SLAC-R-826

Measurements of the Branching Fraction, CP Asymmetry And Isospin Asymmetry of the Radiative Penguin B0 ---> K*0 Gamma And a Search for the Decays B0 ---> Rho Gamma And B0 ---> Omega Gamma at BaBar

Patrick Michael Spradlin

Stanford Linear Accelerator Center Stanford University Stanford, CA 94309

SLAC-Report-826

Prepared for the Department of Energy under contract number DE-AC02-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

This document, and the material and data contained therein, was developed under sponsorship of the United States Government. Neither the United States nor the Department of Energy, nor the Leland Stanford Junior University, nor their employees, nor their respective contractors, subcontractors, or their employees, makes an warranty, express or implied, or assumes any liability of responsibility for accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use will not infringe privately owned rights. Mention of any product, its manufacturer, or suppliers shall not, nor is it intended to, imply approval, disapproval, or fitness of any particular use. A royalty-free, nonexclusive right to use and disseminate same of whatsoever, is expressly reserved to the United States and the University.

UNIVERSITY OF CALIFORNIA SANTA CRUZ

MEASUREMENTS OF THE BRANCHING FRACTION, CP ASYMMETRY AND ISOSPIN ASYMMETRY OF THE RADIATIVE PENGUIN DECAY $B^0 \rightarrow K^{*0}\gamma$ AND A SEARCH FOR THE DECAYS $B \rightarrow \rho\gamma$ AND $B^0 \rightarrow \omega\gamma$ AT BABAR

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

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Patrick Michael Spradlin

September 2005

The Dissertation of Patrick Michael Spradlin is approved:

Professor Bruce Schumm, Chair

Professor Thomas Banks

Professor Robert Johnson

Lisa C. Sloan Vice Provost and Dean of Graduate Studies \bigodot 2005 Patrick Michael Spradlin

Contents

List of Figures vii			
List of Tables	x		
Abstract	xii		
1 Motivation 1.1 Measured observables 1.2 $ V_{el}/V_{el} $ and the unitarity triangle	$ \begin{array}{c} 1 \\ 2 \\ 5 \end{array} $		
1.2 [\theta] \theta] and the underley shangle 1.3 Theoretical framework 1.4 State of the measurements	7 11		
2 The BABAR experiment2.1 PEP-II2.2 BABAR detector overview2.3 BABAR subsytem: SVT2.4 BABAR subsystem: DCH2.5 Tracking performance2.6 BABAR subsystem: DIRC2.7 BABAR subsystem: EMC2.8 Simulation	 13 14 17 19 20 23 26 27 33 		
3 Analysis overview 3.1 Candidate identification 3.1.1 Reconstruction 3.1.2 Final states 3.2 Optimization of selection criteria 3.3 Definition of signal 3.4 Backgrounds 3.5 Data sets and involvement	37 38 39 39 41 43 46		
4 Event reconstruction 4.1 BtoXGamma selection criteria 4.2 B daughter photon selection 4.2.1 Measurement quality 4.2.2 π^0 and η suppression	49 49 51 52 53		

	4.3	Charged particle selection
		.3.1 GoodTracksLoose criteria
		.3.2 Particle identification
		.3.3 Kaon selection $(B \to K^* \gamma)$
		.3.4 pion selection $(B \to K^* \gamma)$
		.3.5 pion selection $(B \to \rho \gamma)^{\prime\prime}$
	4.4	π^0 reconstruction \ldots
	4.5	Intermediate meson selection
		.5.1 mass windows
		5.2 Vertexing 62
		5.3 K_{\perp}^0 veto for ω candidates 63
	4.6	R reconstruction and fit region 64
	4.0	Single B candidate selection 64
	4.7	Single D candidate selection $\ldots \ldots \ldots$
5	Ba	ground suppression variables 66
	5.1	Event shape variables
	0.2	1.1 Thrust angle Θ_T
		$12 B'_{-} \qquad \qquad$
		$13 \text{Energy cones} \qquad \qquad$
		$14 L \text{ moments} \qquad \qquad$
	59	R Tagging variables 74
	0.2	D Tagging variables 14 9.1 Not flower 76
		$2.1 \text{Net flavor} \dots \dots \dots \dots \dots \dots \dots \dots \dots $
		2.2 Standard BABAR B tagging variables
	F 0	.2.3 Number of Kaons and $P_{K,\max}^*$
	5.3	Physics variables
		$.3.1 \Delta z \dots \dots \dots \dots \dots \dots \dots \dots \dots $
		$3.2 B \text{ direction } \cos \Theta_B^* \dots \dots \dots \dots \dots \dots \dots \dots \dots $
		.3.3 Helicity angle Θ_H
		.3.4 Dalitz angle Θ_D
0	C	• • • • • • • • • • • • • • • • • • • •
0	C 1	inuum suppression with neural networks
	0.1	Introduction to neural networks
		.1.1 Topology
		$1.2 \text{Training} \dots \dots \dots \dots \dots \dots \dots \dots \dots $
		.1.3 Final note
	6.2	Networks for continuum suppression
		2.1 Mode $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$
		.2.2 Modes $B \to \rho \gamma$ and $B^0 \to \omega \gamma$
_	ЪĒ	
7	BB	background suppression and discrimination 107
	7.1	Mode $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ 108
		$.1.1 B \to X_s \gamma, \text{ down-feed } \dots $
		1.2 Other $B \to K^* \gamma$ decays, cross-feed $\dots \dots \dots$
		.1.3 Other B decays $\ldots \ldots \ldots$
		$.1.4 \text{Strategy} \dots \dots \dots \dots \dots \dots \dots \dots \dots $
	7.2	Fisher discriminants
	7.3	Mode $B^0 \to \rho^0 \gamma (\rho^0 \to \pi^+ \pi^-)$ 116
		$.3.1 \qquad B^0 \to K^{*0}\gamma . \dots \dots$
		.3.2 $B^0 \to \rho^0(\eta/\pi^0)$

		7.3.3	Strategy	9
	7.4	Mode	$B^+ \to \rho^+ \gamma (\rho^+ \to \pi^+ \pi^0)$	1
		7.4.1	$B^+ \to K^{*+} \gamma$	1
		7.4.2	$B^+ \to \rho^+(\eta/\pi^0)$	1
		7.4.3	Strategy	1
	7.5	Mode	$B^0 \to \omega \gamma (\omega \to \pi^+ \pi^- \pi^0)$	4
		7.5.1	$B \to K^{'*\gamma}$	5
		7.5.2	$B^0 \to \omega(n/\pi^0)$	5
		7.5.3	Strategy	5
				-
8	Mo	nte Ca	lo efficiencies of selection criteria 12	7
	8.1	$B^0 \rightarrow$	$K^{*0}\gamma, K^{*0} \to K^+\pi^-$	7
	8.2	$B^0 \rightarrow$	$\rho^0\gamma$	0
	8.3	$B^+ \rightarrow$	$\rho^+\gamma$	3
	8.4	$B^0 \rightarrow$	$\omega\gamma$	6
9	\mathbf{Sig}	nal ext	action for individual modes 13	9
	9.1	Likelił	ood fit variables	:0
	9.2	Mode	$B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$	-1
		9.2.1	Fit components	1
		9.2.2	Toy Monte Carlo studies	:6
		9.2.3	Fit results on data	5
	9.3	Modes	$B \to \rho \gamma \text{ and } B^0 \to \omega \gamma \dots \dots \dots \dots \dots \dots \dots \dots \dots $	0
		9.3.1	Signal components	0
		9.3.2	Continuum background components	2
		9.3.3	$B \to K^* \gamma$ background components	0
		9.3.4	π^0/η background components	3
		9.3.5	Combinatoric B background components $\ldots \ldots \ldots$	4
		9.3.6	Summary of likelihood functions	7
		9.3.7	Toy Monte Carlo studies	0
		9.3.8	Signal yields	6
10	a			_
10	5ys		ting 10	5 6
	10.1	D COU.	10 afficience 10	0
	10.2	DID of	Ig eniciency	10
	10.3	π^0 on	circle photon officionay	0
	10.4	π° and Swatan	single photon enclency	1
	10.0	Nouna	atic uncertainty for photon quanty selection	1
	10.0	Theura	ieuwork systematic	2
	10.7	Fit en	ciency systematic	13
	10.8	FIL SYS	$\mathcal{D}_{\mathcal{D}} = \mathcal{L}_{\mathcal{T}} * 0 $	4
		10.8.1	$B^* \to \Lambda^{-*}\gamma \dots \dots \dots \dots \dots \dots \dots \dots \dots $	10 10
	10.0	10.8.2	$B \to \rho \gamma$ and $B^{\circ} \to \omega \gamma$ modes	9
	10.9	System	atic bias of the best candidate selection $\dots \dots \dots$	3
	10.10	J INUCIEE	r interaction effects in $A_{CP}(B^* \to K^*\gamma)$	0. 77
	10.1	I Summ	iry of systematics and efficiency	(
	10.12	2 Indivio	ual mode measurements $\dots \dots \dots$. (0
		10.12.1	$B^{\circ} \to \Lambda^{\circ} \gamma$ branching fraction	9
		10.12.2	$B^{\circ} \to K^{\circ} \gamma \text{ CP}$ asymmetry 21	9

		10.12.3 $B \to (\rho/\omega)\gamma$ branching fraction limits $\ldots \ldots \ldots \ldots \ldots \ldots \ldots$	220
11	Coi	mbined mode measurements	224
	11.1	$B \to K^* \gamma \text{ modes}$	225
		11.1.1 Statistical formalism	226
		11.1.2 Correlation of systematic branching fraction uncertainties	230
		11.1.3 Combined branching fractions	233
		11.1.4 Isospin asymmetry	235
		11.1.5 CP asymmetry	236
	11.2	$B \to \rho \gamma, B^0 \to \omega \gamma \text{ modes}$	238
		11.2.1 Combined likelihood fit	239
		11.2.2 Fit results	240
		11.2.3 Systematic uncertainties	240
		11.2.4 Combined limit	243
	11.3	$ V_{td}/V_{ts} $	245
12	Sur	nmary and conclusions	247
\mathbf{A}	Fur	nctions	251
	A.1	Gaussian function	251
	A.2	Novosibirsk function	251
	A.3	Crystal Ball function	252
	A.4	Argus function	252
	A.5	KEYS non-parametric distribution	253
		-	
В	Acr	conyms	255
Bi	bliog	raphy	258

List of Figures

1.1	The leading order Feynman diagram for the $b \to s\gamma, b \to d\gamma$ process
1.2	CKM unitarity triangle
1.3	CKMfitter winter 2004 results
1.4	Penguin diagram contribution to O_7 10
1.5	Local operators of the OPE 10
2.1	Illustration of the SLAC linac and PEP-II storage ring
2.2	BABAR detector longitudinal section
2.3	BABAR detector end view
2.4	BABAR SVT longitudinal section
2.5	BABAR SVT transverse section
2.6	BABAR DCH longitudinal section
2.7	BABAR DCH cell layout
2.8	BABAR DCH dE/dx distribution
2.9	BABAR DCH dE/dx resolution
2.10	BABAR DIRC longitudinal section
2.11	BABAR DIRC longitudinal section
2.12	Variation of $\cos \theta_c$ with particle momentum $\ldots \ldots 28$
2.13	DIRC π -K separation
2.14	BABAR EMC longitudinal section
2.15	BABAR EMC crystal diagram 32
2.16	Distribution of $E_{\gamma,\text{lab}}$ in $B \to K^* \gamma$
2.17	BABAR EMC energy and angular resolution
3.1	Distributions of m_{ES} and ΔE^*
3.2	Distribution of E_{γ}^* in $B \to K^* \gamma$
5.1	Cartoons of continuum background processes
5.2	Distributions of $ \cos \Theta_T $
5.3	Distributions of R_2
5.4	Diagrammatic definition of energy cones
5.5	Distributions of energy cones
5.6	Distributions of L -moments $\ldots \ldots \ldots$
5.7	Distributions of $N_{\mathcal{F}}$
5.8	Distributions of standard BABAR tagging variables
5.9	Distributions of K variables $\ldots \ldots \ldots$

5.10	Distributions of Δz	85
5.11	Distributions of $\cos \Theta_B$	85
5.12	Distributions of Θ_H	86
5.13	Distributions of Θ_D	87
61	Diagram of a simple neural network	00
6.9	Neural network training plota	101
0.2 6.2	Neural network training plots	101
0.5	Neural network performance plots $\dots \dots \dots$	102
0.4	Optimization of cut on neural network output, $K^{**} \rightarrow K^{*}\pi^{*}$	103
0.5	Neural network distribution plots	106
7.1	m_{ES} and ΔE^* distributions of $B \to X_s \gamma$	110
7.2	m_{ES} and ΔE^* distributions of cross-feed	111
7.3	m_{ES} and ΔE^* distributions of combinatoric <i>B</i> backgrounds	112
7.4	$m_{ES}, \Delta E^*, \cos(\theta_B), \text{ and } \cos(\theta_H) \text{ distributions of the } B^0 \to \rho^0 \gamma \text{ signal MC.}$	118
7.5	$m_{ES}, \Delta E^*, \cos(\theta_B), \text{ and } \cos(\theta_H) \text{ distributions for the } B^0 \to \rho^0 \gamma \text{ signal MC} \dots$	119
7.6	$m_{ES}, \Delta E^*, \cos(\theta_B), \text{ and } \cos(\theta_H) \text{ distributions for the } B^+ \to \rho^+ \gamma \text{ signal MC}$	122
7.7	$m_{ES}, \Delta E^*, \cos(\theta_B), \text{ and } \cos(\theta_H) \text{ distributions of the } B^0 \to \rho^0 \gamma \text{ signal MC}$	123
7.8	m_{ES} , ΔE^* , $\cos(\theta_B)$, $\cos(\theta_D)$, and $\cos(\theta_H)$ distributions for the $B^0 \to \omega \gamma$ signal	
	$MC \dots \dots$	126
9.1	$K^{*0} \rightarrow K^+ \pi^-$ Component Fit to Monte Carlo · Signal	149
0.2	$K^{*0} \rightarrow K^+ \pi^-$ Component Fit to Monte Carlo : Continuum Background	142
9.2 0.3	$K^{*0} \rightarrow K^+ \pi^-$ Component Fit to Monte Carlo : Continuum Dackground	145
9.5	$K^{*0} \rightarrow K^{+} \pi^{-}$ Full Fit to SP4 Monte Carle: Weighted sum of all components	140
9.4	$K \to K^+ \pi^-$ Pure toy Monte Carlo: Distributions of N=2, and $\alpha = \pi^-$	151
9.5 9.6	$K^{*0} \to K^+\pi^-$ Pure toy Monte Carlo : Enlarged distributions of $\sigma(N_{BBkg})$ and $\sigma(N_{BBkg})$ and	101
	$\sigma(a_{CP,ContBkg})$	152
9.7	$K^{*0} \rightarrow K^+ \pi^-$ Embedded MC : Component Yields for High Statistics Toys	157
9.8	$K^{*0} \to K^+ \pi^-$: CP fit to onpeak RunI + RunII data	158
9.9	A_{CP} Goodness of fit test for $K^{*0} \to K^+ \pi^-$	159
9.10	PDFs for the signal in $B^0 \to \rho^0 \gamma$ decay mode	163
9.11	PDFs for the signal in $B^+ \to \rho^+ \gamma$ decay mode	164
9.12	PDFs for the signal in $B^0 \rightarrow \omega \gamma$ decay mode	165
9.13	1 D D D D D D D D D D D D D D D D D D D	100
	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167
9.14	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168
$\begin{array}{c} 9.14\\ 9.15\end{array}$	PDFs for continuum background in $B^0 \rightarrow \rho^0 \gamma$ decay mode	167 168 169
$9.14 \\ 9.15 \\ 9.16$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171
$9.14 \\ 9.15 \\ 9.16 \\ 9.17$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172
9.14 9.15 9.16 9.17 9.18	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172 174
$\begin{array}{c} 9.14 \\ 9.15 \\ 9.16 \\ 9.17 \\ 9.18 \\ 9.19 \end{array}$	PDFs for the $B \to K^* \gamma$ background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172 174 175
$\begin{array}{c} 9.14 \\ 9.15 \\ 9.16 \\ 9.17 \\ 9.18 \\ 9.19 \\ 9.20 \end{array}$	PDFs for the $B \to K^* \gamma$ background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172 174 175 176
$\begin{array}{c} 9.14\\ 9.15\\ 9.16\\ 9.17\\ 9.18\\ 9.19\\ 9.20\\ 9.21\\ \end{array}$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172 174 175 176
$\begin{array}{c} 9.14\\ 9.15\\ 9.16\\ 9.17\\ 9.18\\ 9.19\\ 9.20\\ 9.21\\ 9.22\\ \end{array}$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 168 169 171 172 174 175 176 178
9.14 9.15 9.16 9.17 9.18 9.19 9.20 9.21 9.22	PDFs for the $B \to K^* \gamma$ background in $B^0 \to \rho^0 \gamma$ decay mode	167 167 168 169 171 172 174 175 176 178 179
9.14 9.15 9.16 9.17 9.18 9.19 9.20 9.21 9.22 9.22	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 167 168 169 171 172 174 175 176 178 179 185
$\begin{array}{c} 9.14\\ 9.15\\ 9.16\\ 9.17\\ 9.18\\ 9.19\\ 9.20\\ 9.21\\ 9.22\\ 9.22\\ 9.23\\ 9.24 \end{array}$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 167 168 169 171 172 174 175 176 178 178 179 185 187
$\begin{array}{c} 9.14\\ 9.15\\ 9.16\\ 9.17\\ 9.18\\ 9.19\\ 9.20\\ 9.21\\ 9.22\\ 9.22\\ 9.23\\ 9.24\\ 9.25\end{array}$	PDFs for continuum background in $B^0 \to \rho^0 \gamma$ decay mode	167 167 168 169 171 172 174 175 176 178 179 185 187 188

9.26	Projection plots of the unblinded results for $B^0 \to \rho^0 \gamma$ decay mode
9.27	Projection plots of the unblinded results for $B^+ \to \rho^+ \gamma$ decay mode
9.28	Projection plots of the unblinded results for $B^0 \rightarrow \omega \gamma$ decay mode $\ldots \ldots \ldots 193$
9.29	$-\log L$ of the unblinded fit compared to $-\log L$ from toy MC for each decay mode 194
10.1	Multiple-candidate events in with $m_{ES} < 5.26 \text{ GeV}/c^2$ fitted with a Gaussian plus
	2nd order polynomial
10.2	Events in the fit region with $m_{ES} < 5.26 \text{ GeV}/c^2$ fitted with a Gaussian plus 2nd
	order polynomial
10.3	Plots of systematics-convolved likelihood functions
11.1	Results of combined $B \to \rho \gamma$, $B^0 \to \omega \gamma$ fit to data
12.1	$ V_{td}/V_{ts} $ limit on CKM fitter winter 2004 results $\dots \dots \dots$

List of Tables

$1.1 \\ 1.2$	Next to leading order calculation of measured parameters $\ldots \ldots \ldots \ldots 3$ Current experimental measurements for $B \to K^* \gamma$, $B \to \rho \gamma$, and $B^0 \to \omega \gamma \ldots 4$
2.1	Production cross-sections at $\sqrt{s} = M(\Upsilon(4S))$
$3.1 \\ 3.2 \\ 3.3$	Exclusive decay modes39Monte Carlo samples used in this analysis with total amounts used47Release-12 series SP5 Monte Carlo and off-peak data used in this analysis48
6.1	Summary of the neural net configurations
7.1 7.2 7.3	Background yields for $K^{*0} \rightarrow K^+\pi^-$
$7.4 \\ 7.5 \\ 7.6$	Yields of the peaking background components for the $B^0 \rightarrow \rho^0 \gamma$ decay mode 116 Yields of the peaking background components for the $B^+ \rightarrow \rho^+ \gamma$ decay mode 120 Yields of the peaking background components for the $B^0 \rightarrow \omega \gamma$ decay mode 124
$\begin{array}{c} 8.1 \\ 8.2 \\ 8.3 \\ 8.4 \\ 8.5 \\ 8.6 \\ 8.7 \\ 8.8 \end{array}$	Efficiency of the cuts for $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ signal
9.1	$K^{*0} \to K^+\pi^-$ PDF parameters fit to fully simulated Monte Carlo and off resonance data
9.2 9.3	$K^{*} \rightarrow K^{+}\pi^{-}$ values of Fit Farameters for CF Asymmetry in from pure toy Monte Carlo Study
$9.4 \\ 9.5$	Correlations between m_{ES} and ΔE^*

9.6	$K^{*0} \to K^+ \pi^-$ Toy Monte Carlo yield efficiencies
9.7	Comparison of results from branching fraction and A_{CP} fits to data
9.8	Treatment of the normalization and shape parameters in the fit
9.9	Functions used to build the final PDF 161
9.10	All fixed or floating parameters in the final likelihood fit for $B^0 \rightarrow \rho^0 \gamma$ decay mode181
9.11	All fixed or floating parameters in the final likelihood fit for $B^+ \to \rho^+ \gamma$ decay mode 182
9.12	All fixed or floating parameters in the final likelihood fit for $B^0 \rightarrow \omega \gamma$ decay mode 183
9.13	Expected yields of each component
9.14	Mean and sigma of the pull distributions
9.15	Estimated sensitivity from the likelihood fit
9.16	$B^0 \rightarrow \rho^0 \gamma$ Signal-embedded toy parameters
9.17	$B^+ \rightarrow \rho^+ \gamma$ Signal-embedded toy parameters
9.18	$B^0 \rightarrow \omega \gamma$ Signal-embedded toy parameters
9.19	Fitted yields for Run1-4 data set and the MC expectations
10.1	$K^{*0} \to K^+ \pi^-$ Branching Fraction $B\bar{B}$ Systematics: Toy MC fit result summary 207
10.2	$K^{*0} \to K^+\pi^-$ Branching Fraction $\alpha_{\Delta E^*}$ Systematics: Toy MC fit result summary 208
10.3	Summary of all systematic uncertainties associated with the likelihood fit for each
	decay mode
10.4	Summary of all systematic uncertainties associated with the likelihood fit \ldots 211
10.5	Average candidate multiplicity
10.6	Efficiency corrections and systematic uncertainties for $B^0 \to K^{*0}\gamma$ 217
10.7	Systematic uncertainties on $\mathcal{B}(B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-))$ and on $A_{CP}(B^0 \to K^+\pi^-)$
	$K^{*0}\gamma(K^{*0} \to K^+\pi^-)).$ \ldots 218
10.8	Summary of all systematic uncertainties for each decay mode
10.9	Summary of $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ measurement results
11.1	Unblinded signal yield and the branching fraction for each decay mode on Run1
	and Run2 data
11.2	Unblinded A_{CP} for each decay mode on Run1 and Run2 data
11.3	Systematic uncertainties on the branching fractions $\mathcal{B}(B \to K^* \gamma)$ and on CP
	asymmetries $A_{CP}(B \to K^* \gamma)$
11.4	Correlated systematics of the four $B \to K^* \gamma$ modes $\dots \dots \dots$
11.5	Covariance matrix of the $B \to K^* \gamma$ measurements
11.6	Systematic covariance matrix of the $B \to K^* \gamma$ measurements
11.7	Correlated systematics of the three A_{CP} modes $\ldots \ldots 237$
11.8	Covariance matrix of the A_{CP} measurements
11.9	Systematic covariance matrix of the A_{CP} measurements $\ldots \ldots \ldots \ldots \ldots 237$
11.1(OSummary of the continuum and combinatoric B background yields in $B \to (\rho/\omega)\gamma$ 240
11.11	Break down of the systematic uncertainties on the signal efficiency for each decay
	mode
11.12	2 Summary of all the systematic uncertainties for the combined effective signal yield. 244
12.1	Experimental measurements presented in this thesis
12.2	Next to leading order calculation of measured parameters
	· ·

Abstract

Measurements of the branching fraction, CP asymmetry and isospin asymmetry of the radiative penguin decay $B^0 \to K^{*0}\gamma$ and a search for the decays $B \to \rho\gamma$

and $B^0 \to \omega \gamma$ at BABAR

by

Patrick Michael Spradlin

Radiative penguin decays of B mesons provide a fertile ground for precision tests of the Standard Model. Because such decays must proceed through 1-loop or higher processes in Standard Model perturbation theory, they are quite rare and their amplitudes are particularly susceptible to interference from physics beyond the Standard Model. This thesis presents measurements carried out at the *BABAR* experiment of the branching fraction \mathcal{B} , CP asymmetry parameter A_{CP} and isospin asymmetry parameter Δ_{0-} of the radiative penguin decay $B^0 \to K^{*0}\gamma$. The results of these measurements are

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = (3.92 \pm 0.20 \pm 0.24) \times 10^{-5}$$
$$A_{CP}(B \to K^*\gamma) = -0.013 \pm 0.036 \pm 0.010$$
$$\Delta_{0-}(B \to K^*\gamma) = 0.050 \pm 0.045 \pm 0.037.$$

The measurements are consistent with Standard Model predictions, but do not rule out future discovery of non-Standard Model deviations with an enlarged data set. This thesis also presents a related search for the radiative penguin decays $B^0 \rightarrow \rho^0 \gamma$, $B^+ \rightarrow \rho^+ \gamma$ and $B^0 \rightarrow \omega \gamma$. These decays have yet to be observed at any experiment. This analysis sets 90% confidence upper limits on their branching fractions of

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

where $\mathcal{B}(B \to (\rho/\omega)\gamma)$ is a combined limit from a search for decays in any of the three modes. The $\mathcal{B}(B \to (\rho/\omega)\gamma)$ measurement assumes theoretical isospin relationships among $\mathcal{B}(B^0 \to \rho^0 \gamma), \mathcal{B}(B^+ \to \rho^+ \gamma), \text{ and } \mathcal{B}(B^0 \to \omega \gamma)$. The ratio of the branching fractions $\mathcal{B}(B \to (\rho/\omega)\gamma)$ and $\mathcal{B}(B \to K^*\gamma)$ are related to the CKM matrix ratio $|V_{td}/V_{ts}|$. The measurements presented in this thesis give a 90% confidence level upper limit of

$$|V_{td}/V_{ts}| < 0.190.$$

Chapter 1

Motivation

This thesis details measurements of the decays $B \to K^*\gamma$, $B \to \rho\gamma$, and $B^0 \to \omega\gamma$ carried out at the BABAR experiment and published in [1] and [2]. The physics of these decays are closely related. In the Standard Model (SM), the leading order contributions to their amplitudes are flavor changing neutral quark currents $b \to s\gamma$ or $b \to d\gamma$. Such transitions cannot proceed through tree-level SM weak processes; they must involve one loop or higher order diagrams, the leading of which is depicted in Figure 1.1. Such one loop diagrams have become commonly referred to as penguin diagrams. Decays dominated by such processes, such as $B \to K^*\gamma$ and $B \to \rho(\omega)\gamma$, are called radiative penguin decays.



