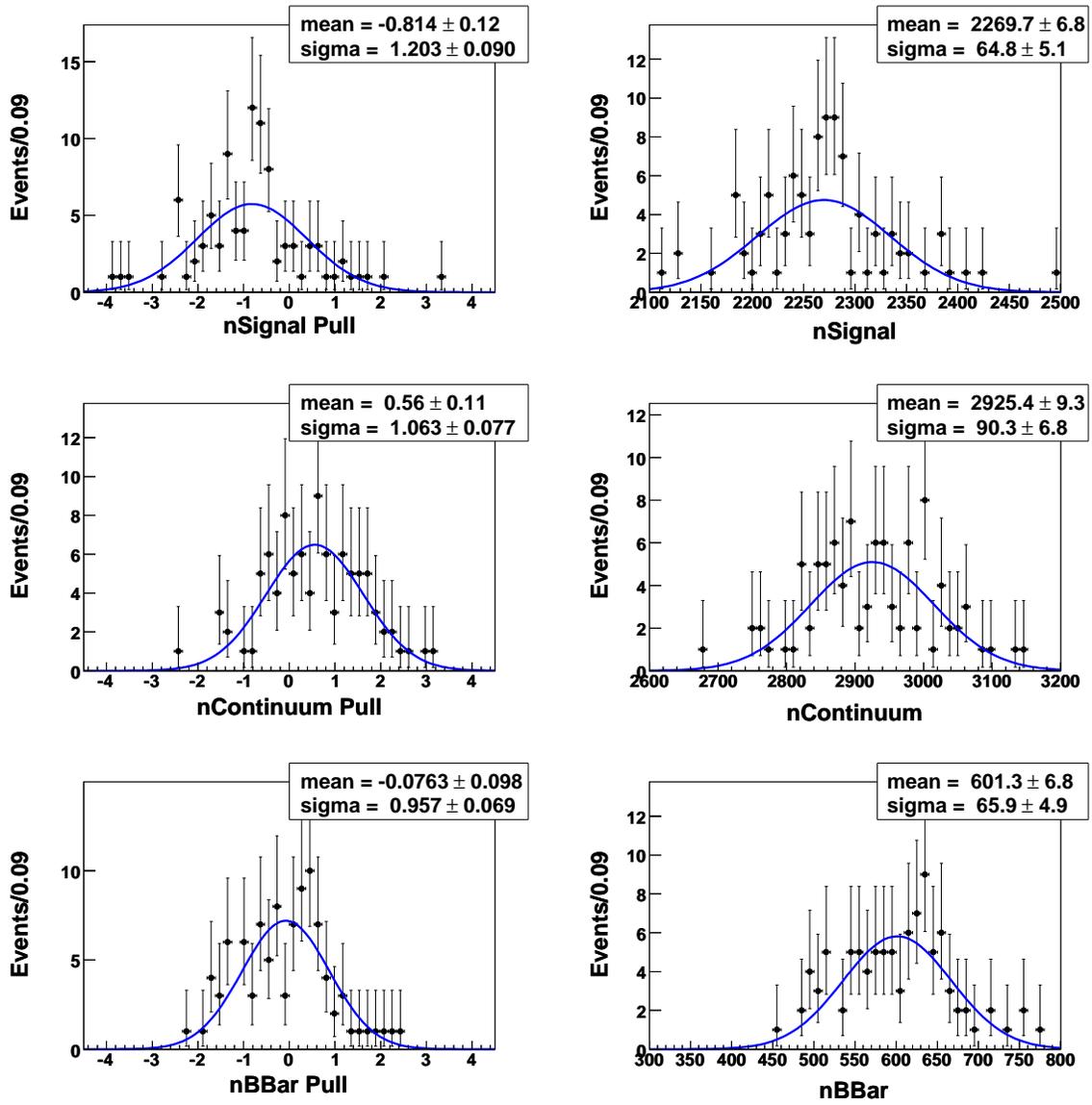


Figure 6.10:  $K^{*0} \rightarrow K^+\pi^-$  signal embedded toy results for the 3 different component yields.

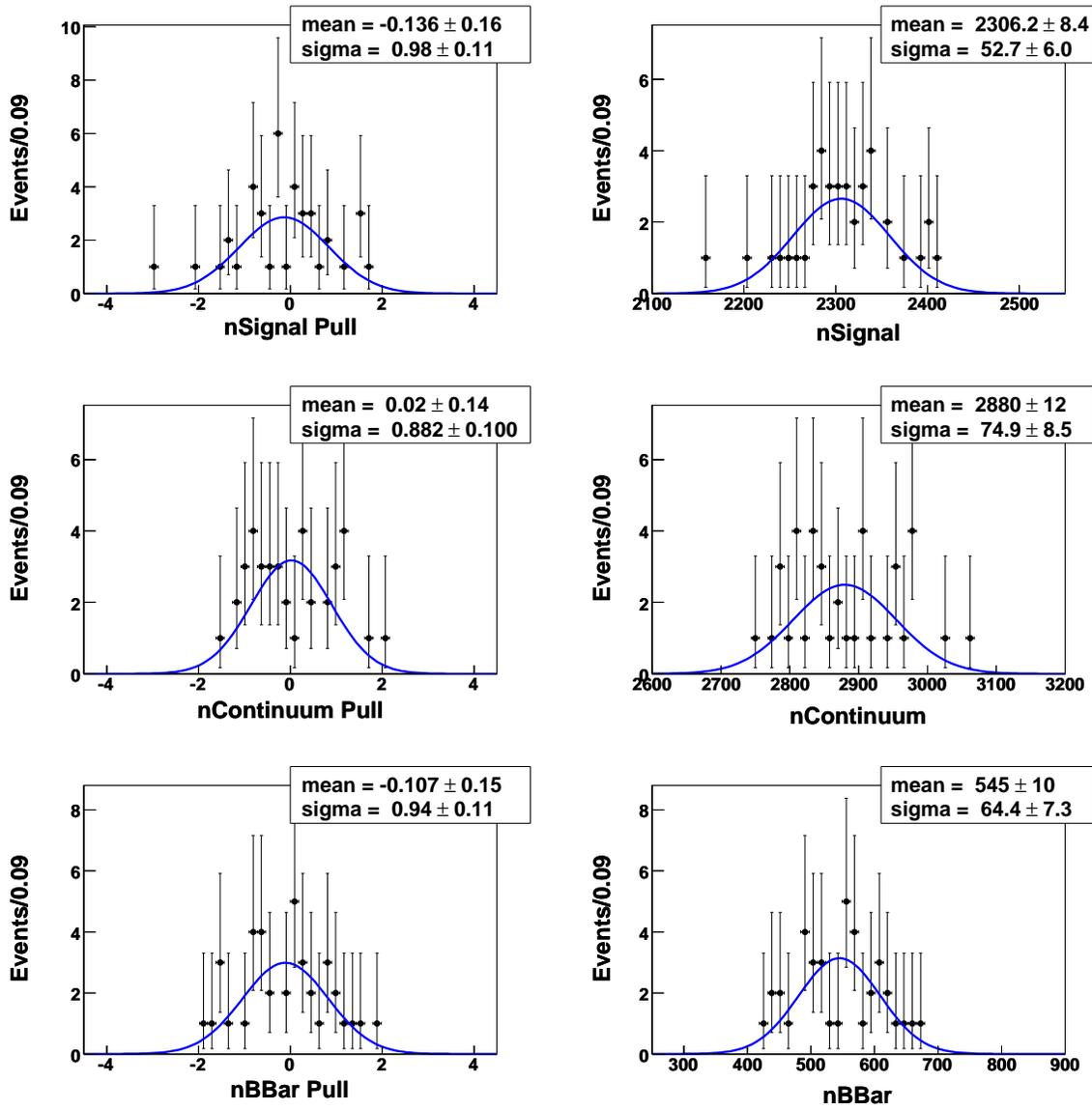


Figure 6.11:  $K^{*0} \rightarrow K^+ \pi^- B\bar{B}$  embedded toy results for 39 experiments for the 3 different component yields.

Table 6.9: The major  $B\bar{B}$  backgrounds of the  $B \rightarrow K^*\gamma$  modes. Shown in parentheses is the percentage of the total  $B\bar{B}$  background.

$B \rightarrow K^*\gamma$ mode	$B\bar{B}$ backgrounds (% of total $B\bar{B}$ background)
$B^0 \rightarrow K^{*0}\gamma, K^{*0} \rightarrow K^+\pi^-$	$B \rightarrow X_s^+\gamma$ (81%) $B^0 \rightarrow K^{*0}\eta, K^{*0} \rightarrow K^+\pi^-$ (5%) $B^+ \rightarrow K^{*+}\gamma, K^{*+} \rightarrow K^+\pi^0$ (5%)
$B^0 \rightarrow K^{*0}\gamma, K^{*0} \rightarrow K_S\pi^0$	$B \rightarrow X_s^+\gamma$ (75%) $B^+ \rightarrow K^{*+}\gamma, K^{*+} \rightarrow K_S\pi^+$ (11%)
$B^+ \rightarrow K^{*+}\gamma, K^{*+} \rightarrow K^+\pi^0$	$B \rightarrow X_s^+\gamma$ (72%) $B^+ \rightarrow K^{*+}\eta, K^{*+} \rightarrow K^+\pi^0$ (4%) $B^0 \rightarrow K^{*0}\gamma, K^{*0} \rightarrow K^+\pi^-$ (11%)
$B^+ \rightarrow K^{*+}\gamma, K^{*+} \rightarrow K_S\pi^+$	$B \rightarrow X_s^+\gamma$ (79%) $B^+ \rightarrow K^{*+}\eta, K^{*+} \rightarrow K_S\pi^+$ (6%) $B^0 \rightarrow K^{*0}\gamma, K^{*0} \rightarrow K_S\pi^0$ (8%)

Table 6.10: The pulls of the yields for each component for nominal fits (parameters behaving as in table 6.6 and 6.7). All fits are  $B\bar{B}$  embedded.

	$B^0 \rightarrow K^{*0}\gamma$		$B^+ \rightarrow K^{*+}\gamma$	
	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$	$K^{*+} \rightarrow K^+\pi^0$	$K^{*+} \rightarrow K_S\pi^+$
Nominal Pull				
Sig. Pull	$-0.136 \pm 0.16$	$0.27 \pm 0.17$	$0.15 \pm 0.15$	$0.17 \pm 0.18$
Cont. Pull	$0.020 \pm 0.14$	$0.40 \pm 0.14$	$0.46 \pm 0.13$	$0.36 \pm 0.17$
$B\bar{B}$	$-0.107 \pm 0.15$	$-0.807 \pm 0.14$	$-0.824 \pm 0.14$	$-0.570 \pm 0.16$

## 6.5 Branching Fraction Systematic Errors

The  $B \rightarrow K^*\gamma$  branching fraction from each mode is calculated as

$$\mathcal{B}(B \rightarrow K^*\gamma) = \frac{N_{sig}}{2N_B \cdot \epsilon \cdot I \cdot \mathcal{B}(\Upsilon(4S) \rightarrow B\bar{B})} \quad (6.8)$$

where  $N_{sig}$  is the signal yield,  $N_B$  is the number of  $B$  mesons,  $\epsilon$  is the signal efficiency, and  $I$  is the isospin factor. The systematic uncertainty on the branching fraction is derived from the uncertainty on the signal efficiency, the uncertainty on  $\mathcal{B}(\Upsilon(4S) \rightarrow B\bar{B})$ , and the error on  $N_B$ . All of the systematic uncertainties are summarized in Table 6.12. There are two categories that the uncertainties fall into: signal reconstruction efficiency and signal extraction. The latter are composed of “ $B$  Counting” ( $N_B$ ) and both systematic uncertainties pertaining to the fit. The signal reconstruction efficiency systematic uncertainties

Table 6.11: The correction factor for signal efficiency for each decay mode. Included are the signal and  $B\bar{B}$  embedded toy studies.

	$B^0 \rightarrow K^{*0} \gamma$		$B^+ \rightarrow K^{*+} \gamma$	
	$K^{*0} \rightarrow K^+ \pi^-$	$K^{*0} \rightarrow K_S \pi^0$	$K^{*+} \rightarrow K^+ \pi^0$	$K^{*+} \rightarrow K_S \pi^+$
Expected Sig. Yield	2313	253	842	754
Signal Embedded Toys				
Sig. Yield	$2269.7 \pm 6.8$	$241.4 \pm 1.2$	$815.1 \pm 5.4$	$732.4 \pm 3.9$
Correction Factor	0.981	0.954	0.969	0.971
$B\bar{B}$ Embedded Toys				
Sig. Yield	$2306.1 \pm 8.4$	$256.9 \pm 3.7$	$850.5 \pm 7.6$	$759.7 \pm 6.7$
Correction Factor	1.00	1.02	1.01	1.01

Table 6.12: This table describes the factors needed to correct the Monte Carlo efficiency and their systematic errors. Shown in paranthesis is the section number that provides a description of the associated systematic. The final estimated efficiency is also given.

	$B^0 \rightarrow K^{*0} \gamma$				$B^+ \rightarrow K^{*+} \gamma$			
	$K^{*0} \rightarrow K^+ \pi^-$		$K^{*0} \rightarrow K_S \pi^0$		$K^{*+} \rightarrow K^+ \pi^0$		$K^{*+} \rightarrow K_S \pi^+$	
Monte Carlo Eff.	0.229		0.145		0.167		0.215	
Descriptions (section)	Factor	Rel. Err.	Factor	Rel. Err.	Factor	Rel. Err.	Factor	Rel. Err.
$B\bar{B}$ Production (6.5.1)	1.000	1.2%	1.000	1.2%	1.000	1.2%	1.000	1.2%
$B$ Counting (6.5.2)	1.000	1.1%	1.000	1.1%	1.000	1.1%	1.000	1.1%
Tracking Eff. (6.5.3)	1.000	1.2%	-	-	1.000	0.6%	1.000	0.8%
Particle Identification (6.5.4)	1.002	0.6%	-	-	1.003	0.6%	0.999	0.2%
Photon Eff. (6.5.6)	0.993	0.7%	0.993	0.7%	0.993	0.7%	0.993	0.7%
$\pi^0$ Eff. (6.5.7)	-	-	0.968	3.0%	0.968	3.0%	-	-
Photon Quality (6.5.8)	1.000	2.1%	1.000	2.1%	1.000	2.1%	1.000	2.1%
$\pi^0(\eta)$ veto(6.5.9)	0.993	1.0%	0.993	1.0%	0.998	1.0%	0.997	1.0%
$K_s$ (6.5.5)	-	-	0.992	0.7%	-	-	0.989	0.7%
Neural network (6.5.10)	0.983	1.5%	0.974	1.0%	0.976	1.0%	0.976	1.0%
Fitting (6.5.11)	1.000	0.8%	1.000	5.6%	1.000	3.1%	1.000	1.7%
Fit bias (6.4.1.2)	0.981	0.9%	0.973	2.2%	0.979	1.6%	0.980	1.4%
Total Correction	0.953	3.8%	0.897	7.4%	0.919	5.6%	0.936	3.9%
Corrected Efficiency	0.218	3.8%	0.130	7.4%	0.153	5.6%	0.201	3.9%
Sub-Mode BR	0.667		0.114		0.329		0.231	

associated with the photon are common to all four modes. In general, the rest of the efficiency systematic uncertainties depend on the mode. All systematic uncertainties are described in more detail below.

### 6.5.1 $B\bar{B}$ Production

The  $B\bar{B}$  production uncertainty comes from the measured  $\Upsilon(4S)$  branching fraction to  $B^0\bar{B}^0, B^+B^-$ . We use [4]

$$\begin{aligned}\mathcal{B}(\Upsilon(4S) \rightarrow B^0\bar{B}^0) &= 0.484 \pm 0.006 \\ \mathcal{B}(\Upsilon(4S) \rightarrow B^+B^-) &= 0.516 \pm 0.006.\end{aligned}\tag{6.9}$$

### 6.5.2 B Counting Systematic

The total number of  $\Upsilon(4S)$  mesons is obtained by counting the number of hadronic events in the on-resonance data set [59], and subtracting the number of muon pairs in the on-resonance data set multiplied by the ratio of hadronic events to muon pairs in an off-resonance sample. The error obtained on the number of produced  $\Upsilon(4S)$  mesons is 1.1%, which comes primarily from systematic uncertainties in the reconstruction of the control sample to determine the efficiency of  $B\bar{B}$  pairs.

### 6.5.3 Tracking Efficiency

The tracking efficiency systematic uncertainty is determined from two different sources, which are added in quadrature with each other to obtain the final uncertainty. One study uses  $\tau$  pair decays [60]. The tau pair is identified by one tau being reconstructed as the decay  $\tau^\pm \rightarrow l^\pm \bar{\nu} \nu$  ( $l = e, \mu$ ), while the second tau decay is  $\tau \rightarrow h^\pm h^\pm h^\pm \nu_\tau$  ( $h = \pi, K$ ). These particular  $\tau$  decays are utilized in order to constrain the event to a specific multiplicity  $n$ . The tracking efficiency is determined by requiring  $n - 1$  tracks in the event, and then finding the probability of reconstructing the  $n$ th track. This study yields an overall systematic uncertainty, obtained from the data/MC difference in track reconstruction, per track of 0.45%. The second method is SVT-based [61]. The efficiency of both data and MC simulation of the DCH is determined by taking the ratio of the number of GTL tracks with 10 SVT hits to the number of GTL tracks with 10 SVT hits without the requirement of 12 DCH hits. The correction is determined by taking the data/MC ratio

of the efficiencies. The result is in bins of transverse momentum, track multiplicity, polar angle, and azimuthal angle. An overall systematic uncertainty is obtained by applying the correction on a track-by-track basis to the signal Monte Carlo simulation. Adding the results from both studies in quadrature gives a variation of the signal track systematic uncertainty of 0.6% to 0.8%.

#### 6.5.4 Particle Identification Efficiency

The particle identification efficiency systematic uncertainty associated with the kaon `Tight` selection is studied in Ref. [62]. A control sample of  $B \rightarrow K J/\psi$  and  $B \rightarrow K^* J/\psi$  decays is used to determine the data/MC efficiencies. The  $K$  correction and systematic uncertainty is determined by removing the  $K$  identification from the  $B \rightarrow K J/\psi$  control sample and dividing it into two datasets, one where the  $K$  is positively identified as passing the kaon `Tight` selection criteria and one where it is not. The efficiency is then calculated as the number that were positively identified over the sum of the two datasets. A similar procedure is used for the  $B \rightarrow K^* J/\psi$  control sample for the pion. The systematic uncertainty is less than 1% for both the  $\pi$  and  $K$ , and comes from the statistics of the control samples.

#### 6.5.5 $K_S$ Efficiency

The  $K_S$  efficiency systematic uncertainty is determined from an inclusive sample of well-reconstructed  $K_S$  candidates, which is documented in Ref. [63]. Using the  $K_S$  selection criteria specified in Section 5.3.2, data/MC ratios are made in bins of transverse momentum of the  $K_S$ , polar angle of the  $K_S$ , and transverse distance between the primary vertex and the  $K_S$  vertex. An overall correction is obtained by applying the correction per  $K_S$  to the signal Monte Carlo simulation, while the systematic uncertainty is given by calculating the differences in the average correction on a control sample of  $B^0 \rightarrow \pi^+ D^- (D^- \rightarrow K_S \pi^-)$  signal Monte Carlo simulation for different selection criteria for the  $K_S$ . For the  $K^{*0} \rightarrow K_S \pi^0$  mode we obtain a correction of 0.992 with an uncertainty of 0.7%, while for the  $K^{*+} \rightarrow K_S \pi^+$  mode a correction of 0.989 with an uncertainty of 0.7% is obtained.

### 6.5.6 Single Photon Efficiency

Using a sample of  $e^+e^- \rightarrow \mu^+\mu^-\gamma$  events, a single photon efficiency correction and systematic uncertainty is determined [64]. This sample is used because it is large, has relatively low backgrounds, and can provide high energy photons. The correction is given as the data/MC ratio of the number of events observed divided by the total number of events, while the systematic uncertainty is given by systematic uncertainties in the reconstruction of the control sample. The correction is 0.993 with a systematic uncertainty of 0.7%.

### 6.5.7 $\pi^0$ Efficiency

The  $\pi^0$  systematic uncertainty is obtained from a sample of  $\tau$  pair events [65]. One  $\tau$  is selected to decay by  $\tau^\pm \rightarrow e^\pm\nu\bar{\nu}$ , while the other is required to decay by  $\tau^\pm \rightarrow h^\pm\nu_\tau$  ( $h = \pi, \rho$ ) ( $\rho^\pm \rightarrow \pi^\pm\pi^0$ ). The  $\pi^0$  efficiency is obtained by taking the ratio of the  $\tau \rightarrow \rho$  to  $\tau \rightarrow \pi$  yields. The  $\tau \rightarrow \rho$  yield is roughly proportional to reconstructing the  $\pi^\pm$  and the  $\pi^0$ , while the  $\tau \rightarrow \pi$  yield is proportional to the  $\pi^\pm$  efficiency. The systematic error of 3% is primarily due to the discrepancy between data and MC simulation of low energy photons and fluctuations from hadronic showers which fake a bump in the calorimeter. These phenomena are not well understood.

### 6.5.8 Photon Quality Selection

There are two systematic uncertainties associated with the photon quality selection. One is the data/MC difference of requiring the photon to be 25 cm away from the closest charged or neutral EMC bump, while the other is the data/MC difference of requiring the second moment, defined in Section 5.3.5, to be less than 0.002.

The efficiency of the photon isolation requirement is sensitive to any discrepancy in the MC simulation that pertains to the number of calorimetric objects in the event. For example, if there are more calorimetric objects in the MC simulation than in data, then it is more likely there will be one close to the signal photon. This effect is studied by “embedding” the photon in the MC simulation and in data [66]. The response of the EMC to photons from radiative Bhabba events is added to MC simulated  $B\bar{B}$  events and a data sample of reconstructed  $B \rightarrow D\pi$  events. The event is then reconstructed again with the new photon and the ratio of the data/MC efficiencies is determined. The systematic uncertainty is 2%.

The systematic uncertainty related to the second moment is determined by a control sample of  $e^+e^- \rightarrow \mu^+\mu^-\gamma$  events [52]. The energy of the  $\gamma$  in the control sample is weighted to match the photon energy spectrum of  $B \rightarrow K^*\gamma$  events, which is between 1.5 and 3.5 GeV. The systematic error is found to be 0.6%.

### 6.5.9 $\pi^0/\eta$ Veto

The systematic uncertainty of the  $\pi^0/\eta$  veto is determined by embedding a photon in  $B \rightarrow D\pi$  Monte Carlo and data events [52]. Again, the energy of the embedded photon matches the energy spectrum of the high energy gamma in  $B \rightarrow K^*\gamma$  events. The efficiency of the likelihood ratios are compared for data and MC simulation, and a correction factor and systematic uncertainty are determined. The systematic uncertainty obtained is 1%.

## 6.5.10 Neural Net systematic

### 6.5.10.1 Overview

The neural network systematic uncertainty is evaluated on a sample of exclusively reconstructed  $B \rightarrow D\pi$  events. The  $B \rightarrow D\pi$  sample is used because it has a similar topology to  $B \rightarrow K^*\gamma$  events; the signal side of the event has a heavy and light particle. To evaluate the systematic uncertainty, the neural network that is constructed in Section 5.4 is applied to the  $B \rightarrow D\pi$  sample. However, the input variables are calculated by treating the bachelor pion as the high energy photon. The modes that are reconstructed are  $B^- \rightarrow D^0\pi^- (D^0 \rightarrow K^+\pi^-)$  and  $B^0 \rightarrow D^-\pi^+ (D^- \rightarrow K^+\pi^-\pi^-)$ . To closely resemble the ROE in the signal  $B \rightarrow K^*\gamma$  modes, the  $B^- \rightarrow D^0\pi^- (D^0 \rightarrow K^+\pi^-)$  mode is used to validate the  $K^{*+} \rightarrow K^+\pi^0$  and  $K^{*+} \rightarrow K_S\pi^+$  modes, while the  $B^0 \rightarrow D^-\pi^+ (D^- \rightarrow K^+\pi^-\pi^-)$  mode is used to validate the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*0} \rightarrow K_S\pi^0$  modes. The neural network output distribution between  $B \rightarrow D\pi$  MC simulation and data is compared and a correction factor and systematic uncertainty is extracted.

### 6.5.10.2 Validation Outline

The following steps are followed to validate the neural network used in  $B \rightarrow K^*\gamma$  decays:

- The  $B \rightarrow D\pi$  decays mentioned above are reconstructed and the neural network

variables used in  $B \rightarrow K^*\gamma$  decays are calculated by treating the bachelor pion as the photon.

- Additional event selections are applied on the signal side of the event to suppress the background.
- After background rejection, we select  $B \rightarrow D\pi$  events in the signal region from the on-peak data and make a bin-by-bin comparison of the neural network shape with the corresponding signal Monte Carlo. Background subtraction is performed on on-peak data before this comparison. Based on this bin-by-bin comparison between data and Monte Carlo, we calculate the efficiency ratio as a function of neural network output between data and MC simulation.

### 6.5.10.3 Selections and Background Levels

**Signal Region and Sidebands** We define the following regions in the  $m_{ES}-\Delta E$  plane for the control sample.

- Signal region:  $[5.27, 5.29] \text{ GeV}/c^2 \times [-0.05, 0.05] \text{ GeV}$ .
- Fitting region:  $[5.20, 5.29] \text{ GeV}/c^2 \times [-0.05, 0.05] \text{ GeV}$ .
- Upper sideband:  $[5.20, 5.29] \text{ GeV}/c^2 \times [0.1, 0.2] \text{ GeV}$ .
- Lower sideband:  $[5.20, 5.27] \text{ GeV}/c^2 \times [-0.05, 0.05] \text{ GeV}$ .

**Event Selections** The following selections are used on the  $B \rightarrow D\pi$  decays to suppress the background and to ensure that a good event is reconstructed:

- The GTL criteria is used on all charged tracks.
- The kaon `veryTight` LH selection is applied on charged kaons.
- The probability of the  $\chi^2$  for the  $D$  vertex fit is required to be greater than 0.0001.
- A selection of  $\pm 20 \text{ MeV}/c^2$  around the PDG [4] central value of the  $D$  mass.
- A selection of  $\pm 50 \text{ MeV}$  around 0 of  $\Delta E$ .

After all selections are applied, we choose the best candidate in each event by selecting the one with the minimum  $\Delta E$ .

Table 6.13: The signal efficiency and expected yields for the  $B \rightarrow D\pi$  control samples in the on-peak dataset. The error on the expected yield contains the statistics error on the signal efficiency only. The uncertainties on the branching fraction are not included.

Description	eff.( $\pm$ stat) (fit region)	eff.( $\pm$ stat) (signal region)	Expected Yields (347 fb $^{-1}$ )
$B^- \rightarrow D^0 \pi^-, D^0 \rightarrow K^+ \pi^-$	$0.337 \pm 0.001$	$0.336 \pm 0.001$	$17373 \pm 52$
$B^0 \rightarrow D^- \pi^+, D^- \rightarrow K^+ \pi^- \pi^-$	$0.247 \pm 0.001$	$0.246 \pm 0.001$	$17276 \pm 70$

**Background** The continuum background events contained in the signal region after the selection criteria have been applied are dealt with as described in Sec 6.5.10.5.

After all of the selection criteria have been applied, the  $B$  background is approximately 2% in the signal region. This is determined by  $B\bar{B}$  MC simulation. Most of the  $B$  background is from  $B \rightarrow D^* \pi$  events, in which the slow pion is missing; therefore, the event is treated as signal, since it has a neural network distribution identical to signal decays. The remaining  $B$  background fraction is less than 1% .

#### 6.5.10.4 Efficiency and Yields

**Signal Efficiency and Expected Yield** Table 6.13 shows the signal efficiency after applying all the selection criteria and the expected yields in the on-peak dataset.

**Signal Yields and Background Estimation in the On-Peak Dataset** We estimate the signal yields and the continuum background in the signal region by a one dimensional fit to  $m_{ES}$ . The signal shape is modeled using a Gaussian and the continuum background is modeled by an ARGUS function [58]. Figure 6.12 shows the likelihood fit of the  $B \rightarrow D\pi$  on-peak data. The number of events in the lower sideband and the signal region, the fitted signal yields, and the estimated background in the signal region are shown in Table 6.14. The expected yields shown in Table 6.13 are calculated using the branching fractions from PDG [4].

#### 6.5.10.5 Background Subtraction

The on-peak data in the signal region contains a small amount of continuum background, which is subtracted using events in the lower sideband. Table 6.14 gives the

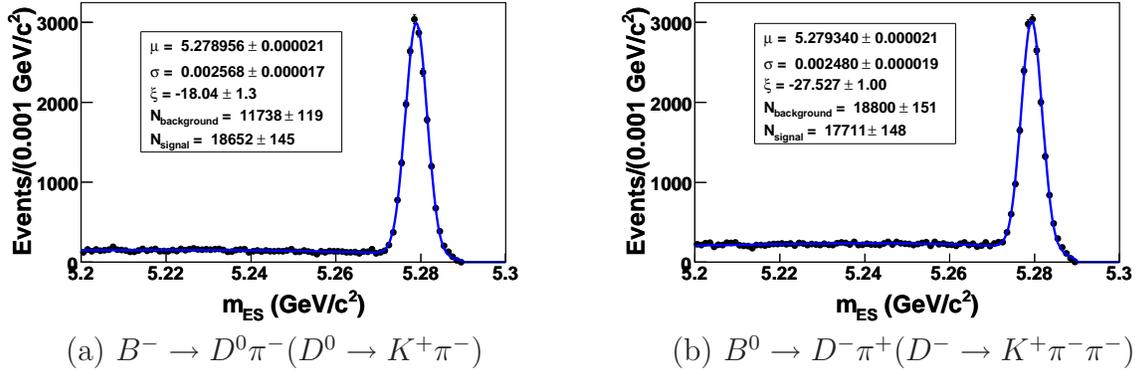


Figure 6.12: The likelihood fit on the (a)  $B^- \rightarrow D^0 \pi^- (D^0 \rightarrow K^+ \pi^-)$  and (b)  $B^0 \rightarrow D^- \pi^+ (D^- \rightarrow K^+ \pi^- \pi^-)$  on-peak data.

Table 6.14: The number of events in the lower sideband and the signal region, the fitted signal yields, and the estimated background in the signal region for the on-peak data set.

Description	Events (lower sideband)	Events (sig. region)	Sig. yield	Background (sig. region)
$B^- \rightarrow D^0 \pi^-, D^0 \rightarrow K^+ \pi^-$	10130	20259	$18652 \pm 145$	$1629 \pm 17$
$B^0 \rightarrow D^- \pi^+, D^- \rightarrow K^+ \pi^- \pi^-$	15795	20715	$17711 \pm 148$	$2979 \pm 24$

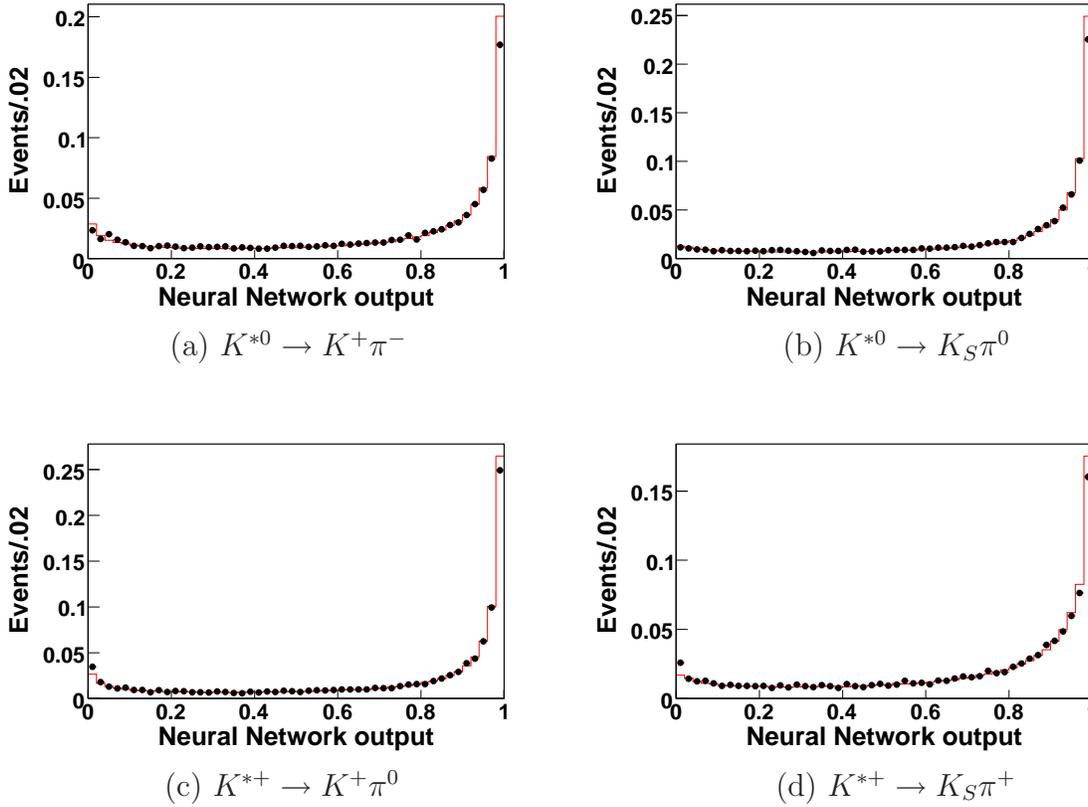


Figure 6.13: The comparison of the neural network distributions for the Monte Carlo simulation (red line) and the background subtracted on-peak data (black dots) for the  $B \rightarrow D\pi$  sample. The distributions, which are normalized to the same area, correspond to the a)  $K^{*0} \rightarrow K^+ \pi^-$ , b)  $K^{*0} \rightarrow K_S \pi^0$ , c)  $K^{*+} \rightarrow K^+ \pi^0$ , and d)  $K^{*+} \rightarrow K_S \pi^+$  modes.

estimated number of continuum events in the signal region. The  $i$ -th bin of the background subtracted distribution for the on-peak data is

$$h_{\text{sub}}(i) = h_{\text{total}}(i) - N_{\text{bkg}}/N_{\text{lower}} \cdot h_{\text{lower}}(i)$$

where  $h_{\text{sub}}$  is the background subtracted distribution, and  $h_{\text{total}}$  and  $h_{\text{lower}}$  are the distributions for all events in the signal region and lower sideband respectively.  $N_{\text{bkg}}$  and  $N_{\text{lower}}$  are, respectively, the estimated number of background events in the signal region and the total number of events in the lower sideband.

Figure 6.13 shows the comparison of the neural network distributions between the  $B \rightarrow D\pi$  signal Monte Carlo simulation and the background-subtracted on-peak data.

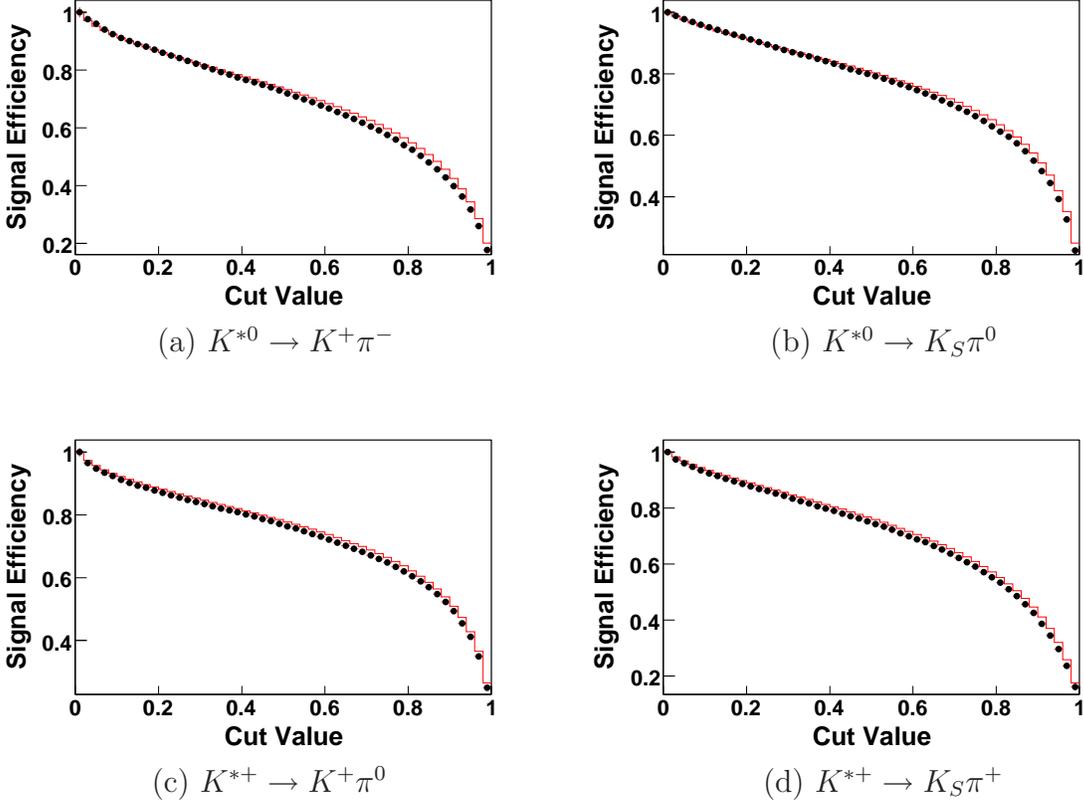


Figure 6.14: The comparison of the signal efficiency vs. cut value between Monte Carlo simulation (red line) and background subtracted on-peak data (black dots) for the a)  $K^{*0} \rightarrow K^+\pi^-$ , b)  $K^{*0} \rightarrow K_S\pi^0$ , c)  $K^{*+} \rightarrow K^+\pi^0$ , and d)  $K^{*+} \rightarrow K_S\pi^+$  modes.

#### 6.5.10.6 Efficiency Correction

Using the background subtracted on-peak data, we calculate the signal efficiency in data. Then, we compare this efficiency with the efficiency of the corresponding signal Monte Carlo simulation as shown in Figure 6.14. The bin-by-bin efficiency ratio,  $e_{\text{Data}}/e_{\text{MC}}$  is calculated and shown Figure 6.15. A slight inefficiency is visible in the data when using the neural network.

#### 6.5.10.7 Result Discussion

The systematic uncertainty is evaluated using two different methods. The first method uses the efficiency ratio at the neural network selection as the correction to the signal efficiency. The signal efficiency corrections associated with the neural network selection

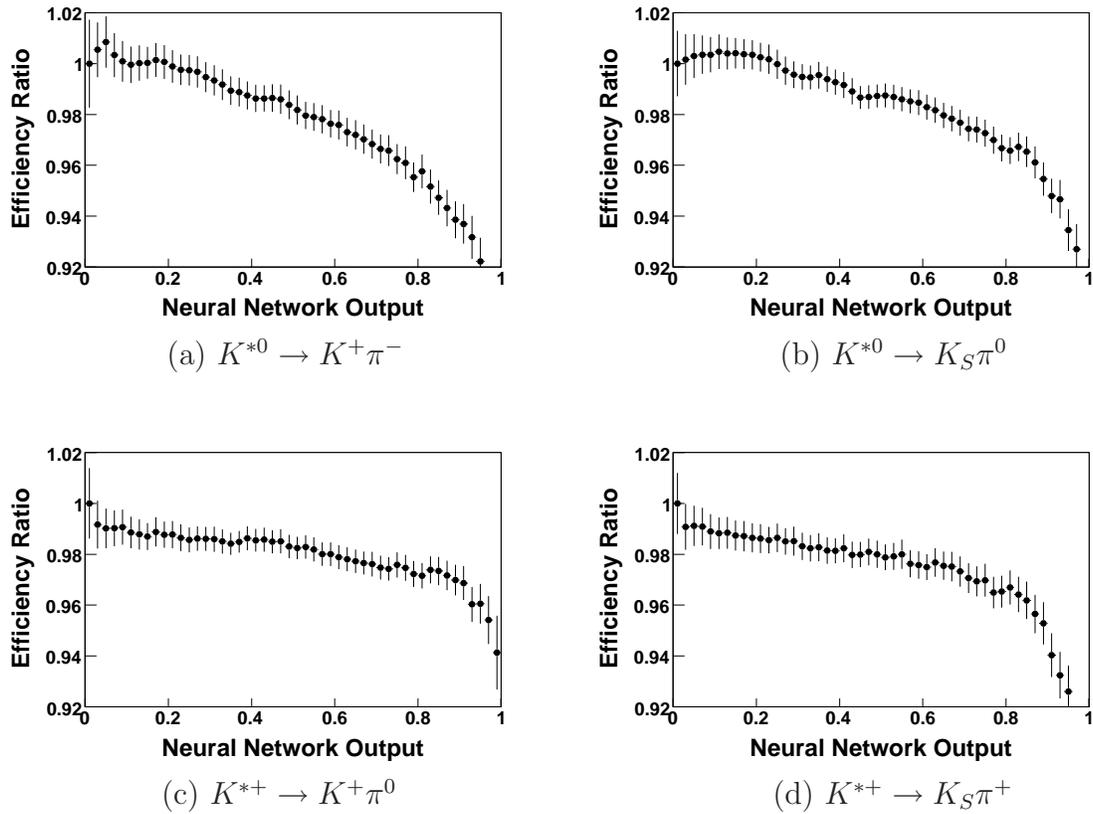


Figure 6.15: The bin-by-bin  $e_{\text{Data}}/e_{\text{MC}}$  for the a)  $K^{*0} \rightarrow K^+ \pi^-$ , b)  $K^{*0} \rightarrow K_S \pi^0$ , c)  $K^{*+} \rightarrow K^+ \pi^0$ , and d)  $K^{*+} \rightarrow K_S \pi^+$  modes.

are 96.8%, 97.0%, 97.4%, and 97.1% for the  $K^{*0} \rightarrow K^+\pi^-$ ,  $K^{*0} \rightarrow K_S\pi^0$ ,  $K^{*+} \rightarrow K^+\pi^0$ , and  $K^{*+} \rightarrow K_S\pi^+$  decay modes respectively. The second method uses the neural network selection which gives the same Monte Carlo efficiency as the  $B \rightarrow K^*\gamma$  signal efficiency. This method gives 98.3%, 97.4%, 97.6%, and 97.6% for the  $K^{*0} \rightarrow K^+\pi^-$ ,  $K^{*0} \rightarrow K_S\pi^0$ ,  $K^{*+} \rightarrow K^+\pi^0$ , and  $K^{*+} \rightarrow K_S\pi^+$  decay modes respectively. The statistical error associated with both methods is  $\sim 0.5\%$ . Taking into account the differences between the two methods, we assign a 1.5% systematic to  $K^{*0} \rightarrow K^+\pi^-$  and a 1.0% to the other three modes, and the corrections using the second method.

### 6.5.11 Fitting Systematic

There are two kinds of systematic error related to the fitting. One uncertainty is that the  $B\bar{B}$  background is not well described in the Monte Carlo simulation, while the other pertains to the variation of the fixed parameters in the fit. We describe both in more detail below.

#### 6.5.11.1 $B\bar{B}$ parameterization systematic

Since the  $B\bar{B}$  shape is fixed in the fit, a difference between the shape obtained from Monte Carlo simulation and the actual shape in data can cause a bias in the signal yield. In particular, the inclusive  $b \rightarrow s\gamma$  background is not well described in the Monte Carlo simulation. In the  $b \rightarrow s\gamma$  Monte Carlo simulation, there are no resonances above the  $K^*$ . Therefore, the particles only decay through phase space with no angular dependence. We vary the normalization of the inclusive  $b \rightarrow s\gamma$  to understand what effect this component has on the signal yield. We use two different methods to do this.

The first method varies the amount inclusive  $b \rightarrow s\gamma$  Monte Carlo events embedded into the fit. The experimental  $b \rightarrow s\gamma$  branching fraction measurement has an uncertainty of  $\sim 10\%$  [67]. We perform independent embedded toys corresponding to two scenarios: one where the  $b \rightarrow s\gamma$  yield is reduced by 20%, and another where it is increased by 20%. The signal and continuum data are retrieved from the PDF shapes, while the  $B\bar{B}$  data is from actual Monte Carlo. In both scenarios we fit with the nominal likelihood function. The results are shown in Table 6.15.

The second method performs pure toy experiments for three different scenarios. The default scenario uses the nominal fit, and is a reproduction of the pure toy experiments performed in Section 6.4.1.1. The second scenario decreases the luminosity weight of the inclusive  $b \rightarrow s\gamma$  sample such that the yield decreases by 20%, while the third scenario

Table 6.15: Signal yields in which the  $b \rightarrow s \gamma$  yield has been varied by  $\pm 20\%$ . The signal yield has been extracted with the nominal fit using the  $B\bar{B}$  embedded toys.

$B \rightarrow K^* \gamma$ mode	Yield from $B\bar{B}$ embedded toys	$-20\%$ $b \rightarrow s \gamma$ yield	$+20\%$ $b \rightarrow s \gamma$ yield
$K^{*0} \rightarrow K^+ \pi^-$	$2306.1 \pm 8.4$	$2317.6 \pm 7.7$	$2305 \pm 10$
$K^{*0} \rightarrow K_S \pi^0$	$256.9 \pm 3.7$	$255.8 \pm 4.0$	$257.4 \pm 6.8$
$K^{*+} \rightarrow K^+ \pi^0$	$850.5 \pm 7.6$	$855.6 \pm 5.2$	$861.1 \pm 7.0$
$K^{*+} \rightarrow K_S \pi^+$	$759.7 \pm 6.7$	$760.6 \pm 4.7$	$763.5 \pm 5.8$

increases the weight such that the inclusive  $b \rightarrow s \gamma$  sample increases by 20%. To both of the latter scenarios, the  $B\bar{B}$  PDF is fit to the new distribution. Then, pure toys are generated using this new PDF (in which only the  $B\bar{B}$  parameters have changed), but fit with the nominal fit. The  $B\bar{B}$  fits for the  $K^{*0} \rightarrow K^+ \pi^-$  mode for the three cases is shown in Figure 6.16, and the results are shown in in Table 6.16. To determine the systematic, we take the largest average relative difference as determined by either method using the central values. We find a systematic 0.3%, 0.9%, and 0.3% for the  $K^{*0} \rightarrow K^+ \pi^-$ ,  $K^{*+} \rightarrow K^+ \pi^0$ , and  $K^{*+} \rightarrow K_S \pi^+$  modes respectively. We assign no systematic for the  $K^{*0} \rightarrow K_S \pi^0$  mode.

Table 6.16: Signal yields in which the  $b \rightarrow s \gamma$  yield has been varied by  $\pm 20\%$ . The signal yield has been extracted with the nominal fit using pure toys. However, the pure toys were generated by varying the  $b \rightarrow s \gamma$  yield.

$B \rightarrow K^* \gamma$ mode	Yield from pure toy study	$-20\%$ $b \rightarrow s \gamma$ yield	$+20\%$ $b \rightarrow s \gamma$ yield
$K^{*0} \rightarrow K^+ \pi^-$	$2315.5 \pm 1.8$	$2311.9 \pm 1.7$	$2314.0 \pm 1.7$
$K^{*0} \rightarrow K_S \pi^0$	$253.2 \pm 0.7$	$253.3 \pm 0.7$	$252.5 \pm 0.7$
$K^{*+} \rightarrow K^+ \pi^0$	$842.4 \pm 1.2$	$844.6 \pm 1.2$	$842.1 \pm 1.2$
$K^{*+} \rightarrow K_S \pi^+$	$752.9 \pm 1.0$	$755.2 \pm 1.0$	$754.7 \pm 1.1$

### 6.5.11.2 Fixed signal parameters

We fix the tail parameters of the signal  $m_{ES}$  and  $\Delta E$  PDFs for all four modes to the MC simulation. Since the tail parameters are fixed, we vary them to determine a systematic uncertainty to account for differences in the Monte Carlo simulation and real data. We vary each parameter by one sigma according to the individual PDF fit. The average

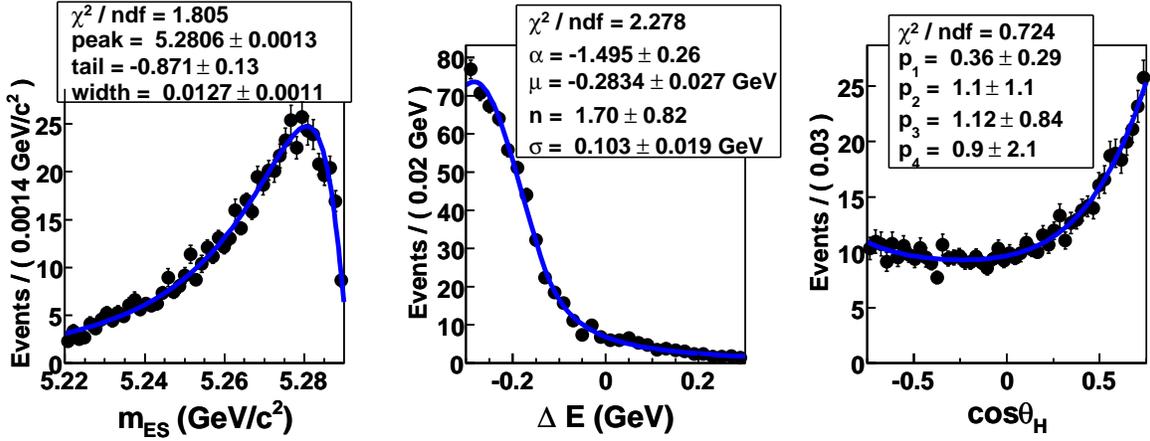
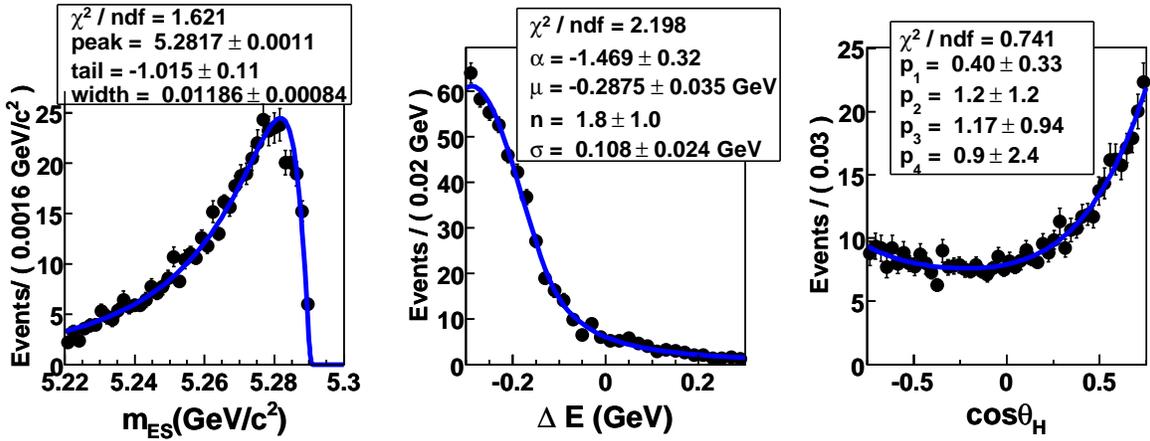
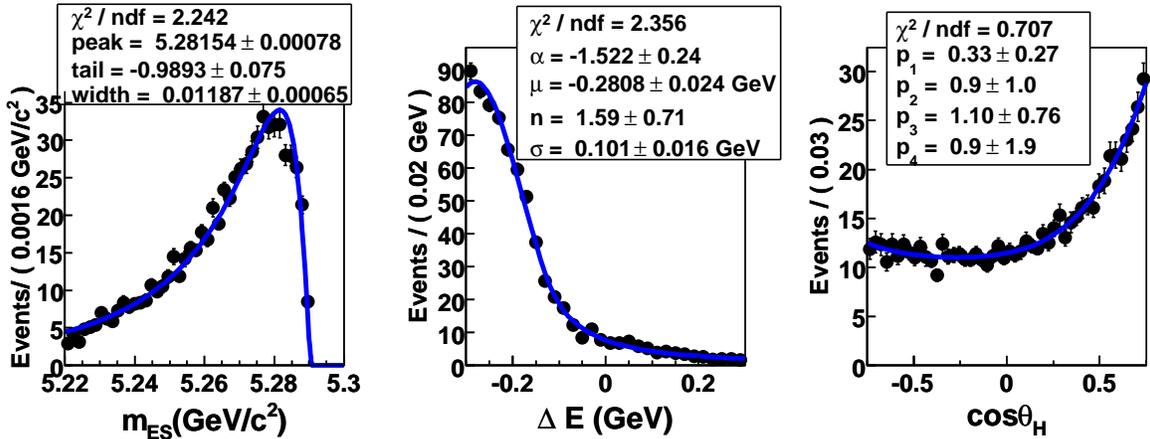
a) Nominal  $B\bar{B}$  parametersb)  $b \rightarrow s\gamma$  yield decreased by 20%c)  $b \rightarrow s\gamma$  yield increased by 20%

Figure 6.16:  $B\bar{B}$  distributions and fit parameters for the three dimensions of the fit for a) the nominal fit, b) when the  $b \rightarrow s\gamma$  yield is decreased by 20%, and when c) the  $b \rightarrow s\gamma$  yield is increased by 20%. The mode shown is  $K^{*0} \rightarrow K^+\pi^-$ .

Table 6.17: Values of  $\mathcal{A}_{CP}$  measured in onpeak  $m_{ES}$  sideband data.

	$\mathcal{A}_{CP}$
$K^{*0} \rightarrow K^+ \pi^-$	$-0.012 \pm 0.005$
$K^{*+} \rightarrow K^+ \pi^0$	$-0.005 \pm 0.006$
$K^{*+} \rightarrow K_S \pi^+$	$+0.007 \pm 0.008$

difference of the signal yield is taken as the systematic uncertainty. Taking into account correlations among the fit parameters, we find a systematic uncertainty of 0.8%, 5.6%, 3.0%, and 1.7% for the  $K^{*0} \rightarrow K^+ \pi^-$ ,  $K^{*0} \rightarrow K_S \pi^0$ ,  $K^{*+} \rightarrow K^+ \pi^0$ , and  $K^{*+} \rightarrow K_S \pi^+$  modes respectively.

The total fitting systematic is found by adding in quadrature the uncertainty derived from varying the  $b \rightarrow s \gamma$  background and that from the fixed parameters. These numbers are displayed in Table 6.12.

## 6.6 $CP$ Asymmetry

The time-independent  $CP$  asymmetry is measured using

$$\mathcal{A}_{CP} = \frac{\Gamma(\bar{B} \rightarrow \bar{K}^* \gamma) - \Gamma(B \rightarrow K^* \gamma)}{\Gamma(\bar{B} \rightarrow \bar{K}^* \gamma) + \Gamma(B \rightarrow K^* \gamma)}. \quad (6.10)$$

Section 6.3.3.1 gave an overview of the method of extraction of the  $CP$  asymmetry. In this section, a discussion of the systematic uncertainties will be given.

### 6.6.1 $CP$ Asymmetry Systematic Errors

To look for any detector-induced asymmetry, we look in the  $m_{ES}$  sideband ( $5.22 < m_{ES} < 5.29$ ,  $-0.3 < \Delta E < 0.3$ ), remove the neural net selection and form  $\mathcal{A}_{CP}$  by counting the number of events with different  $C$  quantum numbers. Table 6.17 lists the results. The  $K^{*0} \rightarrow K^+ \pi^-$  and  $K^{*+} \rightarrow K^+ \pi^0$  modes are consistent with an asymmetry of 0.5%. Due to hadronic interactions in the detector, there is a detector-induced asymmetry for  $K^\pm$ 's and  $\pi^\pm$ 's. We divide these interactions into before and after the DCH, and describe them both in the following sections.