Figure 1.1: The leading order Feynman diagram for the $b \to s \gamma, b \to d \gamma$ process.

Radiative penguin decays of B mesons have garnered increasing theoretical attention as measurements of them have become more precise. Most interesting is the potential for discovery of effects in the $b \to s\gamma$ and $b \to d\gamma$ transition amplitudes that cannot be accounted for in the Standard Model (new physics contributions). Many theoretical extensions to the Standard Model, including supersymmetric theories, involve new particles and couplings that lead to one-loop contributions similar to that in Figure 1.1 but with non-SM particles in the loop. The absence of a dominant tree level contribution and the large mass of the loop propagators make the amplitudes of penguin transitions highly susceptible to interference from such new physics interactions. A less exotic, but nonetheless valuable, use of radiative penguin measurements provides independent determinations of Standard Model parameters. Most important among the SM parameters measureable in radiative penguin physics is the ratio of CKM matrix elements $|V_{td}/V_{ts}|$, which is currently known only indirectly from unitarity constraints on the CKM elements.

1.1 Measured observables

Ideally, one would make direct measurements of $b \to s\gamma$ or $b \to d\gamma$ transitions. Quark confinement makes this impossible. One is forced, instead, to make measurements of hadronic manifestations of quark transitions such as $B \to K^*\gamma$ or $B \to \rho\gamma$. Theoretic calculations on such hadronic decays are complicated by non-perturbative QCD effects leading to large theoretical uncertainties. Since most of most of the theoretical uncertainty involves hadronization, or the arrangement of the quarks into bound states after the $b \to s(d)$ transition, it can be significantly reduced if one integrates hadronization effects out by considering inclusive decays, e.g. $B \to X_s\gamma$, where X_s represents all allowable hadronic final states. Such calculations and measurements have been performed (e.g. [3], [4]), but this thesis is restricted to exclusive decays.

The basic experimental observables of interest in studies of exclusive B decays are the branching fractions $\mathcal{B}(B \to X)$, or the fractions of all B mesons that will decay to the final

Calculations			
$\mathcal{B}(B^0 \to K^{*0}\gamma)(\times 10^{-5})$	$7.09^{+2.47}_{-2.27}$ [5]	$7.6^{+3.5}_{-3.0}$ [6]	7.0 ± 2.7 [7]
$\mathcal{B}(B^+ \to K^{*+} \gamma) (\times 10^{-5})$	$7.45^{+2.47}_{-2.27}$ [5]	$8.1^{+3.5}_{-3.0}$ [6]	7.4 ± 2.7 [7]
$A_{CP}(B \to K^* \gamma)$	$< 1\% \ [5, 8]$		
$\Delta_{0-}(B \to K^* \gamma)$	$(8.0^{+2.1}_{-3.2})\% \times (0.3/T_1^{B \to K^*})$ [9]		
$\mathcal{B}(B^0 \to \rho^0 \gamma) (\times 10^{-6})$	$0.49 \pm 0.18 (\text{th}) \pm 0.04 (\text{ex})$ [7]	$0.66 \pm 0.20 \; [10]$	$0.76^{+0.26}_{-0.23}$ [5]
$\mathcal{B}(B^+ \to \rho^+ \gamma) (\times 10^{-6})$	$0.90 \pm 0.33(\text{th}) \pm 0.10(\text{ex})$ [7]	1.35 ± 0.42 [10]	$1.58^{+0.53}_{-0.46}$ [5]

Table 1.1: Next to leading order theoretical predictions of parameters for the $B \to K^* \gamma$ and $B \to \rho \gamma$ decays. Columns represent different published calculations for each observable. The source of each value is referenced in square brackets [] beside it. The value of $\mathcal{B}(B^0 \to \omega \gamma)$ is theoretically predicted by isospin symmetry to be the same as $\mathcal{B}(B^0 \to \rho^0 \gamma)$. The factor $T_1^{B \to K^*}$ in the prediction for Δ_{0-} is a form factor which is the dominant source of the theoretical error. Determinations of $T_1^{B \to K^*}$ include $0.32^{+0.04}_{-0.02}$ ([11]), 0.38 ± 0.06 ([12]), and 0.27 ± 0.04 ([6]).

states X. Branching fractions are simply related to the partial widths or partial transition rates $\Gamma(B \to X)$, which are the more commonly theoretically calculated observables:

$$\mathcal{B}(B \to X) = \frac{\Gamma(B \to X)}{\Gamma(B)} = \tau_B \Gamma(B \to X) , \qquad (1.1)$$

where $\Gamma(B) = 1/\tau_B$ is the *B* full width or decay rate, and τ_B is the mean *B* lifetime. As evidence of the challenge posed by theoretical calculations of these values, measurements of the branching fraction $\mathcal{B}(B^0 \to K^{*0}\gamma)$ are already more than twice as precise as theoretical calculations for the same quantity (Tables 1.1 and 1.2).

No measurement of any of the branching fractions $\mathcal{B}(B^0 \to \rho^0 \gamma)$, $\mathcal{B}(B^+ \to \rho^+ \gamma)$, or $\mathcal{B}(B^0 \to \omega \gamma)$ has yet produced a statistically significant result. This has led to combined measurements on a union of the individual decay modes. The hope is that, although the number of events observed for any single decay mode is not statistically significant, the sum of the number of events in the three modes together may be. The result of such a combined measurement is usually quoted as a combined branching fraction defined in terms of the three individual branching fractions as

$$\mathcal{B}(B \to (\rho, \omega)\gamma) \equiv \frac{1}{2} \cdot \left[\mathcal{B}(B^+ \to \rho^+ \gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot \left(\mathcal{B}(B^0 \to \rho^0 \gamma) + \mathcal{B}(B^0 \to \omega\gamma)\right)\right].$$
(1.2)

Experiment	Previous BABAR[13]	BELLE $[14]$	This thesis [1]
$B\overline{B}$ pairs	$23 imes 10^6$	$85 imes 10^6$	88×10^6
Branching Fract	ions (×10 ⁻⁵)		
$\mathcal{B}(B^0 \to K^{*0}\gamma)$	$4.23 \pm 0.40 \pm 0.22$	$4.01 \pm 0.21 \pm 0.17$	$3.92 \pm 0.20 \pm 0.24$
$\mathcal{B}(B^+ \to K^{*+}\gamma)$	$3.83 \pm 0.62 \pm 0.22$	$4.25 \pm 0.31 \pm 0.24$	$3.87 \pm 0.28 \pm 0.26$
Asymmetries			
$A_{CP}(B \to K^* \gamma)$	$-0.044 \pm 0.076 \pm 0.012$	$-0.015\pm0.044\pm0.012$	$-0.013 \pm 0.036 \pm 0.010$
$\Delta_{0-}(B \to K^* \gamma)$	N/A	$0.012 \pm 0.044 \pm 0.026$	$0.050 \pm 0.045 \pm 0.037$
Experiment	Previous BABAR[15]	BELLE [16]	This thesis [2]
$B\overline{B}$ pairs	84×10^6	274×10^6	211×10^6
Branching Fract	ions (×10 ⁻⁶)		
$\mathcal{B}(B^0 o ho^0 \gamma)$	< 1.2	< 0.8	< 0.4
$\mathcal{B}(B^+ \to \rho^+ \gamma)$	< 2.1	< 2.2	< 1.8
$\mathcal{B}(B^0\to\omega\gamma)$	< 1.0	< 0.8	< 1.0
Measurement	Limit	from $\Delta M_{B_d} / \Delta M_{B_s}$ [17]	This thesis $[2]$
$ V_{td}/V_{ts} $		< 0.25	< 0.195

Table 1.2: Current experimental measurements for $B \to K^*\gamma$, $B \to \rho\gamma$, and $B^0 \to \omega\gamma$. Values for $B \to K^*\gamma$ are quoted as (central value)±(statistical uncertainty)±(systematic uncertainty). The upper limits for $\mathcal{B}(B \to \rho\gamma)$ and $\mathcal{B}(B^0 \to \omega\gamma)$ are quoted at 90% confidence level. The publications in which the measurements are published are cited in square brackets []. The rightmost column lists the results of the measurements described in this thesis.

Isospin symmetry implies the relationship

$$\Gamma(B^0 \to \rho^0 \gamma) = \Gamma(B^0 \to \omega \gamma) = \frac{1}{2} \Gamma(B^+ \to \rho^+ \gamma) , \qquad (1.3)$$

which gives $\mathcal{B}(B \to (\rho, \omega)\gamma)$ a theoretical value equal to $\mathcal{B}(B^+ \to \rho^+\gamma)$. Unfortunately, even this combined approach has failed to yield statistically significant evidence for $b \to d\gamma$ transitions.

The impact of the large relative magnitudes of the theoretical branching fraction uncertainties can be reduced by considering quantities that are ratios of transition rates of exclusive decays. Many of the uncertain terms associated with the non-perturbative physics are approximately factorizable and largely cancel in appropriate ratios. Two of these ratios—the CP-asymmetry parameter A_{CP} and the isospin violation parameter Δ_{0-} —have been particularly studied by theorists as possible signatures of new physics ([8, 9]). For the set of decays $B \to K^*\gamma$, these quantities are defined as

$$A_{CP} \equiv \frac{\Gamma(\overline{B} \to \overline{K}^* \gamma) - \Gamma(B \to K^* \gamma)}{\Gamma(\overline{B} \to \overline{K}^* \gamma) + \Gamma(B \to K^* \gamma)}$$
(1.4)

and

$$\Delta_{0-} \equiv \frac{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) - \Gamma(B^- \to K^{*-} \gamma)}{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) + \Gamma(B^- \to K^{*-} \gamma)}.$$
(1.5)

1.2 $|V_{td}/V_{ts}|$ and the unitarity triangle

The cancellation of theoretical errors in ratios also improves the chance for a direct measurement of the ratio of CKM elements $|V_{td}/V_{ts}|$. The CKM ratio is related to branching fractions for exclusive radiative penguin decays of *B* mesons by ([7])

$$\frac{\Gamma(B^+ \to \rho^+ \gamma)}{\Gamma(B^+ \to K^{*+} \gamma)} = \frac{\mathcal{B}(B^+ \to \rho^+ \gamma)}{\mathcal{B}(B^+ \to K^{*+} \gamma)} = \left|\frac{V_{td}}{V_{ts}}\right|^2 \left(\frac{1 - m_{\rho}^2 / M_B^2}{1 - m_{K^*}^2 / M_B^2}\right)^3 \zeta^2 [1 + \Delta R] , \qquad (1.6)$$

where m_{ρ} , m_{K^*} and M_B are respectively the masses of the ρ^+ , K^{*+} and B^+ mesons, ζ is the ratio of the transition form factors and ΔR parameterizes the remaining small calculable dynamical differences. This relation assumes that interference from physics beyond the Standard Model is negligible. The same relationship is expected to hold for $2\Gamma(B^0 \to \rho^0 \gamma)/\Gamma(B^0 \to K^{*0} \gamma)$, where the additional factor of 2 is the isospin factor from Equation 1.3.

A direct measurement of $|V_{td}/V_{ts}|$ is important for confirming the unitarity of the CKM matrix. A brief digression is required to explain the presentation conventions for CKM unitarity tests in *B* physics (see [17] for a more complete review). Unitarity in the *B* sector is most commonly presented in terms of the Wolfenstein approximate parameterization of the CKM matrix *V* ([18])

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.7)$$

Unitarity of V implies, among other relations, that

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 , (1.8)$$

which can be graphically represented by a triangle in the complex plane, as shown in Figure 1.2 (a). Usually, relation 1.8 is divided by $V_{cd}V_{cb}^*$ to yield the similar triangle in Figure 1.2 (b), where $\overline{\rho} = \rho(1 - \lambda^2/2)$ and $\overline{\eta} = \eta(1 - \lambda^2/2)$. *B* physics literature usually refers to Figure 1.2 (b) as the unitarity triangle.

Independently measuring each of unitarity triangle's angles and sides is a test of the unitarity of the CKM matrix. The CKMfitter group [19] maintains a global fit of the unitarity triangle to relevent measurements. Their results at the time of the publication of [2] appear in Figure 1.3. Still missing from this picture is a direct determination of the length of \overline{AB} , which can be provided by a direct measurement of $|V_{td}/V_{ts}|$. Conventional wisdom suggests that the leading experimental method for directly measuring this quantity is a ratio of the neutral B meson mass splittings $\Delta M_{B_s} = m_{B_s^0} - m_{\overline{B}_s^0}$ and $\Delta M_{B_d} = m_{B_d^0} - m_{\overline{B}_d^0}$. However, the B_s mixing

measurement to determine ΔM_{B_s} has proven difficult and has yielded only a limit on $|V_{td}/V_{ts}|$ [17] (see Table 1.2). The branching fraction ratio in Equation 1.6 represents an independent and increasingly promising method for a direct measurement.



Figure 1.2: (a) Representation in the complex plane of the triangle formed by the CKM matrix elements $V_{ud}V_{ub}^*$, $V_{td}V_{tb}^*$, and $V_{cd}V_{cb}^*$. (b) Rescaled triangle with vertices A, B and C at $(\overline{\rho}, \overline{\eta})$, (1,0), and (0,0) respectively. Figure and caption taken from [17].

1.3 Theoretical framework

Theoretical values for each of the observables have been calculated to next to leading order (NLO) in the Standard Model. For immediate reference, NLO calculations for the branching fraction $B \to K^* \gamma$ are presented in [5, 6, 7]. The branching fractions $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ are



Figure 1.3: Winter 2004 results of the CKM fitter group's global CKM fit of relevant experimental results to the unitarity triangle. Figure taken from [19].

calculated in [5, 10]. The paper [9] calculates the isospin violation parameter $\Delta_{0-}(B \to K^*\gamma)$ to NLO and analyzes its potential as a signature of non-Standard Model physics. The new physics potential for the CP asymmetry parameter $A_{CP}(B \to X_s\gamma)$ is similarly treated with an NLO calculation in [8]. This section subsequently presents a brief outline of the theoretic branching fraction calculations as presented in [5].

The basic calculational framework is an operator product expansion (OPE) of the Standard Model. The OPE for B meson decays is an effective field theory that replaces non-local Standard Model interactions with highly virtual propagators—which necessarily take place over small distances—with a set of effective local operators O_n . The details of the short distance physics thus replaced—most significantly W^{\pm} and Z exchange, t quark propagation, and hard gluon exchange—are contained within the effective coupling strengths C_n , called Wilson coefficients, of these new local operators. For example, the t quark penguin loop (Figure 1.1), which is the leading order contributor to the $b \to s\gamma$ decay, is absorbed into the OPE operator O_7 as shown in Figure 1.4. The set of B decay OPE operators relevant to the $B \to K^*\gamma$ calculations are depicted in Figure 1.5 and listed below:

$$O_{1}^{p} = \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})p_{\beta}\bar{p}_{\beta}\gamma^{\mu}(1-\gamma_{5})b_{\alpha}$$

$$O_{2}^{p} = \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})p_{\alpha}\bar{p}_{\beta}\gamma^{\mu}(1-\gamma_{5})b_{\beta}$$

$$O_{3} = \sum_{q} \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})b_{\alpha}\bar{q}_{\beta}\gamma^{\mu}(1-\gamma_{5})q_{\beta}$$

$$O_{4} = \sum_{q} \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})b_{\beta}\bar{q}_{\beta}\gamma^{\mu}(1-\gamma_{5})q_{\alpha}$$

$$O_{5} = \sum_{q} \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})b_{\alpha}\bar{q}_{\beta}\gamma^{\mu}(1+\gamma_{5})q_{\beta}$$

$$O_{6} = \sum_{q} \bar{s}_{\alpha}\gamma_{\mu}(1-\gamma_{5})b_{\beta}\bar{q}_{\beta}\gamma^{\mu}(1+\gamma_{5})q_{\alpha}$$

$$O_{7} = \frac{e}{8\pi^{2}}m_{b}\bar{s}\sigma^{\mu\nu}(1+\gamma_{5})F_{\mu\nu}$$

$$O_{8} = \frac{g}{8\pi^{2}}m_{b}\bar{s}\sigma^{\mu\nu}(1+\gamma_{5})T^{a}bG_{\mu\nu}^{a},$$
(1.9)



Figure 1.4: Penguin diagram contribution to O_7



Figure 1.5: Local operators of the OPE effective theory for B meson decay.

where α, β are color indices, $p \in u, c$ in $O_{1,2}$, and the sum in O_{3-6} over quark flavors $q \in u, d, s, c, b$. Assembling these components, the effective Hamiltonian for $B \to K^* \gamma$ decays can be written

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{ps}^* V_{pb} \left[\sum_{p=u,c} (C_1 O_1^p + C_2 O_2^p) \sum_{i=3,\dots,8} C_i O_i \right].$$
(1.10)

The effective theory for $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ decays is obtained from Equations 1.9 and 1.10 by the substitution $s \to d$ in the operator quark currents and CKM factors. Operators O_{1-6} are often called *four-fermion* operators with $O_{1,2}$ additionally termed *current-current* operators. O_7 and O_8 are respectively called the *electromagnetic* and *chromomagnetic penguin* operators.

The Wilson coefficients C_i are independent of the initial and final state wave functions and have been calculated perturbatively to next to leading order. They obey renormalization group equations that allow matching to the scale at which operator matrix elements are calculated. These mesonic matrix elements of O_i necessarily involve nonperturbative physics. Bosch and Buchalla [5] demonstrate that the remaining perturbative physics in the operator matrix elements can be further factorized from the nonperturbative physics in the heavy quark $(m_b \to \infty)$ limit by

$$\langle V\gamma|O_i|\overline{B}\rangle = \left[F^{B\to V}(0)T_i^I + \int_0^1 d\xi \,dv \,T_i^{II}(\xi,v)\Phi_B(\xi)\Phi_V(v)\right] \cdot \epsilon , \qquad (1.11)$$

where ϵ is the photon polarization four-vector. The long distance nonperturbative physics is contained in the $B \to V$ transition form factor $F^{B\to V}$ and the leading twist light-cone distributions of the mesons Φ_B and Φ_V . These have been calculated using QCD sum rules in [12]. The remaining perturbative physics contribute to the hard-scattering kernels T_i^I and T_i^{II} . Equation 1.11 is valid to corrections of the order Λ_{QCD}/m_b .

Recent calculations in [8] and [9] have explored possible deviations from the Standard Model values for A_{CP} and Δ_{0-} respectively in some of the more popular supersymmetric extensions to the Standard Model. Such models do not expand the set of operators in the OPE of the model as the additional supersymmetric particles have large masses that are integrated out of the expansion like the t quark and W. However, they do modify the Wilson coefficients C_i . In these models the differences most likely to affect the asymmetries A_{CP} and Δ_{0-} are modifications to the coefficients for the electromagnetic penguin operator C_7 and the chromomagnetic penguin operator C_8 . Such new physics effects can have quite a dramatic effect on A_{CP} and Δ_{0-} , creating $A_{CP} \sim 10\%$ in one scenario calculated in [8].

1.4 State of the measurements

The first evidence for any radiative penguin decay of a B meson was an observation of $B \rightarrow K^* \gamma$ published by the CLEO collaboration in 1993 [20]. Since then, the measurements have been and continue to be refined with increasingly large data sets. The latest measurements from

the BABAR and Belle collaborations, which are still continuing to take data, are summarized in Table 1.2.

The decays $B \to K^* \gamma$ are experimentally well-established and their branching ratios $\mathcal{B}(B^0 \to K^{*0}\gamma)$ and $\mathcal{B}(B^+ \to K^{*+}\gamma)$ are fast becoming precision measurements, with uncertainties of < 10%. Their central values appear lower than the theoretically calculated values (Table 1.1), however the theoretical uncertainties are still large enough to accommodate the measurements. Measurements of $A_{CP}(B \to K^*\gamma)$ and Δ_{0-} are also consistent with the Standard Model calculations and constrain new physics effects. It should be noted that the uncertainties of A_{CP} and Δ_{0-} are still limited by the statistics of their data sets and can significantly improve, thus further constraining new physics models.

Experimental evidence for the decays $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ remains elusive. However, the current upper bounds are beginning to encroach upon the Standard Model predictions. The next iteration of measurements from Belle and *BABAR* are greatly anticipated and have an excellent probability of establishing first evidence for the existence of the decays or forcing a reevaluation of the calculations.

The analysis presented in this thesis represents the most recent results published by the BABAR collaboration. It is organized as follows: Chapter 2 describes the BABAR experiment, which provides the data for the measurements. Chapter 3 provides an overview of the measurement strategy, which is subsequently detailed in Chapters 4-11. The final results of the measurements from Chapters 10 and 11 are recapitulated with concluding remarks in Chapter 12.

Chapter 2

The B_{AB} experiment

The data upon which the present measurements were based were collected by the BABAR experiment. BABAR is a general purpose solenoidal particle detector observing positronelectron (e^+e^-) collisions in the PEP-II (Positron- Electron Project II) [21] at the Stanford Linear Accelerator Center (SLAC). BABAR also refers to the international collaboration of universities and research institutions that operate the detector and collect and analyze its data.

BABAR has published a complete description of the detector as [22]. The bulk of this chapter is condensed from that excellent paper. The design goals of the detector and the initial prospective physics program for the experiment appear in [23].

The principal goal of the *BABAR* experiment is a comprehensive study of CP-violation in *B* meson decay processes, represented by unitarity triangle (Figure 1.2). PEP-II provides an abundant supply of *B* mesons by colliding positrons and electrons at the $\Upsilon(4S)$ resonance energy in the collision center-of-momentum (CM) reference frame ($\sqrt{s} = 10.58$ GeV). $\Upsilon(4S)$ decays almost exclusively to a pair of *B* mesons (B^+B^- or $B^0\overline{B}^0$). To facilitate CP-asymmetry measurements, the CM frame of the e^+e^- collision is boosted with respect to the laboratory frame of the detector. Because the *B* daughters of the $\Upsilon(4S)$ are produced almost at rest in the CM frame, this boost transforms differences in their decay times in the CM frame into displacements of their decay points (termed decay vertices) in the lab frame. This is significant because CP-violation in a $B \to X$ decay manifests as a difference between the decay amplitudes of $B \to X$ and the conjugate decay process $\overline{B} \to \overline{X}$. Asymmetric e^+ and e^- beam energies accomplish the relative boost between the collision CM frame and the laboratory frame.

Although CP-violation studies motivated its construction, the BABAR experiment is an excellent environment in which to study a broad range of τ lepton and b and c quark physics. The rich supply of B mesons allows the studies of very rare B decays, such as the radiative penguin decays studied in this thesis. At the time the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses were published in [2], BABARhad recorded approximately $230 \times 10^6 \ B\overline{B}$ events. This large set of data is coarsely divided into periods of data collection called Runs, which represent many months of continuous data collection separated from each other by periods of scheduled detector maintenance and improvement. The $B^0 \to K^{*0}\gamma$ measurement presented in this thesis is based on the $88 \times 10^6 \ B\overline{B}$ events of Runs I and II. The $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ measurements are based on $211 \times 10^6 \ B\overline{B}$ events from Runs I-IV. As this thesis is written, Run V is under way; BABAR continues to take data that will lead to future refinements of the measurements presented here.

2.1 PEP-II

Figure 2.1 depicts the PEP-II storage ring and the linear electron/positron accelerator (linac), which gives SLAC its name. Electrons and positrons are accelerated to their collision energies—9.0 GeV for e^- and 3.1 GeV for e^+ —in the linac and injected into the PEP-II rings where the beams are stored, tuned, and brought into collision. The electron and positron beams are stored in separate counterrotating rings. The asymmetric energies of the particles give their head–on collisions the specified 10.58 GeV CM energy and the CM frame boost $\beta\gamma = 0.56$ relative to the laboratory frame.

While in the PEP-II storage rings, the beams are tuned to minimize the spatial extent of the interaction region, to ensure that its position remains stable, and to optimize the luminosity. In the BABAR coordinate system, the typical RMS Cartesian extent of the luminous region is $\sigma_{L_x} = 120 \,\mu\text{m}, \, \sigma_{L_y} = 5.6 \,\mu\text{m}, \text{ and } \sigma_{L_z} = 9 \,\text{mm}$ [22]. Figures 2.2 and 2.3 diagram the BABAR detector and define its coordinate system. The positive z direction is along the cylindrical axis of the detector in the direction of the e^- momentum. The positive y direction is vertically upward, and the x axis completes the right-handed coordinate system.

Through continued improvements to the hardware and operations, the instantaneous luminosity supplied by PEP-II has steadily improved since operations began in 2001. At the time of the publication of the $B \rightarrow \rho \gamma$ measurements presented in this thesis ([2]), PEP-II regularly supplied over twice the original design luminosity of $3 \times 10^{33} \text{ cm}^{-2} \text{s}^{-1} = 3 \text{ nb}^{-1} s^{-1}$. Its total time-integrated supplied luminosity had reached 256 fb⁻¹.

Table 2.1 lists the quark and leptonic production cross-sections for e^+e^- collisions at the $\Upsilon(4S)$ resonance. Production of the $b\overline{b}$ quark pairs, which leads to the production of the desired $B\overline{B}$ meson pairs, competes with the production of light and charm quark pairs and with the production of charged leptons. The production of charmed and light quark pairs together with the production of $\tau^+\tau^-$ pairs are called collectively continuum production among *BABAR* physicists and in subsequent chapters of this thesis. Because of their relative abundance, continuum events are a significant background for almost any analysis of *B* mesons. To facilitate the characterization and suppression of these backgrounds, approximately 10% of the events produced by PEP-II (and recorded by *BABAR*) are produced with a collision CM energy approximately 10 MeV below the $\Upsilon(4S)$ resonance mass. This energy is beneath the $B\overline{B}$ production threshold, yet near enough resonance so that the continuum cross-sections are approximately equal to their values at the $\Upsilon(4S)$ resonance. This sample of off-resonance events also allows



Figure 2.1: Illustration of the SLAC linac and PEP-II storage ring.



Figure 2.2: BABAR detector longitudinal section [22]



Figure 2.3: BABAR detector end view [22]

the independent measurement of the number of B decays recorded at BABAR that is described in Chapter 10.1. Although the complete set of BABAR data includes both the off-resonance and on-resonance samples, unless otherwise stated, the term BABAR data in this thesis refers only to the on-resonance subset.

2.2 BABAR detector overview

BABAR is a cylindrically symmetric general-purpose solenoidal particle detector. Figures 2.2 and 2.3 depict the detector and indicate its major subsystems and dimensions. Because the CM frame (and hence all of the $\Upsilon(4S)$ and subsequent decay products) are boosted with respect to the laboratory frame, the longitudinal center of the entire detector is set 0.37 m behind (in the direction opposite of the boost to) the interaction point in order to maximize its geometric acceptance in the CM frame. In addition, most of the component subdetectors

Product	Cross-section
$e^+e^- \rightarrow$	(nb)
$b\overline{b}$	1.05
$c\overline{c}$	1.30
$u\overline{u}, d\overline{d}, s\overline{s}$	2.09
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 2.1: Production cross-sections at $\sqrt{s} = M(\varUpsilon(4S)) = 10.58 \; {\rm GeV} \; [23]$

have a forward-backward asymmetry with a more comprehensive laboratory frame polar angle acceptance in the boost (+z) direction.

The charged particle tracking and vertexing system occupies the volume closest to the interaction point (IP). This system is composed of two major subsystems: the silicon vertex tracker (SVT), which lies closest to the IP within a mechanical support tube, and the drift chamber (DCH), which surrounds the mechanical support tube. Proceeding radially outward, the next major detector subsystem is a novel ring-imaging Cherenkov detector (DIRC) followed by a CsI calorimeter (EMC). These are surrounded by the superconducting solenoid, which suffuses the inner detector with a magnetic field of 1.5T. The magnetic field bends the paths of charged particles into helices allowing measurement of the particles' momenta. The outer steel flux return for the solenoid field (IFR), a last major detector hardware subsystem, is instrumented for the detection of muons and neutral hadrons.

With the exception of the IFR, each of these detector subsystems is described with more detail in subsequent sections of this chapter. The primary purposes of the IFR are identifying muons and detecting neutral hadrons (mostly neutrons and K_L^0 mesons). Neither muons nor neutrons are part of the decays under study, nor do they produce significant backgrounds. While K_L^0 detection could be used in $B \to K^* \gamma$ analyses, measurements using K_L^0 would be much less precise than the same measurements using K_S^0 due to the significantly lower detection efficiency and less precisely measured momenta of the K_L^0 candidates. In contrast to those of the other subdetectors, IFR measurements are used only indirectly in the current analyses.

2.3 BABAR subsytem: SVT

The Silicon Vertex Tracker (SVT) is the innermost active section of the BABAR detector. It occupies a cylindrical radius of approximately 32 mm to 144 mm relative to the collision axis. Its main purpose is to provide precise measurements of charged particle positions very close to the interaction point. This role is crucial in the time-dependent CP-asymmetry measurements that drove many aspects of the BABAR design. These measurements probe the time structure of B meson decays by utilizing the CM frame boost, which transforms meson lifetimes to laboratory flight distances. With a boost of $\beta\gamma = 0.56$ and a mean lifetime $c\tau_{B^0} = 460 \pm 4 \,\mu\text{m}$ [17], the mean flight distance of a B^0 meson is approximately 260 μm . Analysis requires the determination of a B decay vertex from its charged decay products with a precision significantly smaller than this mean flight distance. The BABAR SVT was designed to provide estimates of reconstructed B decay vertices with resolutions better than 80 μm in z direction, and better than 100 μm in the transverse plane. In addition, measurements of the energy deposited by charged particles in the silicon of the SVT (dE/dx) contribute to charged particle identification.

The analyses described in this thesis employs this remarkable vertexing power of the SVT for the more prosaic purpose of background suppression. For the analysis modes with two or more charged final state particles, a reconstructed B candidate is analyzed to determine if the charged particles are consistent with the hypothesis that they emerge from the same decay vertex. The displacement of the calculated vertex is also analyzed for consistency with the flight distance of a B meson. The use of decay vertexes in the analysis are detailed in Chapter 4.5.2.

Figures 2.4 and 2.5 show the layout of the SVT. The active sensor components are 300 μ m double-sided silicon strip detectors, which are described briefly below. These are arranged in five coaxial layers covering a range of polar angles $\theta = [20^{\circ}, 150^{\circ}]$ with respect to the interaction point.

Many textbooks on detector physics (e.g., [24]) describe the theory and operation of semiconductor particle detectors in more detail than the account given here. The SVT silicon wafers are biased to deplete the conduction band of charge carriers. Interactions of charged particles with the silicon material create electron-hole pairs in the silicon. These drift with the applied electric field to the planar faces of the detector, where the charge is collected on detection strips. The strips on opposite sides of the wafers are orthogonal to give two-dimensional position measurements with a single detector wafer. The wafers are positioned so that the strips on one side are oriented parallel to the detector axis and measure the azimuthal cylindrical angle ϕ coordinate at which a charged particle crossed the detector (ϕ strips). The strips on the other side run transversely to the detector axis and measure the z position of the charged particle "hit" (z strips). The pitch (distance between centers of adjacent strips) of the z strips is 100 μ m on the innermost 3 layers and 210 μ m for the outer two layers. The ϕ strip pitch varies among the layers in the range 50-110 μ m. Hit position measurements in the SVT achieve sub-pitch resolutions, with z resolution 10-60 μ m and ϕ resolution 10-40 μ m depending on the layer and angle of the charged particle momentum with respect to the detector.

2.4 BABAR subsystem: DCH

The BABAR drift chamber (DCH) is the primary source of charged particle momentum and angle measurements. It occupies the cylindrical region from a radius of 236 mm, just outside the support tube which houses the SVT, to 809 mm, just inside the cylindrical DIRC



Figure 2.4: BABAR Silicon Vertex Tracker (SVT) longitudinal section. [22]



Figure 2.5: BABAR Silicon Vertex Tracker (SVT) transverse section. [22]
barrel. The DCH also provides measurements of the energy lost by charged particles passing through the DCH, dE/dx, which provides information about the particle's identity. The DCH measurements are the most powerful discriminant of charged particle types for particles with momenta below 700 MeV. Higher momenta charged hadrons rely more on measurements from the DIRC.

Almost all *BABAR* analyses rely on precise measurements of the momenta of charged particles, especially analyses, like those of this thesis, that proceed by reconstructing decay chains from a set of detected final state particles.

Figure 2.6 depicts a longitudinal cross section of the DCH. The detector volume is comprised of hexagonal drift cells, which span the entire length of the detector. These cells are arranged in 40 cylindrical layers, which are grouped into ten superlayers. Figure 2.7 shows a section of the innermost four superlayers. Each cell is delimited by six field wires and has a sense wire at its center. The detector is filled with a gas mixture of 80:20 helium:isobutane at a constant pressure of 4 mbar over environmental air pressure.

The operation is typical of a gas drift detector (see, e.g., [24]). A charged particle passing through a drift cell locally ionizes the gas. The sense wire of each cell is held at an electrical potential of 1960V relative to the field wires. As the charge carriers accelerate toward the sense wire, they collide with the gas causing an ionization avalanche, which is detected on the sense wire. The amount of time required for the avalanche to reach the sense wire, called drift time, and total charge of the avalanche are directly measured on the sense wire. The drift time is converted to a measurement of the radius with respect to the sense wire at which the original ionization event occurred. The total charge provides a measurement of the amount of energy transferred to the gas in that ionization event (dE/dx).

To allow for measurement of the z positions of passing charged particles, the cell axes of several superlayers have angles roatated with respect to the detector axis for stereo observation.



Figure 2.6: BABAR Drift Chamber (DCH) longitudinal section with principal dimensions in mm. The chamber center is offset by 370 mm from the interaction point. [22]

Superlayers of axial cells (A) alternate with pairs of stereo superlayers (U, V) with opposite-sign offsets in the order AUVAUVAUVA from innermost to outermost superlayer. The stereo angles vary between ± 45 mrad and ± 76 mrad. This structure is indicated in Figure 2.7, which shows a section of the innermost four superlayers (AUVA).

The precise position information provided by the DCH allows precise fits of the helical path a charged particle traces through the detector's solenoid field, and hence precise measurements of its momentum. Figure 2.8 shows the relationship between dE/dx and particle momentum used for identifying particles below the DIRC's Cherenkov threshold. As shown in Figure 2.9, the measured RMS dE/dx resolution for Bhabha events is typically 7.5%.

2.5 Tracking performance

A sophisticated and evolving system of software and electronic hardware reconstructs the paths (tracks) of charged particles from the observed interactions with the tracking subsystems (the SVT and DCH). The hardware component resides mostly in the electronic triggering system, which determines which detector events are recorded for physics analysis. Among



Figure 2.7: BABAR Drift Chamber (DCH) schematic layout of the drift cells for the four innermost superlayers. Lines have been added between field wires to aid in visualization of the cell boundaries. The numbers on the right side give the stero angles (mrad) of sense wires in each layer. The 1 mm thick beryllium inner wall is shown inside of the first layer. [22]



Figure 2.8: Measurement of dE/dx in the DCH as a function of track momenta. The data include large samples of beam background triggers, as evident from the high rate of protons. The curves show the Bethe-Bloch predictions derived from selected control samples of particles of different masses [22].



Figure 2.9: Difference between the measured and expected energy loss dE/dx for e^{\pm} from Bhabha scattering, measured in the DCH at an operating voltage of 1900 V. The curve represents a Gaussian fit to the data with a resolution of 7.5% [22].

many other things less relevant to tracking, the trigger system performs a very fast association of drift chamber interactions into track segments. In subsequent software processing of recorded events, these track segments and the other observed SVT and DCH interactions are combined into helical paths. The track finding and fitting procedures rely on Kalman filter algorithms that produce estimates and error matrices for the track parameters. Five parameters define the path of a track: d_0 , ϕ_0 , z_0 , ω and $\tan \lambda$. These parameters are measured at the point of closest approach (POCA) of the helix to the z axis. The parameters z_0 , d_0 and ϕ_0 are the cylindrical coordinates of the POCA relative to the origin of the coordinate system: z_0 is the z displacement of the POCA, d_0 is the distance between the POCA and the z axis, and ϕ_0 is the azimuthal angle of the POCA. The angle λ is the dip angle of the helix relative to the x - y plane, and $\omega = 1/p_t$ is its curvature, where p_t is the magnitude of the transverse (x - y) momentum of the charged particle. The resolution with which the tracking system measures helix parameters for charged tracks is given by [22] as

$$\sigma_{d_0} = 23 \ \mu \text{m} \qquad \sigma_{z_0} = 29 \ \mu \text{m}$$

$$\sigma_{\phi 0} = 0.43 \ \text{mrad} \qquad \sigma_{\tan \lambda} = 0.53 \times 10^{-3}$$

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

The efficiency of the charged particle tracking depends on the selection criteria suitable to a given measurement. The efficiencies for the analyses described in this thesis are discussed in Chapter 10.2.

2.6 BABAR subsystem: DIRC

The Detector of Internally Reflected Cherenkov light (DIRC) is the main source of information for charged particle identification (PID) for hadrons with momenta from 700 MeV up to 4.2 GeV. Discrimination for particles below 700 MeV relies primarily on dE/dx measurements in the DCH.

The DIRC is a ring-imaging Cherenkov detector with a novel design that is shown in Figure 2.10. This design minimizes the amount of material between the DCH and EMC. Optically flat fused silica bars are arranged in a thin 12-sided polygonal barrel between the outer radius of the DCH and the inner radius of the EMC. These bars are the Cherenkov radiators, with an index of refraction n = 1.473 at optical frequencies. Charged particles with velocities vexceeding that of light in the fused silica ($\beta = v/c > 1/n$) radiate a cone of of light with polar angle $\cos \theta_c = 1/n\beta$ with respect to the particle's velocity \vec{v} . The highly-polished fused silica bars also act as light pipes for Cherenkov photons trapped by total internal reflection. Mirrors at the forward ($+\hat{z}$) ends of the bars reflect forward-traveling photons back to the rear of the bars so that only one end of the barrel needs to be instrumented. The internally reflected photons exit the rear of the bar and travel through a toroidal reservoir of purified water before their



Figure 2.10: Elevation view of the nominal DIRC system geometry. All dimensions are given in mm. [22]

detection by a densely packed array of array of photomultiplier tubes (PMTs) (Figure 2.11). The photon's emission angle is preserved in its transmission allowing for reconstruction of Cherenkov rings at the PMT array and for measurement of the track's Cherenkov angle θ_c . This gives a measurement of the charged particle's velocity, which in turn, when considered with the its measured momentum, determines the particle's mass and identity. Figure 2.12 shows the variation of $\cos \theta_c$ with particle momentum p for several charged particle types. Figure 2.13 shows the expected π -K separation. The typically quoted DIRC performance benchmark is $4.2\sigma \pi$ -K separation for tracks of momentum 3 GeV. The uses of the BABAR PID system in the analyses of this thesis are detailed in Chapter 4.3.2.

2.7 BABAR subsystem: EMC

The BABAR EMC is a total absorption calorimeter for measuring electromagnetic showers. It occupies cylindrical volume at a radial interval of 920 mm to 1375 mm, between the



Figure 2.11: Elevation view of the nominal DIRC system geometry. All dimensions are given in mm. [22]



Figure 2.12: Variation of $\cos\theta_c$ with particle momentum.



Figure 2.13: Expected π -K separation in $B^0 \to \pi^- \pi^+$ events versus track momentum inferred from the measured Cherenkov angle resolution and number of Cherenkov photons per track in dimuon events [22].

DIRC and the superconducting solenoid. As the main detector of photons, its excellent energy and angular resolution is vital to any analysis relying on the detection of photons or the reconstruction of π^0 or η decays. The EMC is designed to measure the large range of shower energies from 20 MeV to 9 GeV with energy resolutions approaching 1-2%. Thallium-doped cesium iodide (CsI(Tl)) crystals compose the scintillator material of the EMC. These are arranged in a finely segmented series of axial rings as depicted in Figure 2.14. The scintillation light is read by silicon photodiodes optically coupled to each crystal.

The decays studied in this thesis all involve a photon with a lab frame energy in the range range 1.5 GeV to 3.5 GeV (Figure 2.16) from a decay $B \to X\gamma$, where X is one of (K^*, ρ, ω) . This prominent photon is a key signature of such decays, and the analyses of this thesis exercise the full capabilities of the EMC to maximally exploit this signature. Details of the role of this high energy photon in the analyses appear in Chapter 4.2. These analyses rely on the excellent detection efficiency and angular coverage of the EMC to find the *B*-daughter photon. The resolution of the measured energy of the *B*-daughter photon dominates the resolution of the estimated energy of the reconstructed *B* meson. The EMC's resolution performance allows the reconstructed *B* energy to be used as a good discriminating variable for background suppression. The full range of detectable photon energies is used to find the sister photons for high energy photons from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays that would otherwise mimic signal photons. The EMC's energy resolution allows a relatively narrow veto for such photons, reducing the number true signal decays eliminated by spurious association.

The 9 GeV upper bound of photon detection design range derives from the kinematic limit for photons from $e^+e^- \rightarrow \gamma\gamma$ and from asymmetric laboratory frame decays of π^0 s from $B^0 \rightarrow \pi^0 \pi^0$. The lower limit of 20 MeV is largely determined by backgrounds and the amount of material between the interaction point and the calorimeter. There is just 0.3-0.6 radiation lengths of material before most of the calorimeter. The CsI(Tl) crystals have a trapezoidal cross-section, as depicted in Figure 2.15. They are arranged in 56 rings around the beam axis (Figure 2.14) with the axis of each crystal aimed at the interaction point. Readout diodes are attached to the rear face of each crystal. The EMC covers the lab frame polar angles from 15.8° to 141.8° , corresponding to approximately 90% coverage in the CM frame, with full azimuthal coverage. The dimensions of the crystals vary among the rings to achieve the proper orientation of the crystals with minimal gaps between them, though typical dimensions are $4.7 \times 4.7 \text{ cm}^2$ at the front face of the crystal (side nearest the interaction point) increasing to typically $6.1 \times 6.0 \text{ cm}^2$ at the rear face. This gives the front faces of the crystals an angular extent of 2.3 to 5.1 mrad in each dimension. Individual crystal size is a key factor in the angular resolution of the calorimeter. Most electromagnetic showers will not be confined to a single crystal, but will also register in adjacent crystals. Clustering algorithms are required to agglomerate signals associated with a single particle impacting on the calorimeter and to separate the contributions from closely spaced particles. This extended 'bump' structure allows for an angular resolution smaller than crystal size.

Figure 2.17 shows EMC's relative energy resolution σ_E/E and angular resolution $\sigma_{\theta} = \sigma_{\phi}$. The energy resolution generally conforms to the empirical relation

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E}(\text{ GeV})} \oplus (1.85 \pm 0.12) , \qquad (2.1)$$

where \oplus indicates a sum in quadrature. For photons in the signal region, $\sigma_E/E \in [2.51, 2.80]\%$. The angular resolution has been fit to the empirical relation

$$\sigma_{\theta} = \sigma_{\phi} = \left(\frac{3.87 \pm 0.07}{\sqrt{E} (\text{ GeV})} + (0.00 \pm 0.04)\right) \text{mrad.}$$
(2.2)

For signal photons this gives, $\sigma_{\theta} \in [2.07 \pm 0.08, 3.16 \pm 0.10]$ mrad.



Figure 2.14: A longitudinal cross-section of the EMC (only the top half is shown) indicating the arrangement of the 56 crystal rings. The detector is axially symmetric around the z-axis. All dimensions givin in mm. [22]



Figure 2.15: A schematic of the wrapped CsI(Tl) crystal and the front-end readout package mounted on the rear face. Also indicated is the tapered, trapezoidal CFC compartment, which is open at the front. This drawing is not to scale. [22]



Figure 2.16: Distribution of $E_{\gamma,\text{lab}}$ in $B \to K^*\gamma$; based on fully simulated $B \to K^*\gamma$ Monte Carlo decays.

2.8 Simulation

The BABAR collaboration maintains a complete and highly detailed Monte Carlo (MC) simulation of the experiment for use in detector studies and physics analyses. By comparing the expected properties of the data from a detailed simulation of the detector to those of the actual data, BABAR physicists have achieved a deeper understanding of the detector and the subtleties of its data. Data-Monte Carlo comparison studies have significantly improved the quality of BABAR physics results and have proven invaluable for the evaluation of systematic sources of uncertainty in measurements. Understanding gained from such studies is incorporated into the simulation, iteratively refining the understanding of the detector. The detector itself and the way the experiment operates has and continues to change. The simulation has helped to guide its evolution.