### 6.6.1.1 Hadronic Interactions before the DCH

As mentioned in Section 5.1.1.1, we use **GEANT** to simulate the passage of particles through the detector. However, the accuracy to which **GEANT** handles the asymmetry is not well understood. Therefore, we perform an independent calculation that uses hadronic interaction data from PDG [4] and a model of the *BABAR* detector. This calculation is described in this section.

The nuclear collision length  $\lambda_T$ , given in units of g/cm<sup>2</sup>, is related to the total nuclear cross-section  $\sigma_T$  by

$$\lambda_T^{-1} = \frac{N_A}{A_r} \sigma_T(p), \quad (6.11)$$

where  $N_A$  is Avogadro's number,  $A_r$  is the atomic weight, and  $\sigma_T(p)$  is a function of momentum. However, to a good approximation  $\sigma_T \propto A^{0.7}$  [4], where  $A$  is the mass number. We use the kaon-deuteron cross-section  $\sigma_{K-D}$  as a reference, so that

$$\sigma_T(p) = \sigma_{K-D}(p) \frac{A^{0.7}}{2^{0.7}}. \quad (6.12)$$

This equation is normalized so that when  $A = 2$ ,  $\sigma_T(p) = \sigma_{K-D}$ . The interaction probability is given by

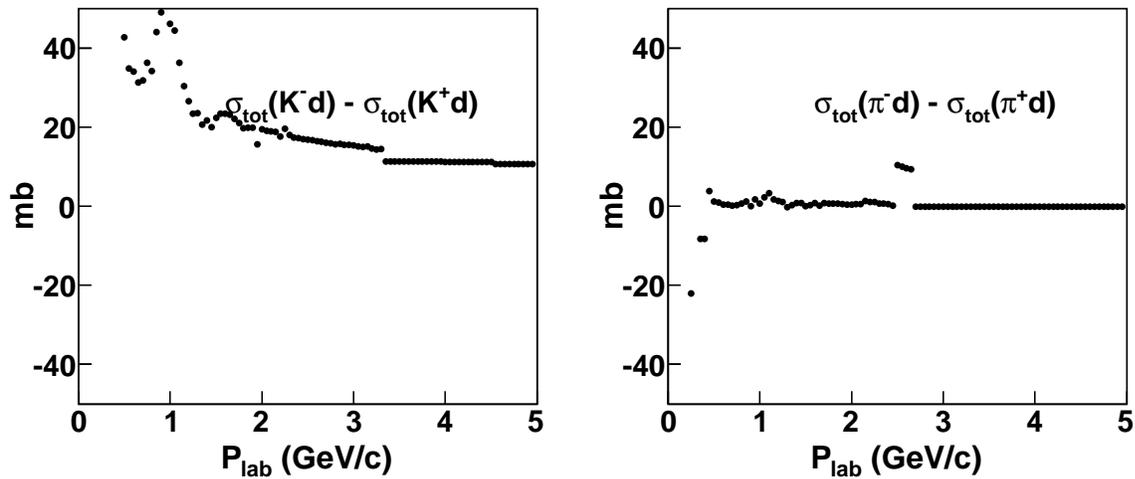
$$P = \sum_i \frac{L_i}{\lambda_{Ti}} = \sum_i L_i \cdot \frac{N_A \sigma_{K-D}}{A_i^{0.3} 2^{0.7}}, \quad (6.13)$$

where  $i$  is an index over the detector materials that were crossed,  $L_i$  is the thickness, and we have used the approximation  $A_r = A$ . Table 6.18 shows the material model that was used. The kaon-deuteron cross-section is a function of momentum. Therefore, it must be evaluated using existing hadronic interaction data folded in with the signal momentum spectrum.

Toward this end, we evaluate the cross-section asymmetry for  $K$ 's and  $\pi$ 's from Ref. [4], and the momentum distribution for  $K$ 's and  $\pi$ 's from truth-matched  $K^{*0} \rightarrow K^+ \pi^-$  signal. The results are shown in Figures 6.17 and 6.18. We convolute these two distributions together and divide by the  $\sin(\theta)$ , where  $\theta$  is the polar angle of the track, on an event-by-event basis to find the average asymmetry. The  $\sin(\theta)$  correction is to account for the amount of material encountered by the track. We find  $\Delta\mathcal{A}_{CP} = -0.33\%$  for  $K$ 's and  $\Delta\mathcal{A}_{CP} = -0.03\%$  for  $\pi$ 's.

Table 6.18: Material model used for hadronic interaction asymmetry calculation

Detector Element	Material	$l(g/cm^2)$	A
Beam Pipe	Au	0.008	196.967
	Be	0.157	9.012
	Ni	0.008	58.693
	$H_2O$	0.147	18.015
	Ni	0.009	58.693
Gas Bag	Be	0.094	9.012
	Al	0.014	26.982
	C	0.010	12.011
SVT	Si	0.350	28.086
Support Tube	C	0.460	12.011
DCH Support	Be	0.185	9.012

Figure 6.17: Total cross section difference for K-deuteron and  $\pi$ -deuteron interactions from the PDG

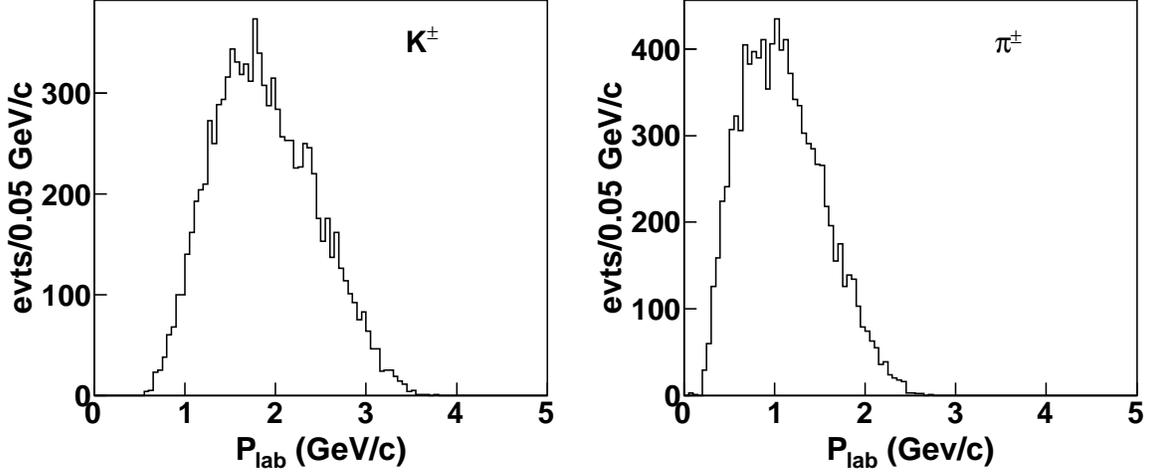


Figure 6.18:  $p_{lab}$  distribution for K's and  $\pi$ 's for truth-matched  $K^{*0} \rightarrow K^+\pi^-$  signal decays.

### 6.6.1.2 Hadronic Interactions after the DCH

The material between the DCH and the DIRC may also absorb mesons and cause an asymmetry in the particle identification efficiency, which is not accounted for above. To understand this effect, we use control samples of  $D^{*+} \rightarrow D^0\pi^+$  ( $D^0 \rightarrow K^-\pi^+$ ) and its charge conjugate to determine data efficiencies binned in momentum, polar angle, and azimuthal angle. We fold this in with the momentum and  $\theta$  distributions from signal Monte Carlo simulation. We find an  $\mathcal{A}_{CP}$  shift of  $-0.38\%$  for  $K^\pm$  and  $-0.02\%$  for  $\pi^\pm$ .

### 6.6.1.3 Total $\mathcal{A}_{CP}$ Shift and Systematic

Adding together the two results obtained above, we find total  $\mathcal{A}_{CP}$  shifts of  $-0.66\%$  for  $K^{*0} \rightarrow K^+\pi^-$ ,  $-0.71\%$  for  $K^{*+} \rightarrow K^+\pi^0$  and  $-0.05\%$  for  $K^{*+} \rightarrow K_S\pi^+$ . We apply shifts of  $+0.7\%$  to the data for the first two modes and apply no shift to the third mode. In addition, we adopt a systematic equal to the largest total shift ( $0.7\%$ ) for all three modes.

## 6.7 Results from the Maximum Likelihood Fit

As described in Sections 6.4.1 and 6.5, we understand the biases from the maximum likelihood fit, in addition to data/MC differences. Thus, we are able to correct and assign systematic errors to them. We also able to parameterize the detector-induced  $CP$

asymmetry. Therefore, we perform a fit to the on-peak data, which is  $347 \text{ fb}^{-1}$ , to extract the branching fractions and  $CP$  asymmetries.

## 6.7.1 Branching Fractions

### 6.7.1.1 Fit Projections

The fit projections for each dimension are shown in Figures 6.19 through 6.22, while the extracted yields are shown in Tables 6.19 through 6.22. From these fits, we obtain the individual branching fractions according to Eq. 6.8. The branching fractions are  $4.45 \pm 0.10 \pm 0.17 \times 10^{-5}$  ( $K^{*0} \rightarrow K^+ \pi^-$ ),  $4.66 \pm 0.37 \pm 0.35 \times 10^{-5}$  ( $K^{*0} \rightarrow K_S \pi^0$ ),  $4.38 \pm 0.19 \pm 0.26 \times 10^{-5}$  ( $K^{*+} \rightarrow K^+ \pi^0$ ), and  $4.13 \pm 0.18 \pm 0.16 \times 10^{-5}$  ( $K^{*+} \rightarrow K_S \pi^+$ ). The results are displayed in Table 6.23.

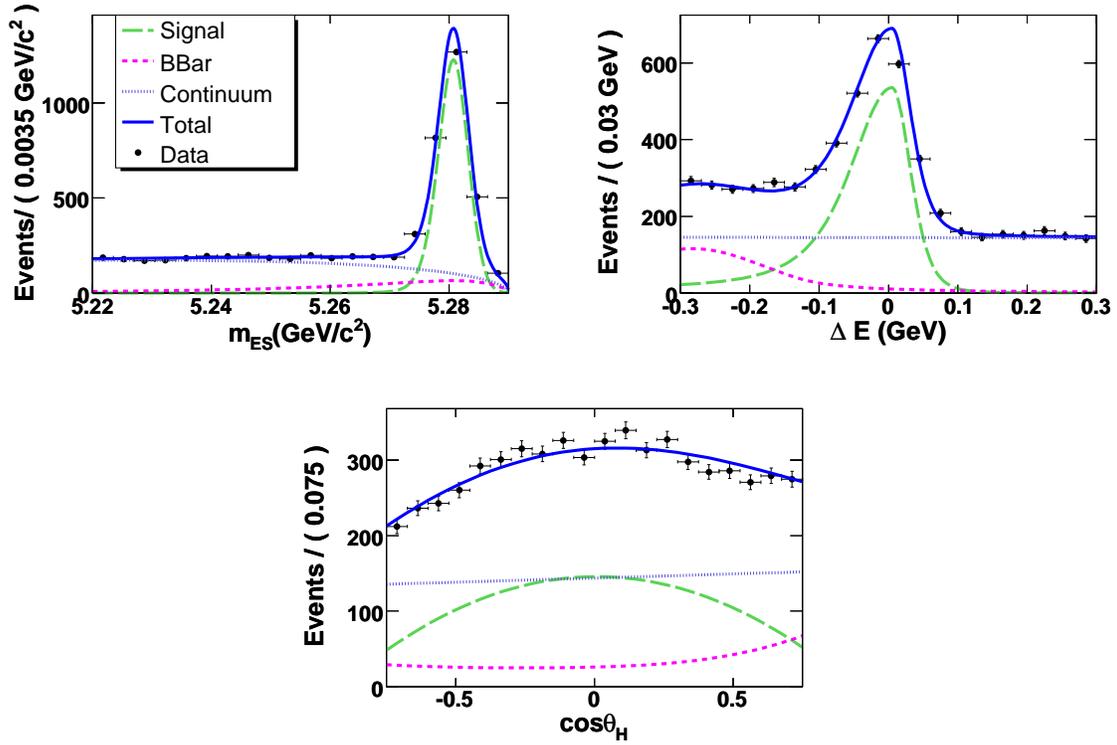


Figure 6.19:  $K^{*0} \rightarrow K^+\pi^-$  projection plots of the full fit to on-peak data. For each projection, the signal region selections ( $5.27 < m_{ES} < 5.29$ ,  $-0.2 < \Delta E < 0.1$ ) have been applied, except for the distribution shown.

Table 6.19: Component yields for the  $K^{*0} \rightarrow K^+\pi^-$  mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure 6.19.

Yield	Value
$N_{sig}$	$2400.0^{+55.4}_{-54.9}$
$N_{cont}$	$3071.3^{+87.7}_{-86.3}$
$N_{B\bar{B}}$	$563.0^{+70.1}_{-69.9}$

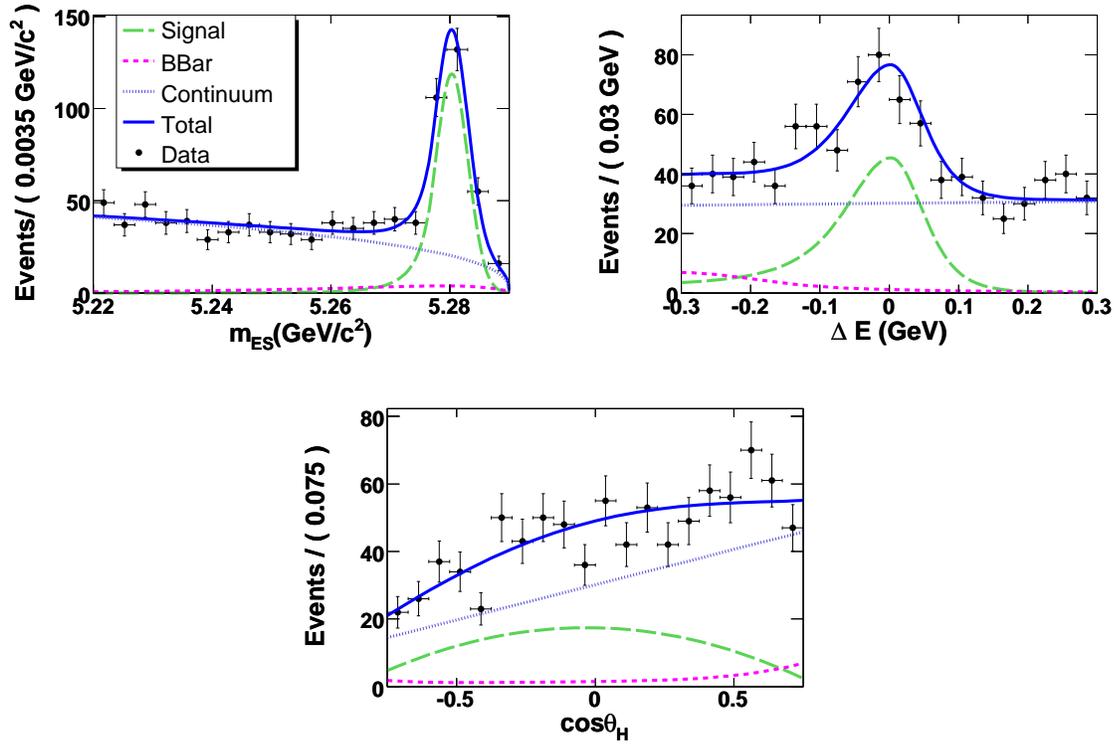


Figure 6.20:  $K^{*0} \rightarrow K_S \pi^0$  projection plots of the full fit to the on-peak data. For each projection, the signal region selections ( $5.27 < m_{ES} < 5.29$ ,  $-0.2 < \Delta E < 0.1$ ) have been applied, except for the distribution shown.

Table 6.20: Component yields for the  $K^{*0} \rightarrow K_S \pi^0$  mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure 6.20.

Yield	Value
$N_{sig}$	$256.0^{+20.6}_{-19.8}$
$N_{cont}$	$603.1^{+43.0}_{-41.1}$
$N_{B\bar{B}}$	$42.9^{+32.9}_{-32.2}$

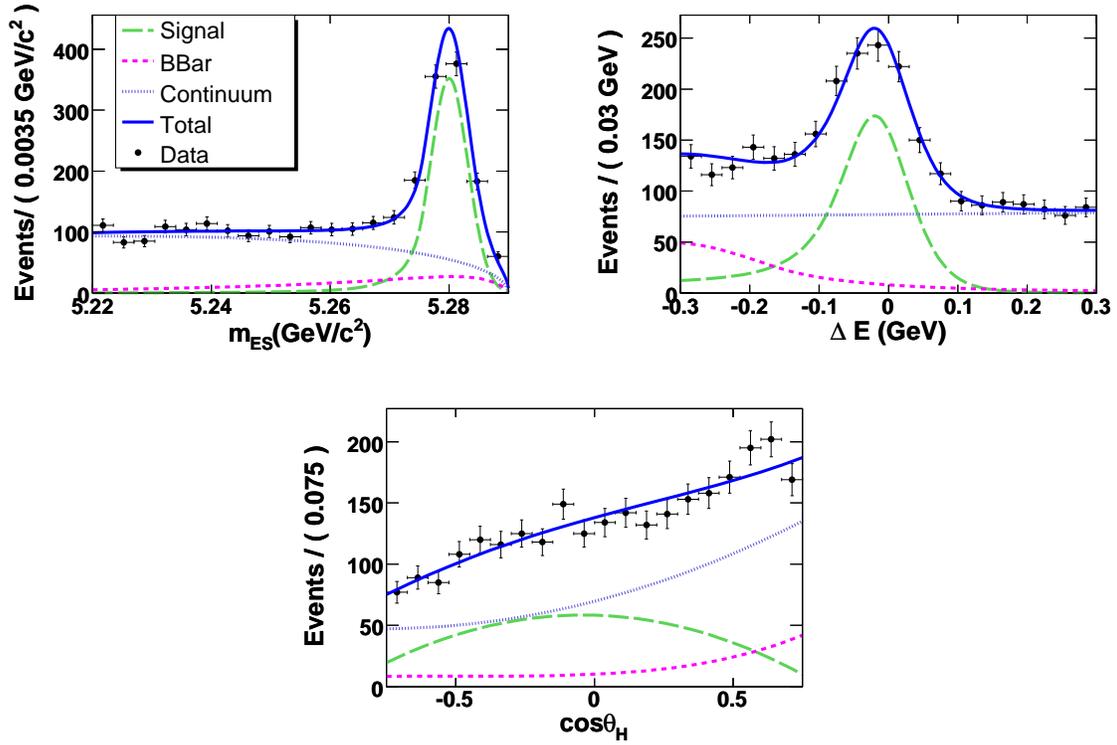


Figure 6.21:  $K^{*+} \rightarrow K^+\pi^0$  projection plots of the full fit to the on-peak data. For each projection, the signal region selections ( $5.27 < m_{ES} < 5.29$ ,  $-0.2 < \Delta E < 0.1$ ) have been applied, except for the distribution shown.

Table 6.21: Component yields for the  $K^{*+} \rightarrow K^+\pi^0$  mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure 6.21.

Yield	Value
$N_{sig}$	$872.7^{+37.6}_{-36.7}$
$N_{cont}$	$1538.3^{+78.4}_{-75.9}$
$N_{B\bar{B}}$	$298.0^{+67.6}_{-67.7}$

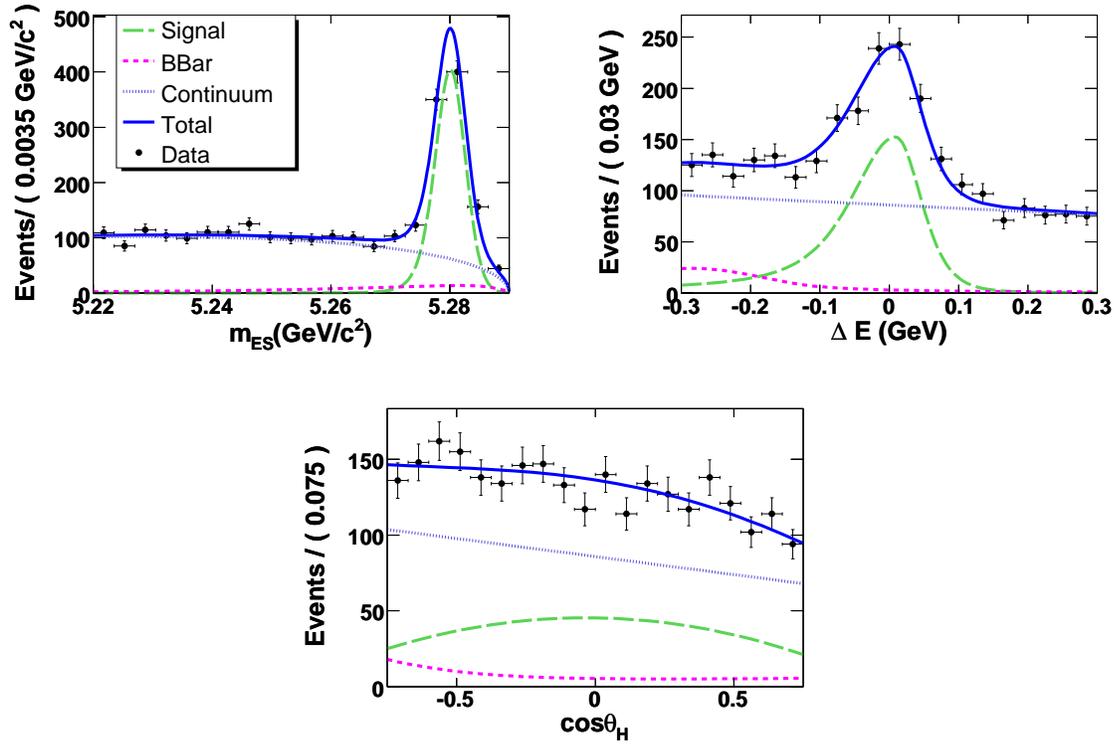


Figure 6.22:  $K^{*+} \rightarrow K_S \pi^+$  projection plots of the full fit to the on-peak data. For each projection, the signal region selections ( $5.27 < m_{ES} < 5.29$ ,  $-0.2 < \Delta E < 0.1$ ) have been applied, except for the distribution shown.

Table 6.22: Component yields for the  $K^{*+} \rightarrow K_S \pi^+$  mode from the full fit to the on-peak data. The projection plots for each dimension are shown in figure 6.22.

Yield	Value
$N_{sig}$	$759.1^{+33.8}_{-32.9}$
$N_{cont}$	$1716.5^{+67.8}_{-66.2}$
$N_{B\bar{B}}$	$141.1^{+52.1}_{-51.1}$

Table 6.23: The signal yields from the fit to the on-peak data, signal efficiencies and total systematic uncertainty, isospin factors, and branching fractions.

Mode	Yield	Efficiency ( $\sigma_{sys}$ )	Isospin factor	BF $\pm$ stat $\pm$ sys ( $\times 10^{-5}$ )
$K^{*0} \rightarrow K^+\pi^-$	$2400.0 \pm 55.4$	0.218 (3.6%)	0.667	$4.45 \pm 0.10 \pm 0.17$
$K^{*0} \rightarrow K_S\pi^0$	$256.0 \pm 20.6$	0.130 (7.1%)	0.114	$4.66 \pm 0.37 \pm 0.35$
$K^{*+} \rightarrow K^+\pi^0$	$872.7 \pm 37.6$	0.153 (5.3%)	0.329	$4.38 \pm 0.19 \pm 0.26$
$K^{*+} \rightarrow K_S\pi^+$	$759.1 \pm 33.8$	0.201 (3.7%)	0.231	$4.13 \pm 0.18 \pm 0.16$

### 6.7.1.2 Comparison of Data and Monte Carlo simulation

After performing a fit to the on-peak data, we compare the results to MC simulation two different ways, which are described below. This is to ensure that we understand the parameters and yields obtained from the on-peak data fit.

- Negative Log Likelihood (NLL) comparison: In this method of comparison, we fix the parameter values of each PDF to the ones obtained from the fit to on-peak data. Then, we randomly draw events from the signal MC simulation, while obtaining the events from the other two components from the PDFs. We do not fit for these events, but instead form the log likelihood, and repeat the procedure one thousand times. The results are show in Figure 6.23, where red line indicates the log likelihood from the on-peak fit. A large difference between the onpeak NLL and the mean of the distribution would indicate a large disagreement between the onpeak data fit and the signal MC simulation.
- Parameter comparison: Tables E.1 and E.2 in Appendix E compare the parameters obtained the fit to onpeak data to those obtained from the full fit to Monte Carlo simulation. Even though there is never more than a  $3\sigma$  difference, the continuum yield is consistently higher than the expected continuum yield. This is expected from the studies done with off-peak data, shown in Table 6.1. In addition the,  $B\bar{B}$  yield is consistently lower than the expected value, though this difference never more than  $3\sigma$ .

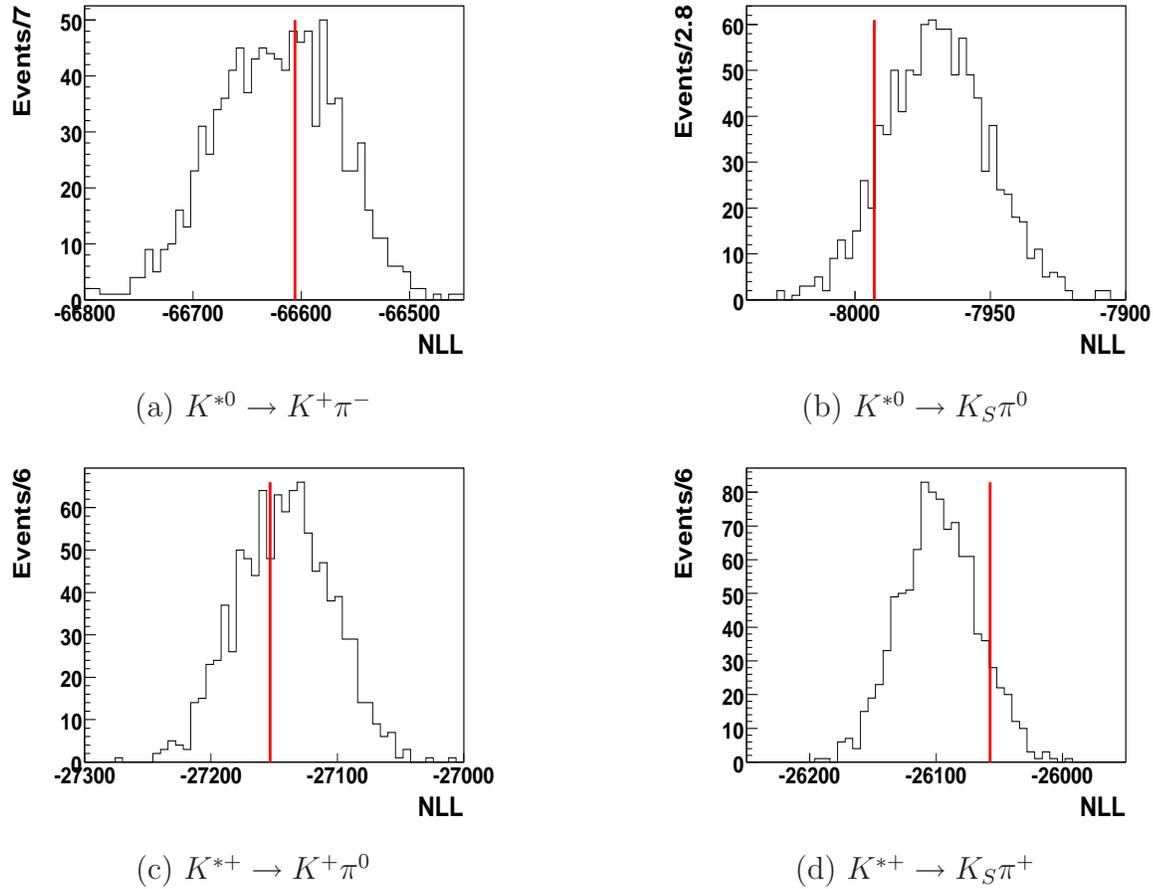


Figure 6.23: Negative maximum likelihood distributions of a signal embedded studies for a)  $K^{*0} \rightarrow K^+ \pi^-$ , b)  $K^{*0} \rightarrow K_S \pi^0$ , c)  $K^{*+} \rightarrow K^+ \pi^0$ , d)  $K^{*+} \rightarrow K_S \pi^+$ , obtained as described in the text. The negative maximum likelihoods from the on-peak data fit are marked by a red line. There are one thousand entries in each plot.

### 6.7.1.3 $K^*$ line shape

To further validate that we have selected true  $B \rightarrow K^*\gamma$  events, we fit a relativistic P-wave Breit-Wigner to the  $m_{K\pi}$  distribution. To form the  $m_{K\pi}$  distribution, we first widen the  $m_{K\pi}$  mass selection to be  $[0.7,1.1]$ , and fit to the on-peak dataset. Finally, we make an sPlot [68] of the  $m_{K\pi}$  distribution. The results of the fit for each mode are shown in Figure 6.24. For the  $K^{*0} \rightarrow K_S\pi^0$  and  $K^{*+} \rightarrow K^+\pi^0$  modes, we convolve the Breit-Wigner line shape with a Gaussian with a width of 10 MeV (determined from MC simulation) to account for detector resolution. For the  $K^{*0} \rightarrow K^+\pi^-$  and the  $K^{*+} \rightarrow K_S\pi^+$  modes, the detector resolution is negligible. The results are consistent with the signal events containing only P-wave  $K^*$  mesons and no other  $K\pi$  resonances. We estimate the contribution from the  $K^*(1430)$  to the invariant mass regions  $m_{K^+\pi^-}$ ,  $m_{K^+\pi^0}$ , and  $m_{K_S\pi^+}$  defined in Section 5.3.4 by using the measured values of the branching fractions of  $B^0 \rightarrow K^{*0}(1430)\gamma$  and  $B^+ \rightarrow K^{*+}(1430)\gamma$  [69]. We find that the contribution is  $\sim 1$  event or less.

A comparison of the fitted values of the mass and width to the PDG [4] values are shown in Table 6.24. For the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K^+\pi^0$  modes, both the mass and width are compatible with the PDG values. However, for the  $K^{*0} \rightarrow K_S\pi^0$  mode, the width is about 10 MeV below the PDG value.

Table 6.24: The fitted results of the  $m_{K\pi}$  spectrum are compared with PDG [4] values for all four decay modes.

Mode	Data		PDG Value	
	m (MeV)	$\Gamma$ (MeV)	m (MeV)	$\Gamma$ (MeV)
$K^{*0} \rightarrow K^+\pi^-$	$895.33 \pm .69$	$50.3 \pm 1.6$	$896.00 \pm .25$	$50.3 \pm 0.6$
$K^{*0} \rightarrow K_S\pi^0$	$894.4 \pm 1.8$	$34.3 \pm 3.6$	$896.00 \pm .25$	$50.3 \pm 0.6$
$K^{*+} \rightarrow K^+\pi^0$	$892.1 \pm 1.2$	$46.5 \pm 2.6$	$891.66 \pm .26$	$50.8 \pm 0.9$
$K^{*+} \rightarrow K_S\pi^+$	$894.9 \pm 1.1$	$44.4 \pm 2.4$	$891.66 \pm .26$	$50.8 \pm 0.9$

## 6.7.2 CP Asymmetry

We also fit for  $\mathcal{A}_{CP}$  to the on-peak dataset for each mode. After applying the shift discussed in Section 6.6.1.3, the results are  $-0.016 \pm 0.022 \pm 0.007$  ( $K^{*0} \rightarrow K^+\pi^-$ ),  $0.040 \pm 0.039 \pm 0.007$  ( $K^{*+} \rightarrow K^+\pi^0$ ), and  $-0.006 \pm 0.041 \pm 0.007$  ( $K^{*+} \rightarrow K_S\pi^+$ ). They are listed in Table 6.25.

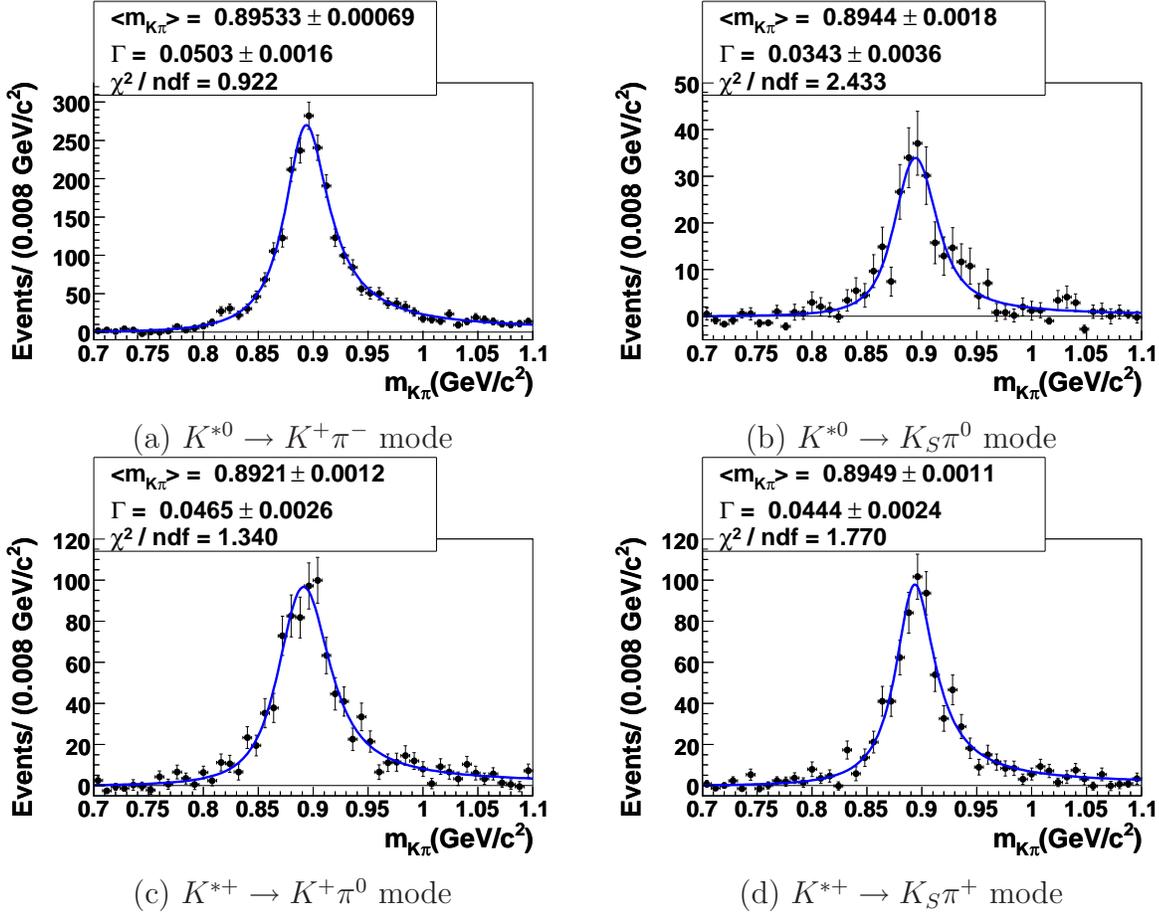


Figure 6.24: Relativistic P-wave Breit-Wigner line shape fit to the  $K\pi$  invariant mass distribution of the sPlot of the on-peak dataset for the a)  $K^{*0} \rightarrow K^+\pi^-$ , b)  $K^{*0} \rightarrow K_S\pi^0$ , c)  $K^{*+} \rightarrow K^+\pi^0$ , d)  $K^{*+} \rightarrow K_S\pi^+$  modes.

Table 6.25: The individual  $\mathcal{A}_{CP}$  values from the fit to the on-peak data.

Mode	$\mathcal{A}_{CP}$
$K^{*0} \rightarrow K^+\pi^-$	$-0.016 \pm 0.022 \pm 0.007$
$K^{*+} \rightarrow K^+\pi^0$	$0.040 \pm 0.039 \pm 0.007$
$K^{*+} \rightarrow K_S\pi^+$	$-0.006 \pm 0.041 \pm 0.007$

## 6.8 Combined Results

After fitting to the on-peak data for the branching fractions and the  $CP$  asymmetry in each mode, we now want to combine the individual results to form the branching fractions  $B^0 \rightarrow K^{*0}\gamma$  and  $B^+ \rightarrow K^{*+}\gamma$ , and also a combined  $CP$  asymmetry. In addition, from the combined branching fractions, we calculate the isospin asymmetry defined by Eq. 4.32. Toward this end, we construct a four by four covariance matrix constructed from the uncertainties on the branching fractions in which the diagonal elements correspond to the total errors of each branching fraction, while the off-diagonal elements are from correlated systematic errors. This method is described in more detail below.

### 6.8.1 Covariance Matrix

The covariance matrix of the four modes is constructed by examining which systematic uncertainties are common between any two modes. Each systematic contribution is calculated using the definition of correlation coefficient

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y} \quad (6.14)$$

where  $x$  and  $y$  each correspond to a systematic uncertainty. In the case in which  $x$  and  $y$  are fully correlated with each other, this reduces to  $V_{xy} = \sigma_x \sigma_y$ . Each matrix element is then determined by adding each systematic contribution together in quadrature. Table 6.26 lists the common systematic uncertainties. An error is considered correlated if it is determined by the same control sample. The calculated covariance matrix is shown in Table 6.27.

### 6.8.2 Branching Fractions

To extract a combined branching fraction for the neutral and charged modes, we minimize a  $\chi^2$  defined by [70]

$$\chi^2(\mathcal{B}(B \rightarrow K^*\gamma)) = \sum_{i,j=1}^N (y_i - \mathcal{B}(B \rightarrow K^*\gamma))(V^{-1})_{ij}(y_j - \mathcal{B}(B \rightarrow K^*\gamma)) \quad (6.15)$$

where the indices  $i, j$  label an individual mode,  $y_{i,j}$  corresponds to an individual branching fraction, and  $V^{-1}$  is the inverse of the covariance matrix. By minimizing the  $\chi^2$ , a weight

Table 6.26: The correlated systematic uncertainties of the four modes. “Common 5” refers to the five systematics that are common to all modes (B Counting, Photon Efficiency, Photon Quality,  $\pi^0/\eta$  veto, and production uncertainty of  $\Upsilon(4S) \rightarrow B\bar{B}$ ). The diagonal entries correspond to the total uncertainty for each mode. The matrix is symmetric, so all entries below the diagonal are omitted

Mode	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$	$K^{*+} \rightarrow K^+\pi^0$	$K^{*+} \rightarrow K_S\pi^+$
$K^{*0} \rightarrow K^+\pi^-$	-	Common 5	Common 5 Tracking Particle ID	Common 5 Tracking Particle ID
$K^{*0} \rightarrow K_S\pi^0$	-	-	Common 5 $\pi^0$	Common 5 $K_s$
$K^{*+} \rightarrow K^+\pi^0$	-	-	-	Common 5 Tracking
$K^{*+} \rightarrow K_S\pi^+$	-	-	-	-

Table 6.27: The calculated covariance matrix for all four modes.

	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$	$K^{*+} \rightarrow K^+\pi^0$	$K^{*+} \rightarrow K_S\pi^+$
$K^{*0} \rightarrow K^+\pi^-$	$3.84 \times 10^{-12}$	$1.78 \times 10^{-12}$	$1.31 \times 10^{-12}$	$1.22 \times 10^{-12}$
$K^{*0} \rightarrow K_S\pi^0$	$1.77 \times 10^{-12}$	$2.60 \times 10^{-11}$	$3.00 \times 10^{-12}$	$1.19 \times 10^{-12}$
$K^{*+} \rightarrow K^+\pi^0$	$1.32 \times 10^{-12}$	$3.00 \times 10^{-12}$	$0.96 \times 10^{-11}$	$1.64 \times 10^{-12}$
$K^{*+} \rightarrow K_S\pi^+$	$1.23 \times 10^{-12}$	$1.19 \times 10^{-12}$	$1.64 \times 10^{-12}$	$6.10 \times 10^{-12}$

can be defined

$$w_i = \frac{\sum_{j=1}^N (V^{-1})_{ij}}{\sum_{k,l=1}^N (V^{-1})_{kl}} \quad (6.16)$$

where the index  $i$  corresponds to an individual mode, and the sum is over the  $N$  elements of the inverse of the covariance matrix. Table 6.28 lists the results. The combined branching fractions are then extracted from

$$\mathcal{B}(B \rightarrow K^* \gamma) = \sum_{i=1}^N w_i y_i \quad (6.17)$$

The variance of the combined branching fraction is determined as

$$V[(B \rightarrow K^* \gamma)] = \sum_{i,k=1}^N w_i V_{ij} w_j. \quad (6.18)$$

The results are

$$\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = (4.47 \pm 0.10 \pm 0.16) \times 10^{-5}$$

$$\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = (4.22 \pm 0.14 \pm 0.16) \times 10^{-5}.$$

Table 6.28: The weights of the individual modes.

Mode	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$	$K^{*+} \rightarrow K^+\pi^0$	$K^{*+} \rightarrow K_S\pi^+$
Weight	0.9219	0.0781	0.3557	0.6443

### 6.8.3 Isospin Asymmetry

Using the combined branching fractions and the covariance matrix, we obtain the isospin asymmetry defined in Eq. 4.32. We rewrite the asymmetry as

$$\Delta_{0-} = \frac{I \frac{\tau^+}{\tau^0} - 1}{I \frac{\tau^+}{\tau^0} + 1},$$

where  $\tau^+$  is  $B^+$  lifetime,  $\tau^0$  is the  $B^0$  lifetime, and  $I$  is

$$I = \frac{\mathcal{B}(B^0 \rightarrow K^{*0}\gamma)}{\mathcal{B}(B^{*-} \rightarrow K^{*-}\gamma)}.$$

Under the assumption that  $I$  and  $\frac{\tau^+}{\tau^0}$  are approximately 1, we have,

$$\Delta_{0-} = \frac{1}{2} \left( I \frac{\tau^+}{\tau^0} - 1 \right).$$

Inserting the combined branching fractions, we find, taking into account correlated errors,

$$I = 1.016 \pm 0.040 \pm 0.044.$$

After accounting for  $\tau_+/\tau_0 = 1.071 \pm 0.009$  [4], we have

$$\Delta_{0-} = 0.066 \pm 0.021 \pm 0.022.$$

The 90% confidence interval for  $\Delta_{0-}$  is

$$0.017 < \Delta_{0-} < 0.116.$$

### 6.8.4 $CP$ Asymmetry

The systematic error as quoted in Section 6.6.1 is the same for each mode and fully correlated between modes. We therefore take a weighted average using only statistical errors to calculate the combined  $\mathcal{A}_{CP}$  and find  $\mathcal{A}_{CP} = -0.003 \pm 0.017 \pm 0.007$ . We also set a 90% confidence level by combining the statistical and systematic errors in quadrature and obtain  $-0.033 < \mathcal{A}_{CP} < 0.028$ . We also calculate the weighted average of the two  $K^{*\pm}$  modes and find  $\mathcal{A}_{CP}(B^\pm \rightarrow K^\pm \gamma) = 0.018 \pm 0.028 \pm 0.007$ .

# Chapter 7

## Measurements of $B^+ \rightarrow \rho^+ \gamma$ , $B^0 \rightarrow \rho^0 \gamma$ , and $B \rightarrow \omega \gamma$ Decays

This chapter will outline the measurement of the decays  $B \rightarrow (\rho/\omega)\gamma$ . To first order, these decays are suppressed relative to  $B \rightarrow K^*\gamma$  decays by a factor of  $|V_{td}/V_{ts}|^2 \sim 0.04$ . In addition, the  $\rho$  mass has a width ( $\sim 150$  MeV) that is three times greater than the width of the  $K^*$  ( $\sim 50$  MeV). Both of these facts lead to a dramatically lower signal-to-background ratio than the  $B \rightarrow K^*\gamma$  analysis. It was for this reason that the neural network described in Section 5.4 for the  $B \rightarrow (\rho/\omega)\gamma$  analysis was designed to be more statistically sensitive than the  $B \rightarrow K^*\gamma$  analysis. The maximum likelihood fit is also more sensitive, which will be described in this chapter.

The background composition is also different. The continuum and  $B \rightarrow X_s \gamma$  background are still prominent; however,  $B \rightarrow K^*\gamma$  events also play a significant role, despite their suppression by the DIRC kaon/pion separation abilities. The  $B \rightarrow (\rho/\omega)\gamma$  analysis deals with the difference in background composition in the maximum likelihood fit. This will be described below.

### 7.1 Fit Overview

In Chapter 5, a series of selection criteria for the  $B \rightarrow (\rho/\omega)\gamma$  analysis was given. The optimized selections are displayed in Tables 5.9 through 5.11. However, as mentioned in Section 5.5, we relax the limit on  $\cos\theta_H$  and the neural network output because these variables are used in the fit. We include the neural network output as a fit dimension to increase the sensitivity. For the  $B \rightarrow \omega\gamma$  mode, we include as a fit dimension the cosine

of the Dalitz angle  $\cos\theta_D$ , which is defined as the angle between the  $\pi^+$  and  $\pi^0$  in the  $\pi^+\pi^-$  rest frame (Sec. 4.3.1.2).

Table 7.1: Number of Monte Carlo events in the fit region scaled to  $316 \text{ fb}^{-1}$  using the MC luminosities for the  $B \rightarrow (\rho/\omega)\gamma$  analysis.

	$B^0 \rightarrow \rho^0\gamma$	$B^+ \rightarrow \rho^+\gamma$	$B \rightarrow \omega\gamma$
Signal	$22 \pm 1$	$36 \pm 1$	$13 \pm 1$
Continuum + $\tau^+\tau^-$	$1658 \pm 35$	$1852 \pm 34$	$599 \pm 19$
$B\bar{B}$	$54 \pm 3$	$150 \pm 6$	$21 \pm 2$

Table 7.1 gives the number of events according to MC simulation in the fit region after all the selection criteria have been applied. The fit region is defined as  $m_{\text{ES}} > 5.22 \text{ GeV}/c^2$ ,  $|\Delta E| < 0.3 \text{ GeV}$ , and  $\cos\theta_H < 0.75$ . In addition, the neural network output is constrained to be greater than 0.85, 0.80, and 0.90 for the  $B^0 \rightarrow \rho^0\gamma$ ,  $B^+ \rightarrow \rho^+\gamma$  and  $B \rightarrow \omega\gamma$  modes respectively. The numbers in Tables 7.1 can be compared to the number of events in the  $B \rightarrow K^*\gamma$  analysis, which is given in Table 6.1. For the  $B^0 \rightarrow \rho^0\gamma$  and  $B^+ \rightarrow \rho^+\gamma$  modes, the  $B \rightarrow K^*\gamma$  modes comprises  $\sim 20\%$  of the  $B\bar{B}$  background. To extract the signal yield, we perform a maximum likelihood fit to each of the three modes individually. The likelihood function is defined as in Eq. 6.5. The number of dimensions and components differ for each mode. The  $B\bar{B}$  background is further segregated into individual contributions, which differ for each mode. Whether or not the normalizations of these  $B\bar{B}$  components are fixed or free in the fit also differs for each mode. The signal and  $B\bar{B}$  shapes are fixed in the fit, while the parameters related to the continuum background are free. Differences in the  $\Delta E$  PDF between data and MC simulation are corrected by using samples of  $B \rightarrow K^*\gamma$  decays. In order to fit to the neural network output, we make a transformation as

$$\text{Transform} = \tanh^{-1}(c_1 \cdot \text{Output} - c_2), \quad (7.1)$$

where ‘‘Transform’’ is the transformed neural network output, ‘‘Output’’ is the neural network output, and  $c_1$ ,  $c_2$  are constants. The constants are determined by requiring ‘‘Output’’ to lie in the arbitrary range  $[-4.0, 8.7522]$ . The descriptions of each fit are given in the following sections.

Table 7.2:  $B\bar{B}$  background components used for the  $B^0 \rightarrow \rho^0\gamma$  fit.

	$B\bar{B}$ (excluding $B \rightarrow K^*\gamma$ )	$B^0 \rightarrow K^{*0}\gamma$	$B^+ \rightarrow K^{*+}\gamma$
Expected yield	$N_{B\bar{B}} = 40.27$	$N_{K^{*0}\gamma} = 9.41$	$N_{K^{*+}\gamma} = 4.65$
Fixed fractional yields	$N_{B\bar{B}}/N_{tot}$	$N_{K^{*0}\gamma}/N_{tot}$	$N_{K^{*+}\gamma}/N_{tot}$
Floated yield	$N_{tot} = N_{B\bar{B}} + N_{K^{*0}\gamma} + N_{K^{*+}\gamma}$		

### 7.1.1 $B^0 \rightarrow \rho^0\gamma$

The  $B^0 \rightarrow K^{*0}\gamma$  background is kinematically almost the same as signal, except that the kaon is incorrectly identified as a pion. Therefore, the  $\Delta E$  distribution is shifted downward with respect to signal. The  $B^+ \rightarrow K^{*+}\gamma$  background has a different distribution in  $\Delta E$ , as well as  $\cos\theta_H$ . Taking into account all of this information, the  $B^0 \rightarrow \rho^0\gamma$  fit contains the following five components:

1. Signal
2. Continuum +  $\tau^+\tau^-$
3.  $B\bar{B}$  (excluding  $B \rightarrow K^*\gamma$ )
4.  $B^0 \rightarrow K^{*0}\gamma$
5.  $B^+ \rightarrow K^{*+}\gamma$

The last three components on the above list have one common yield in the fit model, which is floated. However, the ratios between the three  $B\bar{B}$  components are fixed from MC simulation. To find the systematic uncertainty related to fixing these ratios, we vary the components by  $\sqrt{N_{expected}}$  in the on-peak data fit. This has the effect of shifting the signal yield, which is taken as the systematic uncertainty. The treatment of the individual  $B\bar{B}$  backgrounds is summarized in Table 7.2. The PDF shapes used in the fit are given in Table 7.3. The description of each PDF shape is given in Section 6.3.1.

### 7.1.2 $B^+ \rightarrow \rho^+\gamma$

Similar to the  $B^0 \rightarrow \rho^0\gamma$  mode, the  $B^+ \rightarrow \rho^+\gamma$  mode also suffers from  $B \rightarrow K^*\gamma$  background. This is dealt with in the same manner. The  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  mode is

Table 7.3: A summary of the PDFs used for each component and dimension in the  $B^0 \rightarrow \rho^0\gamma$  fit.

	$\Delta E$	$m_{ES}$	$NN_{transform}$	$\cos\theta_H$
Signal	Cruijff	Crystal Ball	Cruijff	$2^{nd}$ -order polynomial
Continuum + $\tau^+\tau^-$	$2^{nd}$ -order polynomial	ARGUS	Cruijff	$2^{nd}$ -order polynomial
$B\bar{B}$ (excluding $B \rightarrow K^*\gamma$ )	Cruijff	ARGUS + Crystal Ball	Cruijff	$2^{nd}$ -order polynomial
$B^0 \rightarrow K^{*0}\gamma$	Cruijff	Crystal Ball	Cruijff	$2^{nd}$ -order polynomial
$B^+ \rightarrow K^{*+}\gamma$	Cruijff	ARGUS + Gaussian	Cruijff + Gaussian	$2^{nd}$ -order polynomial

kinematically almost the same as signal; the  $\Delta E$  distribution is again shifted downward with respect to signal. This background is treated as one component. The remaining  $B \rightarrow K^*\gamma$  background is again treated as one component because the distribution in  $\Delta E$  and  $\cos\theta_H$  differs from the rest of the  $B\bar{B}$  background. The complete set of components is as follows:

1. Signal
2. Continuum +  $\tau^+\tau^-$
3.  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$
4.  $B \rightarrow K^*\gamma$  (excluding  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$ )
5.  $B \rightarrow X_s\gamma$  (excluding  $B \rightarrow K^*\gamma$ )
6.  $B\bar{B}$  (excluding  $B \rightarrow X_s\gamma$  and  $B \rightarrow K^*\gamma$ )

Unlike the  $B^0 \rightarrow \rho^0\gamma$  mode, the  $B\bar{B}$  plays a prominent role, and so is separated into a different component. The last three components have one common yield in the fit model which is floated. However, like in the previous case, the ratios between the three  $B\bar{B}$  components are fixed from MC simulation and varied in the same manner. The  $B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$  component is fixed to the MC simulation expectation of four events. The yield is varied by 50% in the on-peak data fit to determine the systematic uncertainty. Table 7.4 details the treatment of the  $B\bar{B}$  backgrounds, while Table 7.5 gives the PDF shapes that were used.

Table 7.4:  $B\bar{B}$  background components used for the  $B^+ \rightarrow \rho^+\gamma$  fit.

Type of $B$ -background	other $B\bar{B}$	other $B \rightarrow X_s\gamma$	other $B^0 \rightarrow K^{*0}\gamma$	$B^+ \rightarrow K^{*+}\gamma$ , $K^{*+} \rightarrow K\pi^0$
Expected yield	$N_{B\bar{B}} = 72.40$	$N_{X_s\gamma} = 50.00$	$N_{K^{*0}\gamma} = 23.25$	$N_{K\pi^0\gamma} = 4.48$
Fixed fractional yields	$N_{B\bar{B}}/N_{tot}$	$N_{X_s\gamma}/N_{tot}$	$N_{K^{*0}\gamma}/N_{tot}$	$N_{K\pi^0\gamma}$
Floated yield	$N_{tot} = N_{B\bar{B}} + N_{X_s\gamma} + N_{K^{*0}\gamma}$			

Table 7.5: A summary of the PDFs used for each component and dimension in the  $B^+ \rightarrow \rho^+\gamma$  fit.

	$\Delta E$	$m_{ES}$	$NN_{transform}$	$\cos\theta_H$
Signal	Cruijff	Crystal Ball	Cruijff	$2^{nd}$ -order polynomial
Continuum + $\tau^+\tau^-$	$2^{nd}$ -order polynomial	ARGUS	Cruijff	$4^{th}$ -order polynomial
$B^+ \rightarrow K^{*+}\gamma(K^{*+} \rightarrow K^+\pi^0)$	Gaussian	ARGUS + Crystal Ball	Cruijff	$2^{nd}$ -order polynomial
$B \rightarrow K^*\gamma$ (excluding $B^+ \rightarrow K^{*+}\gamma$ ( $K^{*+} \rightarrow K^+\pi^0$ ))	Cruijff	Novosibirsk	Cruijff	$2^{nd}$ -order polynomial
$B \rightarrow X_s\gamma$ (excluding $B \rightarrow K^*\gamma$ )	Cruijff	Gaussian + ARGUS	Cruijff	$4^{th}$ -order polynomial
$B\bar{B}$ (excluding $B \rightarrow X_s\gamma$ and $B \rightarrow K^*\gamma$ )	Cruijff	Crystal Ball	Cruijff	Novosibirsk

Table 7.6: A summary of the PDFs used for each component and dimension in the  $B \rightarrow \omega\gamma$  fit.

	Signal	Continuum + $\tau^+\tau^-$	$B\bar{B}$
$\Delta E$	Cruijff	$2^{nd}$ -order polynomial	Gaussian
$m_{ES}$	Crystal Ball	ARGUS	Novosibirsk
$NN_{transform}$	Cruijff	Cruijff	Cruijff
$\cos\theta_H$	$2^{nd}$ -order polynomial	$2^{nd}$ -order polynomial	$2^{nd}$ -order polynomial
$\cos\theta_D$	$2^{nd}$ -order polynomial	$2^{nd}$ -order polynomial	$2^{nd}$ -order polynomial

### 7.1.3 $B \rightarrow \omega\gamma$

The  $B \rightarrow \omega\gamma$  has negligible  $B \rightarrow K^*\gamma$  contamination. The MC simulation expectation is  $\sim 4$  events. Therefore, all of the  $B\bar{B}$  background is combined into one component. The complete set of components is identical to the  $B \rightarrow K^*\gamma$  analysis, and is as follows:

1. Signal
2. Continuum +  $\tau^+\tau^-$
3.  $B\bar{B}$

Table 7.6 gives the PDF shapes that were used.