More relevant to this thesis, the full simulation of the BABAR experiment played a vital role in developing the analyses presented herein. Details of the use of the Monte Carlo



Figure 2.17: The energy resolution for the EMC measured for photons and electrons from various processes, and the angular resolution of the EMC for photons from π^0 decays. [22]

simulation pervade the description of the analysis in subsequent chapters. The analyses were developed 'blind'. This means that all of the selection criteria and measurement procedures were fully developed on the Monte Carlo simulated data before being applied to the real data. By avoiding reference to the real data in the development of the selection criteria, analysts avoid biasing the final result by, consciously of unconsciously, tuning selection criteria toward preferred values of the final parameters. This removal of personal bias is especially important for the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ modes, which have not yet been observed. Blind analyses are only possible with accurate simulated data. The simulated data is also important for dealing with backgrounds. As a concrete example from this analysis, $B \to \rho \pi^0$ decays can closely resemble $B\to \rho\gamma$ decays if one of the photons from the $\pi^0\to\gamma\gamma$ is undetected or otherwise allows the other photon to pass the π^0 veto (Chapter 4.2.2). $B \to \rho \pi^0$ occurs at a rate sufficient to make it a significant background that may mistakenly raise the estimate of the number of signal decays unless it is dealt with carefully. However, its rate is small enough that no clean, statistically significant sample of such events can be extracted from the real data. The ability to generate such events in large quantities with the detailed simulation led to the identification of variables that discriminate signal decays from this class of background (Chapter 7.3).

The simulation encompasses the complete sequence of events from the e^+e^- collision to reconstruction of candidates and physics quantities identical in form to the real data products upon which analysis is performed. The first step is a generation of a physics event without detector interactions. These start with the initial e^+e^- collision and simulate the subsequent particle-level interactions. This step is performed by EvtGen [25], a package developed by *BABAR* physicists. EvtGen is a suite of evolving event generators representing the best understanding of the underlying physics. Most of the *B* decay generators have been developed by *BABAR* physicists, however, third party event generators have been incorporated for specific types of interactions. For example, events evolving from an initial interaction of $e^+e^- \rightarrow q\bar{q}$, where $q \in u, d, s$, are simulated with JETSET [26].

After EvtGen has created the physics event, the resulting particulate interactions with the detector are simulated with a detailed representation of the physical detector based on the GEANT4 [27] simulation toolkit. This step produces a sequence of interactions between the physics particles and the material of the detector, e.g. scattering and energy deposition, pair conversion, etc. Another layer of simulation is applied to these interactions to give the detector response and the corresponding electronic signals. At this point the simulated Monte Carlo events are in exactly the same form as the raw data from the detector itself and can be submitted to the same suite of reconstruction software for the extraction of physics information. Throughout the simulation chain, an association with the original particles produced by Evt-Gen is maintained so that it is possible to identify which generated particle produced a given electronic signal or higher level structure.

The full BABAR simulation requires a great deal of computing power. To prevent redundant and wasteful use of limited computing resources in simulation, and to provide uniformity in the quality of the simulated data used in the many BABAR analyses, production of most of the Monte Carlo data used by BABAR is generated by a single dedicated subgroup. They produce data sets for a wide variety of physics processes including continuum $e^+e^- \rightarrow q\bar{q}$ interactions, generic ensembles of *B* decays in accordance with the best current branching fraction measurements, and sets exclusively containing signal decays for many modes and analyses. So that sets of simulated signal and background sets are based on single instance of the evolving event and detector models, the Simulation Production group organizes the data sets they produce with the same detector representation. SP runs are referred to by a serial number, i.e. SP4. The $B \rightarrow K^*\gamma$ analysis described in this thesis was developed with SP4 Monte Carlo simulated events. The $B \rightarrow \rho\gamma$ and $B^0 \rightarrow \omega\gamma$ analyses were developed with SP6 Monte Carlo.

Chapter 3

Analysis overview

The goal of this analysis is to measure branching fractions of the specific (signal) decay processes $B \to K^*\gamma$, $B \to \rho\gamma$, and $B^0 \to \omega\gamma$ in data collected by the *BABAR* experiment. The basic procedure achieving this goal first identifies candidate signal decays in the data, then exploits calculable physics quantities to discriminate true signal decays from various processes that can mimic signal decays (backgrounds). Last is the calculation of the branching fractions from the resulting number of signal events and an estimate of the efficiency of the selection criteria.

3.1 Candidate identification

Functionally, the current analysis has employed a two-stage process to identify candidate signal decays. First, the events are coarsely filtered to remove quark jet or continuum backgrounds (see Chapter 5) and to ensure that the requisite final state particles were detected. This first stage filter is computationally very fast and efficiently reduces the data to a manageable volume. The second stage is a reconstruction of the complete signal B decay chain from detected final state particles.

3.1.1 Reconstruction

Reconstruction is experimental particle physics jargon by which we mean the process of identifying hypotheses for a series of particle interactions linking the detected evidence of a final state with the prepared initial state.

For example, part of this analysis searches for evidence of the decay $B^+ \to \rho^+ \gamma$. The ρ^+ decays almost immediately and is too short-lived to be detected directly in the BABAR detector. ρ^+ decays almost uniformly to $\pi^+\pi^0$. The π^0 will also decay before it is detected into a pair of photons ($\mathcal{B}(\pi^0 \to \gamma \gamma) = 98.798 \pm 0.032$) [17].

Consider a BABAR event with (at least) three detected neutral clusters that are consistent with photon interactions with the calorimeter and one or more charged track that is consistent with the passage of a π^+ through the tracking volume. Reconstruction of the signal decay in this event means determining whether the decay chain $B^+ \to \rho^+ \gamma, \rho^+ \to \pi^+ \pi^0, \pi^0 \to \gamma \gamma$ is a valid hypothesis leading to the detected final state. This analysis uses a rather literal 'construction' technique common in BABAR analyses.

First the photons are examined pairwise to find a couple consistent with a π^0 decay. The resulting π^0 candidate, provided one is found, is paired with a charged π^+ candidate for consideration in the $\rho^+ \to \pi^+\pi^0$ hypothesis. A valid ρ^+ candidate is then combined with a γ candidates to test the hypothesis $B^+ \to \rho^+\gamma$. A set of detected final state particles with a decay hypothesis with which it is found consistent is termed a reconstructed candidate, e.g. a reconstructed B.

This process of reconstruction can lead to a combinatoric explosion of valid candidates unless the hypothesis criteria are very specific. Even when the hypotheses are very narrowly defined, random combinations of final state candidates will still appear valid. This leads to

			Fraction of
		Final state	$B \to V \gamma$
$B^0 \to K^{*0} \gamma$	$K^{*0} \to K^+ \pi^-$	$\gamma_B K^+ \pi^-$	0.667
	$K^{*0} \to K^0_S(\pi^+\pi^-)\pi^0(\gamma\gamma)$	$\gamma_B(\pi^+\pi^-)_{K^0_S}(\gamma\gamma)_{\pi^0}$	0.114
$B^+ \to K^{*+} \gamma$	$K^{*+} \to K^+ \pi^0(\gamma \gamma)$	$\gamma_B K^+ (\gamma \gamma)_{\pi^0}$	0.333
	$K^{*+} \to K^0_S(\pi^+\pi^-)\pi^+$	$\gamma_B(\pi^+\pi^-)_{K^0_S}\pi^+$	0.229
$B^+ \to \rho^+ \gamma$	$\rho^+ \to \pi^+ \pi^0 (\gamma \gamma)$	$\gamma_B \pi^+ (\gamma \gamma)_{\pi^0}$	1.000
$B^0 \to \rho^0 \gamma$	$\rho^0 \to \pi^+\pi^-$	$\gamma_B \pi^+ \pi^-$	1.000
$B^0 \to \omega \gamma$	$\omega \to \pi^+ \pi^- \pi^0 (\gamma \gamma)$	$\gamma_B \pi^+ \pi^- (\gamma \gamma)_{\pi^0}$	0.891

Table 3.1: Exclusive decay modes and final states used in published measurements. Conjugate modes are implied.

combinatoric backgrounds. Further, each event may have several reconstructed B candidates passing the reconstruction criteria, necessitating a strategy to select a 'best' candidate in an event.

3.1.2 Final states

The K^* , ρ and ω mesons are too short-lived to be directly detected in BABAR. Instead, they are reconstructed from a set of decay hypotheses as described in Chapter 3.1.1.

Table 3.1 lists the decay hypotheses measured for publication.

3.2 Optimization of selection criteria

Of course, the best set of selection criteria for B candidate reconstruction does not spring fully formed from anyone's head. Optimization is usually an iterated process. The key is finding physics quantities with clearly different distributions for true signal decays and for backgrounds that provide information not already present in the existing selection criteria.

The analyses described in this report were carried out 'blind'. This means that the measurement procedures were completely determined–all selection criteria, cross-checks, and statistical methods-before they were applied to recorded *BABAR* data. In lieu of true data, the analyses were developed using data from *BABAR*'s detailed GEANT4-based Monte Carlo simulation.

The BABAR computing group produces this detailed Monte Carlo data in large batches or series labeled by a Simulation Production (SP) serial number. The $B \to K^* \gamma$ analyses were developed using the SP-4 series of BABAR Monte Carlo, which represents a luminosity-weighted sample consistent with the detector configuration in collecting the actual Run I + Run II data. The $B \to \rho \gamma$ analyses were developed with SP-6 series Monte Carlo data, which gives a sample that is luminosity-weighted to the Run I-IV detector conditions. Table 3.2 lists the Monte Carlo samples used for these analyses.

Physics quantities that show clear differences in their distributions for signal decays and background may be used either directly as selection variables in a cut or as a domain variable in a fit, or included in a multivariate combination–a Fisher discriminant (linear combination) or neural network (non-linear combination)–to construct a composite variable with more discrimination power than any of its components. A cut is usually applied when there is a clear separation between the signal and background distributions, or when either the signal or background is confined to a relatively narrow region in the domain of the variable. Multivariate techniques are applied when there are clear differences in distribution but no clear separation or confinement that would make a cut an efficient selector, or to analyze a set of correlated variables.

When applying a set of cuts (\vec{C}) to a set of variables (\vec{x}) , an optimized placement for these cuts is sought by maximizing the significance score function f(S, B) on Monte Carlo data:

$$f(S,B) = \frac{S(\vec{x};C(\vec{x}))}{\sqrt{S(\vec{x};C(\vec{x})) + B(\vec{x};C(\vec{x}))}}$$

where $S(\vec{x}; \vec{C(x)})$ and $B(\vec{x}; \vec{C(x)})$ are the numbers of signal and background candidates expected

in the measurement luminosity given the cuts $C(\vec{x})$. S and B are calculated from the efficiency $(N_{S(B),passed}/N_{S(B),tested})$ of the cuts on Monte Carlo events, and from estimates of the signal and background branching fractions. Assuming a branching fraction for the signal during cut optimization should not significantly bias the final measurement of the true branching fraction. At worst, it should lead to sub-optimal selection criteria and a smaller signal significance for the final measurement than might have been obtained.

Complete details of the cuts and multivariate techniques used in the present analyses can be found in Chapters 5 and 7.

3.3 Definition of signal

After hypothetical decay candidates have been reconstructed and selection criteria optimized and applied to refine the hypothesis and reduce the number of background events, the next step in the analysis is to measure how many of the surviving B candidates are truly signal decays and how many are misreconstructed background events. Two strategies have commonly been applied to this task: a cut-and-count method and a likelihood fit.

The cut-and-count strategy is the simpler of the two, but has recently been uniformly replaced by a likelihood fit for BABAR exclusive radiative penguin analyses. In this strategy, only cuts are used to eliminate backgrounds. The candidate hypotheses are refined by paring away segments of variable distributions to optimize the significance score on Monte Carlo data. When the cuts are applied to actual data, the number of surviving background candidates is estimated either from Monte Carlo projections or, more commonly, extrapolations from sidebands or control samples in data. This estimated background is subtracted from the total number of passing candidates giving a signal yield, which, in turn, is used with an estimate of the total number of B events in the observed data and the efficiency of selection criteria determined with signal Monte Carlo events to provide an estimate of the branching fraction.

In the likelihood fit method, one or more variables (\vec{x}) are reserved for use in a fit rather than cut. A fit to the distribution of \vec{x} can directly estimate simultaneously the number of signal candidates and the number of candidates in several classes of background. There is no need for the post-hoc background subtraction of the cut-and-count method. The additional uncertainty in the signal measurement resulting from a background subtraction is almost entirely eliminated. For the signal and each class of background, parameterized probability density functions (PDFs) $\mathcal{P}_{class}(\vec{x}; \vec{\alpha}_{class})$ are constructed from their Monte Carlo distributions in \vec{x} $(\vec{\alpha}_{class}$ are free parameters determining the shape of \mathcal{P}_{class}). These PDFs are used to construct a likelihood function over the N_{cand} candidates surviving selection:

$$\mathcal{L}(\vec{n},\vec{\alpha}) = \exp\left(-\sum_{i=1}^{N_{class}} n_i\right) \cdot \left[\prod_{j=1}^{N_{cand}} \left(\sum_{i=1}^{N_{class}} n_i \mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right)\right] , \qquad (3.1)$$

where n_i is the estimated yield in each class. The N_{cand} data set is fit by maximizing the likelihood score over $(\vec{n}, \vec{\alpha})$, giving estimated yields and errors on those yields for each signal and background class.

The key to a success with a likelihood fit is choosing an appropriate set of discriminating variables \vec{x} . Experience in exclusive radiative penguin analyses has yielded excellent results with the beam-energy-substituted mass:

$$m_{ES} \equiv \sqrt{E_{\rm beam}^{*2} - p_B^{*2}} ,$$

where E_{beam}^* is the center-of-mass (CM) beam energy and p_B^* is the CM momentum of the *B* candidate. All quantities calculated in the CM frame are denoted with an asterisk (*). In *BABAR* data, E_{beam}^* is known with a much better precision than E_{γ}^* , the CM energy of the photon from $B \to K^* \gamma$ and $B \to \rho \gamma$. The narrow distribution of m_{ES} for true signal events make it an excellent discriminator. Figure 3.1 shows the distribution of m_{ES} for $B^0 \to K^{*0} \gamma$.



Figure 3.1: Distributions of the fit variables m_{ES} and ΔE^* for signal $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ and background components for fully simulated BABAR Monte Carlo. Black lines denote signal distributions, red lines continuum background, green line $B\overline{B}$ backgrounds, and blue lines the sum of the 3 components. Parameters values in plots represent a trial maximum likelihood fit to the Monte Carlo data.

The PDFs for both the $B \to \rho \gamma$ and the $B \to K^* \gamma$ analyses presented in this report rely on m_{ES} . The $B \to K^* \gamma$ analysis uses a slightly modified version denoted the beamconstrained energy-substituted mass, $m'_{ES} \equiv \sqrt{E_{\text{beam}}^{*2} - p'_B^{*2}}$, where the B momentum p'_B is modified by scaling the photon momentum to make $E'_{\gamma}^* + E_{K^*}^* - E_{\text{beam}}^* = 0$. The value of m'_{ES} is largely independent of the calorimetric energy measurement of the photon and reduces the asymmetry in the signal distribution systematic of energy leakage in the calorimeter. In addition to m'_{ES} , the $B \to K^* \gamma$ analysis uses $\Delta E^* \equiv E_B^* - E_{\text{beam}}^*$ as a fit variable. The $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses are based on a four-dimensional likelihood fit incorporating m_{ES} , ΔE^* , a neural network output variable for continuum background discrimination (Chapter 5), and a Fisher discriminant variable for discrimination from similar B decay modes (Chapter 7).

3.4 Backgrounds

This is a short introduction to the background processes leading to false decay candidates. More complete discussions of the backgrounds and their suppression appear in Chapters 5 and 7.

The primary signature of photonic radiative penguin decays is the high energy photon daughter of the of B mesons. Figure 3.2 shows the expected observed distribution of the CMS energy of the photon E_{γ}^* for $B \to K^* \gamma$ decays as modeled in BABAR Monte Carlo. The primary backgrounds for the analyses are then those that can produce or fake photons in the energy range [1.5, 3.5]GeV.

With that, the dominant sources of high energy photons are non-resonant $e^+e^- \rightarrow q\overline{q}$ $(q \in u, d, s, c)$ 'continuum' processes (Chapter 5). Continuum interactions produce high energy photons predominantly through initial state radiation (ISR) or through π^0 and η decays. In ISR, the photon is emitted by either the e^- or e^+ prior to the $e^+e^- \rightarrow q\overline{q}$ interaction. A $\pi^0,\eta\to\gamma\gamma$ decay in one of the quark jets may proceed asymmetrically in the CM frame with one of the daughter photons carrying most of the energy. In both cases, the $q\bar{q}$ system produces jets rich in final state particles from which a K^* , ρ , or ω candidate might be reconstructed, either because the jet really contained such a meson or simply by a combinatoric coincidence. This jet-like character is also the best discriminator to use against this class of background, and suppression is based largely on event shape variables. The combinatoric nature of these background candidates also gives them flat distributions in the fit variables m_{ES} and ΔE^* allowing additional discrimination from signal candidates. Backgrounds from π^0 and η decays can be further suppressed by eliminating high energy photon candidates that, when combined with another photon candidate, are consistent with a π^0 or η daughter. This strategy fails when the secondary photon is beneath the detection threshold or when it falls outside of the calorimeter acceptance.

Though far less numerous than continuum background events, background events from $B\overline{B}$ decays pose a more significant complication to the analysis. Decay modes that are similar enough to signal decays to pass the selection criteria in significant numbers may also have

distributions that peak in m_{ES} and ΔE^* .

Monte Carlo studies indicate that non-resonant $b \to s\gamma$ penguin transitions supply the dominant peaking background in $B \to K^*\gamma$ analyses, and are also a significant background for $B \to \rho\gamma$. These transitions manifest as $B \to X_s\gamma$ decays, where the X_s is any non- K^* system of unit strangeness. Final states of the X_s system that differ from the signal final states of Table 3.1 (p. 39) by one or a small number of low energy particles easily mimic signal decays with a peaking structure in m_{ES} that overlaps the signal distribution. Because of the missing mass in the reconstructed B this class background peaks at a lower value of ΔE^* than the signal. The $B \to X_s\gamma$ backgrounds are controlled with relatively strict selection criteria for the reconstructed intermediate K^* , ρ , or ω . Yet even after strict cuts, enough of these background candidates remain that they must be accounted in the final yield calculation. Neither the overall $b \to s\gamma$ branching fraction, nor the fractional hadronization of the X_s system have yet been well measured. This imparts a relatively large systematic uncertainty to the signal yield if an estimated background level is subtracted. Chapters 7 and 9 will demonstrate that the fit variable distributions of $B \to X_s\gamma$ backgrounds differ from the signal distributions enough to allow independent estimation of their yield in the likelihood fit.

The $B \to K^* \gamma$ decays themselves form a dangerous peaking background for the $B \to \rho \gamma$ analyses. The kinematics for the decays are very similar with final states differing only by a substitution of a pion for a kaon. The *BABAR* charged particle identification (PID) group has developed tools optimized to discriminate between charged kaons and pions. These tools will misidentify a few percent of the charged kaons as a pions, with the rate of misidentification highly dependent on the particle's momentum. Considering that the branching fraction for $B \to K^* \gamma$ is expected to be about forty times that of $B \to \rho \gamma$, even a few percent PID inefficiency results in a significant background. The misassignment of a pion mass to a kaon in these cases results in m_{ES} and ΔE^* peaking structure different enough to allow independent yield estimation in



Figure 3.2: Distribution of E_{γ}^* in $B \to K^* \gamma$ Monte Carlo.

the likelihood fit (Chapters 7 and 9).

Decays of the form $B \to \rho \pi^0$ and $B \to \rho \eta$ comprise a final important class of background in $B \to \rho \gamma$ analyses. As in continuum processes, the danger lay in a daughter photon carrying most of the energy of the π^0 or η in the CM frame. The $B \to \rho \pi^0(\eta)$ branching fractions are similar to the expected $B \to \rho \gamma$ branching fractions making this a potentially significant class of background. The missing π^0/η daughter photon in the reconstruction of the *B* candidate leads to a different peaking structure in ΔE^* , which can be exploited to independently measure this class of background yield in the likelihood fit (Chapters 7 and 9).

3.5 Data sets and involvement

The work detailed in this report was performed by the author in direct collaboration with several other BABAR physicists (see author lists of [28] and [29]).

Complete documentation of the measurements of the $B \to K^* \gamma$ branching fractions and associated asymmetries appears in [28]. These measurements have been published ([1]) and are based on an estimated $88.2 \times 10^6 \pm (1.1\%) B\overline{B}$ decays, equivalent to an integrated luminosity

Mode	Amount (events)	Integrated luminosity (fb^{-1})	Assumed branching ratio (10^{-5})
$B^0 \to K^{*0} \gamma, \ K^{*0} \to K^+ \pi^- \ + \ c.c.$	72,000	2,571.4	4.0 ± 0.4
$B^0 \to K^{*0} \gamma, \ K^{*0} \to K^0_s \pi^0 \ + \ c.c.$	16,000	3,341.7	4.0 ± 0.4
$B^+ \to K^{*+} \gamma, \ K^{*+} \to K^+ \pi^0 \ + \ c.c.$	18,000	1,285.7	4.0 ± 0.4
$B^+ \to K^{*+} \gamma, \ K^{*+} \to K^0_s \pi^+ \ + \ c.c.$	16,000	1,663.5	4.0 ± 0.4
$e^+e^- \rightarrow c\bar{c}$	56,817,800	42.1	
$e^+e^- ightarrow u ar{u}, d ar{d}, s ar{s}$	83, 392, 000	39.9	
$e^+e^- \rightarrow \tau^+\tau^-$	42,958,300	45.7	
$B^0\overline{B}{}^0$ generic	155, 287, 100	295.8	
B^+B^- generic	150, 426, 700	286.5	
$B^0 \to K^{*0} \gamma \ generic \ + \ c.c.$	108,000	2571.4	4.0 ± 0.4
$B^+ \to K^{*+}\gamma \ generic + c.c.$	114,000	2714.3	4.0 ± 0.4
$B^0 \to X_{su}\gamma + c.c. \ (m_b = 4.80 \text{GeV})^{1)}$	106,000	280.4	36 ± 3
$B^0 \rightarrow X_{su} \gamma + c.c. \ (m_b = 4.65 \text{GeV})$	72,000	190.5	36 ± 3
$B^0 \rightarrow X_{su} \gamma + c.c. \ (m_b = 4.75 \text{GeV})$	18,000	47.6	36 ± 3
$B^0 \rightarrow X_{su} \gamma + c.c. \ (m_b = 4.95 \text{GeV})$	76,000	201.1	36 ± 3
$B^+ \to X_{sd}\gamma + c.c. \ (m_b = 4.80 \text{GeV})^{1)}$	110,000	291.0	36 ± 3
$B^+ \to X_{sd}\gamma + c.c. \ (m_b = 4.65 \text{GeV})$	78,000	206.3	36 ± 3
$B^+ \to X_{sd}\gamma + c.c. \ (m_b = 4.75 \text{GeV})$	18,000	47.6	36 ± 3
$B^+ \to X_{sd}\gamma + c.c. \ (m_b = 4.95 \text{GeV})$	76,000	201.1	36 ± 3
Off–Peak data	117,041,128	9.49	
On–Peak data	1,093,418,668	81.9	

Table 3.2: Monte Carlo samples used in this analysis with total amounts used for $B \to K^* \gamma$ analyses (Second column: Total numbers; Third column: Corresponding integrated luminosity; Fourth column: Assumed production cross-section at the $\Upsilon(4S)$ resonance. [23]; Fifth column: Assumed branching ratios for the decaying B meson). ¹Standard samples.