## 7.2 Systematic Errors

Table 7.7 gives an overview of the contributions to the systematic uncertainties. Similar to the  $B \rightarrow K^*\gamma$  analysis, there are two types of systematic uncertainties. There are uncertainties associated with the signal reconstruction efficiency, in addition to ones pertaining to the the signal extraction. The latter includes the  $B$  counting uncertainty and the uncertainties associated with the shape parameters of the PDFs. Unlike the  $B \rightarrow K^*\gamma$  analysis, there are systematic errors pertaining to modeling of the signal component because these parameters are obtained from MC simulation and are fixed in the on-peak data fit. The  $B$  counting, tracking efficiency, photon selection,  $\pi^0$  reconstruction,  $\pi^0(\eta)$  vetoes, and neural network systematic uncertainties are evaluated in the same way as in Section 6.5. The particle identification systematic is determined from a sample of  $D^{*+} \rightarrow D^0\pi^+(D^0 \rightarrow K^-\pi^+)$  decays. To estimate the uncertainty related to the modeling of the signal and  $B$  background in the MC simulation, we vary the parameters of the

Table 7.7: Fractional systematic errors (in %) of the measured branching fractions.

Descriptions	$B^+ \rightarrow \rho^+\gamma$	$B^0 \rightarrow \rho^0\gamma$	$B \rightarrow \omega\gamma$	$B \rightarrow (\rho^+, \rho^0)\gamma$	$B \rightarrow (\rho, \omega)\gamma$
$B$ Counting	1.1%	1.1%	1.1%	1.1%	1.1%
Tracking Eff.	1.0%	2.0%	2.0%	1.4%	1.5%
Particle Identification	2.0%	4.0%	2.0%	2.9%	2.7%
Photon selection	1.9%	2.6%	1.7%	2.2%	2.1%
$\pi^0$ reconstruction	3.0%	-	3.0%	1.9%	2.5%
$\pi^0(\eta)$ veto	2.8%	2.8%	2.8%	2.8%	2.8%
Neural Network Eff.	1.0%	1.0%	1.0%	1.0%	1.0%
Neural Network shape	0.4%	0.3%	2.3%	0.4%	0.7%
Signal PDF shapes	4.8%	3.3%	2.4%	3.1%	2.6%
$B\bar{B}$ background PDFs	3.9%	2.9%	9.7%	3.2%	3.1%
$BF(\omega \rightarrow \pi^+\pi^-\pi^0)$	-	-	0.8%	-	0.1%
Combined	8.1%	7.4%	11.6%	7.0%	6.9%

PDFs that are fixed in the fit within their errors. The uncertainty related to the choice of a specific functional form for the shape of the neural network distribution is evaluated by using a binned PDF as an alternative description. All relative and absolute normalizations of  $B$  background components that are fixed in the fit are varied within their errors. For all these variations, the corresponding change in the fitted signal yield is taken as a systematic uncertainty.

## 7.3 Results

### 7.3.1 Branching Fractions

We perform a fit to the on-peak data assuming  $\mathcal{B}(\Upsilon(4S) \rightarrow B^0\bar{B}^0) = \mathcal{B}(\Upsilon(4S) \rightarrow B^+B^-) = 0.5$  to extract the branching fractions. The fit projections for each dimension are shown in Figs. 7.1 through 7.3. For  $B \rightarrow \omega\gamma$ , we also compute the 90% C.L. upper limit  $\mathcal{B}(B \rightarrow \omega\gamma) < 0.78 \times 10^{-6}$  using a Bayesian technique. We determine the branching fraction upper limit  $\mathcal{B}_l$  such that

$$\int_0^{\mathcal{B}_l} \mathcal{L} d\mathcal{B} / \int_0^\infty \mathcal{L} d\mathcal{B} = 0.90 \quad (7.2)$$

where  $\mathcal{L}$  is the likelihood function. We assume a flat prior for the branching fraction and

take into account the systematic uncertainty by convoluting the likelihood function with a Gaussian with a width equal to that of the variance due to the systematic uncertainty. We extract a combined branching fraction defined as in Eq. 4.35 by performing a simultaneous fit to all three decay modes using the isospin constraints on the width of the decay modes:  $\Gamma_{B \rightarrow \rho^+\gamma} = 2\Gamma_{B \rightarrow \rho^0\gamma} = 2\Gamma_{B \rightarrow \omega\gamma}$ . The result is

$$\mathcal{B}(B \rightarrow (\rho/\omega)\gamma) = 1.25_{-0.24}^{+0.25} \pm 0.09 \times 10^{-6}. \quad (7.3)$$

The significance of the signal is  $6.4\sigma$ , including systematic uncertainties. If we exclude the  $B^0 \rightarrow \omega\gamma$  mode from the simultaneous fit, we obtain  $\mathcal{B}(B \rightarrow \rho\gamma) = 1.36_{-0.27}^{+0.29} \pm 0.10 \times 10^{-6}$ . The results are shown in Table 7.8.

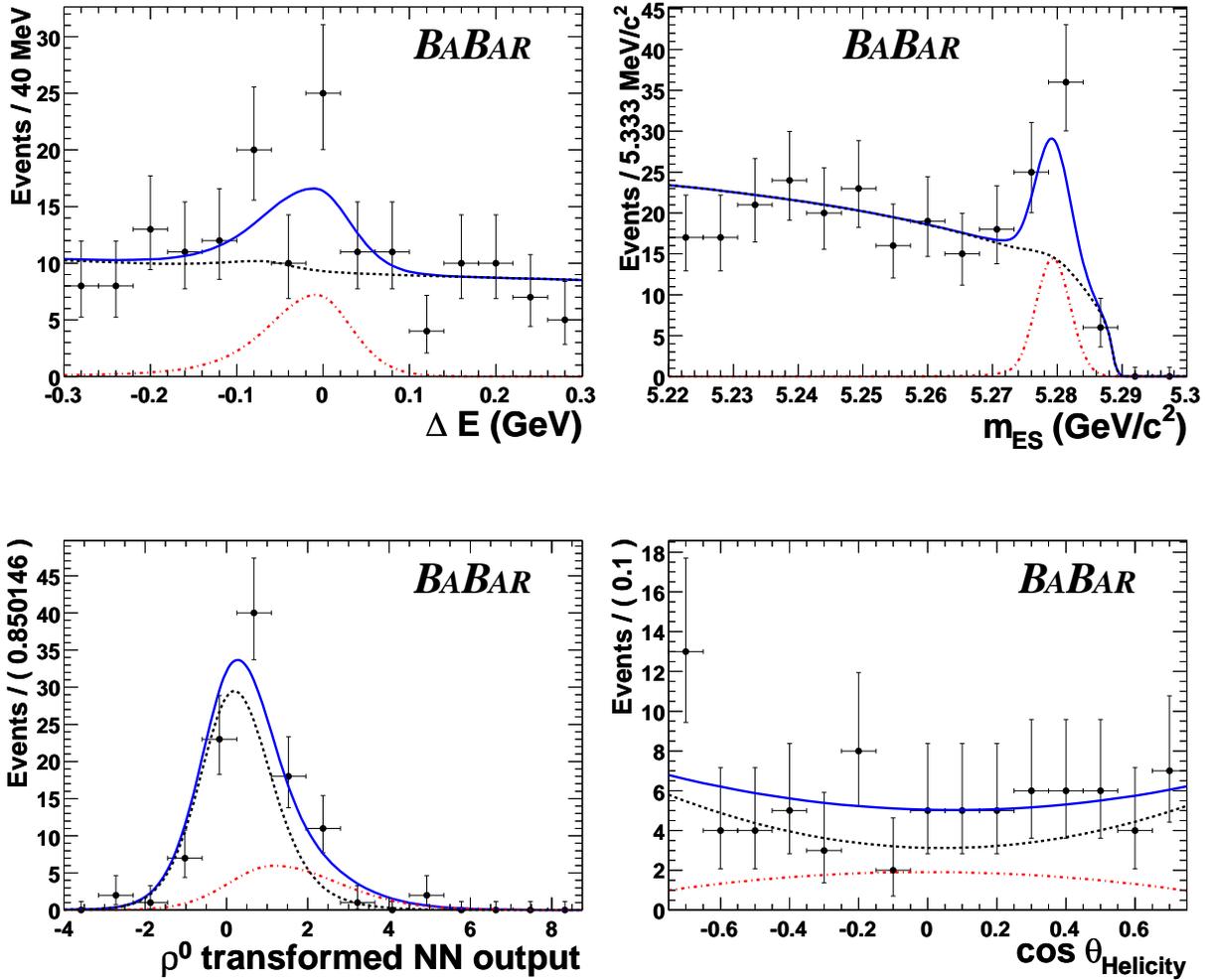


Figure 7.1: Projections for the  $\rho^0$  mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed-dotted curve is the signal component.

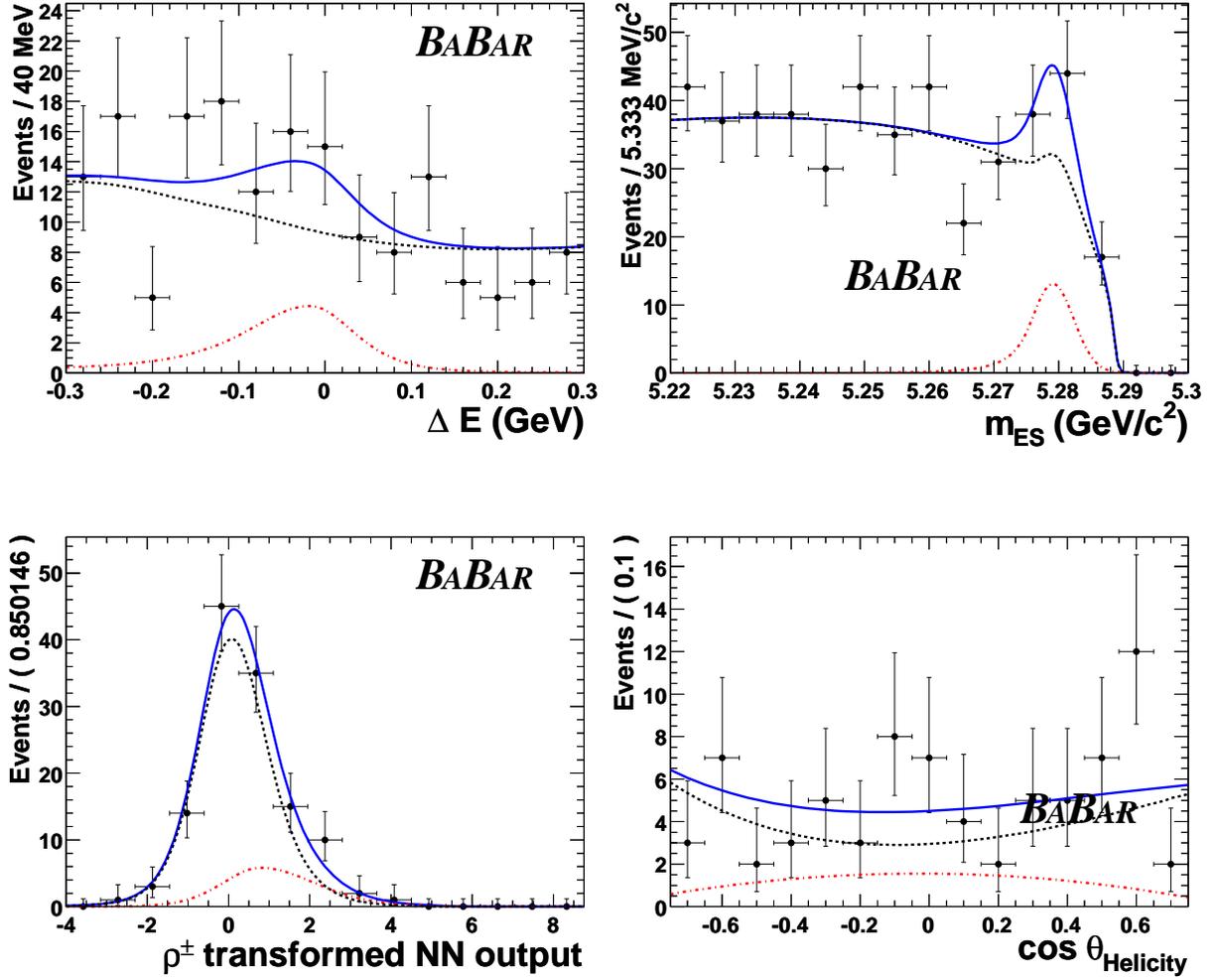


Figure 7.2: Projections for the  $\rho^+$  mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed-dotted curve is the signal component.

### 7.3.2 Isospin Symmetry

We test the hypothesis of isospin symmetry by measuring the quantity defined in Eq. 4.33. We obtain  $\Delta = -0.35 \pm 0.27$ . The result is consistent with the theoretical expectation [32] described in Section 4.4.3.2.

### 7.3.3 $|V_{td}/V_{ts}|$

Using the world average value of  $\mathcal{B}(B \rightarrow K^*\gamma)$  [48], we calculate  $\mathcal{B}(B \rightarrow (\rho/\omega)\gamma)/\mathcal{B}(B \rightarrow K^*\gamma) = 0.030 \pm 0.006$ . This result can be used to calculate the ratio  $|V_{td}/V_{ts}|$  using

Table 7.8: The signal yield ( $n_{\text{sig}}$ ), significance ( $\Sigma$ ) in standard deviations including systematic errors, efficiency ( $\epsilon$ ), and branching fraction ( $\mathcal{B}$ ) for each mode. The errors on  $n_{\text{sig}}$  are statistical only, while for the branching fraction the first error is statistical and the second systematic.

Mode	$n_{\text{sig}}$	$\Sigma$	$\epsilon(\%)$	$\mathcal{B}(10^{-6})$
$B^+ \rightarrow \rho^+\gamma$	$42.0^{+14.0}_{-12.7}$	$3.8\sigma$	11.0	$1.10^{+0.37}_{-0.33} \pm 0.09$
$B^0 \rightarrow \rho^0\gamma$	$38.7^{+10.6}_{-9.8}$	$4.9\sigma$	14.1	$0.79^{+0.22}_{-0.20} \pm 0.06$
$B \rightarrow \omega\gamma$	$11.0^{+6.7}_{-5.6}$	$2.2\sigma$	7.9	$0.40^{+0.24}_{-0.20} \pm 0.05$
$B \rightarrow (\rho/\omega)\gamma$		$6.4\sigma$		$1.25^{+0.25}_{-0.24} \pm 0.09$
$B \rightarrow \rho\gamma$		$6.0\sigma$		$1.36^{+0.29}_{-0.27} \pm 0.10$

Eq. 4.36. We obtain

$$|V_{td}/V_{ts}| = 0.200^{+0.021}_{-0.020} \pm 0.015, \quad (7.4)$$

where the first error is experimental and the second is theoretical. The result is consistent with  $B_d/B_s$  mixing.

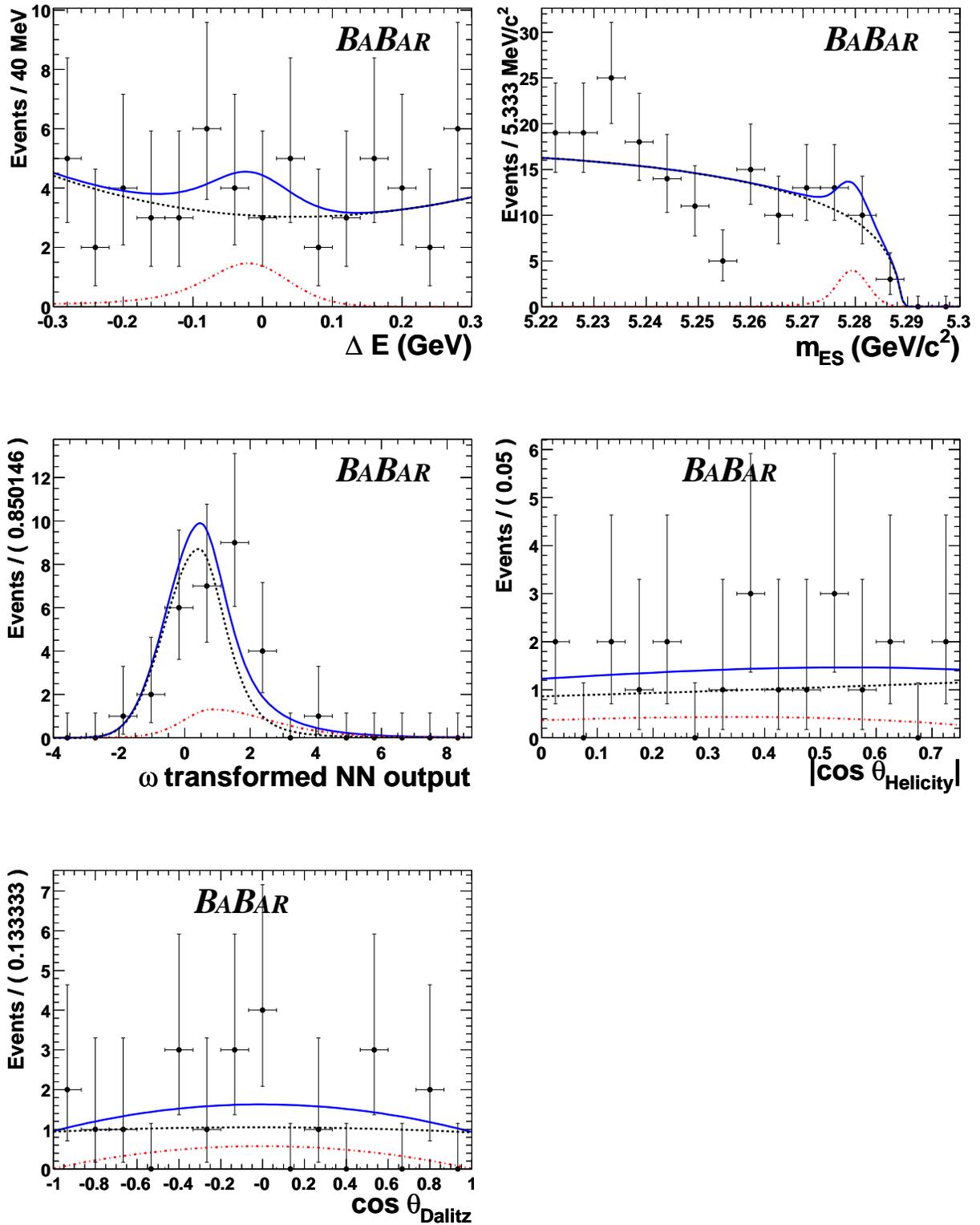


Figure 7.3: Projections for the  $\omega$  mode. The blue curve is the full fit, the black dashed curve are all background components combined and the red dashed-dotted curve is the signal component.

# Chapter 8

## Search for the Rare Decay $B \rightarrow \phi\gamma$

This chapter describes the procedure to measure the branching fraction of  $B \rightarrow \phi\gamma$ . Like the  $B \rightarrow \rho\gamma$  analysis, the signal-to-background ratio is significantly affected by the width of the  $\phi$  mass, as well as the suppression due to the annihilation diagram of the decay. However, it is the small width of the  $\phi$  that leads to a negligible amount of  $B$  background contamination. This observation, along with the rareness of the decay in the SM, leads to a simpler fitting procedure than the  $B \rightarrow K^*\gamma$  and  $B \rightarrow \rho\gamma$  analyses. The reconstruction, neural network, and selection criteria related to  $B \rightarrow \phi\gamma$  has already been described in Chapter 5. The fitting procedure, along with the description of the  $B$  background, is described in this chapter.

### 8.1 B background

As mentioned above, the  $B$  background is significantly lower as compared to the  $B \rightarrow K^*\gamma$  and  $B \rightarrow (\rho/\omega)\gamma$  analyses. According to MC simulation, Table 8.1 lists the modes that are found in the signal region ( $5.27 < m_{ES} < 5.29 \text{ GeV}/c^2$ ,  $-0.2 < \Delta E < 0.1 \text{ GeV}$ ). After scaling to  $113.1 \text{ fb}^{-1}$ , there are  $0.073 \pm 0.004 B\bar{B}$  events in the signal region.

### 8.2 Fitting Procedure

To make a determination of the upper limit on the number of signal events, we make a background estimate in the signal region using real data, and perform a consistency check using MC simulation. We divide the fit region into four separate regions as follows:

**Region 1**  $0.1 \leq \Delta E \leq 0.5 \text{ GeV}$  and  $5.1 \leq m_{ES} \leq 5.29 \text{ GeV}/c^2$

Table 8.1: Origin of  $B\bar{B}$  events that fall in the signal region.

Mode	Number of events in signal region (scaled to $113.1 \text{ fb}^{-1}$ )
$B_d^0 \rightarrow \phi\pi^0$	$0.04 \pm 0.002$
$B_d^0 \rightarrow \phi\eta$	$0.02 \pm 0.0009$
$B_d^0 \rightarrow \phi K_L^0$	$0.003 \pm 0.001$
$B_d^0 \rightarrow \phi K_S^0$ ( $K_S^0 \rightarrow \pi^0\pi^0$ )	$0.01 \pm 0.003$

**Region 2**  $-0.2 \leq \Delta E \leq -0.1 \text{ GeV}$  and  $5.1 \leq m_{ES} \leq 5.27 \text{ GeV}/c^2$

**Region 3**  $-0.5 \leq \Delta E \leq -0.2 \text{ GeV}$  and  $5.1 \leq m_{ES} \leq 5.29 \text{ GeV}/c^2$

**Region 4**  $-0.2 \leq \Delta E \leq 0.1 \text{ GeV}$  and  $5.27 \leq m_{ES} \leq 5.29 \text{ GeV}/c^2$

These regions are illustrated in Fig. 8.1. Regions 1 and 3 will be used to estimate the number of background events in region 4 (signal region), while region 2 will be used to validate the procedure.

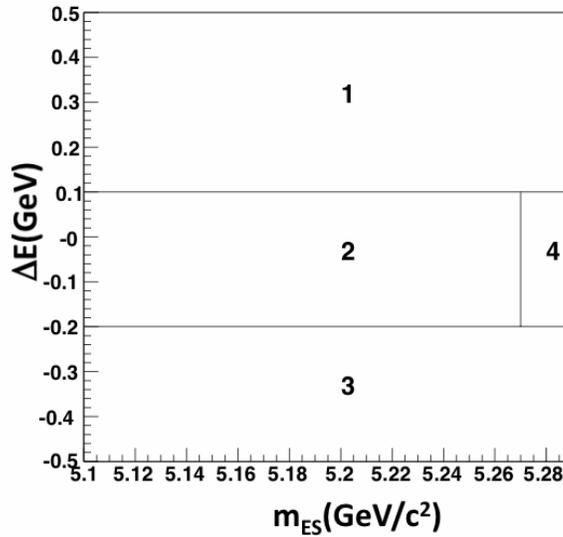


Figure 8.1: The four different regions used in estimating the number of background events in the signal region.

### 8.2.1 Monte Carlo Consistency

We first determine how consistent the background distributions in each region are with each other by fitting an Argus function to regions 1, 3, 1+3, and 2+4. This is shown in Fig. 8.2. The Argus parameters are within  $\approx 1\sigma$  of each other.

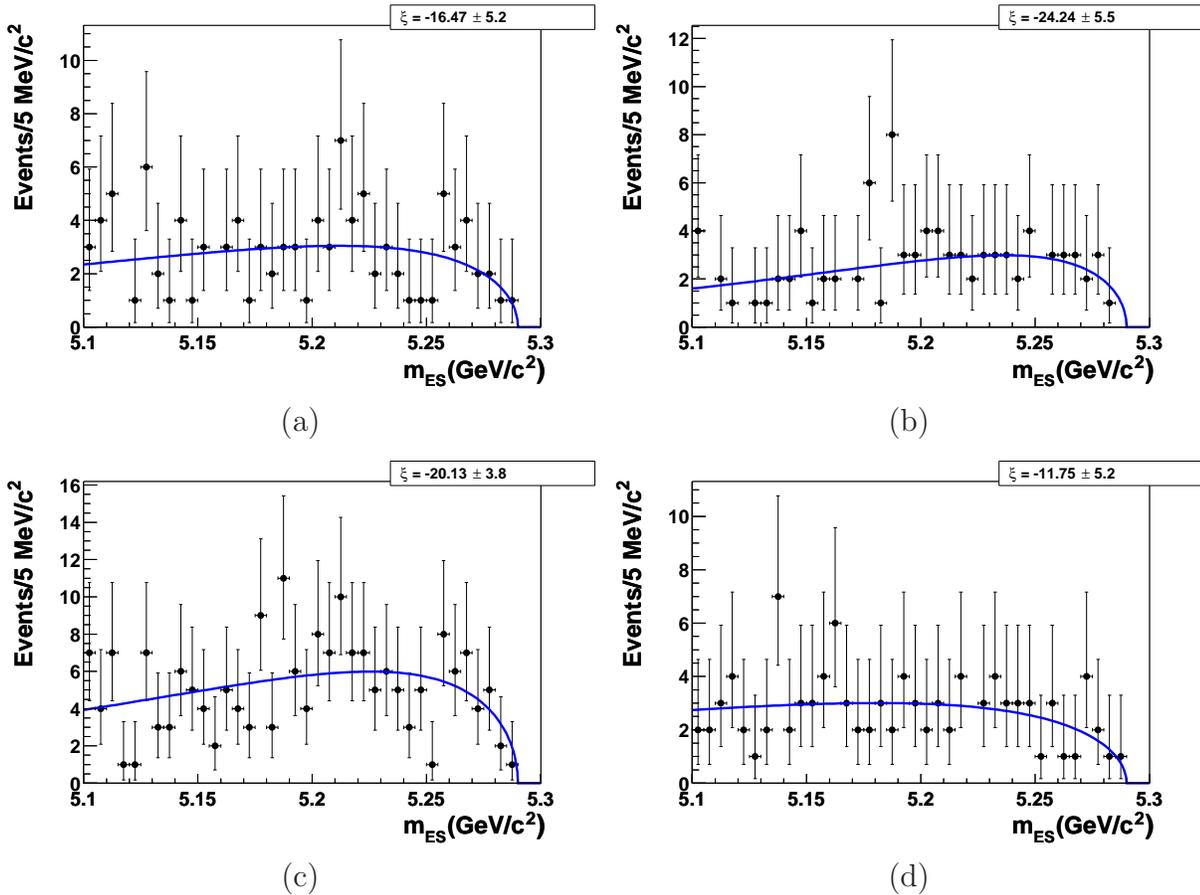


Figure 8.2: Argus fits to the  $m_{ES}$  distributions in a) region 1, b) region 3, c) regions 1 and 3, d) regions 2 and 4. The data set used is continuum MC simulation.

### 8.2.2 Data Consistency and Background Estimation

We validate the procedure of fitting the  $m_{ES}$  and  $\Delta E$  distributions in regions 1 and 3, and then extrapolating into region 4, by first extrapolating into region 2 and comparing with the actual number. We fit an Argus function to the  $m_{ES}$  distribution and a 1<sup>st</sup> order polynomial to the  $\Delta E$  distribution. Fig. 8.3 shows the  $m_{ES}$  and  $\Delta E$  distributions for all regions, with the fits to regions 1 and 3 superimposed.

The fitted number obtained upon extrapolating into region 2 is  $79.4 \pm 6.3$ . The actual

number is  $71 \pm 8.4$ . Thus, the fitting procedure is validated and we use this to estimate the background number in the signal region. The fitted number obtained is  $6.0 \pm 0.8$ . This is consistent with the number obtained by performing the same procedure on MC simulation, which is  $4.7 \pm 0.7$ .

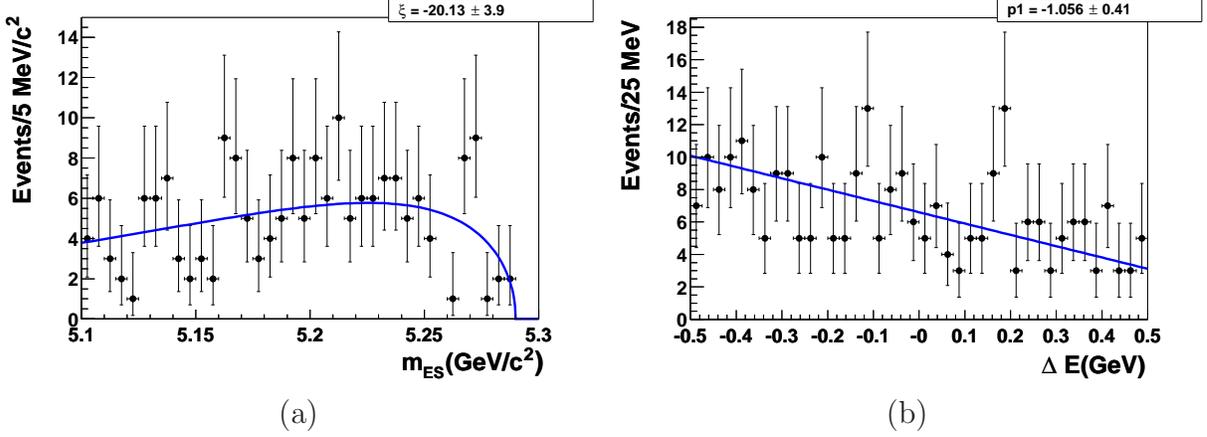


Figure 8.3: The (a)  $m_{ES}$  and (b)  $\Delta E$  distributions of on-resonance data. Superimposed are (a) an Argus function and (b) a 1<sup>st</sup> order polynomial that are obtained from fits to the data in regions 1 and 3. The different regions and fitting method are described in the text.

## 8.3 Systematic Errors

Table 8.2 gives an overview of the contributions to the systematic uncertainties and corrections factors. Similar to the previous two analysis, there are uncertainties associated with the signal reconstruction efficiency and the signal extraction. In this analysis, however, the  $B$  counting systematic is the only uncertainty associated with the signal extraction. The  $B$  counting, tracking efficiency, photon distance, and neural network systematic uncertainties are evaluated in the same manner as in Section 6.5. The particle identification systematic is determined from a sample of  $D^{*+} \rightarrow D^0 \pi^+ (D^0 \rightarrow K^- \pi^+)$  decays. From this sample, the data/MC efficiencies are binned in momentum, polar angle, and azimuthal angle. The systematic uncertainty is determined by varying these bins to account for the different momentum spectra of the  $D^{*+} \rightarrow D^0 \pi^+ (D^0 \rightarrow K^- \pi^+)$  control sample and the signal momentum distribution. The systematic uncertainty is determined in Ref. [71], where a similar kaon momentum spectra is expected, to be 1%. The  $\pi^0/\eta$  systematic uncertainty is determined by “embedding” a signal photon into continuum MC simulation and off-resonance data. This photon is then paired with every other photon

in the event and the veto applied. The discrepancy between the efficiency between data and MC simulation is taken as the systematic uncertainty. This is evaluated in Ref. [66] and the systematic uncertainty is 1%. The photon efficiency systematic uncertainty is determined by a sample of  $e^+e^- \rightarrow \tau^+\tau^-$  events. One  $\tau$  is selected to decay to  $e\nu_e\nu_\tau$ , while the other is required to decay to a charged hadron, a neutrino, and one or two  $\pi^0$ 's. The photon efficiency systematic uncertainty is given as half of the  $\pi^0$  systematic uncertainty, which is determined by the errors on the  $\tau$  branching fractions, as well as the uncertainties mentioned in Section 6.5.7. The systematic is 2.5% with a correction of 0.997.

Table 8.2: This table shows the factors needed to correct the Monte Carlo efficiency and their systematic errors.

Description	Factor	$\sigma$
B Counting	1.00	1.0%
Tracking Eff	0.984	2.6%
PID	1.00	1.0%
Distance Cut	1.00	2.0%
$\pi^0/\eta$ veto	1.00	1.0%
Photon Eff	0.997	2.5%
Neural Network	0.985	2.7%
Signal Efficiency	0.144	0.1%
Total Correction	0.966	5.2%
Corrected Efficiency	0.139	5.2%

## 8.4 Results

The systematic uncertainties described in the previous section are incorporated into upper limit on the branching fraction using the technique of Cousins and Highland [72]. The upper limit is obtained as

$$\mathcal{B}(B \rightarrow \phi\gamma) = \frac{N_{UL}^{90}}{2\epsilon_s \times L \times \sigma_{B^0\bar{B}^0} \times \mathcal{B}(\phi \rightarrow K^+K^-)} (1 + 1.15\sigma_S^2) \quad (8.1)$$

where  $N_{UL}^{90}$  is the upper limit on the expected number of signal events at the 90% confidence interval,  $\epsilon_s$  is the signal efficiency,  $L$  is the luminosity,  $\mathcal{B}(\phi \rightarrow K^+K^-) = 0.491 \pm 0.006$ ,  $\sigma_{B^0\bar{B}^0}$  is the cross-section for  $\Upsilon(4S)$  to decay into  $B^0\bar{B}^0$  pairs, and  $\sigma_S$  is the total systematic uncertainty. The factor of 1.15 in front of  $\sigma_S$  comes from calculating the confidence

interval by averaging over the probability density function of the systematic [72].

The number of events in the signal region is 8 events as displayed in Figure 8.4. With a mean of 6 events, the Poisson consistency of 8 events or more being consistent with background is 0.15. From Table 8.3, which shows 90% confidence intervals given by Ref. [56], one can see that the upper limit of the number of signal events is 7.99. The upper limit on the branching fraction given by Eq. 8.1 is  $9.4 \times 10^{-7}$ . In Fig. 8.4, the density of points seems to be significantly less in the upper right half of the plane. To see whether this is a statistical effect, the  $\Delta E$  slope parameter is determined for  $5.21 < m_{\text{ES}} < 5.25 \text{ GeV}/c^2$  and  $5.25 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$  separately. The slope parameters agree at a little over  $2\sigma$ , which makes it marginally a statistical effect.

Table 8.3: Confidence intervals for an expected background of 6 events

Number of observed events	Confidence Interval
0	[0.00, 0.97]
1	[0.00, 1.14]
2	[0.00, 1.57]
3	[0.00, 2.14]
4	[0.00, 2.83]
5	[0.00, 4.07]
6	[0.00, 5.47]
7	[0.00, 6.53]
8	[0.00, 7.99]
9	[0.00, 9.30]

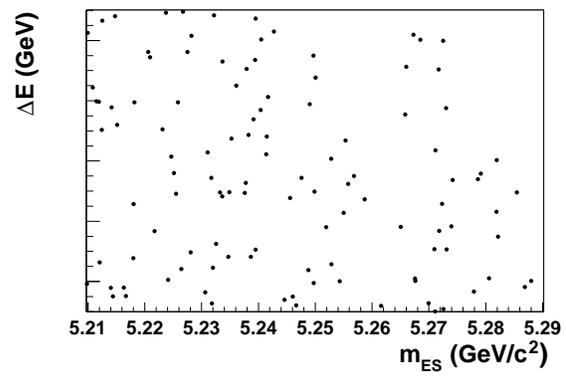


Figure 8.4: The 2D  $\Delta E$  vs.  $m_{\text{ES}}$  plane of the fitting region using real data. The box indicates the boundaries of the signal region, where 8 events lie.

# Chapter 9

## Conclusions

The results of  $B \rightarrow K^*\gamma$  analysis are shown in Table 9.1, while the results of the  $B \rightarrow (\rho/\omega)\gamma$  analysis are displayed in Table 9.3. Also shown in these tables are previous results from CLEO and *Belle*. Tables 9.2 and 9.4 show the theoretical predictions. These tables are identical to Tables 4.2 through 4.5, except that the current measurements have superseded the previous *BABAR* measurements. As mentioned in Section 4.4.1, the theoretical predictions for the  $B \rightarrow \phi\gamma$  branching fraction are  $\sim 10^{-11}$  [35, 36]. The result from this thesis is  $\mathcal{B}(B \rightarrow \phi\gamma) < 9.4 \times 10^{-7}$  [74].

Table 9.1: Previous measurements of the branching ratios and asymmetries for  $B \rightarrow K^*\gamma$ , as well as the measurements of this thesis

	CLEOII [29] 9.2 fb <sup>-1</sup>	<i>Belle</i> [31] 78 fb <sup>-1</sup>	<i>BABAR</i> [73] 347 fb <sup>-1</sup>
$B^0 \rightarrow K^{*0}\gamma$ ( $\times 10^{-5}$ )	$4.55^{+0.72}_{-0.68} \pm 0.34$	$4.01 \pm 0.21 \pm 0.17$	$4.47 \pm 0.10 \pm 0.16$
$B^+ \rightarrow K^{*+}\gamma$ ( $\times 10^{-5}$ )	$3.76^{+0.89}_{-0.83} \pm 0.28$	$4.25 \pm 0.31 \pm 0.24$	$4.22 \pm 0.14 \pm 0.16$
$\mathcal{A}_{CP}$	$+0.08 \pm 0.13 \pm 0.03$	$-0.015 \pm 0.044 \pm 0.012$	$-0.003 \pm 0.017 \pm 0.007$
Isospin asymmetry	N/A	$+0.012 \pm 0.044 \pm 0.026$ (measured parameter: $\Delta_{0+}$ )	$0.066 \pm 0.021 \pm 0.022$ (measured parameter: $\Delta_{0-}$ )

### 9.1 $B \rightarrow K^*\gamma$

The  $B \rightarrow K^*\gamma$  measurements described in this thesis represent an improvement in the uncertainty of  $\sim 50\%$  over the previous *BABAR* measurements [30]. These improvements

Table 9.2: Current theoretical calculations of the  $B \rightarrow K^*\gamma$  branching fraction

$\mathcal{B}(B \rightarrow K^*\gamma)$ (NLO)	Reference
$(7.25 \pm 2.6) \times 10^{-5}$	Ali and Parkhomenko [24]
$(7.9_{-3.0}^{+3.5}) \times 10^{-5}$	Beneke, Feldmann and Seidel [25]
$(7.09_{-2.27}^{+2.47}) \times 10^{-5}$	Bosch and Buchalla [26]
$(5.8 \pm 2.9) \times 10^{-5}$ ( $B^0 \rightarrow K^{*0}\gamma$ ) $(6.0 \pm 3.0) \times 10^{-5}$ ( $B^+ \rightarrow K^{*+}\gamma$ )	Matsumori, Sanda, and Keum [27]
$(4.3 \pm 1.4) \times 10^{-5}$ ( $B^0 \rightarrow K^{*0}\gamma$ ) $(4.6 \pm 1.4) \times 10^{-5}$ ( $B^+ \rightarrow K^{*+}\gamma$ )	Ali, Pecjak, and Greub [28]

Table 9.3: Previous measurements of the branching ratios for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$ , where the limits are given by the 90% confidence level, as well the measurements of this thesis. The errors on the *Belle* measurements are statistical and systematic, respectively, while the entry in paranthesis refers to the significance of the measurement.

	CLEOII [29] 9.2 fb <sup>-1</sup>	<i>Belle</i> [34] 350 fb <sup>-1</sup>	<i>BABAR</i> [75] 316 fb <sup>-1</sup>
$B^+ \rightarrow \rho^+\gamma(\times 10^{-6})$	< 13	$0.55_{-0.36-0.08}^{+0.42+0.36}$ (1.6 $\sigma$ )	$1.10_{-0.33}^{+0.37} \pm 0.09$
$B^0 \rightarrow \rho^0\gamma(\times 10^{-6})$	< 17	$1.25_{-0.33-0.06}^{+0.37+0.07}$ (5.2 $\sigma$ )	$0.79_{-0.20}^{+0.22} \pm 0.06$
$B \rightarrow \omega\gamma(\times 10^{-6})$	< 92	$0.56_{-0.27-0.10}^{+0.34+0.05}$ (2.3 $\sigma$ )	$0.40_{-0.20}^{+0.24} \pm 0.05$
$B \rightarrow (\rho/\omega)\gamma(\times 10^{-6})$		$1.32_{-0.31-0.09}^{+0.34+0.10}$ (5.1 $\sigma$ )	$1.25_{-0.24}^{+0.25} \pm 0.09$

arise not only from an increase in statistics, but also an improvement of the of systematic uncertainty. The branching fractions from the three experiments are compatible with each other. In addition, in all three cases, direct  $CP$  violation is not observed. The theoretically predicted branching fractions are shown in Table 9.2, and the trend in these predictions is discussed in Section 4.4.1.

The most interesting observable arises from the isospin asymmetry. The previous measurements have not been precise enough to constrain this observable to a definite sign. However, this measurement has constrained this observable to be positive, and have a value of  $0.017 < \Delta_{0-} < 0.116$  at the 90% C.L. This in ageement with the standard model predictions in Table 9.5. The reason this constraint is interesting is because of possible new physics contributions to the isospin asymmetry. According to Ref. [42], the contribution of the operator  $Q_6$  to the isospin asymmetry is  $\sim 9\%$ , so that the isospin asymmetry is mainly sensitive to the magnitude and sign of the ratio  $C_6/C_7$  of Wilson

Table 9.4: Theoretical predictions of the branching fractions for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$ .

Mode	Branching fraction ( $\times 10^{-6}$ )		
	Ref. [32]	Ref. [26]	Ref. [23]
$B^+ \rightarrow \rho^+\gamma$	$1.37 \pm 0.28$	$1.58^{+0.53}_{-0.46}$	$1.16 \pm 0.26$
$B^0 \rightarrow \rho^0\gamma$	$0.65 \pm 0.12$		$0.55 \pm 0.13$
$B \rightarrow \omega\gamma$	$0.53 \pm 0.12$		$0.44 \pm 0.10$

coefficients. When considering the minimal supersymmetric standard model for low  $\tan\beta$ , where  $\tan\beta$  is the ratio of Higgs vacuum expectation values,  $\text{Re}(C_7)$  is negative as in the standard model. However, the large  $\tan\beta$ ,  $\text{Re}(C_7)$  can be positive, thus flipping the sign of the isospin asymmetry to be negative. Therefore, this measurement excludes that portion of parameter space.

Table 9.5: Current theoretical calculations of isospin violation in  $B \rightarrow K^*\gamma$  decays within the SM.  $F^{B \rightarrow V}(q^2 = 0)$  is the  $B \rightarrow K^*$  form factor whose estimates range from  $0.23 \pm 0.06$  [27] to  $0.38 \pm 0.06$  [76].

Isospin violation	Reference
$(+8.0^{+2.1}_{-3.2})\% \times 0.3/T_1^{B \rightarrow K^*} (\Delta_{0-})$	Kagan and Neubert [42]
$(+2.7 \pm 0.8)\% (\Delta_{0+})$	Matsumori, Sanda, and Keum [27]

## 9.2 $B \rightarrow (\rho/\omega)\gamma$

The measurement described in this thesis found evidence for the decays of  $B^0 \rightarrow \rho^0\gamma$  and  $B^+ \rightarrow \rho^+\gamma$ . For  $B^+ \rightarrow \rho^+\gamma$ , this is the first such observation. As can be seen in Table 9.3, the *BABAR* measurement is consistent with isospin symmetry (described in Section 4.4.3.2), which requires that  $B(B^+ \rightarrow \rho^+\gamma)$  be twice the value of  $B(B^0 \rightarrow \rho^0\gamma)$ . However, the central values of the *Belle* measurement are not consistent with this symmetry. Both measurements of the branching ratios described in this thesis are also consistent with theoretical predictions, which are described in Table 9.5. For the  $B \rightarrow \omega\gamma$  mode, only an upper limit was obtained.

By dividing the  $B(B \rightarrow (\rho/\omega)\gamma)$  by  $B(B \rightarrow K^*\gamma)$ , the magnitude of  $|V_{td}/V_{ts}|$  can be obtained. We calculate  $|V_{ts}/V_{ts}| = 0.200^{+0.021}_{-0.020} \pm 0.015$ . This can be compared to the value

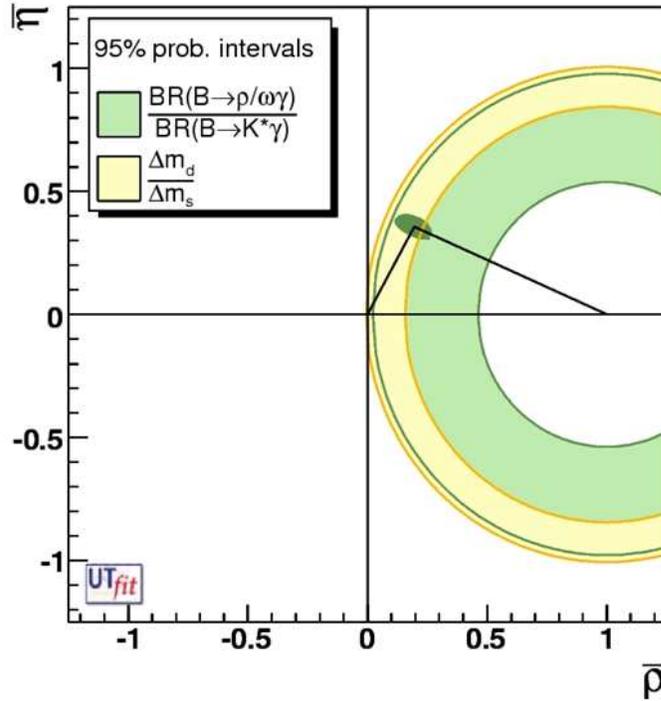


Figure 9.1: The measurement of  $B(B \rightarrow (\rho/\omega)\gamma)$ , when combined with  $B(B \rightarrow K^*\gamma)$ , provides a measurement of the magnitude of one length of the unitary triangle  $R_t$  (described in Section 4.4.4). This is outlined in green, and extends from the inner green circle to the outer green area. The measurement of  $B_d/B_s$  mixing also gives a measurement of  $R_t$ , and extends across the yellow shaded region.

found from  $B_d/B_s$  mixing, which is considerably more precise. This method of extraction yields  $|V_{td}/V_{ts}| = 0.2060 \pm 0.0007(\Delta m_s)^{+0.0081}_{-0.0060}(\Delta m_d + \text{theor.})$  ( $\Delta m_d$  and  $\Delta m_s$  denote the  $B_d$  and  $B_s$  mixing frequencies respectively, while “theor.” stands for the theoretical uncertainty.). Fig 9.1 gives a graphical representation of the results in the  $\bar{\rho} - \bar{\eta}$  plane defined in Section 4.4.4. While the  $B_d/B_s$  mixing result defines the length of the triangle, the  $B \rightarrow (\rho/\omega)\gamma$  result provides a consistency check.

### 9.3 $B \rightarrow \phi\gamma$

Because of the rarity of this decay, only an upper limit of  $\mathcal{B}(B \rightarrow \phi\gamma) < 9.4 \times 10^{-7}$  was set. This is consistent with the theoretical prediction of  $\sim 10^{-11}$ , but is still not sensitive enough to provide a measurement of the branching fraction.

## 9.4 Future Prospects

The dataset for the  $B \rightarrow K^*\gamma$  analysis is  $347 \text{ fb}^{-1}$ . An updated analysis would use the entire dataset, which is  $\sim 20\%$  more data ( $423 \text{ fb}^{-1}$ ). This corresponds to  $\sim 10\%$  improvement in statistical precision. The systematic uncertainty is limited by the photon isolation study of data/MC differences. This study was performed in Ref. [66], which was done eight years before this analysis was completed. An updated study is desirable. However, even so, it is not clear the total uncertainty on the branching fractions can be reduced to levels that will warrant the extra effort. The branching fractions are already well measured, and the theoretical predictions are greater than the experimental uncertainties. However, the isospin and  $CP$  asymmetries could still contain new physics effects. The leading systematic uncertainty for the isospin asymmetry is the fitting systematic, which can be reduced by having a more well understood fitting procedure. The leading systematic uncertainty for the  $CP$  asymmetry is the charge-asymmetry of the hadronic interactions of the kaons with the detector material. In this thesis, the systematic uncertainty is  $0.7\%$ . A reduction in this uncertainty would most likely involve developing a kaon control sample.

Table 9.6: The measurements of the branching ratios for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$  which were performed after the measurements described in this thesis. They are described in Ref. [77]. The entry in paranthesis refers to the significance of the measurement.

	<i>BABAR</i> [77] $423 \text{ fb}^{-1}$
$B^+ \rightarrow \rho^+\gamma(\times 10^{-6})$	$1.20_{-0.37}^{+0.42} \pm 0.20 (3.2\sigma)$
$B^0 \rightarrow \rho^0\gamma(\times 10^{-6})$	$0.97_{-0.22}^{+0.24} \pm 0.06 (5.4\sigma)$
$B \rightarrow \omega\gamma(\times 10^{-6})$	$0.50_{-0.23}^{+0.27} \pm 0.09 (2.2\sigma)$
$B \rightarrow (\rho/\omega)\gamma(\times 10^{-6})$	$1.73_{-0.32}^{+0.34} \pm 0.17 (6.0\sigma)$

The  $B \rightarrow (\rho/\omega)\gamma$  analysis described in this thesis has already been updated to the full dataset (Ref. [77]). The results are in Table 9.6. The  $90\%$  C.L. for  $\mathcal{B}(B \rightarrow \omega\gamma) < 0.9 \times 10^{-6}$ , while  $|V_{td}/V_{ts}| = 0.233_{-0.024-0.021}^{+0.025+0.022}$ . These results are all consistent with the results of this thesis. The goal of the analysis described in Ref. [77] was to discover the  $B \rightarrow \omega\gamma$  mode by using the full dataset, along with new background suppression techniques. Since the main limitation of this analysis is statistics, a different more robust classifier, the bootstrap-aggregated decision tree [78], was used. Even though only an

upper limit was set, it is possible the new analysis techniques that were developed can be carried over into a future analysis.

Finally, the  $B \rightarrow \phi\gamma$  analysis suffers severe statistical limitations. The full *BABAR* dataset is sensitive to branching fractions  $\sim 10^{-8}$ , so that it is unlikely that using the full dataset will help without significant improvement in analysis techniques. At the very least, a tighter upper limit can be achieved.

In all likelihood, these three analyses will only be attempted again at another experiment, such as SuperB or LHCb. At SuperB, there is projected to be an improvement in calorimeter energy resolution as compared to *BABAR*. A large portion of the continuum background originates from  $\pi^0$ 's and  $\eta$ 's, in which one of the photons fake the high energy gamma. Having a smaller energy resolution will mean that the  $\pi/\eta$  veto will be more effective. In addition, a smaller  $\Delta E$  resolution will offer a stronger discrimination between signal and background. However, the major difference is that the instantaneous luminosity will be at least 100 times that of *BABAR*. The huge increase in statistics will make the measurements systematics limited. At the LHCb, there will be  $\sim 2fb^{-1}$  of data taken every year. Within the first year, a measurement of  $B \rightarrow \omega\gamma$  is expected, along with a measurement of the  $CP$  asymmetry of  $B \rightarrow K^*\gamma$  with a statistical error of  $< 1\%$  [79]. The future prospects of these decays have a bright outlook.

# Appendix A

## OffPeak Data-Monte Carlo agreement

### A.1 $K^{*0} \rightarrow K^+\pi^-$ mode

Figures A.1 through A.3 show the comparison between offpeak data and continuum Monte Carlo for the neural net inputs, while figure A.4 shows the comparison for the cuts that are optimized.

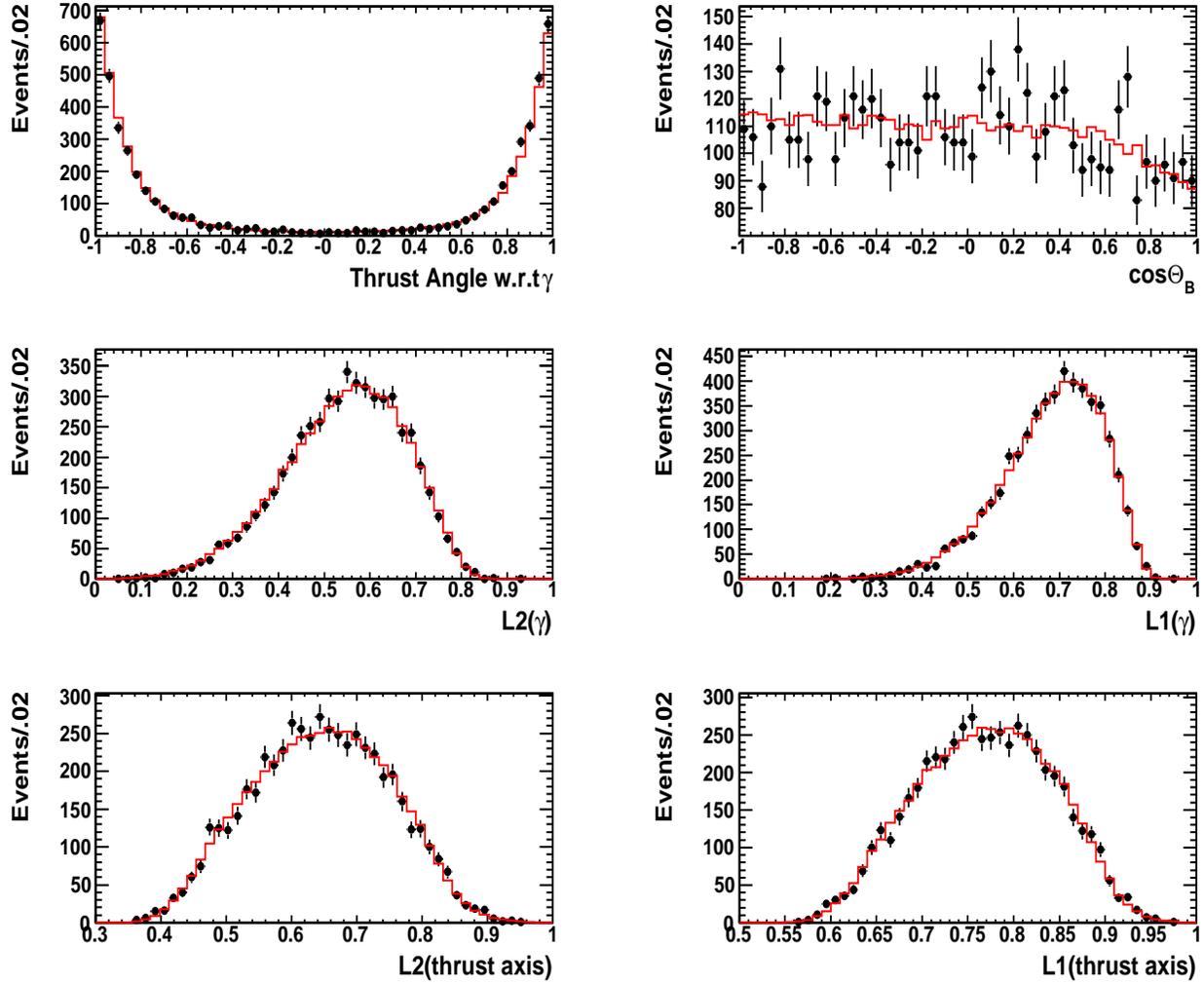


Figure A.1: Thrust angle,  $\cos\theta_B$ , and Legendre moments for offpeak data (black dots) and continuum Monte Carlo (red line). The mode shown is  $K^{*0} \rightarrow K^+\pi^-$  and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data.