Description	Events	Luminosity (fb^{-1})	Comments
$B^0 o ho^0 \gamma$	118,000	224,762	$\mathcal{B} = 0.5 \times 10^{-6}$
$B^+ \to \rho^+ \gamma$	67,000	63,810	$\mathcal{B} = 1.0 \times 10^{-6}$
$B^0 \to \omega \gamma$	67,000	127,619	$\mathcal{B} = 0.5 \times 10^{-6}$
$B^0 \to K^{*0} \gamma$	130,000	3,158	$\mathcal{B} = 39.2 \times 10^{-6}$
$B^+ \to K^{*+} \gamma$	156,000	3,839	$\mathcal{B} = 38.7 \times 10^{-6}$
$B^+ \to \rho^+ \pi^0$	220,000	19,222	$\mathcal{B} = 10.9 \times 10^{-6}$
$B^0 \to \rho^0 \pi^0$	$215,\!000$	102,381	$\mathcal{B} = 2.00 \times 10^{-6}$
generic $B^0 \overline{B^0}$	75,180,000	143.20	$\sigma = 1.05\mathrm{nb}$
generic B^+B^-	82,880,000	157.86	$\sigma = 1.05\mathrm{nb}$
$e^+e^- \rightarrow u\overline{u}, \ d\overline{d}, s\overline{s}$	80,100,000	38.33	$\sigma = 2.09 \mathrm{nb}$
$e^+e^- \to c\overline{c}$	54,750,000	42.12	$\sigma = 1.30 \mathrm{nb}$
$e^+e^- \to \tau^+\tau^-$	13,626,000	14.50	$\sigma = 0.94\mathrm{nb}$
off-peak	n/a	11.59	n/a

Table 3.3: Release-12 series SP5 Monte Carlo and off-peak data used in the $B \rightarrow \rho \gamma$ analysis.

of approximately 81.9 fb⁻¹ recorded at *BABAR*The present author had primary responsibility for the measurement of the branching fraction $\mathcal{B}(B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-)$.

Complete documentation of the measurements of the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ branching fractions and associated asymmetries appears in [29]. These measurements have been published ([2]) and are based on an estimated $211 \times 10^6 \pm (1.1\%) B\overline{B}$ decays, equivalent to an integrated luminosity of approximately 191 fb⁻¹ recorded at *BABAR*.

The remainder of this thesis will cover the areas of the analyses with which the author was involved. Coverage of the $B \to K^* \gamma$ analysis will be restricted to the $B^0 \to K^{*0} \gamma, K^{*0} \to K^+ \pi^-$ decay mode until Chapter 11, where the measurements from the various individual modes are combined into composite branching fraction and asymmetry measurements. The $B \to \rho \gamma$ analyses will be fully covered.

Chapter 4

Event reconstruction

This chapter covers in more detail the initial process of candidate reconstruction for the four exclusive decay modes: $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$), $B^0 \to \rho^0\gamma$ ($\rho^0 \to \pi^+\pi^-$), $B^+ \to \rho^+\gamma$ ($\rho^+ \to \pi^+\pi^0$), and $B^0 \to \omega\gamma$ ($\omega \to \pi^+\pi^-\pi^0$). As described in 3.1, this procedure first involves a coarse efficient filtering (detailed in Subsection 4.1) followed by a step-by-step construction of the decay hypothesis from detected final state candidates (covered in the rest of the chapter). Subsequent chapters (Chapters 5 and 7) will further refine the selection criteria to suppress specific remaining backgrounds.

4.1 BtoXGamma selection criteria

The calculations and combinatorics involved in the reconstruction of signal candidates and background suppression are very compute-intensive. With a data set of 1.457 billion recorded events (and still counting), it is imperative to discard as many uninteresting events as possible prior to reconstruction. Physicists at *BABAR* have developed a set of selection criteria to very rapidly eliminate a majority of continuum events and ensure the presence of the necessary final state particles for radiative penguin analyses. These selection criteria are referred to by BABAR physicists and this thesis as the 'BtoXGamma skim'.

The BtoXGamma skim consists of three simple criteria:

•
$$1.5 \text{GeV} < E_{\gamma, \text{max}}^* < 3.5 \text{GeV}$$

 $E_{\gamma,\max}^*$ is the largest CM energy of any neutral calorimeter energy deposit or 'bump' (CalorNeutral candidate in *BABAR* reconstruction jargon). A neutral bump is one that not been matched with the trajectory of a detected charged track. The reconstruction of bumps in the EMC and their association with charged tracks is described in [22]. The 3-momentum of each neutral bump is estimated in the lab frame from its angle in the detector coordinate system and its measured energy, assuming a photon mass. The estimated 4-momentum is then transformed into the known interaction CM frame and the largest resultant energy tagged as $E_{\gamma,\max}^*$.

As emphasized in 3.4, the primary signature of photonic radiative penguin decays is the high energy photon daughter of the B mesons. For the two-body exclusive B decays of the present analysis, this photon is almost monoenergetic with an energy approximately half of the B rest mass ($m_B = 5.279$ GeV) in the B rest frame. The B rest frame is almost coincident with the CM frame, hence this criterion requires the existence of a photon candidate consistent with a penguin primary photon in the event.

• $N_{GTL} \ge 2$

 N_{GTL} is the number of charged tracks meeting the 'Good Tracks Loose' criteria established by *BABAR* physicists to identify charged tracks with a high probability of representing the path of a charged product of the e^-e^+ interaction (rather than that of an environmental charged particle or a result of detector noise). The Good Tracks Loose criteria are described in Chapter 4.3.1. All but two of the final states listed in Table 3.1 involve two charged particles. Even for the $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) and $B^+ \to \rho^+\gamma$ ($\rho^+ \to \pi^+\pi^0$) modes, at least one more charged particle is expected from the decay of the non-signal or 'tag side' B^{\mp} . This criterion enforces the minimum charged content required for reconstruction of a signal candidate.

• $R_2 < 0.9$

 R_2 is the ratio of the second to the zeroth Fox-Wolfram (FW) moment calculated in the CM frame of the charged content of the event. The FW moments are rotationally invariant moments of angular energy distribution described in [30]. The momentum for each charged track in an event is transformed into the CM frame with the assumption that it has the mass of a pion. These momenta are then used to calculate the Fox-Wolfram moments of the event. The normalized second FW moment is a measure of the 'jettiness' of an event. For spherically symmetric distributions of momentum, it has a value of 0. For axial events, it has a value close to 1.

This is the first of the event shape variables used to suppress continuum backgrounds by exploiting their jet-like topology. More are discussed in Chapter 5. This R_2 cut also significantly reduces the number of radiative Bhabha scattering events. B mesons from the decay of an $\Upsilon(4S)$ are produced almost at rest in the $\Upsilon(4S)$ rest frame (BABAR CM frame). Since the decay axes of the two B mesons are uncorrelated, we expect signal decays to have a spherically symmetric shape.

4.2 *B* daughter photon selection

Care is taken to guaranty the quality and verity of the signature photon. The selection of photon candidates for reconstruction begins with the set of CalorNeutral candidates in the event with a CM energy in the range [1.5, 3.5]GeV (see Chapter 4.1). To these basic requirements additional criteria of two types are applied to photon candidates: criteria related to the quality and completeness of the detector's measurements, and criteria eliminating photons from π^0 and η decays.

4.2.1 Measurement quality

An enumeration of criteria:

• $-0.74 < \cos \theta_{\gamma} < 0.93$

The EMC provides full azimuthal coverage and a polar coverage of about

 $\cos \theta_{\gamma} \in [-0.774, 0.962]$ ([22]) in the lab frame, taking the interaction point (IP) as the origin and the beam axis as the polar axis. The *B* daughter photon candidates are restricted to a tighter polar region for two resons. First, to ensure that a charged particle making an energy deposit at the bump location would need to pass through the tracking volume of the detector. This eliminates photon candidates faked by untracked charged particles. Second, to ensure that the photon showers are completely contained within the EMC. Interactions near the edge of the detector can produce showers that extend beyond the fiducial calorimetric volume. This results in an underestimation of the particle's energy.

• $N_{crystal} > 4$

 $N_{crystal}$ is the number of calorimeter crystals contributing to the cluster containing the photon candidate bump (see [22] for a description of clusters and bumps). A bump created by a noisy electronic channel will generally contain just one or a small number of channels while a true photon bump will contain many more.

• No bad crystals/channels

Photon candidates carry a status flag to indicate certain types of electronic errors that can affect the measured energy ([31]). Candidates are flagged 'bad' if any of the crystals contained in the bump results from a noisy or dead channel, or if the electronic event readout was somehow incomplete.

• 25 cm bump isolation

Requiring that the high energy photon is well isolated from any other calorimeter bump simplifies the calorimeter systematic error and eliminates several classes of background. Nearly adjacent or overlapping EMC showers may simultaneously deposit energy in one or more of the same channels-frequently called energy sharing. Energy sharing systematically degrades the measured energy and position resolution of a calorimeter bump. Isolating the bump avoids this degradation.

Isolation also eliminates bumps that are indirect products of a high energy hadron. High energy π^0 decays form a key class of such events. The bumps corresponding to the photon daughters of high energy π^0 s are frequently close together as they are effectively collimated by the π^0 boost.

The position centroid of the B daughter photon candidate's EMC bump is required to be at least 25 cm from the the centroid of the EMC bump of every other CalorNeutral candidate and from that of every charged track (ChargedTracks in *BABAR* jargon) candidate associated with a calorimeter bump.

4.2.2 π^0 and η suppression

To suppress photons from π^0 and η decays:

• Lateral shower profile: second moment

Typically the EMC shower from a photon is well collimated and cylindrically symmetric with respect to the photon direction. The photon daughters of highly energetic π^0 may hit the calorimeter so near each other that the clustering algorithm cannot resolve the interactions into two bumps. However, in this case we expect the EMC shower of the merged bump to be elongated along the decay plane of the π^0 rather than cylindrically symmetric.

The second moment (L_2) of the shower shape is a useful discriminating variable:

$$L_2 = \sum_{\text{crystals } i} \frac{E_i[(\theta_i - \theta_{\text{centroid}})^2 + (\phi_i - \phi_{\text{centroid}})^2]}{\sum_i E_i}$$

where $(\theta_{\text{centroid}}, \phi_{\text{centroid}})$ are the angular coordinates of the bump's centroid in the laboratory system and (θ_i, ϕ_i) are the angular coordinates of center of the *i*th crystal in the bump. Well collimated cylindrically symmetric bumps have small values of L_2 .

To reduce backgrounds from 'merged π^0 s' the second moment is required to satisfy $L_2 < 0.002$.

• Explicit π^0/η veto

High energy photon candidates (γ_B) consistent with the final state of a $\pi^0 \to \gamma\gamma$ or a $\eta \to \gamma\gamma$ decay are explicitly excluded from signal reconstruction. The estimated laboratory 4-momentum of γ_B is added in turn to that of each of a set of CalorNeutral candidates, $\{\gamma_i\}$. If the invariant mass $(m_{\gamma_B\gamma_i})$ of any one of the resulting combinations is consistent with that of a π^0 or a η meson, then γ_B is vetoed from signal reconstruction.

The set of CalorNeutral candidates $\{\gamma_i\}$ consists of the set of all CalorNeutral candidates excepting γ_B that have a lab energy E_{γ_i} greater than a threshold value $E_{\gamma,\min}$: $\{\gamma_i | \gamma_i \neq \gamma_B \land E_{\gamma_i} > E_{\gamma,\min}\}$. Application of this threshold reduces the frequency with which valid γ_B are vetoed by spurious combinations with calorimeter backgrounds. The mass windows and threshold $E_{\gamma,\min}$ for the vetoes are set by an optimization procedure as described in Chapter 3.2, and differ between the $B \to K^* \gamma$ and $B \to \rho \gamma$ analyses.

The optimization for $B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$ appears in [32] and results in vetoing γ_B if for any γ_i :

- (π^0 veto) $E_{\gamma_i} > 0.050 \text{GeV}$ and $m_{\gamma_B \gamma_i} \in (0.115, 0.155) \text{GeV}$
- (η veto) $E_{\gamma_i} > 0.250 \text{GeV}$ and $m_{\gamma_B \gamma_i} \in (0.507, 0.587) \text{GeV}$

The optimization for $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ was performed for [29] and results in vetoing γ_B if for any γ_i :

-
$$(\pi^0 \text{ veto}) E_{\gamma_i} > 0.030 \text{GeV} \text{ and } m_{\gamma_B \gamma_i} \in [0.105, 0.155] \text{GeV}$$

- (η veto) $E_{\gamma_i} > 0.250 \text{GeV}$ and $m_{\gamma_B \gamma_i} \in [0.500, 0.590] \text{GeV}$

The wider veto windows for the $B \to \rho \gamma$ and analyses are in part due to the presence of significant $B \to \rho \pi^0$ and $B \to \rho \eta$ backgrounds (see Chapter 7).

4.3 Charged particle selection

Each of the final states reconstructed for this analysis includes at least one charged particle. Physicists at *BABAR* have extensively studied the problems of charged particle tracking, identification, and momentum estimation in the *BABAR* data. This analysis relies on the tools they have developed for identifying good tracks with well estimated momenta, discriminating between charged particles of various types, and estimating the systematic uncertainties of calculated values and efficiencies.

The basics of tracking and of the formation of charged track candidates is described in [22]. In a BABAR event, the broadest class of tracks identified by the tracking system is termed the list of 'ChargedTracks' candidates. This analysis requires that ChargedTracks candidates pass more restrictive quality and identification criteria described below prior to their use in signal reconstruction.

4.3.1 GoodTracksLoose criteria

The BABAR Charged Particle Tracking Group has developed a hierarchy of quality criteria sets to assure that the tracks are well defined, their momenta well measured, and they likely originate from a BABAR physics event (e^-e^+ collision) rather than from a background source. These criteria sets are dubbed, in increasing stringency, GoodTracksVeryLoose, Good-TracksLoose, and GoodTracksTight. Each ChargedTracks candidate has a helical fit to the detector hits, which provides many of the values used in the quality classification, such as the momentum estimate, the χ^2 score of the fit, and the distance of closest approach (DOCA) to the interaction point.

Of the levels of criteria, GoodTracksLoose represents the best compromise of efficiency and purity for the analyses described in this thesis. Each candidate used in signal reconstruction is required to pass the GoodTracksLoose criteria.

A summary of the GoodTracksLoose criteria [33]:

- successful helical fit: $\operatorname{Prob}(\xi^2) > 0$
- minimum transverse momentum: $p_T > 0.1 \text{GeV}$
- maximum total momentum: $|\vec{p}| < 10 \text{GeV}$
- minimum number of DCH hits ([22]): $N_{\text{DCH}} \ge 12$
- maximum DOCA in XY plane: doca(xy) < 1.5cm
- range of DOCA along Z axis: -10cm < doca(z) < 10cm

4.3.2 Particle identification

The BABAR Particle Identification (PID) Group has developed many tools for classifying charged tracks by the charged particle type most likely to have generated the track. This analysis relies on several of the tools, usually called PID hadron selectors, to discriminate tracks from charged pions and those from charged kaons. A general description of these tools follows.

Three detector subsytems contribute information directly to the hadron PID selector algorithms: SVT, DCH, and DIRC. The key values from the SVT and DCH are the measurements of the energy deposited by the charged particle in the detector elements dE/dx. Two key measurements are provided by the DIRC: the measured Cherenkov angle θ_C and the number of detected Cherenkov photons N_{phot} . N_{phot} is dependent on dE/dx in the DIRC. Each of these values are related to the speed β of the charged track, which can be used with the measured momentum of the track to estimate the candidate's mass. However, a direct estimation of a candidate's mass, or even β , is not the optimal use of the data.

The PID Group has developed two principal methods of using the detector measurements to classify hadrons: calculation of probabilistic likelihoods for the various particle hypotheses, and combination in a neural network. These two methods have led to two classes of hadron selectors termed in *BABAR* Likelihood (LH) selectors and SMS or Micro selectors respectively.

The Likelihood selectors use the measured charged track data to calculate the probabilistic likelihood that the track matches each of five charged particle hypotheses: electron (e^{\pm}) , muon (μ^{\pm}) , pion (π^{\pm}) , kaon (k^{\pm}) , and proton (p^{\pm}) . The selectors then apply to these likelihoods and their ratios to produce a binary classification. For example, the PidKaonLHSelector will analyze a candidate's likelihoods and tag it either as a kaon or as not a kaon.

The SMS selectors combine a candidate's data in a single-output neural network trained
to recognize a specific particle type. Information about neural networks can be found in Chapter 5. The output of the neural network is a single floating point value between 0.0, indicating a poor match with the particle hypothesis, and 1.0, indicating good agreement with the particle hypothesis. A cut on this output value provides the binary classification of the selector. For example, the PidKaonSMSSelector will analyze the candidate with its neural network, and deliver a classification of 'kaon' or 'not a kaon' based on the output value.

Each of these selectors has a hierarchy of criteria, essentially levels of agreement with the given particle hypothesis. Usually, at least four levels of agreement are defined for each selector: (in increasing level of agreement with particle hypothesis) VeryLoose, Loose, Tight, and VeryTight. So in *BABAR* parlance, a KLHTight candidate is a charged track candidate that satisfies the Tight criteria of the PidKaonLHSelector selector.

It is worthwhile to reiterate that the selectors only deliver binary decisions-does this candidate look like a given particle type or doesn't it. The selectors for the various hypotheses are not mutually exclusive. It is quite possible for a candidate to satisfy both the KLHTight and piLHTight criteria.

With that short description of the BABAR PID tools out of the way, on to their use in the present analyses.

4.3.3 Kaon selection $(B \rightarrow K^* \gamma)$

Only charged candidates satisfying the Tight criterion of the PidKaonSMSSelector selector are considered as the final state K^{\pm} in the reconstruction of $B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$. This is a selection criteria inherited from the previous iteration of the measurement ([13]), where it was found to maximize signal significance among the selectors considered, and was not reconsidered for this analysis.

4.3.4 pion selection $(B \rightarrow K^* \gamma)$

The final state π^{\pm} in the reconstruction of $B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$ is taken from the set of charged candidates that *fail* the Tight criteria of the PidKaonSMSSelector. This is also a requirement inherited from [13], where it was found to maximize signal significance, and was not reconsidered for this analysis.

4.3.5 pion selection $(B \rightarrow \rho \gamma)$

Because $B \to K^* \gamma$ decays form a particularly dangerous class of background for $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ decays, much more care is applied to pion/kaon discrimination for these modes.

To determine the best selection criteria for the final state π^{\pm} candidates in these modes, several PID selectors at several level of criteria were considered in an optimization study. First the basic reconstruction detailed throughout this Chapter 4, excepting any PID requirements for the final state π^{\pm} , was applied to *BABAR* Monte Carlo samples of simulated signal decays for each mode and to a background of simulated $B \to K^* \gamma$ decays. Then each of the PID selection criteria was applied to these samples, and for each signal mode the significance score $S/\sqrt{S+B}$ is calculated. In the calculation of $S/\sqrt{S+B}$, the signal and background MC $B \to K^* \gamma$ yields are weighted to represent their expected relative abundances in real data.

The result for both the $B^+ \to \rho^+ \gamma \ (\rho^+ \to \pi^+ \pi^0)$ and $B^0 \to \rho^0 \gamma \ (\rho^0 \to \pi^+ \pi^-)$ data sets the VeryTight criteria of the PidPionLHSelector as the optimal choice. However, the Likelihood selectors do not make optimal use of the number of DIRC photons detected in the DIRC. This is a known weakness of the Likelihood selectors. A brief addendum to the optimization study indicates that the significance of the π^{\pm} selection criteria can be further improved if a cut is applied on the probabilistic consistency of the number of detected DIRC photons $N_{phot,meas}$ with the number expected for the passage of a charged pion $N_{phot,exp}(\pi)$. $N_{phot,exp}$ is assumed to be Poisson-distributed. The final set of PID requirements for the π^{\pm} candidates used in $B \rightarrow \rho \gamma$ reconstruction:

- satisfies the PidPionLHSelector VeryTight criteria and
- if the candidate momentum $|\vec{p}| > 0.6$ GeV, the consistency $C(N_{phot,meas}; N_{phot,exp}(\pi)) > 0.001$

For the $B^0 \to \omega \gamma$ mode, the less stringent Tight criteria of the PidPionLHS elector is found to be optimal. The addition of a DIRC photon consistency requirement degrades the signal significance and is not applied for $B^0 \to \omega \gamma$.

4.4 π^0 reconstruction

Stepping away from the final state, consider next the reconstruction of π^0 candidates for the $B^+ \to \rho^+ \gamma \ (\rho^+ \to \pi^+ \pi^0)$ and $B^0 \to \omega \gamma \ (\omega \to \pi^+ \pi^- \pi^0)$ modes. The reconstruction of π^0 candidates is vital to many *BABAR* analyses. The *BABAR* Neutral Reconstruction Analysis Working Group ([34]) has invested significant effort in studying this problem and this analysis uses their prescription for pi0DefaultMass candidates as the π^0 candidates in the signal reconstruction.

Reconstruction of pi0DefaultMass candidates begins with the set of GoodPhotonLoose candidates. These are CalorNeutral (Chapter 4.2) candidates that satisfy the following additional criteria:

- Measured energy: $E_{\gamma} > 0.030 \text{GeV}$
- Lateral moment: LAT < 0.8.

LAT is a shower shape variable similar to the second moment L_2 described in 4.2.

For every pair (γ_i, γ_j) of GoodPhotonLoose candidates, the 4-momentum of the hypothetical π^0 is estimated in each of two ways. First, by a direct addition of the estimated Good-PhotonLoose candidate 4-momenta. If the resulting invariant mass $m_{\gamma\gamma} \in [0.115, 0.150]$ GeV, then p_{π^0} is recalculated from a fit constraining the invariant mass of the π^0 to its PDG ([17]) value. This 'mass constrained fit' improves the resolution of the energy measurement E_{π^0} . The pi0DefaultMass selection criteria are finalized by requiring that this fit converge and that the resulting $E_{\pi^0, \text{fit}} > 0.200$ GeV. In summary:

- $m_{\gamma\gamma} \in [0.115, 0.150] \text{GeV}$
- convergence of mass constrained fit
- $E_{\pi^0, \text{fit}} > 0.200 \text{GeV}$

For the $B^0 \to \omega \gamma$ mode, the pi0DefaultMass candidates are used in signal reconstruction without further qualification.

For the $B^+ \to \rho^+ \gamma$ mode, two additional criteria are applied to the pi0DefaultMass candidates:

- mass window: $m_{\gamma\gamma} \in [0.117, 0.145]$ GeV
- photon opening angle: $\cos \theta_{\gamma\gamma} > 0.6$

The $\theta_{\gamma\gamma}$ (calculated in the lab frame) cut helps to reduce reconstructed candidate multiplicity by suppressing the low momentum π^0 s, which are frequently produced in *B* decays.

4.5 Intermediate meson selection

The next layer of reconstruction identifies candidates for the *B*-daughter (intermediate) mesonic resonances: K^{*0} , ρ^0 , ρ^{\pm} , and ω . The final state charged particle (K^{\pm}, π^{\pm}) candidates selected according to Chapter 4.3 and the π^0 candidates from Chapter 4.4 are combined according to the sought decay modes: $K^{*0}(K^+\pi^-)$, $\rho^0(\pi^+\pi^-)$, $\rho^+(\pi^+\pi^0)$, and $\omega(\pi^+\pi^-\pi^0)$. Each such combination of candidates passing the previous selection criteria is considered and additional requirements applied to the combination (reconstructed intermediate meson).

At this stage, the selection criteria for the intermediate meson are very simple. The 4momentum of the candidate meson is computed from a simple addition of the measured daughter momenta. The resulting invariant mass is required to fall within a window consistent with the PDG [17] values of the hypothetical meson. Then final states with at least two charged particles (every one except $\rho^+ \rightarrow \pi^+ \pi^0$) are 'vertexed' to determine the probability that the charged candidates originated from the same point.

These selection criteria do not fully exploit the physics of the mesons' decays. The chapters describing background suppression, Chapters 5 and 7, include refinements to the meson selection criteria reliant on the mesons' decay physics.

4.5.1 mass windows

The invariant mass of the reconstructed 4-momentum of intermediate meson candidates are required to fall within the intervals:

- K^{*0} : $m_{K\pi} \in [0.8, 1.0] \text{GeV}$
- ρ^{\pm}/ρ^0 : $m_{\pi\pi} \in [0.55, 1.05] \text{GeV}$
- $\omega: m_{\pi\pi\pi} \in [0.752, 0.812] \text{GeV}$

4.5.2 Vertexing

Vertexing is the estimation of an intersection point or common decay vertex for two or more trajectories. The algorithms commonly employed in *BABAR* analyses for this task are described in [35].

This analysis applies the *BABAR* calculation routine VtxLeastChiVertexer described in [35] to determine the likelihood that the two charged tracks in the reconstruction of K^{*0} , ρ^0 and ω candidates originate from a common point. As its name suggests, the VtxLeastChiVertexer estimates a common origin for the tracks via a χ^2 minimization. The probability of the resulting χ^2 is used as a measure of the consistency of the single-decay-origin hypothesis. The requirements:

- K^{*0} : $Prob(\chi^2) > 0.01$
- ρ^0 : $Prob(\chi^2) > 0.0001$
- ω : $Prob(\chi^2) > 0.0001$.

These cuts require little more than the convergence of the vertex fit.

4.5.3 K_S^0 veto for ω candidates

Although this item contextually belongs in the discussion of B backgrounds (Chapter 7), mechanically it more naturally fits here.

The decay chain $B^0 \to K^{*0}\gamma$, $K^{*0} \to K_s^0\pi^0$, $K_s^0 \to \pi^+\pi^-$ shares a final state with the $B^0 \to \omega\gamma$ mode and represents for it a dangerous background. $\pi^+\pi^-$ combinations consistent with K_s^0 decays must be eliminated. The estimated position of the decay vertex of the ω candidates has proven an excellent way to do this.

A K_s^0 produced with $\beta \approx 1$ has a flight distance of ≈ 2.7 cm, whereas an ω will decay almost at its production point. To eliminate K_s^0 combinations, the ω vertex is required to be < 0.3cm of the 'primary vertex' of the interaction. The primary vertex is the estimated position of the primary e^-e^+ interaction for an event.

4.6 *B* reconstruction and fit region

Combining the high energy photon candidates from Chapter 4.2 with the intermediate meson candidates from Chapter 4.5 completes the reconstruction of the B decay hypothesis.

The two principal variables associated with the complete decay chain are m_{ES} and ΔE^* described in Chapter 3.3. The subsequent steps of the analysis use two rectangular regions in the $m_{ES} : \Delta E^*$ plane:

• signal box

$$(m_{ES}, \Delta E^*) \in [5.27, 5.29] \text{GeV} \times [-0.2, 0.1] \text{GeV}$$

The signal box represents a relatively tight cut on the signal peak in m_{ES} : ΔE^* . This region is used in the optimization of a selection criteria as described in Chapter 3.2. The goal of optimization is to enhance the prominence of the signal peak to improve the yield estimates ultimately obtained from the likelihood fit.

• fit region

$$(m_{ES}, \Delta E^*) \in [5.20, 5.29] \text{GeV} \times [-0.3, 0.3] \text{GeV}$$

The fit region represents a relatively wide window in m_{ES} : ΔE^* . This region becomes the domain for the likelihood fit. It is deliberately extended well beyond the signal peak (signal box) so that the shapes of the background distributions in these variables can be better estimated.

4.7 Single *B* candidate selection

The reconstruction process may produce more than one unique B decay candidate in an individual signal mode for a given event. The probability that each B produced by an $\Upsilon(4S)$ will decay in the signal channel is negligible. Multiple candidates almost always result from combinatoric coincidences.

A single 'best' B candidate for each mode for each event is selected for use in the likelihood fit, provided such a reconstructed candidate exists for that event and mode. This selection is made after all of the other criteria, including the background suppression criteria described in Chapters 5 and 7.

The strategy adopted by this analysis for best candidate selection: use the B candidate with $|\Delta E^*|$ closest to 0.

Chapter 5

Background suppression variables

The basic reconstruction of Chapter 4 provides a rather minimal necessary set of criteria for signal event. High background levels, particularly from continuum processes, remain and must be suppressed by further refining the selection criteria. The variables used in this suppression fall into three broad categories:

- Event shape variables that discriminate continuum events from B decays based on the usual jet shape of $e^-e^+ \rightarrow q\overline{q}$ events,
- *B* tagging variables that identify general signatures of *B* decays in the rest of the event (ROE)–the set of charged tracks and neutral candidates that are *not* part of the signal candidate,
- Physics variables of the signal decay with calculable and distinctive distributions.

Very few of these variables lend themselves to direct cuts. Although the signal and background distributions may differ significantly, there is commonly too much overlap to effectively suppress background while maintaining an acceptable signal level with a simple selection range. Additionally, many of these variables may be highly correlated–especially the event shape variables. Cuts are optimal only for independent variables as changing a cut range for one variable can significantly change the distribution of a correlated variable. With these considerations, the current analysis combines these variables via multivariate techniques to optimally exploit more subtle differences in distributions and correlations among the variables. For each mode, a combination of variables in a neural network has been used to suppress continum backgrounds.

As introduced in Chapter 3.4, the dominant source of backgrounds are continuum processes that produce high energy photons or particles, such as merged π^0 s, which can mimic a high energy photon. Monte Carlo simulations of continuum events and previous analyses of radiative penguin decays indicate that the dominant sources of continuum backgrounds are events with asymmetrically decaying high energy π^0 s and initial state radiation (ISR) events (Figure 5.1). The topologies of these two classes of background events is subtly different. Usually the high energy π^0 is part of one of the two back-to-back jets in the CM frame and the axial nature of the event shape is apparent with a simple well-known transformation from the lab to the CM frame. For ISR events, the high energy ISR photon carries away a significant fraction of the energy from one of the leptons before their collision. Thus the center of momentum frame for the e^-e^+ interaction is boosted relative to the experiment's known initial state CM frame (prior to the emission of the ISR photon). The resulting jets no longer appear to be back-to-back in the experiment's CM frame, but rather each jet has its axis bent away from the direction of the ISR photon. The analysis must employ variables that discern both of these topologies.

5.1 Event shape variables

5.1.1 Thrust angle Θ_T

Continuum events in which a high energy photon originates from a π^0 or other particle in one of the two jets naturally show a high degree of correlation between the photon direction



Figure 5.1: Cartoons of continuum background processes.

and the jet axis. For a signal event the two B mesons are created almost at rest in the CM frame and should show almost no correlation between their decay axes. Hence the direction of a true signal photon should be largely uncorrelated with the decay axis of the non-signal or 'tag side' B meson.

The thrust angle Θ_T is the angle between the direction of the *B* candidate daughter photon and thrust axis of the ROE calculated in the *BABAR* CM frame. Each ChargedTracks candidate (Chapter 4.3) that is not part of the reconstructed signal *B* candidate is boosted into the CM frame with an assumed pion mass. Also, each CalorNeutral candidate (Chapter 4.1) that isn't part of the reconstructed *B* boosted into the CM frame with the assumption that it has zero (photon) mass. The union of these two sets constitutes the ROE in the CM frame. The thrust axis \hat{t} is defined as the unit vector that maximizes the thrust *T*:

$$T = \frac{\sum_{i=1}^{N_{ROE}} |\vec{p_i} \cdot \hat{t}|}{\sum_{i=1}^{N_{ROE}} |\vec{p_i}|}$$

up to a two-fold ambiguity as T is equal for \hat{t} and $-\hat{t}$.

Figure 5.2 shows that $|\cos \Theta_T|$ peaks sharply 1.0 for continuum events while remaining uniform (uncorrelated) for signal events. This variable has proven to be one of the most powerful for discriminating between continuum and signal events.

5.1.2 R'_2

As related in Chapter 4.1, the normalized second FW moment $(R_2, [30])$ is an extremely useful measure of the 'jettiness' of an event shape. This utility can be extended to aid in the identification of ISR events by performing the calculations in the recoil frame of the high energy B daughter photon candidate. This variable is represented by R'_2 : the normalized second FW moment of the event (less the B daughter photon candidate) calculated in the recoil frame of the B daughter photon candidate.



Figure 5.2: Distributions of $|\cos \Theta_T|$ for $B^0 \to \rho^0 \gamma$ signal Monte Carlo, continuum Monte Carlo, and off peak data.

The boost vector from the photon recoil frame into the laboratory frame $\vec{\beta}_{recoil}$ is calculated by subtracting the estimated 4-momentum of the photon candidate p_{γ} from the prepared 4-momentum of the collision p_{CM} : $p_{recoil} = p_{CM} - p_{\gamma}$, $\vec{\beta}_{recoil} = \vec{p}_{recoil}/p_{recoil}^0$. The boost is applied to every ChargedTracks candidate in the event with the assumption that each candidate has a pion mass. It is also applied to every CalorNeutral candidate except the photon recoiled against with the assumption that each candidate has a zero (photon) mass. R'_2 is then FW formula applied to the combined set of boosted momenta.

Distributions of R'_2 in BABAR MC for signal events and a generic sample of continuum events appears in Figure 5.3.

5.1.3 Energy cones

The variables R_2 , R'_2 , and $\cos \Theta_T$ are all good individual event shape indicators, but none of them gives a full picture of the event topology. A more complete description of the event shape requires a related ensemble of variables. One of the simplest schemes for such an ensemble is a division of the 4π steradian solid angle into bins and a measurement of the total energy leaving the interaction through each of those bins.



Figure 5.3: Distributions of R_2' for $B^0 \to \rho^0 \gamma$ signal Monte Carlo, continuum Monte Carlo, and off peak data.

The energy cones depicted in Figure 5.4 is just such a scheme. The cones are coaxial about the direction of the high energy B daughter photon in the BABAR CM frame in opening angle intervals of 10°. The volumes between successive cones and within the smallest cones in the forward (aligned with the photon) direction and backward (opposite to the photon) direction make a set of 18 solid angle bins.

Just as in the calculation of $\cos \Theta_T$ (Chapter 5.1.1), each ChargedTracks candidate and CalorNeutral candidate in the ROE is boosted into the *BABAR* CM frame with a pion and a photon mass hypothesis respectively. The energy in the ROE is then binned accordingly into the 18 energy cones based on the candidates' estimated momentum directions.

For non-ISR events, the *B* daughter photon direction should be highly correlated with the thrust axis of the event, so an excess of energy in the few forward most and backward most cones is expected. Even for ISR events a two jet signature is expected with excess energy deposited in a small adjacent group of forward cones and a corresponding group of backward cones. Distributions of the energy deposited in the cones for *BABAR* MC are shown in Figure 5.5.

The discriminating power of the energy cones exists in their correlations–the complete picture of the event shape–rather than in the distributions of any or a small number of them.



Figure 5.4: Diagrammatic definition of energy cones.

To be useful the energy cones can only be used in conjunction in a multivariate technique. The $B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$ analysis uses all of the 18 energy cones as inputs to a neural network for suppression of continuum backgrounds (Chapter 6).

5.1.4 *L*-moments

An expansion of the event shape in moments provides another set of variables giving a complete picture of the event topology. The previously mentioned FW moments ([30]) are one example of such moments. Of rising popularity in *BABAR* analyses are the *L*-moments defined over a set of *N* candidates with momentum \vec{p}_j as:

$$L_{i} = \frac{\sum_{j=1}^{N} |\vec{p}_{j}| \times |\cos \theta_{j}|^{i}}{\sum_{j=1}^{N} |\vec{p}_{j}|}$$

where θ_j is the angle of \vec{p}_j with respect to an axis \hat{a} : $\cos \theta_j = (\vec{p}_j \cdot \hat{a})/|\vec{p}_j|$.

The *L*-moments are used to describe the ROE shape in the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses in lieu of the energy cones. As for the $\cos \Theta_T$ and energy cone calculations, the Charged-



Figure 5.5: Distributions of energy cones for $B^0 \to K^{*0}\gamma$ signal Monte Carlo, continuum Monte Carlo, and off-peak data.

Tracks and CalorNeutral candidates in the ROE are boosted into the *BABAR* CM frame assuming for them pion and photon mass hypotheses respectively. The *L*-moments of the ROE are then calculated in the CM frame with respect to two axes: the direction of the high energy *B* daughter photon candidate, and the thrust axis \hat{t}^* (Chapter 5.1.1) of the ROE.

A previous BABAR analysis [36] has shown, unsurprisingly, that the *L*-moments indeed contain the same event shape information as the energy cones with respect to the same axis. It was discovered in the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses that a neural network with the four *L*-moments $L_2(\gamma_B)$, $L_1(\hat{t}^*)$, $L_2(\hat{t}^*)$, and $L_3(\hat{t}^*)$ could be trained to a continuum-discriminating power equivalent to a network based on the 18 energy cones. This dramatic simplification of the neural network configuration led to the use of these *L*-moments rather than the energy cones for the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses.

5.2 *B* Tagging variables

If the signal B decay is extant in an event and has been properly reconstructed, then the ROE should be the decay product of the other B meson produced by the $\Upsilon(4S)$ decay. A strategy of identifying B decays by their general properties in the ROE is complementary to identification of continuum events by their general properties. *BABAR* and the group working on these analyses have developed several variables employed for this purpose. The net flavor value $N_{\mathcal{F}}$ was used in the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ decay mode. The standard *BABAR* flavor tagging variables were used in the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses. Flavor tagging variables are employed only for separating signal events from continuum background. The additional information they provide on whether the ROE is more likely a B decay or a \overline{B} decay is unused.



Figure 5.6: The L-moment distributions for the $B^0 \to \rho^0 \gamma$ signal MC, continuum MC, and off-peak data.

5.2.1 Net flavor

The net flavor variable $N_{\mathcal{F}}$ analyzes the flavor content of the ROE in an attempt to discover particle or antiparticle excesses indicative of a weak *B* decay. The strong and electromagnetic interactions, which dominate the initial particle production in continuum events, conserve flavor–no muon without a corresponding antimuon, no net strangeness, etc. Conversely, large classes of *B* decays produce flavor excesses. Single leptons are produced in semi-leptonic decays of *B* mesons and in similar decays of charmed daughters of the *B* mesons. Kaons are also produced singly in flavor changing interactions in *B* decay chains.

Toward a classification of the ROE based on these generic ideas, the ChargedTracks candidates in the ROE are subdivided into mutually exclusive sets of kaons, electrons, muons, slow pions (a signature of D^* decays), and unclassified remainder tracks. Each of these sets is further divided by the charge of the track. The classification is hierarchical-a track is considered for inclusion in a set only if it has failed the criteria for all of the prior sets in the following order:

- 1. Kaons: pass SMSKaonTight selector
- 2. Electrons: pass ElectronMicroTight selector and momentum $|\vec{p}_{CM}| > 0.5 \text{GeV}$
- 3. Muons: pass MuonMicroTight selector and momentum $|\vec{p}_{CM}| > 1.0 \text{GeV}$
- 4. Slow pions: fail previous classifications and momentum $|\vec{p}_{CM}| < 0.250$ GeV and $|\cos \Theta_{T,\vec{p}}| > 0.8$ and $d_0 < 0.5$ cm
- 5. Unclassified: fail all previous classifications

where p_{CM} is the momentum of the track after being boosted into the event CM frame with an assumed pion mass, $\cos \Theta_{T,\vec{p}}$ is the angle between the thrust axis \vec{t} (Chapter 5.1.1) and track momentum \vec{p}_{CM} in the CM frame, and d_0 is the track's distance of closest approach to the primary vertex of the ROE.

In addition to these sets of charged tracks, we also count the number of K_S^0 candidates in the ROE. The set of K_S^0 candidates is composed of unique pairs of oppositely charged ROE ChargedTracks candidates that satisfy the following criteria:

- Vertexing: a successful decay vertex calculated Chapter 4.5.2
- Mass: $m_{K_s^0} \in [0.480, 0.516]$ GeV
- Displacement: calculated K_S^0 vertex must be separated from the calculated signal meson vertex $(K^{*0} \to K^+\pi^-)$ by at least 1 mm: $|\vec{x}_{K_S^0} - \vec{x}_{K^{*0}}| > 1.0$ mm
- Direction: the K_S^0 candidate must be headed away from the decay vertex of the signal K^{*0} . The angle $\Theta_{\text{Disp},K_S^0}$ between the K_S^0 displacement vector $\vec{x}_{K_S^0} - \vec{x}_{K^{*0}}$ and its laboratory momentum $\vec{p}_{K_S^0}$ must be: $\cos \Theta_{\text{Disp},K_S^0} > 0.98$

As it was implemented in the analysis, the inclusion of a ChargedTracks candidate in one of the charged candidate classifications does note preclude its inclusion in a K_S^0 candidate. Nor are ChargedTracks candidates prevented from inclusion in multiple K_S^0 candidates.

With these classifications, the net flavor value $N_{\mathcal{F}}$ is calculated as:

$$N_{\mathcal{F}} = |N_{K^+} - N_{K^-}| + |N_{e^+} - N_{e^-}| + |N_{\mu^+} - N_{\mu^-}| + |N_{\mathrm{Sl},\pi^+} - N_{\mathrm{Sl},\pi^-}| + N_{K_{\mathrm{Sl}}^0}$$

Figure 5.7 shows the distributions of $N_{\mathcal{F}}$ for continuum and signal.

5.2.2 Standard BABAR B tagging variables

A complete description of the standard BABAR B tagging variables can be found in [37]. The B tagging algorithm was developed for time-dependent CP-violation analyses to classify the



Figure 5.7: The $N_{\mathcal{F}}$ distribution for $B^0 \to K^{*0}\gamma$ signal MC, continuum MC, and off-peak data.

tag side of the event or ROE into several mutually exclusive physics and B flavor categories. A complete description of the algorithms is well beyond the scope of this thesis. A coarse summary description follows.

In this analysis five of the standard B tagging variables have been employed. Distributions of these variables appear in Figure 5.8.

5.2.2.1 P_{\max}^*

 P_{\max}^* is the maximum CM momentum of the charged tagging candidates in the ROE. The candidates considered are ChargedTracks candidates satisfying:

- $doca_{xy} < 0.1 \text{cm}$
- $doca_z < 4$ cm
- $p_{lab} < 10 \text{GeV}$.

 P_{\max}^* is more useful for tagging the flavor of a tag-side *B* than for discriminating against continuum events. However, when combined with other *B* tagging variables with multivariate techniques, it does add to the overall background suppressing power.

5.2.2.2 electron tag

The electron tag is the output of a neural network trained to recognize direct electrons from a semileptonic decay of a B meson. The inputs to the neural network are three kinematic variables (to be described later) associated with a primary electron candidate. Its output is a floating point value [0.0, 1.0] with 1.0 representing an electron most like a primary electron from a semi-leptonic decay. The neural network is applied to each of the charged candidates in the ROE passing the ElectronMicroVeryTight criteria. The electron tag value is the neural network output value that is most like a primary electron, multiplied by the sign of the charge of the input electron candidate.

The inputs to the neural network follow:

- p^* : the CM momentum of the electron
- E_{90}^W : the energy in the hemisphere defined by the direction of the virtual W^{\pm} in the assumed semi-leptonic *B* decay as calculated in the $\Upsilon(4S)$ frame.
- $\cos \theta_{miss}$: the cosine of the angle between the electron candidate's momentum \vec{p}_e and the missing momentum of the assumed tag side B meson \vec{p}_{miss} .

5.2.2.3 muon tag

The muon tag is calculated in the same way as the electron tag. A neural network trained to identify primary muons from semi-leptonic *B* decays is applied to each muon candidate in the ROE. This neural network takes the same inputs and produces an output with the same range and significance as that as that for the the electron tag. The list of muons is taken from those ChargedCandidates passing the MuonMicroTight criteria in the ROE. The tag value is the neural network output for the muon candidate most like a primary muon signed by the charge of the muon candidate.

5.2.2.4 slow-pion tag

The slow-pion tag is the output of a neural network trained to recognize the slow pion from a D^* daughter of a B meson. Like the electron and muon tags, the neural network is applied to each candidate in a set of SlowPion candidates and the best (most B-like) score is used as the output of the tagger with the sign of the slow-pion charge. The inputs for the neural network:

- p^* : the CM momentum of the pion
- $|\cos \theta_{thrust}|$: the cosine of the angle between the pion momentum and the ROE thrust axis in the $\Upsilon(4S)$ rest frame.
- output value of PidKaonMicroSelector.

The list of SlowPion candidates in the ROE is composed of charged candidates that aren't classified as leptons or kaons by the tagging algorithms and that have a CM momentum $p^* < 0.25$ GeV.

5.2.2.5 kaon slow-pion tag

The kaon slow-poin tag combines information from the kaon tag and the slow-pion tag described above. The kaon tag is just the output of the PidKaonMicroSelector algorithm. The goal of the kaon slow-poin tag is to identify correlated kaons and slow pions resulting from the same B decay chain to improve tagging in these instances. The tag value is the output of a neural network trained to recognize a kaon and slow-pion pair from a single B decay cascade. Kaon candidates and slow-pion candidates are analyzed pairwise, with the neural network output of the pair most consistent with a B decay returned as the tag value with a sign indicating the flavor of the hypothetical B. The inputs to the neural network:

• Kaon tag value of the kaon candidate

- slow-pion tag value of the slow-pion candidate
- $\cos \theta_{K,\pi}$: the angle between the kaon and slow-pion candidate momenta calculated in the $\Upsilon(4S)$ rest frame.

5.2.3 Number of Kaons and $P_{K,\max}^*$

B decays produce kaons with a much higher frequency than the continuum processes $e^-e^+ \rightarrow u\overline{u}, d\overline{d}$. Hence the total number of kaons in the ROE can help to suppress these interactions. As shown in Figure 5.9, when kaons are produced in continuum processes (usually $e^-e^+ \rightarrow s\overline{s}$), the distribution of their momenta tends to larger values than that for generic *B* decays such as those on the tag side of the event.

The total number of kaons N_K and the maximum kaon momentum in the ROE $P_{K,\max}^*$ are included in a neural network for continuum suppression in the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ analyses. The set of kaons are those ChargedTracks candidates in the ROE that fail to satisfy all of the PionLHLoose, PionLHTight, and PionLHVeryTight criteria, and the reconstructed K_S^0 candidates defined as follows:

- pairs of oppositely charged ChargedTracks candidates in the ROE
- vertex χ^2 probability (Chapter 4.5.2) $Prob(\chi^2) > 0.0001$ (convergent fit)
- mass constraint $m_{\pi\pi} \in [0.4827, 0.5127]$
- flight distance from primary vertex $\vec{x}_{prime} | \vec{x}_{K_S^0} \vec{x}_{prime} | > 0.3 \text{cm}$

5.3 Physics variables

A final class of background suppression variables relies on the physics of signal B decay to distinguish a correctly reconstructed signal B from background processes. In this class are



Figure 5.8: The distributions of some standard BABAR tagging variables for the $B^0 \rightarrow \rho^0 \gamma$ signal MC, continuum MC, and off-peak data.



Figure 5.9: The number of kaons (left) and the maximum kaon momentum in the center of mass frame (right) for $B^0 \to \rho^0 \gamma$ signal Monte Carlo, continuum Monte Carlo and offpeak data.

the previously introduced m_{ES} and ΔE^* (Chapter 4.6), which are included in the likelihood fit (Chapter 9). Additionally, the variables described in this section are used for cuts to suppress backgrounds or as inputs to multivariate background discriminators.

5.3.1 Δz

This is another technique borrowed from the time-dependent CP asymmetry analyses for which BABAR is eminently designed. As described in Chapter 2, the CM frame of the primary e^-e^+ collision is boosted relative to the lab frame. B mesons from $\Upsilon(4S)$ decays travel an average 250µm from their production point, which coincides with the e^-e^+ collision point as the $\Upsilon(4S)$ decays almost instantaneously. The e^-e^+ collision point cannot, in general, be determined precisely enough to be useful in estimating a B flight distance. However, since the decay times of the two B daughters of an $\Upsilon(4S)$ are uncorrelated, the distance between their decay points has the same distribution as the distance of a single B meson from its creation point. The B mesons' displacement in the xy plane is small and in general, not measurable with sufficient relative precision to be useful—the collision boost is in the \hat{z} direction. Hence the use of Δz , the estimated z difference in decay vertices of the reconstructed signal B_{sig} and the tag B_{tag} .

$$\Delta z = z_{sig} - z_{tag}$$

For continuum processes most of the final state charged candidates originate from the e^-e^+ collision point. The decay vertices of the (ersatz) reconstructed *B* candidate and the ROE should be coincident. The distribution of Δz is peaked at 0.0 for both continuum and signal decays, however, as shown in Figure 5.10, the distribution for signal events is predictably wider due to the *B* flight distance.