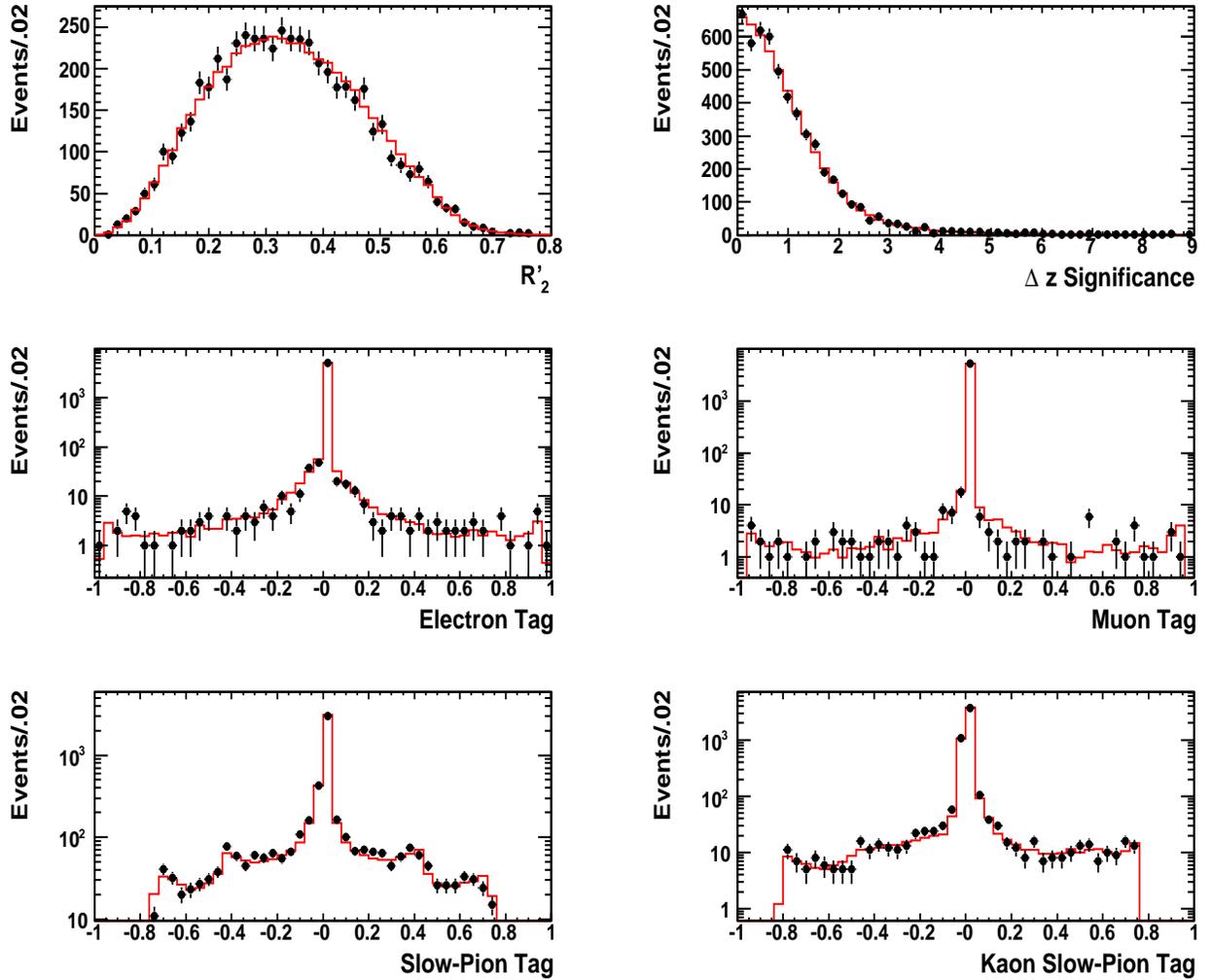


Figure A.2:  $R'_2$ ,  $\Delta z$  significance, and subtagger output for offpeak data (black dots) and continuum Monte Carlo (red line). The mode shown is  $K^{*0} \rightarrow K^+\pi^-$  and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data.

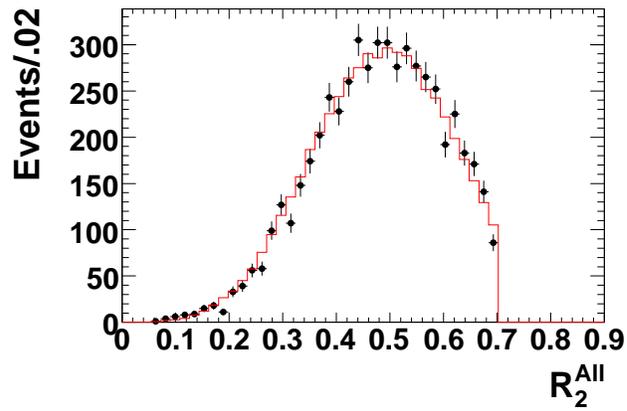


Figure A.3:  $R_2^{\text{All}}$  for offpeak data(black dots) and continuum Monte Carlo (red line). The mode shown is  $K^{*0} \rightarrow K^+\pi^-$  and all cuts are applied except for the Neural Net cut. The Monte Carlo is scaled to the data.

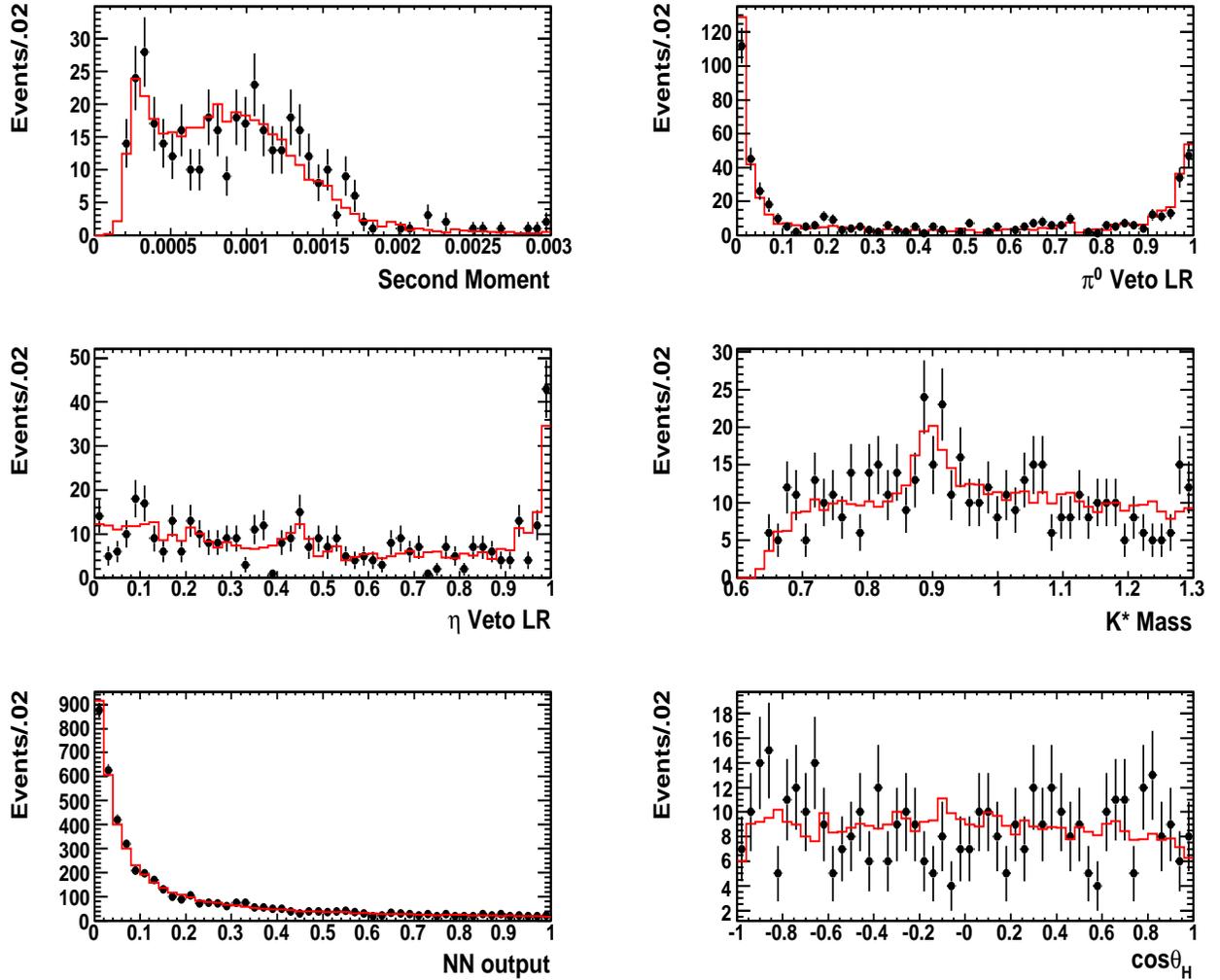


Figure A.4: A comparison of offpeak data(black dots) and continuum Monte Carlo (red line) for all variables that were optimized. The mode shown is  $K^{*0} \rightarrow K^+\pi^-$  and all cuts are applied to each plot except for the variable shown. The Monte Carlo is scaled to the data.

# Appendix B

## Correlation of Fit Observables

### B.1 $K^{*0} \rightarrow K^+\pi^-$ mode

Figures B.1 through B.3 show how the fit observables are correlated with each other for the  $K^{*0} \rightarrow K^+\pi^-$  mode.

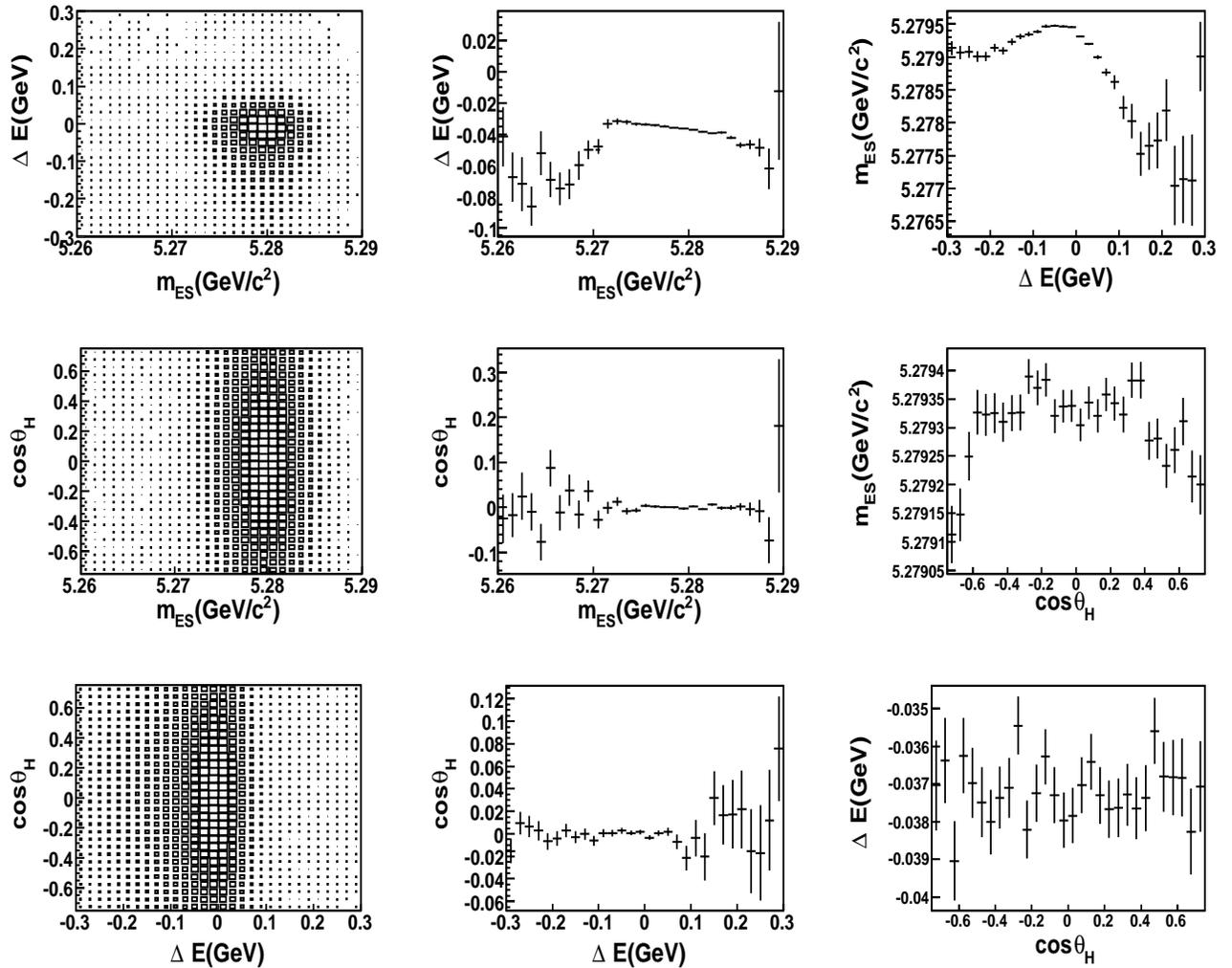


Figure B.1: The left column shows the 2 dimensional scatter plots of the observables for signal Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot.

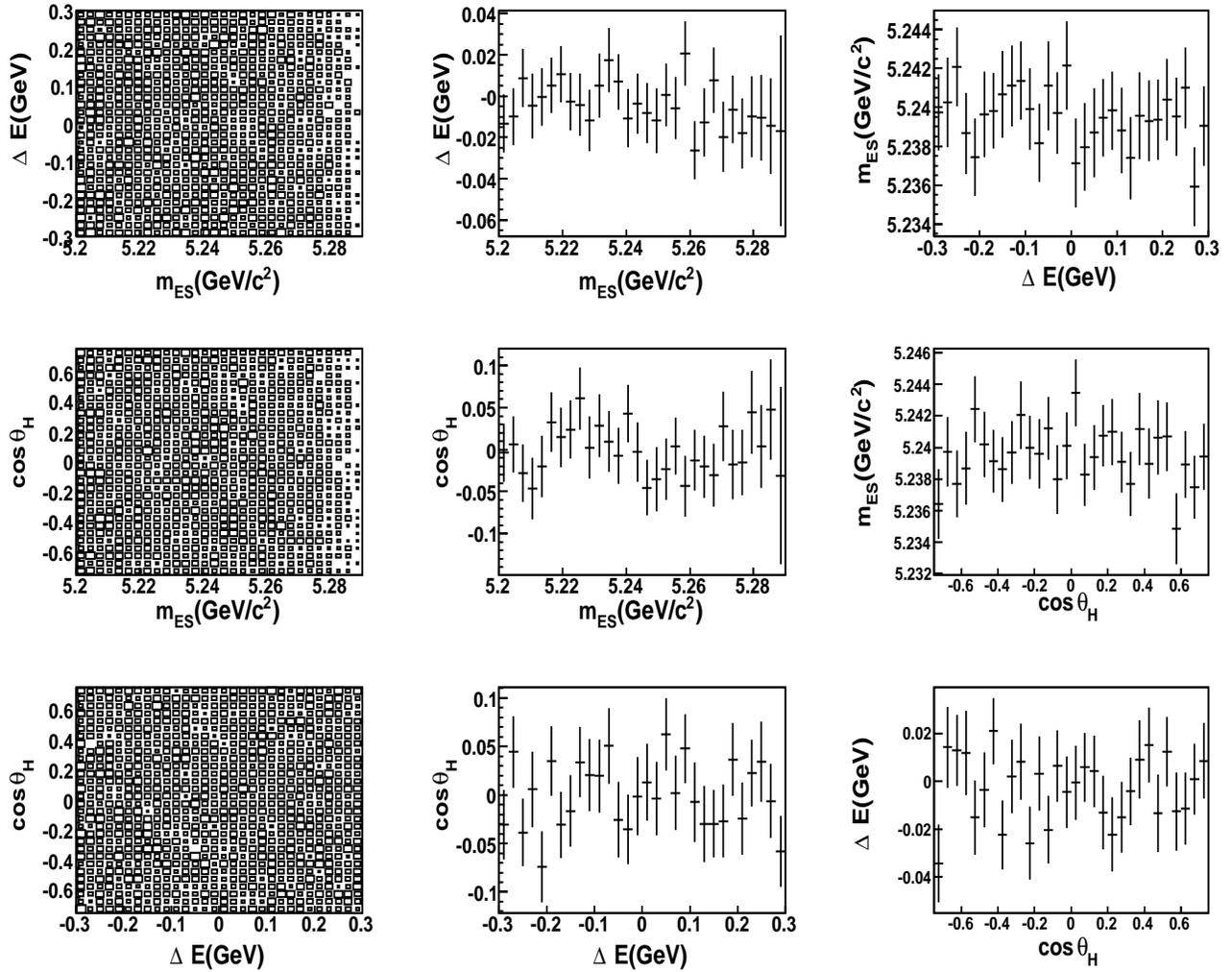


Figure B.2: The left column shows the 2 dimensional scatter plots of the observables for continuum Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot.

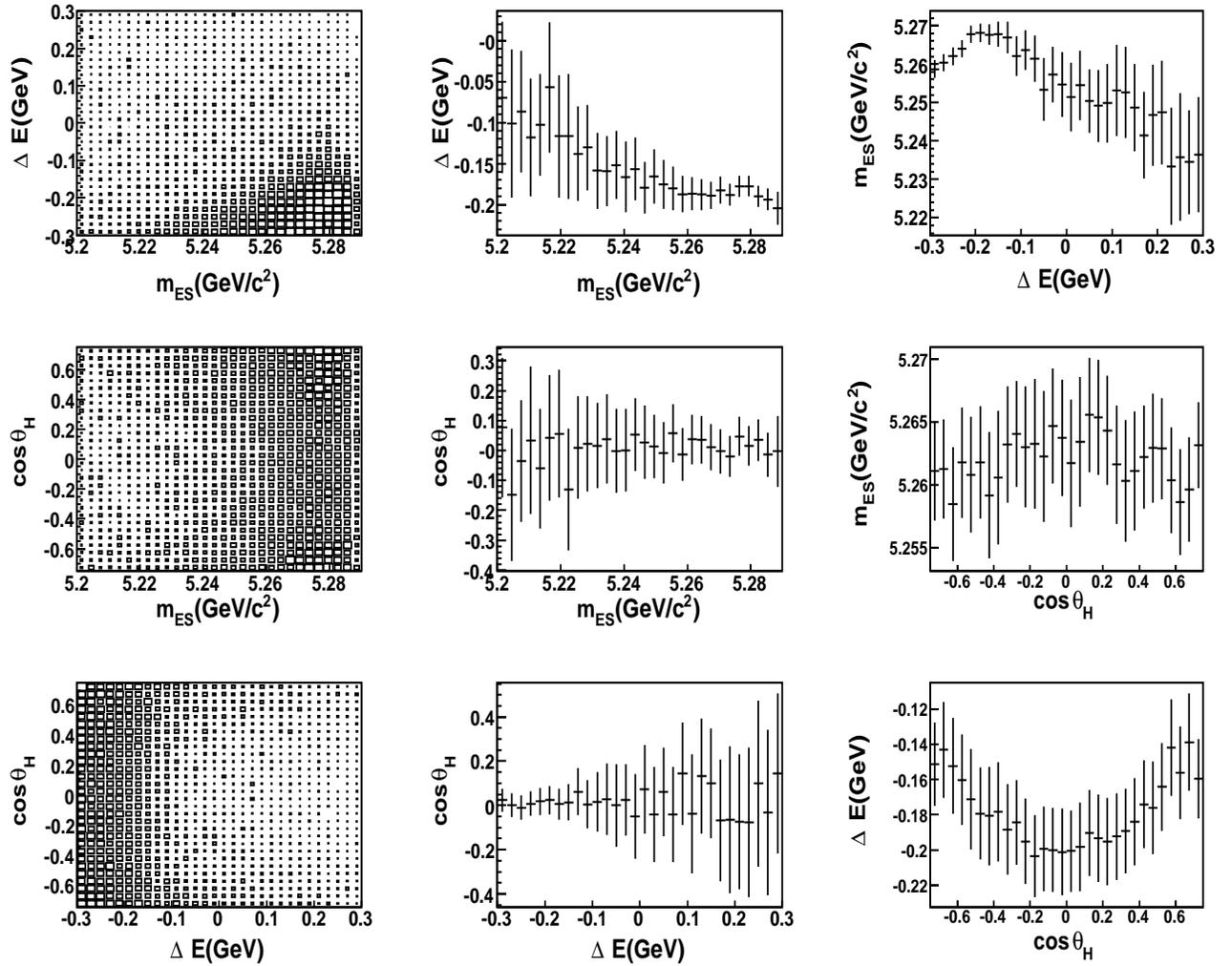


Figure B.3: The left column shows the 2 dimensional scatter plots of the observables for  $BB$  Monte Carlo, while the next 2 columns show the profile plots of both axes of the scatter plot.

# Appendix C

## Neural Network Overview

### C.1 Neural Network Introduction

The neural network combines the input variables (described in Section 5.4.1) into a single variable which is used to separate background from signal events. In short, a neural network has a topological structure which is determined by three main elements: input nodes, hidden layers each with an arbitrary number of nodes, and output nodes. The input nodes are fixed and correspond to the input variables in the problem. Both the output nodes and the nodes in the hidden layers receive a linear combination of inputs. The output of the  $j$ th node is determined as follows:

$$O_j = F\left(\sum_{i=1}^N w_{ji}x_i\right)$$

where  $x_i$  are the inputs to the node,  $w_{ji}$  is the weight of unit  $i$  to unit  $j$ ,  $N$  is the number of nodes that have an output to the  $j$ th node, and  $F$  is a function which determines the shape of the overall output of the node. If this output corresponds to a node in a particular hidden layer, then it is used as an input to another node.

#### C.1.1 Determination of Free Parameters

The free parameters of the problem are the weights  $w_{ji}$ . In order to fix these quantities, we introduce a set of “training” events in which there is a quantity  $T_p(\vec{x}_q)$  that is associated with them and serves as the desired output for the  $p$ th output node in the  $q$ th event ( $\vec{x}_q$  is the input vector). In the present set of analyses, we use only one output node;

$T_1(\vec{x}_q)$  is 0 for continuum background and 1 for signal events. Then, an optimization criterion is chosen. A typical one minimizes the sum squared error (SSE):

$$SSE(w_{ji}) = \sum_p \sum_q [O_p(w_{ji}, \vec{x}_q) - T_p(\vec{x}_q)]^2$$

where  $O_p(w_{ji}, \vec{x}_q)$  is the actual output. Note that the SSE can only apply to output nodes, since these nodes are the only ones that have desired outputs ( $T_p(\vec{x}_q)$ ) available. The errors of the hidden nodes are determined by summing over the errors of the previous layer multiplied by the corresponding weights as follows:

$$\delta_j = \sum_k \delta_k w_{kj}$$

where  $\delta_j$  is the error of the  $j$ th hidden node and  $\delta_k$  is the error of the  $k$ th node of the previous layer. In this manner, the errors are propagated backward through the network. The procedure in which the weights are determined is by an algorithm, typically the “backpropagation” algorithm. For further detail on neural networks, one can refer to [80].

### C.1.2 Implementation of Neural Network

In order to make the above abstraction more concrete, a specific example can be given which the  $B \rightarrow K^* \gamma$  analysis utilizes. All of the event quality cuts to the data sets are applied before they are put into the neural net. The structure used is 13 input nodes, 1 hidden layer with 26 nodes, and 1 output node. We use the standard backpropagation algorithm as the procedure for optimizing the weights. The total data set is divided into two, one that provides a set of training events and the other for validating the neural network. The division is such that the ratio of the number of training events to the number of events in the validation sample is 1:1. The data sets used in training and validation are continuum ( $uds$ ,  $c\bar{c}$ , and  $\tau^+\tau^-$ ) and signal Monte Carlo. In training the network, the relevant quantity is the mean squared error (MSE):

$$\text{MSE} = \frac{\text{SSE}}{\text{Number of Events}}$$

The MSE is evaluated after each cycle (one cycle corresponds to one update of the weights) for the training and validation data sets. The training is stopped when the MSE has

reached its minimum value for the validation data set.

# Appendix D

## Signal Embedded Toy Studies

This section describes understanding the biases of the signal embedded toys observed in section 6.4.1.2. Figure D.1 shows the  $m_{ES}$  distribution for all four modes for  $m_{ES} < 5.27$ . For all four modes, the PDF underestimates the actual distribution of events. However, for the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes, the PDF drops dramatically to zero, while for the other two modes there is a more gradual decrease. This is because the  $\pi^0$  tails tend to broaden out the PDF. For this reason, we treat the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes separately from the  $K^{*0} \rightarrow K_S\pi^0$ ,  $K^{*+} \rightarrow K^+\pi^0$  modes.

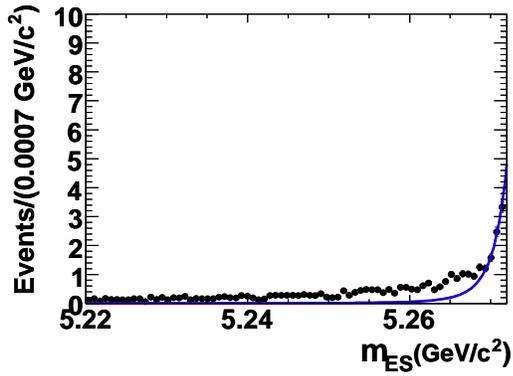
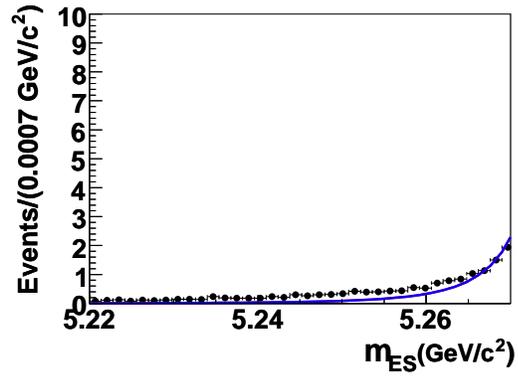
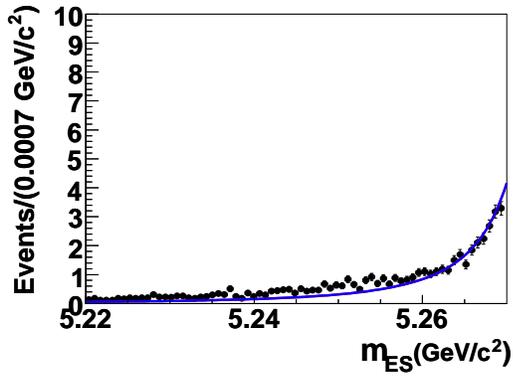
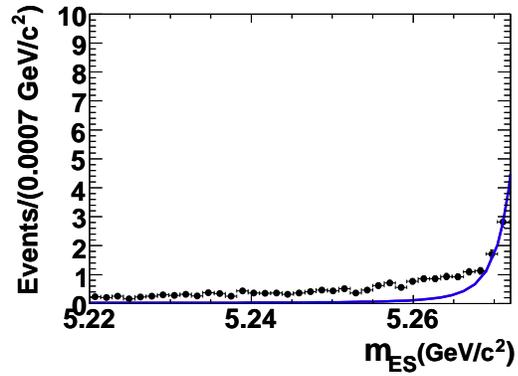
### D.1 $K^{*0} \rightarrow K^+\pi^-$ , $K^{*+} \rightarrow K_S\pi^+$ modes

For these modes, the signal  $m_{ES}$  line shape does not represent the data in the  $m_{ES}$  sideband. Therefore, it is speculated that the cause of the signal bias is due to incorrect PDF modeling in this region. To show this, we eliminate the signal MC in the region  $m_{ES} < 5.27$  and perform the signal embedded toy studies. The results are shown in Table D.1. One can see that the bias has been removed.

Table D.1: The pulls of the yields for each component for fits in which the signal MC has been eliminated for  $m_{ES} < 5.27$ . All fits are signal embedded.

Yield Pull	$K^{*0} \rightarrow K^+\pi^-$	$K^{*0} \rightarrow K_S\pi^0$
Sig. Pull	$-0.105 \pm 0.11$	$0.099 \pm 0.096$
Cont. Pull	$-0.048 \pm 0.11$	$-0.111 \pm 0.098$
$B\bar{B}$	$0.102 \pm 0.090$	$0.205 \pm 0.096$

To understand the origin of events in the region  $m_{ES} < 5.27$  for the  $K^{*0} \rightarrow K^+\pi^-$

(a)  $K^{*0} \rightarrow K^+\pi^-$  mode(b)  $K^{*0} \rightarrow K_S\pi^0$  mode(c)  $K^{*+} \rightarrow K^+\pi^0$  mode(d)  $K^{*+} \rightarrow K_S\pi^+$  modeFigure D.1: Signal MC sidebands for the  $m_{ES}$  distribution for all 4 modes.