Use of Δz requires precise vertexing of the reconstructed signal B. Like the calculated vertex used in reconstruction quality (Chapter 4.5.2), Δz is only used for the modes with at least two charged particles in the final state. Of the modes detailed in this thesis (Table 3.1), this includes every mode save $B^+ \rightarrow \rho^+ \gamma (\rho^+ \rightarrow \pi^+ \pi^0)$.

The calculation of Δz is carried out by a standard *BABAR* analysis routine VtxTagBta-SelFit described in [35]. The vertex calculated in the reconstruction of the intermediate meson candidate in Chapter 4.5.2 is used as the decay point of the reconstructed signal *B* candidate. The routine uses a beam-constrained fit to estimate a vertex for the ChargedTracks candidates in the ROE. Logic in the routine removes candidates that more likely originate from the decays of longer lived particles if the fit fails quality cuts. The VtxTagBtaSelFit routine produces and estimate for Δz and an estimate for its error $\sigma(\Delta z)$.

Some classes of vertexing failures and pathological events appear as unusually large values for Δz or $\sigma(\Delta z)$. Prior to using Δz in a neural network the quality of the estimate is insured by requiring:

- $K^{*0} \to K^+ \pi^-$ mode: $\sigma(\Delta z) < 0.04$ cm
- $\rho^0 \to \pi^+\pi^-$ mode: $|\Delta z| < 0.4 \text{cm}$ and $\sigma(\Delta z) < 0.04 \text{cm}$
- $\omega \to \pi^+ \pi^- \pi^0$ mode: $|\Delta z| < 0.4$ cm and $\sigma(\Delta z) < 0.04$ cm



Figure 5.10: The Δz distribution for $B^0 \to \rho^0 \gamma$ signal MC, continuum MC, and off-peak data.



Figure 5.11: The $\cos \Theta_B$ distribution for $B^0 \to \rho^0 \gamma$ signal MC, continuum MC, and off-peak data.

5.3.2 B direction $\cos \Theta_B^*$

The polar angle of the $\Upsilon(4S)$ decay axis in the event CM frame conforms to a $\sin^2 \Theta = 1 - \cos^2 \Theta$ distribution due to conservation of angular momentum. This distribution is reflected the distribution of the polar angle of the reconstructed signal *B* momentum in the CM frame Θ_B^* . As shown in Figure 5.11, continuum processes result in an almost uniform distribution of $\cos \Theta_B^*$ consistent with the very weak correlation between the event thrust axis and beam axis.



Figure 5.12: The Θ_H distribution for $B^0 \to \rho^0 \gamma$ signal MC, continuum MC, and off-peak data.

5.3.3 Helicity angle Θ_H

Angular momentum conservation provides another useful tool. Each of the radiative penguin B decay modes covered by this thesis involve a pseudoscalar B meson decaying into a photon and vector meson. The transverse polarization of the photon constrains the vector meson to also be transversely polarized relative to its momentum. This transverse polarization manifests as a $\sin^2 \Theta_H$ helicity angle distribution of the vector meson decay products.

For the vector to two pseudoscalar decays $K^{*0} \to K^+\pi^-$, $\rho^+ \to \pi^+\pi^0$, and $\rho^0 \to \pi^+\pi^-$, Θ_H is the angle between the decay axis of the vector meson (momentum of one of the decay products) as calculated in the vector meson candidate rest frame and the momentum of the vector meson as calculated in rest frame of the parent *B* candidate.

For the decay $\omega \to \pi^+ \pi^- \pi^0$, Θ_H is defined as the angle between the normal to the decay plane of the ω as calculated in the ω rest frame and the momentum of the ω as calculated in rest frame of the parent *B* candidate.

Figure 5.12 displays distributions for $\cos\Theta_H$ for the $B^0\to\rho^0\gamma$ mode.



Figure 5.13: The Θ_D distribution for $B^0 \to \omega \gamma$ signal MC, continuum MC, and off-peak data.

5.3.4 Dalitz angle Θ_D

For the $B^0 \to \omega \gamma, \omega \to \pi^+ \pi^- \pi^0$ mode, a degree of discrimination against backgrounds from false or misreconstructed ω candidates can be attained with the Dalitz angle Θ_D . Θ_D is defined as the angle between the π^+ and the π^0 momenta calculated in the rest frame of the $\pi^+\pi^$ system. For real ω mesons, Θ_D follows a $\sin^2 \Theta_D$ distribution required by simultaneous isospin and angular momentum conservation in an ω decay. For false ω candidates the distribution of $\cos \Theta_D$ should be approximately uniform. Figure 5.13 shows distributions for signal and continuum background events. A contribution from real ω s is perceptible in the continuum distribution.

Chapter 6

Continuum suppression with neural networks

The use of neural networks has become commonplace in particle physics analyses to discriminate among classes of events. Each of the analysis modes detailed in this thesis uses a neural network to distinguish continuum background from properly reconstructed signal events based on the variables described in Chapter 5.

A complete discussion of neural networks lies well beyond the scope of this thesis. The User Manual of the Stuttgart Neural Network Simulator (SNNS) package [38] provides more complete description of neural networks with references to excellent publications providing even more detail. Chapter 6.1 below gives a brief introduction to neural networks specific to those used in this analysis. Each of the signal modes uses a neural network with a different architecture and set of input variables. These mode-specific details are covered in Chapter 6.2.

The neural networks for continuum background suppression described in this thesis are implemented and trained with the SNNS [38] software package. SNNS is a software simulator for neural networks with a powerful set of tools for training and testing neural networks. The accompanying C function library makes neural networks created with SNNS easy to integrate into analysis code. Its ease of use has made it one of the more popular implementations of neural networks in *BABAR* analyses.

It should be noted that the contents of this chapter are applicable only to the neural networks developed by the authors of the radiative penguin analyses presented in this thesis for suppressing continuum background. This is not a description of the neural networks used to produce the standard *BABAR B* tagging briefly described in Chapter 5.2.2.

6.1 Introduction to neural networks

A neural network is an algorithmic construct developed to model the operation of a brain or other neurological system. Their basic components are nodes representing neurons, which store single activity levels, and weighted and transformed connections among those nodes representing synapses.

These analyses use a specific type of neural network architecture known as a Feed-Forward network. This means that the outputs of the network are essentially a non-linear combination of the input variables (provided the transfer function of the connections is nonlinear, more on that later). In the present case, the goal is an output with a characteristic value for each class of event (continuum background, signal). The connection weights store the information required to make this discrimination distributed throughout the network. This distribution of information allows the network to exploit rather subtle correlations among the input variables in a powerful way. These connection weights are adjusted in a process termed 'training' so that input values typical of a class of event produce a consistent output. Specifically, this analysis has employed the training method of Back-Propagation using *BABAR* Monte Carlo events.



Figure 6.1: Basic structure of a single hidden layer neural network. Input nodes (red) corresponding to event variables are passed to "hidden" nodes (blue) as linear combinations. The line combination is then transformed by an "activation function,". Linear combinations of the hidden node outputs are then passed to the output node (black), where it is transformed once again by the activation function to give the final neural network output.

6.1.1 Topology

Figure 6.1 shows a simple Feed-Forward neural network. Each circle represents a node that stores an activation value, a single floating point number. The nodes are arranged in layers of three types: input layers, output layers and hidden layers. The input layer receives the input values from their external source. There is a single input layer and it has a number of nodes equal to the number of input variables. The output layer relays the result of the network's calculations to the external user. There is also a single output layer. The number of nodes in an output layer usually depends on the number of classes among which the network is to discriminate. For two types of events, only one output value is necessary. The output value will be in the range [0, 1] with 0 trained to represent continuum background and 1 trained to represent signal. Between the input and output layers are one or more hidden layers. Each of the neural networks used for this analysis have a single hidden layer.

The nodes are connected by unidirectional output-to-input connections. Each connection has a strength or weight associated with it. These weights are fixed in training–only the node activation values vary in the calculations for each set of input values. Each node has a set of input connections, which are used for calculating its activation value, and a set of output connections for nodes that use its output value in their calculations. The output value of a node is calculated from its activation value with a transfer or 'activation' function. The networks of this analysis use an identity transform g(x) = x for the input nodes and a *logistic* function $g(x) = 1/(1 + e^{-x})$ for all other nodes.

In simple Feed-Forward networks, nodes only have input connections from each node in the layer immediately proceeding it or, equivalently, each node has output connections only to each node in the subsequent layer. Indexing the layers Input = 1, Hidden = 2...N - 1, N =Output, the output of each node of layer n is only connected to the input of each node of layer n + 1-not to nodes in any subsequent layers > n + 1, nor to adjacent nodes in layer n, nor to nodes in previous layers < n. There is no feedback and information flows contiguously from input layer to output layer.

Assigning an index i_n to each node in a layer n with N_n nodes, each node has an activation value x_{i_n} and an output value $g_n(x_{i_n})$, where g_n is the transfer function for nodes on layer n (identity for input layer, logistic function for all others). Each connection from the output of node i_n to the input of node j_{n+1} has a weight $w_{i_n j_{n+1}}$. The activation value of the node j_{n+1} is simply the weighted sum of its inputs:

$$x_{j_{n+1}} = \sum_{i_n=1_n}^{N_n} w_{i_n j_{n+1}} \cdot g_n(x_{i_n}).$$
(6.1)

For neural networks with only one hidden layer, the layers of indexing can be simplified. Denote the activation values of the input layer $x_i \equiv x_{i_1}$, the hidden layer $y_j \equiv x_{j_2}$, and the output layer $z_k \equiv x_{k_3}$. Let the connection weights be denoted $a_{ij} \equiv w_{i_1j_2}$ and $b_{jk} \equiv w_{j_2k_3}$. And finally, let g(x) denote the logistic function as the identity function will be explicitly replaced in subsequent formulae. The input values x_i are fed into the input nodes of the network. The activation value of the hidden nodes are easily calculated:

$$y_j = \sum_i a_{ij} \cdot x_i. \tag{6.2}$$

The activation level of the output nodes are

$$z_k = \sum_j b_{jk} \cdot g(y_j) = \sum_j b_{jk} \cdot \frac{1}{1 + e^{-\sum_i a_{ij} x_i}}.$$
(6.3)

Finally, the output values of the network are

$$o_k = g(z_k) = \frac{1}{-\sum_j \frac{b_{jk}}{\left(1 + e^{-\sum_i a_{ij}x_i}\right)}}.$$
(6.4)

6.1.2 Training

6.1.2.1 Training set

Training is the process by which this algorithmic construct becomes useful. In order to adjust the connection weights w_{i_n} to produce the desired output, the first requirement is a 'training set'. This is a set of input vectors and their corresponding desired output vectors. The training set is *a priori* information about events of the various target classes, which will be encoded in the weights of the network by the training process. After it's been trained, the network will provide an *a posteriori* classification for an input of unknown type based on that input's similarities to the features of the training set encoded in the network. Hence it is vital to choose a training set that fully and accurately represents the target classes. The training set must represent the class in the most general way, otherwise the network may be overspecialized. For example, say the network is to distinguish between continuum background and signal events. If continuum events in the training set all have the same jet axis then the network may have difficulty recognizing continuum events aligned significantly differently from the training events and erroneously classify the event as signal. The number of events required for a training set generally depends on the complexity (numbers of nodes, layers, and connections) of the network. In the training process, each training event can be considered a constraint and each connection weight an independent variable. There are no fast rules about the optimal relationship between the relative number of the two. If there are too few training events relative to the network complexity, there is the risk of overspecializing the network to recognize exactly the events in the training set rather than their common features– the exact solutions to the constraint relations become encoded in the network. If there are too many training events for the network complexity (a rare concern), the iterative training process may not converge well due to the overconstrained architecture and the overall performance of the network after training will suffer–the training solution is overconstrained. In general, the complexity of the network needed must be determined empirically, with the goal being the simplest network that sufficiently solves the problem.

The networks used in this analysis have a single output value in the range [0, 1] with 0 indicating a continuum background event and 1 indicating a signal event. There is one network per signal mode. The training sets are drawn from the set of detailed *BABAR* Monte Carlo events. Only MC events that pass all of the cuts listed above the neural network cuts in Tables 8.1, 8.3, 8.5 and 8.7 for the $B^0 \to K^{*0}\gamma$, $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ modes respectively are allowed into the training set. An equal number of continuum MC events and properly reconstructed signal mode MC events are included in the training set for each network. It's important that the two classes of events are equally represented in the training set to produce a network with unbiased discrimination. The largest training set possible for each network is aggregated under these conditions. The complete set of events from the least numerous class, usually continuum MC, is included in the training set. Included events are randomly selected from the more numerous class of event, usually signal mode MC. Then the training set is divided approximately in half with one half the training set proper, and the other half a 'validation set'
(Chapter 6.1.2.3).

6.1.2.2 Back Propagation

With the training set established, the training commences. There are several established methods for training neural networks. The method employed in this analysis is called 'Back Propagation'. Back Propagation is essentially an application of the method of steepest descent to minimize the difference between the desired output and the actual output of the network.

Back Propagation is an iterative process of adjusting the network connection weights. The initial values of the weights are usually randomly assigned to values in the range [-1, 1].

A given event p from the training set $p \in T$ has input values x_{pi} and a set of target output values t_{pk} . For the networks of this analysis, the single target output values $t_p = 0$ for p a continuum background MC events and $t_p = 1$ for p a signal mode MC event. When the values x_{pi} are input to the network, the results are output values o_{pk} . The difference between the actual and target outputs is denoted

$$\delta_{pk} = o_{pk} - t_{pk}.\tag{6.5}$$

The Sum Squared Error E for a subset of the training set $S \subseteq T$ containing N(S) events gives a performance measure for the network on S:

$$E(S) = \sum_{p \in S} \sum_{k} \delta_{pk}^2.$$
(6.6)

The goal of training is to find a set of weights $w_{i_n j_{n+1}}$ that minimizes E(T), or equivalently the Mean Squared Error E(T)/N(T), where N(T) is the number of events in the training set. A single update step of Back Propagation involves changing the weights $w_{i_n j_{n+1}}$ by an amount $\Delta_S w_{i_n j_{n+1}}$ to descend the error surface:

$$\Delta_S w_{i_n j_{n+1}} = -\eta \frac{\partial E(S)}{\partial w_{i_n j_{n+1}}} , \qquad (6.7)$$

where η is a learning rate to adjust the speed and stability of training convergence.

The Back Propagation process comes in three varieties dependent on the size of the size of the subset S used in the update step: Standard, Chunk, and Batch. In Standard Back Propagation the weights are updated after the evaluation of each event-S is a single element. Because the weights are updated for every event, the network learns very quickly with Standard Back Propagation. However, the results can also be unstable as the goal becomes to sequentially minimize individual event errors rather than the error of the entire training set E(T). Batch Propagation uses the entire training set for the update set-S = T. It explicitly pursues the training goal of minimizing E(T), however it converges very slowly. Chunk Back Propagation is the logical compromise between Standard and Batch Back Propagation. An update subset S contains at least several events but significantly smaller than the complete training set T. The inclusion of many events in the update step provides a more stable convergence than Standard Back Propagation, suppressing possible spurious excursions generated by a single event, while the increased frequency of updates speeds convergence over Batch Back Propagation. The optimal size of the update set depends on the network complexity and the size of the training set. It's generally determined empirically.

Regardless of the specific brand of Back Propagation used, a single iteration or training cycle consists of a complete pass through the training set applying the network exactly once to each event. The procedure of a training cycle:

- 1. Partition the training set T into update sets $S_i \subseteq T$.
- 2. For each update set S_i perform an update step:

- (a) Calculate the sum squared error $E(S_i)$ of the network (as it was left by the last update step).
- (b) Calculate the changes to the weights $\Delta_S w_{i_n j_{n+1}}$
- (c) Apply the weight changes $\Delta_S w_{i_n j_{n+1}}$ modifying the network.

For Standard Back Propagation there will be as many update steps in a training cycle as there are events in the training set. Batch Back Propagation will have a single update step per training cycle. The training cycle is iterated until a minimum error or the desired level of performance is attained.

6.1.2.3 Validation set

As a final component, the training process needs a way of evaluating the 'desired level of performance'. Training to a finely convergent minimum of the training set error E(T) is almost never desirable. The problem here is, again, over specialization of the network, termed 'over training' in this context. The ultimate goal of the network is a reliable classification of an event-one that will not be part of the training set-based on the general features of the classes of events as contained in the input variables and their correlations. If the input variables are chosen judiciously, these general features of the class will be 'trained into' the network relatively rapidly. Further training beyond this point risks encoding into the network specific features of individual events or unintended weak correlations accidentally present only in the training set. The network ceases to be a general pattern recognition tool and produces reliable classifications only if an analyzed event closely resembles an event or set of events in the training set.

Use of a validation set helps to prevent this overtraining. The validation set V is a set of events and desired output values similar in composition (and usually in size) to the training set. The validation and training sets must be disjoint and independent in order for the validation set to serve as a measure of the generality of the network. Periodically during the training, the mean squared error of the validation set E(V)/N(V) is evaluated and compared to that of the training set E(T)/N(T). While general properties of the classifications are being trained into the network the mean squared errors of the two sets should be nearly equal. E(V) should decrease as E(T) is minimized by the training. When overtraining begins, E(V)/N(V) will begin to diverge from E(T)/N(T) as the validation set diverges from the idiosyncrasies of the training set. E(T) will continue to decrease, but eventually E(V) will actually begin to increase as the network becomes increasingly specialized to the training set.

Often, the network is optimally trained at the cycle on which E(V) obtains its minimum. Sometimes overtraining is indicated even at this point by significant differences in E(V)/N(V) and E(T)/N(T). In this case, an earlier state of the network is chosen at which the errors are consistent. In practice, neural networks are deliberately over trained in a reproducible way to identify the cycle or range of cycles at which the network is optimally trained. Once the optimal point is identified the network is retrained from the (reproduced) random initialization to this optimal state. Examples of the evolution with training cycles of the MSE for neural networks can be found in Figure 6.2. The upper left hand plot clearly shows the expected rapid decrease in the MSE of the validation set to a minimum followed by divergence from the MSE of the training set.

As indicated in Chapter 6.1.2.1, the validation sets used in this analysis are half of the available BABAR Monte Carlo satisfying the distribution requirements of the training set.

6.1.3 Final note

If an optimized network fails to meet whatever performance goals were set for it prior to its implementation, the fault may lay either in the relative simplicity of the network or just in the lack of discriminating power of the input variables. Determining whether or not a network is too simple or too complex is sometimes possible by closely examining the weights after training. Similarly, how useful an individual input variable is to the final classification can sometimes be determined. For the application to this thesis, the networks are generally quite complex and such judgments are more easily made empirically by experimenting with network topology.

6.2 Networks for continuum suppression

Most of the general properties of the neural networks used in this thesis for the suppression of continuum background are mentioned in context in Chapter 6.1. Summarizing some of the key points:

- Each analysis mode has its own Feed-Forward neural network for continuum suppression.
- Each neural network is trained to discriminate between continuum background events and the events of a specific signal mode.
- The set of input variables is mode-dependent.
- Each network has a single hidden layer, but the number of nodes in that layer is modedependent.
- Each neural network has a single output value in the range [continuum = 0, 1 =signal]
- Each network is trained by a variety of Back Propagation.
- The training and validation events are drawn from the detailed BABAR Monte Carlo simulations.
- MC Events are required to pass reconstruction and basic background cuts to be included in the training and validation sets.

Modes	Variables	Configuration
	$\cos(\Theta_T), \cos(\Theta_H), \cos(\Theta_B^*), R'_2, \Delta z,$	
$B^0 \to K^{*0} \gamma$	Net Flavor $(N_{\mathcal{F}})$,	24:10:1
	18 Energy cones $(10^{\circ} \text{ intervals})$	
	$ \cos(\Theta_T) , R'_2, L_2(\gamma), L_1, L_2, L_3, \Delta z ,$	
$B^0 \to \rho^0 \gamma$	e -tag, μ -tag, π_S -tag, K - π_S -tag, P^*_{max} -tag,	14:5:1
	Num. of Kaons, kaon P^*_{\max}	
	$ \cos(\Theta_T) , R'_2, L_2(\gamma), L_1, L_2, L_3,$	
$B^+ \to \rho^+ \gamma$	e -tag, μ -tag, π_S -tag, K - π_S -tag, P^*_{max} -tag,	13:6:1
	Num. of Kaons, kaon P^*_{\max}	
	$ \cos(\Theta_T) , R'_2, L_2(\gamma), L_1, L_2, L_3, \Delta z ,$	
$B^0 \to \omega \gamma$	e -tag, μ -tag, π_S -tag, K - π_S -tag, P^*_{max} -tag,	14:5:1
	Num. of Kaons, kaon P^*_{\max}	

Table 6.1: Summary of the neural net configuration for each decay mode with the input variables to the neural net in the second column, the configuration of the neural net in the third column, and the learning function in the last column. Due to poor statistics, the **BackpropChunk** learning function is used in the $B^0 \rightarrow \omega \gamma$ decay mode. The configuration shows the number of input nodes, followed by the number of nodes in the (single) hidden layer, followed by the number of output nodes.

The output of a mode's neural network is used either in an explicit cut, as in the $K^{*0} \to K^+\pi^-$ mode, or in the final likelihood fit, as in the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ modes.

The remainder of this chapter includes the details of the neural networks, which are summarized in Table 6.1.

6.2.1 Mode $B^0 \to K^{*0} \gamma (K^{*0} \to K^+ \pi^-)$

The set of input variables used in the $K^{*0} \to K^+\pi^-$ neural network is: thrust angle $\cos(\Theta_T)$, helicity angle $\cos(\Theta_H)$, B direction $\cos(\Theta_B^*)$, second γ recoil Fox-Wolfram moment R'_2 , vertex displacement Δz , net flavor $(N_{\mathcal{F}})$, and the 18 energy cones in 10° intervals. Each of these variables is described in Chapter 5. The energy cone values calculated are divided by 10 prior to their input to the neural network to bring the bulk of their distributions into the range [0, 1]. Neural networks typically learn faster if their inputs are constrained to the range [-1, 1]. The

reason is that large input values will saturate the transfer function at the hidden nodes unless the weight is correspondingly small. Saturation means that the transfer function is deep into one of its asymptotic regimes where its derivative is small. A small derivative typically means a small shift in the weight, which only slowly moves the network out of a regime in which hidden nodes are saturated. None of the other inputs are normalized in this manner-their distributions for the most part already lie in the range [-1, 1].

A total 24 input variables are used. A single hidden layer with 10 nodes and a single output node completes the network architecture.

The training and validation sets each contains 941 signal MC events and 938 continuum MC events selected from the detailed *BABAR* Monte Carlo simulation data sets. The sets contain equal numbers of events reconstructed in the $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ mode and in the implied conjugate mode $\overline{B}^0 \to \overline{K}^{*0}\gamma$, $\overline{K}^{*0} \to K^-\pi^+$. Training proceeded via Chunk Back Propagation with 50 event chunks. Figure 6.2 shows the evolution of the mean squared error (MSE) of the training and validation sets as the network was trained.

One can see in the selection-rejection plot in Figure 6.3 a significant separation between the training and validation set performances for this mode. Such a discrepancy is indicative of overtraining. In this case the overtraining had only a minor impact on the performance of the neural network. The result is still a powerful tool for suppressing continuum background, but one that is slightly less optimal than it might have been.

The output of the neural network for the $K^{*0} \to K^+\pi^-$ mode is used as a simple cut selection criterion. The cut is optimized by the maximization of significance method described in Chapter 3.2. This cut on the neural network output is applied after the reconstruction selection described in Chapter 4. The results of this cut optimization is presented in Figure 6.4. The



Figure 6.2: Neural network training plots, one plot for each mode. Each plot shows the mean squared error (MSE) of the network's performance on both training and validation sets as a function of training cycle. The mode represented in each plot starting with the upper left hand plot and proceeding clockwise: $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$, $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$.



Figure 6.3: Neural network performance plots, one plot for each mode. For a given value x in the output range of the neural network and a given set of background and signal events, the signal selection efficiency is the number of signal events with neural network output $o_{nn} \ge x$ divided by the total number of signal events in the set. The background rejection efficiency is the number of background events with $o_{nn} < x$ divided by the total number of background events in the set. The background rejection efficiency is the number of background events with $o_{nn} < x$ divided by the total number of background events in the set. The plots show the relationship between the signal selection efficiency (horizontal axis) and the background rejection efficiency (vertical axis) for both the training and validation sets. The mode represented in each plot starting with the upper left hand plot and proceeding clockwise: $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$, $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$. The crosses on the plots are drawn automatically by the plotting routine and signify nothing.