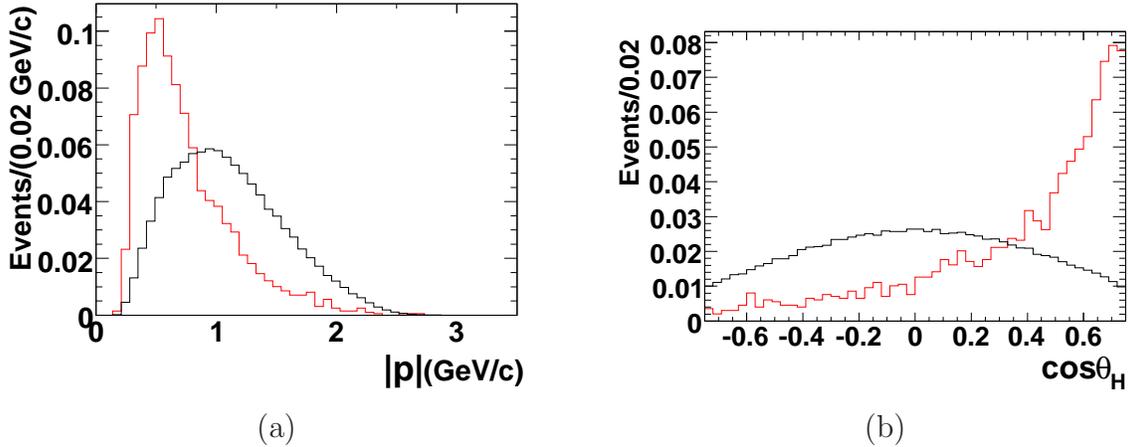


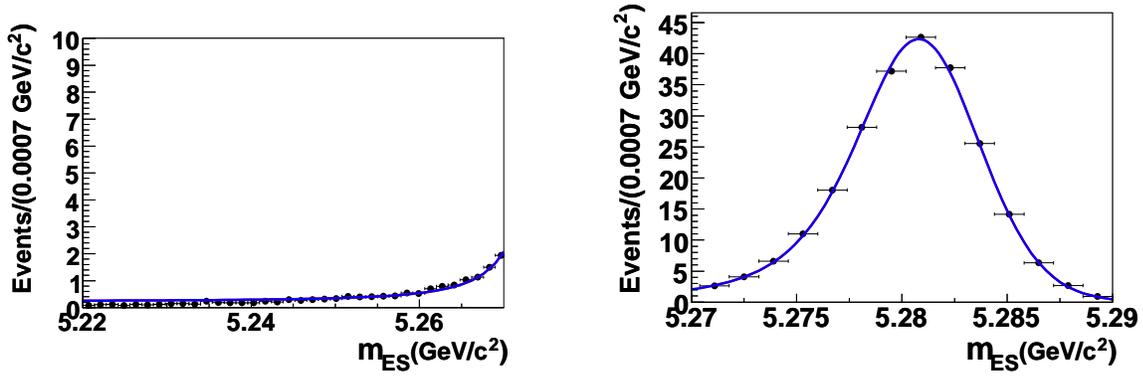
Figure D.2: a)  $K^{*0} \rightarrow K^+\pi^-$  pion lab momentum and b)  $\cos(\theta_H)$  distributions for truth-matched,  $m_{ES} > 5.27$  (black) and untruth-matched,  $m_{ES} < 5.27$  (red)

mode, we perform a study in order to categorize these events. The large majority of events were due to untruth-matched events. Furthermore, approximately 80% of these untruth-matched events were due to the pion in the decay coming from the rest of the event. To attempt to eliminate these events, two variables were considered, the pion lab momentum and the helicity angle. Figure D.2 shows these distributions for truth-matched events with  $m_{ES} > 5.27$  and untruth-matched events with  $m_{ES} < 5.27$ . Since the pion lab momentum did not yield a clear discrimination of the two types of events, placing a cut of  $\cos(\theta_H) < 0.6$  was attempted. This did not significantly reduce the bias. Therefore, it was decided to simply correct for the signal bias.

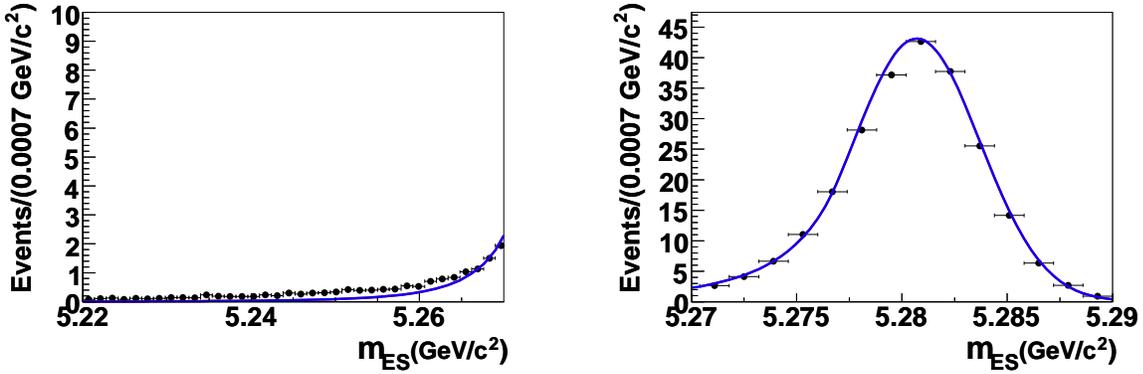
## D.2 $K^{*0} \rightarrow K_S\pi^0, K^{*+} \rightarrow K^+\pi^0$ modes

### D.2.0.1 $K^{*0} \rightarrow K_S\pi^0$

We use a Crystal Ball function to represent the  $m_{ES}$  distribution, but another choice to represent the data would be a Cruijff function. This is shown in Figure D.3. The bias for the Cruijff is displayed in Table D.2. One can see that there is almost no bias for the Cruijff function. Thus, since the  $m_{ES}$  distribution above 5.27 is well represented by both the Crystal Ball and Cruijff functions, one is led to the conclusion that the bias is due to the discrepancy between the two functions below 5.27. We choose to represent the data by a Crystal Ball function because it leads to a lower overall fitting systematic. When the fixed parameters are varied (as shown in Section 6.5.11.2), the variation with respect



(a) Cruijff function



(b) Crystal Ball function

Figure D.3: Signal MC data and PDF fits for the (a) Cruijff and (b) Crystal Ball functions for the  $K^{*0} \rightarrow K_S \pi^0$  mode. The left-hand side shows  $m_{ES} < 5.27$ , while the right-hand side shows  $m_{ES} > 5.27$ . We separate out the two regions for clarity.

to the nominal signal yield is less for the Crystal Ball than for the Cruijff.

### D.2.0.2 $K^{*+} \rightarrow K^+ \pi^0$

Figure D.4 shows the PDF fit to the signal MC data using the Cruijff and Crystal Ball functions. Even though it shows the same characteristics as the  $K^{*0} \rightarrow K_S \pi^0$  mode, the differences between the two PDFs cannot account for full bias. This is show in Table D.3.

According to the linear correlations coefficients in Table 6.4,  $m_{ES}$  and  $\Delta E$  are correlated in the signal MC for the  $K^{*+} \rightarrow K^+ \pi^0$  mode (and to a lesser degree  $m_{ES}$  and  $\cos(\theta_H)$ ). Therefore, we perform one and two dimensional signal embedded toy studies according to Table D.3. The two dimensional toy study with  $m_{ES}$  and  $\Delta E$  gives nearly the same bias as the nominal bias, leading to the conclusion that the majority of the bias

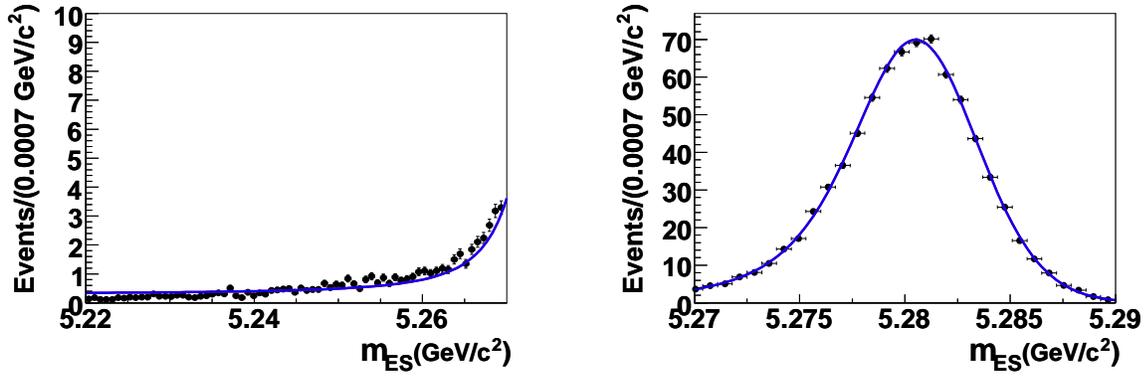
Table D.2: The pulls of the yields for Crystal Ball and Cruijff functions for the  $K^{*0} \rightarrow K_S\pi^0$  mode. All fits are signal embedded.

	Signal	Continuum	$B\bar{B}$
Nominal Pull	$-0.630 \pm 0.062$	$0.092 \pm 0.058$	$0.206 \pm 0.061$
Cruijff	$-0.258 \pm 0.062$	$-0.133 \pm 0.067$	$0.207 \pm 0.061$

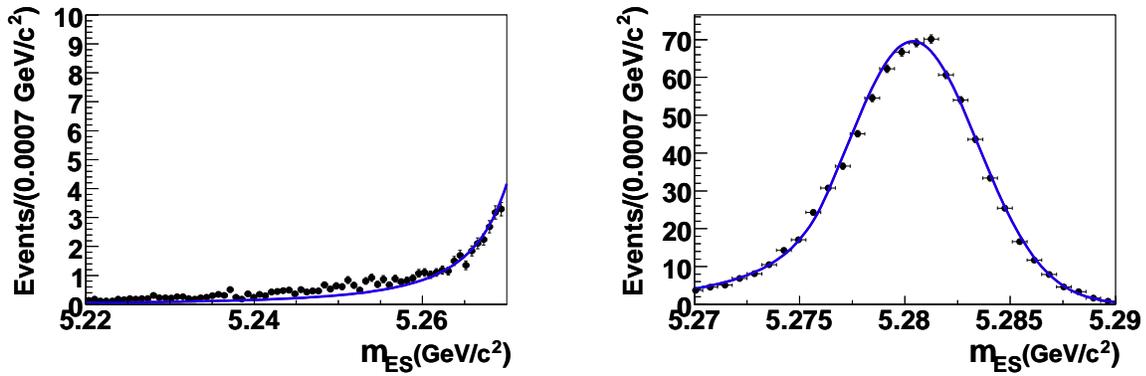
is due to correlations between them.

Table D.3: The pulls of the yields for each component fits of less than 3 dimensions for the  $K^{*+} \rightarrow K^+\pi^0$  mode, with the exception of line labeled ‘‘Cruijff’’, which is in 3 dimensions. This line displays pulls of the yields for the Cruijff function. All fits are signal embedded.

	Signal	Continuum	$B\bar{B}$
Nominal Pull	$-0.751 \pm 0.15$	$0.12 \pm 0.12$	$0.22 \pm 0.12$
Cruijff	$-0.583 \pm 0.15$	$0.01 \pm 0.14$	$0.36 \pm 0.14$
1d ( $m_{ES}$ )	$-0.238 \pm 0.11$	$-0.027 \pm 0.10$	$0.173 \pm 0.079$
1d ( $\Delta E$ )	$-0.264 \pm 0.15$	$0.14 \pm 0.13$	$-0.089 \pm 0.12$
2d ( $m_{ES}$ and $\Delta E$ )	$-0.814 \pm 0.14$	$0.08 \pm 0.15$	$0.35 \pm 0.15$
2d ( $m_{ES}$ and $\cos(\theta_H)$ )	$-0.524 \pm 0.15$	$-0.181 \pm 0.11$	$0.37 \pm 0.12$
2d ( $\Delta E$ and $\cos(\theta_H)$ )	$0.18 \pm 0.12$	$-0.448 \pm 0.11$	$0.57 \pm 0.13$



(a) Cruijff function



(b) Crystal Ball function

Figure D.4: Signal MC data and PDF fits for the (a) Cruijff and (b) Crystal Ball functions for the  $K^{*+} \rightarrow K^+\pi^0$  mode. The left-hand side shows  $m_{ES} < 5.27$ , while the right-hand side shows  $m_{ES} > 5.27$ . We separate out the two regions for clarity.

# Appendix E

## Parameter Comparison

Tables E.1 and E.2 compare the parameters obtained the fit to onpeak data to those obtained from the full fit to Monte Carlo simulation.

Component	Modes Parameter	$K^{*0} \rightarrow K^+\pi^-$		$K^{*+} \rightarrow K_S\pi^+$	
		Data	MC	Data	MC
Signal	Yield	$2394.2 \pm 55.1$	2313	$759.1 \pm 33.3$	754
	$\langle m_{ES} \rangle$	$5.2808 \pm .0000616$	$5.2808 \pm .0000632$	$5.2801 \pm .000119$	$5.2804 \pm .000115$
	$\sigma_{Lm_{ES}}$	$.00248 \pm .0000495$	$.00246 \pm .0000506$	$.00240 \pm .0001$	$.00240 \pm .00009$
	$\langle \Delta E \rangle$	$-.00171 \pm .00310$	$-.00499 \pm .00283$	$-.00908 \pm .00605$	$.00528 \pm .00593$
	$\sigma_{L\Delta E}$	$.04831 \pm .00262$	$.05590 \pm .00265$	$.05965 \pm .00559$	$.05322 \pm .00502$
	$\sigma_{R\Delta E}$	$.03841 \pm .00204$	$.02572 \pm .00176$	$.03543 \pm .00406$	$.02676 \pm .00323$
	$P1_{\cos(\theta_H)}$	$-.06877 \pm .03625$	$.01799 \pm .03516$	$-.05773 \pm .07238$	$-.05471 \pm .06806$
	$P2_{\cos(\theta_H)}$	$-1.0569 \pm .07514$	$-1.1689 \pm .07056$	$-.87457 \pm .16117$	$-1.0573 \pm .13903$
Continuum	Yield	$3077.7 \pm 87.3$	2877	$1716.5 \pm 67.0$	1510
	$\xi_{m_{ES}}$	$-18.532 \pm 3.721$	$-21.699 \pm 3.7675$	$-20.571 \pm 5.025$	$-20.546 \pm 5.2459$
	$P1_{\Delta E}$	$-.01018 \pm .14900$	$-.15908 \pm .15908$	$-.37366 \pm .19264$	$-.07417 \pm .22232$
	$P1_{\cos(\theta_H)}$	$.29817 \pm .04491$	$.07565 \pm .04862$	$-.27718 \pm .05925$	$-.20366 \pm .06702$
B Bkg.	Yield	$562.0 \pm 70.6$	606	$141.1 \pm 51.6$	286

Table E.1: A comparison of the results from the fit to OnPeak data with the full fit to Monte Carlo for the  $K^{*0} \rightarrow K^+\pi^-$  and  $K^{*+} \rightarrow K_S\pi^+$  modes.

Component	Modes Parameter	$K^{*0} \rightarrow K_S \pi^0$		$K^{*+} \rightarrow K^+ \pi^0$	
		Data	MC	Data	MC
Signal	Yield	$256.0 \pm 20.2$	253	$872.7 \pm 37.1$	842
	$\langle m_{ES} \rangle$	$5.2804 \pm 0.0023$	$5.2807 \pm 0.00248$	$5.2800 \pm 0.0012$	$5.2804 \pm 0.0014$
	$\sigma_{m_{ES}}$	$.00272 \pm 0.00021$	$.00286 \pm 0.00022$	$.00315 \pm 0.0013$	$.00297 \pm 0.0013$
	$\langle \Delta E \rangle$	$.00225 \pm 0.01242$	$.00214 \pm 0.00995$	$-.01988 \pm 0.00708$	$-.00908 \pm 0.00647$
	$\sigma_{L\Delta E}$	$.06217 \pm 0.01073$	$.06523 \pm 0.01129$	$.04680 \pm 0.00546$	$.06145 \pm 0.00604$
	$\sigma_{R\Delta E}$	$.04463 \pm 0.00839$	$.03023 \pm 0.00567$	$.04692 \pm 0.00457$	$.03868 \pm 0.00427$
	$P1_{\cos(\theta_H)}$	$-.08808 \pm 0.09853$	$.00938 \pm 0.11718$	$-.10895 \pm 0.05929$	$-.05634 \pm 0.05661$
	$P2_{\cos(\theta_H)}$	$-1.4142 \pm 0.18083$	$-1.1413 \pm 0.24121$	$-1.3275 \pm 0.11861$	$-.97412 \pm 0.18141$
Continuum	Yield	$603.1 \pm 42.0$	502	$1538.3 \pm 77.2$	1484
	$\xi_{m_{ES}}$	$-.92765 \pm 8.8597$	$6.5404 \pm 9.6490$	$-19.211 \pm 5.7281$	$-14.026 \pm 6.0888$
	$P1_{\Delta E}$	$.07388 \pm 0.32869$	$.16871 \pm 0.38249$	$.07501 \pm 0.24155$	$.14711 \pm 0.25184$
	$P1_{\cos(\theta_H)}$	$.69227 \pm 0.09105$	$.27527 \pm 0.11800$	$.84025 \pm 0.08499$	$.55245 \pm 0.07789$
	$P2_{\cos(\theta_H)}$	N/A	N/A	$.54804 \pm 0.21334$	$-.09741 \pm 0.18141$
B Bkg.	Yield	$42.9 \pm 32.5$	107	$298.0 \pm 67.7$	377

Table E.2: A comparison of the results from the fit to OnPeak data with the full fit to Monte Carlo for the  $K^{*0} \rightarrow K_S \pi^0$  and  $K^{*+} \rightarrow K^+ \pi^0$  modes.

# Bibliography

- [1] F. J. Hasert *et al.*, Phys. Lett. 46B (1973) 138, Nucl. Phys. B73 (1974) 800.
- [2] C. Y. Prescott *et al.*, Phys. Lett. 77B (1978) 347.
- [3] R. Brandelik *et al.*, Phys. Lett. 86B (1979) 243.
- [4] C. Amsler *et al.*, Phys. Lett. B667 (2008) 1.
- [5] LEP Collaborations and LEP Electroweak Working Group, arXiv:hep-ex/0612034v2.
- [6] C.T. Hill, Phys. Rev. D 24 (1981) 691.
- [7] M. Carena and C.E.M. Wagner, Nucl. Phys. B452 (1995) 45.
- [8] “PEP-II: An Asymmetric  $B$  Factory”, Conceptual Design Report, SLAC-R-418 (1993).
- [9] W. Menges *et al.*, Nuclear Science Symposium Conference, 2005 IEEE, 3, 1470 (2005).
- [10] EPOTEK-301-2 glue by Epoxy Technol. Inc., Billerica, MA 01821 USA.
- [11] Quantacon XP2020, Philips
- [12] J. Va'vra, “Measurement of EPOTEK-301-2 optical glue refraction index and a reflectivity from EPOTEK-301-2/fused silica interface”, Internal DIRC Note #140, July 2001.
- [13] B. Ratcliff and S. Spanier, Nucl. Instrum. Methods A 433 (1999) 456.
- [14] A. Fasso, A. Ferrari, J. Ranft, and P. Sala. FLUKA99 program CERN. [Online]. Available: <http://fluka.web.cern.ch/fluka>.
- [15] L. L. Chau and W. Y. Keung, Phys. Rev. Lett. 53 (1984) 1039.

- [16] C. Jarlskog, Phys. Rev. Lett. 55 (1985) 1039.
- [17] L. Wolfenstein, Phys. Rev. Lett. 51 (1983) 1945.
- [18] Y. Ushiroda *et al.*, Phys. Rev. Lett. 100 (2008) 021602.
- [19] S. Chung, BNL Preprint, BNL-76975-2006-IR (2008).
- [20] F. Hippel and C. Quigg, Phys. Rev. D 5 (1972) 624.
- [21] D. Decirevic, V. Luicz, and F. Mescia, Nucl. Phys. B 769 (2007) 31.
- [22] P. Colangelo, C. A. Dominguez, G. Nardulli and N. Paver, Phys. Lett. B 317 (1993) 183.
- [23] P. Ball, G.W. Jones, and R. Zwicky, Phys. Rev. D 75 (2007) 054004.
- [24] A. Ali and A. Y. Parkhomenko, Eur. Phys. Jour. C 23 (2002) 89.
- [25] M. Beneke and T. Feldmann D. and Seidel, Nucl. Phys. B 612 (2001) 25.
- [26] S. W. Bosch and G. Buchalla Nucl. Phys. B 621 (2002) 459.
- [27] M. Matsumori and A.I. Sanda and Y.-Y. Keum, Phys. Rev. D 72 (2005) 014013.
- [28] A. Ali and B. Pacjak and C. Greub, arXiv:0709.4422v2 [hep-ph].
- [29] T. E. Coan *et al.*, Phys. Rev. Lett. 84 (2000) 5283.
- [30] B. Aubert *et al.*, Phys. Rev. D 70 (2004) 112006.
- [31] M. Nakao *et al.*, Phys. Rev. D 69 (2004) 112001.
- [32] A. Ali and A. Y. Parkhomenko, DESY 06-187 (2006).
- [33] B. Aubert *et al.*, Phys. Rev. Lett. 94 (2005) 011801.
- [34] D. Mohapatra *et al.*, Phys. Rev. Lett. 96 (2006) 221601.
- [35] Y. Li and C. Lu, Phys. Rev. D 74 (2006) 097502.
- [36] C. Lu, Y. Shen, and W. Wang, Chin. Phys. Lett. 23 (2006) 2684.
- [37] M. Bander, D. Silverman, and A. Soni, Phys. Rev. Lett. 43 (1979) 4.

- [38] A. Kagan and N. Neubert, Phys. Rev. D 58 (1998) 094012.
- [39] D. Atwood, M. Gronau, and A. Soni, Phys. Rev. Lett. 79 (1997) 185.
- [40] B. Aubert *et al.*, Phys. Rev. D 78 (2008) 071102.
- [41] Y. Ushiroda *et al.*, Phys. Rev. D 74 (2006) 111104.
- [42] A. L. Kagan and M. Neubert, Phys. Lett. B 539 (2002) 227.
- [43] M. R. Ahmady and F. Mahmoudi, Phys. Rev. D 75 (2007) 015007.
- [44] P. Ball and R. Zwicky, JHEP 0604 (2006) 046.
- [45] A. Aulencia *et al.*, Phys. Rev. Lett. 97 (2006) 242003.
- [46] “The *BABAR* Physics Book”, SLAC-R-504 (1998).
- [47] S. Agostinelli *et al.*, Nucl. Instrum. Meth. **A506**, 250 (2003).
- [48] W-M Yao *et al.*, 2006 *J. Phys. G: Nucl. Part. Phys.* 33 1.
- [49] S. Eidelman *et al.*, Phys. Lett. B 592, 1 (2004).
- [50] W.D. Hulsbergen, Nucl. Instrum. Meth. **A552**, 566 (2005).
- [51] J. Beringer, BTagger - a multivariate tagging algorithm with categories based on the physics of the  $B_{tag}$  decay, 2002, BAD #315, *BABAR* Internal Analysis Document.
- [52] F. Fang *et al.*, Search for  $B \rightarrow \rho\gamma$  and  $B \rightarrow \omega\gamma$  in Run1 - Run5 data, 2006, BAD#1301, *BABAR* Internal Analysis Document.
- [53] Andres Zell *et al.* SNNS-Stuttgart Neural Network Simulator. User Manual, Version 4.1.
- [54] I. Narsky, *StatPatternRecognition: A C++ Package for Statistical Analysis of High Energy Physics Data*, arXiv:physics/0507143v1 [physics.data-an].
- [55] J. Friedman and N. Fisher, *Bump hunting in high dimensional data*, Statistics and Computing **9**, 123-143 (1999).
- [56] J. Feldman and R. Cousins, Phys. Rev. D **57**, 3873 (1998).

- [57] M. J. Oreglia, Ph.D Thesis, SLAC-236 (1980), J. E. Gaiser, Ph.D. Thesis, SLAC-255 (1982).
- [58] H. Albrecht *et al.*, Z. Phys. C **48**, 543 (1990).
- [59] C. Hearty, Measurement of the Number of  $\Upsilon(4S)$  Meson Produced in Run 1 (B Counting), 2001, *BABAR* Analysis Document #134.
- [60] I. Nugent, M. Roney, and R. Kowalewski, Tau31 Tracking Efficiency Study for 2004, 2004, *BABAR* Analysis Document #931.
- [61] T. Allmendinger *et al.*, Tracking Efficiency Studies in Release 12 and 14, 2004, *BABAR* Analysis Document #867
- [62] G. Eigen, K. Flood, S. Playfer, J. Ritchie, C. Schilling, L. Sun, J. Watson, Measurement of Branching Fractions and Rate Asymmetries in  $B \rightarrow Kl^{+l^{-}}$  and  $B \rightarrow K^{*}l^{+l^{-}}$  Decays, 2008, *BABAR* Analysis Document #1869.
- [63] A. Telnov, A. Gritsan,  $K_S \rightarrow \pi^{+}\pi^{-}$  Reconstruction Efficiency in *BABAR*: Release-10 Data vs. SP4 Monte Carlo, 2003, *BABAR* Analysis Document #677.
- [64] D. Payne, Single photon efficiency using  $\mu^{+}\mu^{-}\gamma$ , 2005, *BABAR* Analysis Document #1110.
- [65] M. Allen, M. Naisbit, A. Roodman, S. Banerjee, A Measurement of  $\pi^0$  Efficiency Using  $\tau \rightarrow \rho\nu$  and  $\tau \rightarrow \pi\nu$  Decays, 2004, *BABAR* Analysis Document #870.
- [66] A. Eisner *et al.*, Measurement of  $\text{Br}(B \rightarrow K^{*}\gamma)$  and Search for Direct CP Violation, 2001, *BABAR* Analysis Document #33.
- [67] B. Aubert *et al.*, Phys. Rev. Lett. 97 (2006) 171803.
- [68] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Meth. A **555**, 356 (2005).
- [69] B. Aubert *et al.* (BABAR Collaboration), Phys. Rev. D **70**, 091105 (2004).
- [70] Glen Cowan, *Statistical Data Analysis*. Oxford University Press Inc., New York, 1998.
- [71] T. Colberg, Measurement of the Decay Fractions  $B^{\pm} \rightarrow K^{*\pm}\gamma(K^{*\pm} \rightarrow \pi^0 K^{\pm})$  and  $B^0 \rightarrow K^{*0}\gamma(K^{*0} \rightarrow \pi^0 K_S^0)$ , 2001, *BABAR* Analysis Document #133.

- [72] R. Cousins and V. Highland, Nuclear Instruments and Methods in Physics Research **A320**, 331 (1992).
- [73] B. Aubert *et al.* (BABAR Collaboration), Phys. Rev. Lett. **103**, 211802 (2009).
- [74] B. Aubert *et al.* (BABAR Collaboration), Phys. Rev. D **72**, 091103 (2005).
- [75] B. Aubert *et al.* (BABAR Collaboration), Phys. Rev. Lett. **94**, 011801 (2005).
- [76] P. Ball and V. M. Braun, Phys. Rev. D 58 (1998) 094016.
- [77] B. Aubert *et al.* (BABAR Collaboration), Phys. Rev. D **78**, 112001 (2008).
- [78] L. Breiman, Mach. Learn. **24**, No. 2, 123 (1996).
- [79] I. Belyaev, Acta Physica Polonica B, Vol. 38, No. 3, (2007).
- [80] Simon Haykin, *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company, Inc., Englewood Cliffs, NF, 1994.