Figure 6.4: Results of an application of cut optimization to the output of the neural network for continuum background suppression, $K^{*0} \to K^+\pi^-$ mode. The network is applied to a marked sample of *BABAR* MC events containing continuum background events, *B* background events, and signal events that have all passed the basic reconstruction criteria. Each event is assigned a relative weight representing the relative abundance of events of that type expected in the real data set. Given a cut level x, S(x) is the weighted number of signal events with neural network output $o_{nn} > x$. Similarly, B(x) is the weighted number of passing background events. The plot shows the squared significance $S(x)^2/(S(x) + B(x))$ as a function of the cut level x. The point of maximum significance (x = 0.425719) represents the optimal selection criteria.

output of the neural network is required to satisfy

$$o_{nn} > 0.425719$$
.

6.2.2 Modes $B \to \rho \gamma$ and $B^0 \to \omega \gamma$

The neural networks for all three exclusive modes $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$ and $B^0 \to \omega \gamma$ take as inputs the following: thrust angle $|\cos(\Theta_T)|$, second γ recoil Fox-Wolfram moment R'_2 , four *L*-moments $L_2(\gamma_B)$, $L_1(\hat{t}^*)$, $L_2(\hat{t}^*)$, and $L_3(\hat{t}^*)$, five standard *BABAR B* tagging variables *e*tag, μ -tag, π_S -tag, K- π_S -tag, and P^*_{\max} , number of kaons N_K , and maximum kaon momentum $P^*_{K,\max}$. In addition, the modes $B^0 \to \rho^0 \gamma$ and $B^0 \to \omega \gamma$ for which a decay vertex can be calculated also use the vertex displacement $|\Delta z|$ as an input. This is a total of 13 inputs for $B^+ \to \rho^+ \gamma$ and 14 inputs each for $B^0 \to \rho^0 \gamma$ and $B^0 \to \omega \gamma$.

There are several important differences between these sets of input variables and the

earlier one used in the $B^0 \to K^{*0}\gamma$ mode. Each of these improvements was the product of experience and study in the time interval between the $B^0 \to K^{*0}\gamma$ analysis and the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ analyses. The first is that these networks only use information from the tag side of the event. The helicity angle $\cos(\Theta_H)$ and signal B direction $\cos(\Theta_B)$ are omitted. In the analysis of the systematic uncertainty associated with the efficiency of the neural network output (see Chapter 10), it was discovered that this interlacing is a major contributor to the uncertainty. By removing these variables from the neural network and incorporating them in a separate Fisher discriminant (Chapter 7) the total systematic uncertainty is reduced for these modes.

A second important difference is the replacement of the primitive net flavor variable $N_{\mathcal{F}}$ with the more powerful suite of standard *BABAR B* tagging variables and with the kaon number N_K and maximum kaon momentum P^*_{max} . The purpose of these new variables is the same as that of $N_{\mathcal{F}}$ —to identify the signature of a *B* decay in the ROE. However, the standard *BABAR B* tagging variables do a much better job than $N_{\mathcal{F}}$.

A last difference is the replacement of the 18 energy cones with the four *L*-moments $L_2(\gamma_B)$, $L_1(\hat{t}^*)$, $L_2(\hat{t}^*)$, and $L_3(\hat{t}^*)$. Several network architectures with various combinations of the *L*-moments and energy cones were constructed and tested for the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ modes. As mentioned in Chapter 5.1.4, for each mode an optimized network with the four listed *L*-moments performs at least as well as an analogous optimized network using the 18 energy cones instead of the *L*-moments. This means a significant simplification of the neural network with no loss of performance.

Each of the three neural networks for $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$ and $B^0 \to \omega \gamma$ has a single hidden layer and a single output node. Tests of various sizes of the hidden layer indicate an optimal 6 nodes in the hidden layer for the $B^+ \to \rho^+ \gamma$ network and 5 nodes in the hidden layers of each of the $B^0 \to \rho^0 \gamma$ and $B^0 \to \omega \gamma$ networks. Simpler networks display a significant degradation of performance while more complex networks show no significant improvement. The training and validation sets come from the set of detailed BABAR Monte Carlo simulation events that pass the basic reconstruction criteria described in Chapter 4. The $B^0 \rightarrow \rho^0 \gamma$ and $B^+ \rightarrow \rho^+ \gamma$ networks are trained using the Standard Back Propagation. The training set for the $B^0 \rightarrow \omega \gamma$ network is small enough that Standard Back Propagation has proven too unstable. This network is trained with Chunk Back Propagation. Figure 6.2 shows the evolution of the mean squared error (MSE) of the training and validation sets as each network was trained.

Figure 6.3 demonstrates the performance of the optimally trained networks. The network performances on the training sets are almost identical to those of the validation sets demonstrating that the networks have not been over-trained. Figure 6.5 shows the distributions of the neural network outputs for continuum background MC, off peak detector data, and signal MC.

Instead of creating an optimized cut as in the $K^{*0} \to K^+\pi^-$ analysis, the output values of the neural networks are used as independent variables to improve the power of the likelihood fit for signal estimation (Chapter 9). The optimization process requires a good estimate of the relative number of continuum and background events. Since there was no evidence of the decays $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ prior to this analysis, there was no way to reliably estimate the expected signal yield in data. Loose cuts are applied to the outputs of the neural networks to suppress the bulk of the continuum background and simplify the fit, but these cuts are far less restrictive than optimized cuts based on optimistic estimates of the signal yields. These loose cuts are $o_{nn} > 0.6$ for the two $B \to \rho \gamma$ modes and $o_{nn} > 0.65$ for the $B^0 \to \omega \gamma$ mode.



Figure 6.5: Plots of the distributions of the output of the fully trained neural network for continuum MC, signal MC, and off-peak data. The distributions are normalized so that all of the distributions on a plot have the same area. The mode represented in each plot starting with the upper left hand plot and proceeding clockwise: $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$, $B^0 \to \rho^0\gamma$, $B^+ \to \rho^+\gamma$, and $B^0 \to \omega\gamma$.

Chapter 7

$B\overline{B}$ background suppression and discrimination

The key sources and features of B background candidates—reconstructed signal decay candidates originating from $\Upsilon(4S)$ decays in which neither B actually decays in the signal mode– are well introduced in Chapter 3.4. Here's a summary of some of the key points made in that introduction:

- Even after a neural network cut meant to suppress them, continuum background candidates still outnumber *B* background candidates by at least an order of magnitude.
- There are two general types of B background candidates:
 - combinatoric accidents like continuum background candidates
 - candidates reconstructed from decays physically similar to the signal decays, usually differing from the signal final states by one or two low energy particles
- Because they may have distributions in the analysis variables similar to those of signal decays, *B* backgrounds are, in a sense, more dangerous than continuum backgrounds to

the measurement of the signal yield.

The detailed simulated BABAR Monte Carlo events are a powerful aid in identifying specific B decays that mimic signal decays with a high frequency. The BABAR Simulation Production Group creates samples of 'generic $B^0\overline{B}^0$, and 'generic B^+B^- ' events, which are well suited to this task. The physics in these generic samples are statistically representative of the best knowledge of natural B decays available at the time of production [17]. An examination of the generic B Monte Carlo events that pass the selection criteria yields an estimate of the frequency with which events of various types ape signal decays. This is not a tool to be used blindly. Decays that are known to be physically similar to the signals must be examined carefully and independently regardless of their survival frequency in the generic samples. However, such an analysis of the generic samples may highlight dangerous backgrounds that may have otherwise been missed.

The details of the B backgrounds and the ways with which they're dealt are covered below.

7.1 Mode
$$B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$$

7.1.1 $B \rightarrow X_s \gamma$, down-feed

The dominant B background sources are the so-called inclusive decays $B \to X_s \gamma$. X_s represents any hadronic system with unit strangeness-either a higher K resonance or a nonresonant collection of kaons and pions. Such decays proceed via precisely the same radiative penguin quark transition $b \to s\gamma$ and hence have a hard photon energy spectrum strongly overlapping with that of $B^0 \to K^{*0}\gamma$. Almost all such background events have a final state for the X_s that differs from the final state of the $K^{*0} \to K^+\pi^-$ decay by a small number of relatively

		All cuts except NN		All cuts with NN	
	Lumi	Raw	Expected	Raw	Expected
	(fb^{-1})	Yield	$81.9 f b^{-1}$	Yield	$81.9 f b^{-1}$
Generic B^+B^-	286.53	62	17.7 ± 2.3	39	11.1 ± 1.8
Generic $B^0 \bar{B}^0$	295.79	82	22.7 ± 2.5	56	15.5 ± 2.1
$B^0 \to X^0_s \gamma$	291.00	321	90.3 ± 5.0	242	68.1 ± 4.4
$B^+ \to X_s^+ \gamma$	280.42	321	93.8 ± 5.2	238	69.5 ± 4.5
Generic $B^0 \to K^{*0} \gamma$	2571.43	1	0.0 ± 0.0	0	0.0 ± 0.0
Generic $B^+ \to K^{*+} \gamma$	2714.29	570	17.2 ± 0.7	344	10.4 ± 0.6
Sum of B bkgnd.			244.4 ± 7.6		179.3 ± 6.9
uds continuum	39.90	3265	6701.8 ± 117.3	639	1311.6 ± 51.9
$c\bar{c}$ continuum	42.09	1937	3769.3 ± 85.6	455	885.4 ± 41.5
$\tau^+\tau^-$	45.70	202	362.0 ± 25.5	10	17.9 ± 5.7
$uds + c\bar{c} + \tau^+\tau^-$			10833.2 ± 147.4		2215.0 ± 66.7
Off-Resonance Data	9.49	1368	11804.8 ± 319.2	306	2640.5 ± 150.9

Table 7.1: Expected background event yields for the mode $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ without and with Neural Net cut. Lumi represents the integrated luminosity of real data in which one would expect the number events produced of that type to be equal to the number of events initially included in the Monte Carlo set. Note about the removal of the $B \to K^*\gamma$ and $B \to X_s\gamma$ modes from the 'Generic' Monte Carlo samples. I'm certain there's much to say about this plot.



Figure 7.1: m_{ES} and ΔE^* distributions of $B \to X_s \gamma$

low energy pions. A typical example is $X_s^+ = \pi^+ K^{*0}$. Omitting a low energy π^+ leaves a decay similar in all save a few respects to a signal decay. Because X_s can represent a higher kaon resonance, this class of background is ofter called 'down-feed'.

Because these background candidates are partially reconstructed real B decays, their m_{ES} and ΔE^* distributions are qualitatively similar to those of signal decays. However, ΔE^* must peak at a negative value because the B is incompletely reconstructed. Also, the wide distribution of the missing momentum broadens the peaks of both the m_{ES} and ΔE^* distributions relative to the distributions of properly reconstructed signal candidates, as seen in Figure 7.1.

7.1.2 Other $B \to K^* \gamma$ decays, cross-feed

The second most abundant source of B background events are other $B \to K^* \gamma$ decaysspecifically, the decay chain $B^+ \to K^{*+} \gamma$, $K^{*+} \to K^+ \pi^0$. Because these decay chains are usually analyzed as part of the same effort (e.g. [1]) background events of this type are often called 'crossfeed' events. This decay is dynamically identical to the signal decay $B^0 \to K^{*0} \gamma$, $K^{*0} \to K^+ \pi^$ and the final states of the two processes differ only in the charge of the pion. If the final state γ and K^+ are combined with a serendipitous π^- from the decay of the tag-side B^- , the resulting



Figure 7.2: m_{ES} and ΔE^* distributions of cross-feed

reconstructed background B can mimic a properly reconstructed signal event well.

The adopted orphan pion in this type candidate has a random momentum. Although the reconstructed K^{*0} candidate must pass a mass window, the variation of pion momentum allowed within this window significantly broadens the m_{ES} and ΔE^* distributions relative to the signal distributions, as seen in Figure 7.2. Despite its claim to the title of second most abundant *B* background source, Table 7.1 clearly shows that candidates of this type very rarely survive all of the selection criteria, befitting of the coincidence required. About 10 events of this type are expected in the actual data set of 81.9 fb⁻¹ compared to approximately 600 signal events expected.

One might expect a similar background from signal $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ decays reconstructed with tag-side orphan pions. However, such a candidate usually has to compete with the correctly reconstructed signal candidate in the best candidate selection method, which rarely selects a misreconstructed candidate over a properly reconstructed one. The combination of coincidences required for a self cross-fed candidate to pass candidate selection make it a negligible source of background.



Figure 7.3: m_{ES} and ΔE^* distributions of combinatoric B backgrounds

7.1.3 Other *B* decays

Most of the remaining background events in the generic $B\overline{B}$ samples are combinatoric coincidences, like the continuum background events. There is a small component from $B \to K^* \eta$ and $B \to K^* \pi^0$, where one of photons from $\eta/\pi^0 \to \gamma \gamma$ is lost or not included in the reconstruction. Like the down-feed component, this component peaks at negative ΔE^* and has broader m_{ES} and ΔE^* distributions than the signal. This class of background is highly suppressed. Firstly, such decays have relatively small branching fractions—on the order of the signal branching fraction or smaller. The $\pi^0 \eta$ veto eliminates this background with a high efficiency. These events have a different K^* helicity structure causing further suppression by the neural network. The cumulative effect reduces this background to negligible levels.

7.1.4 Strategy

The *B* background yield is estimated in the final fit to data by including a representative component distribution. The expected yields of the various types of *B* backgrounds are too small and their distributions in m_{ES} and ΔE^* have too much overlap with each other to allow reliable estimation of the size of each component individually. Instead an aggregate *B* background shape is used to measure the total B background yield. This shape is described in Chapter 9 and is a parameterized fit shape to a weighted sum of the distributions of the B background components. The weight given to each component in constructing the representative distribution is its expected relative yield in data based on estimates from studies in the detailed *BABAR* Monte Carlo data.

The principal complication with this method is a significant systematic uncertainty associated with the relative sizes of the *B* background components. Neither the total branching fraction nor the individual fragmentation rates for the $B \to X_s \gamma$ decays, which dominate the *B* backgrounds, are well measured. The magnitude of this systematic uncertainty is estimated after the likelihood fit to data by varying the shape of the *B* background distribution used in the fit to represent varying proportions of $B \to X_s \gamma$ events. This study is described completely in Chapter 10.

7.2 Fisher discriminants

It was noted in Chapter 6.2 that for the three $b \to d\gamma$ modes, the variables $\cos(\Theta_H)$, $\cos(\Theta_B)$, and $\cos(\Theta_D)$ were left out of the neural network to improve the systematic uncertainty associated with the neural network efficiency. Yet it is these variables (and m_{ES} and ΔE^*) that contain most of the discriminating power against B backgrounds. The event shape and flavor tagging variables, which do serve as neural network inputs, are intended to identify general properties of B decays that will be present in almost any true B event. The variables of the physics of the signal decay are what really distinguish them from background decays with similar structures.

To incorporate these physics variables back into the analysis, they are used to construct a Fisher discriminant. This is a linear combination of the variables optimized to the problem of classification. The value of this Fisher discriminant F is used as an independent variable in the likelihood fit.

Briefly, let populations π_1 and π_2 over independent variables \vec{x} be representative of two categories that are to be discriminated. Let $\vec{\mu}_1$ and $\vec{\mu}_2$ represent the means of \vec{x} over populations π_1 and π_2 respectively, and let V_{ij} be the joint covariance matrix over the two populations. The goal is a set of coefficients \vec{a} in the linear function

$$F(\vec{x}) = \vec{a} \cdot \vec{x}$$

that maximizes the distance d between the two populations

$$d = \frac{(\vec{a} \cdot (\vec{\mu}_1 - \vec{\mu}_2))^2}{\sum_{i,j} a_i V_{ij} a_j}$$

It can be shown by the Cauchy-Schwarz inequality [39] that

$$\frac{(\vec{a} \cdot (\vec{\mu}_1 - \vec{\mu}_2))^2}{\sum_{i,j} a_i V_{ij} a_j} \le \sum_{i,j} (\vec{\mu}_1 - \vec{\mu}_2)_i (V^{-1})_{ij} (\vec{\mu}_1 - \vec{\mu}_2)_j$$

with equality holding only for $a_i = \lambda \sum_j (V^{-1})_{ij} (\vec{\mu}_1 - \vec{\mu}_2)_j$ for arbitrary scalar λ . The optimal $F(\vec{x})$ is then:

$$F(\vec{x}) = \vec{a} \cdot \vec{x} = \sum_{i,j} (\vec{\mu}_1 - \vec{\mu}_2)_i (V^{-1})_{ij} x_j$$

The populations selected to train the Fisher discriminant for each mode are sets of signal and continuum from *BABAR* detailed Monte Carlo. Although it is optimized to discriminate against the dominant continuum backgrounds, it has proven effective against both combinatoric B events, which share the continuum's flat distributions in the Fisher variables, and against $B \rightarrow \rho(\eta/\pi^0)$ backgrounds, which have a significantly different helicity structure that manifests as a significantly different Fisher distribution.

The formulae for the Fisher discriminants used in this analysis are listed in Table 7.2.

Mode	Fisher Coefficients
$B^0 \to \rho^0 \gamma$	$-0.567607 + 0.659776 \cdot \cos(\theta_H) + 0.75146 \cdot \cos(\theta_B) $
$B^+ \to \rho^+ \gamma$	$-0.584202 + 0.761998 \cdot \cos(\theta_H) + 0.64758 \cdot \cos(\theta_B) $
$B^0 \to \omega \gamma$	$-0.533986 - 0.183946 \cdot \cos(\theta_H) + 0.754176 \cdot \cos(\theta_B) + 0.630382 \cdot \cos(\theta_D) $

Table 7.2: The Fisher coefficients for each decay mode. These coefficients have been obtained from Monte Carlo samples.

	Exp. Results	Assumed \mathcal{BF}		
Description	(×10 ⁻⁶)	$(\times 10^{-6})$	Events	$\operatorname{Lumi}(\operatorname{fb}^{-1})$
$B^0 \to K^{*0} \gamma$	$39.2{\pm}2.0{\pm}2.4$ [28]	39.2 ± 3.1	130,000	3,158
$B^+ \to K^{*+} \gamma$	$38.7{\pm}2.8{\pm}2.6$ [28]	38.7 ± 3.8	156,000	$3,\!839$
$B \to K^{*0} \eta$	$18.6 \pm 2.3 \pm 1.2$ [40]	$18.6{\pm}2.6$	70,000	$13,\!638$
$B \to K^{*+} \pi^0$	$4.0^{+1.3}_{-1.2} \pm 0.63$ [41]	$4.0{\pm}1.4$	$213,\!000$	$152,\!295$
$B \to K^{*+} \eta$	$25.6 \pm 4.0 \pm 2.4$ [40]	$25.6{\pm}4.7$	70,000	$19,\!848$
$B \to \rho^0 \pi^0$	< 2.9(90% C.L.) [42]	2	215,000	102,381
$B\to \rho^0\eta$	< 1.5(90% C.L.) [40]	1	68,000	$164,\!370$
$B \to \rho^+ \pi^0$	$10.9 \pm 1.9 \pm 1.9$ [42]	$10.9{\pm}2.7$	220,000	19,222
$B \to \rho^+ \eta$	$9.2{\pm}3.4{\pm}1.0$ [40]	$9.2{\pm}3.5$	70,000	18,391
$B\to\omega\pi^0$	< 1.2 (90% C.L.) [40]	1.0	70,000	$74,\!822$
$B \to \omega \eta$	$4.0{\pm}1.2{\pm}0.4$ [43]	$4.0{\pm}1.3$	70,000	18,705
$B \to \pi^+ \eta$	$5.3 \pm 1.0 \pm 0.3$ [44]	$5.0{\pm}1.0$	70,000	33,840
$B \to K^+ \eta$	$3.4{\pm}0.8{\pm}0.2$ [44]	$3.0{\pm}0.8$	70,000	56,401
$B \to K_S \eta$	$2.9{\pm}1.0{\pm}0.2$ [44]	$3.0{\pm}1.0$	70,000	82,218
$B \to \pi^+ \pi^0$	$5.5 \pm 1.0 \pm 0.6$ [45]	$5.0{\pm}1.2$	$63,\!000$	12,000
$B \to K^+ \pi^0$	$12.8 \pm 1.2 \pm 1.0$ [45]	$13.0{\pm}1.6$	220,000	$16,\!117$
$B \to K_S \pi^0$	$10.7 \pm 2.5 \ [17]$	$10.0{\pm}2.5$	64,000	8,885

Table 7.3: The current experimental results and the assumed branching fractions of the peaking *B* background decay modes. Column 2 shows the mode number used in the *BABAR* simulation production to identify each decay mode. Column 3 shows the fraction of sub-decays produced in the simulation, which consider decays like $\omega \to \pi^+\pi^-\pi^0(0.891)$ or $\eta \to \gamma\gamma(0.394)$ or $K^{*0} \to K^+\pi^-$ (0.667) or $K^{*+} \to K^+\pi^0$ (0.333) or $K_S \to \pi^+\pi^-(0.686)$, with the number in bracket as decay fraction.

		fit region		signal region
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})
$B^0 \to \rho^0 \gamma$	29416	$26.17 {\pm} 0.15$	26839	$23.88 {\pm} 0.14$
$B^0 \overline{B}^0$ combinatoric	138	$192.7{\pm}16.4$	23	$32.1 {\pm} 6.7$
B^+B^- combinatoric	150	$190.0{\pm}15.5$	26	$32.9{\pm}6.5$
$B^0 \to K^{*0} \gamma$	336	21.3 ± 1.2	217	$13.7 {\pm} 0.9$
$B \to K^{*0} \eta$	40	$0.58 {\pm} 0.09$	12	$0.18{\pm}0.05$
$B \to \rho^0 \pi^0$	2948	$5.76 {\pm} 0.11$	2283	$4.46{\pm}0.09$
$B \to \rho^0 \eta$	2631	$3.20{\pm}0.06$	1561	$1.90{\pm}0.04$
$B \to \pi^+ \eta$	1017	$6.01 {\pm} 0.18$	259	1.5 ± 0.1
$B \to K^+ \eta$	34	$0.12{\pm}0.02$	7	$0.024{\pm}0.009$
$B \to K_S \eta$	88	$0.21 {\pm} 0.02$	27	$0.06 {\pm} 0.01$
$B \to \pi^+ \pi^0$	265	$4.41 {\pm} 0.27$	45	$0.75 {\pm} 0.11$
$B \to K^+ \pi^0$	23	$0.28 {\pm} 0.06$	1	$0.01 {\pm} 0.01$
$B \to K_S \pi^0(*)$	1	$0.02{\pm}0.02$	0	0 ± 0

Table 7.4: The yields of the peaking background components for the $B^0 \to \rho^0 \gamma$ decay mode at 200 fb⁻¹. All yields are normalized using the branching fractions assumed in Table 7.3. The error on the yield is the statistical error only and does not include the uncertainty in the branching fractions.

7.3 Mode $B^0 \rightarrow \rho^0 \gamma(\rho^0 \rightarrow \pi^+ \pi^-)$

Table 7.4 lists the various B background modes considered as possibly dangerous for the $B^0 \rightarrow \rho^0 \gamma (\rho^0 \rightarrow \pi^+ \pi^-)$ analysis. Combinatoric coincidences are the primary source of Bbackground for this mode. However, special care must be taken with those specific modes that peak in signal variables m_{ES} and ΔE^* . Because this analysis had the possibility of a first significant evidence of this decay mode these signal-faking backgrounds must be removed properly.

As Table 7.4 indicates, the only non-negligible sources of peaking background events in the signal region are $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ decays, where K^+ is misidentified as π^+ , and events from $B^0 \to \rho^0 \eta$ and $B^0 \to \rho^0 \pi^0$ decays, where one of the photon daughters of the $\eta \to \gamma\gamma$ or $\pi^0 \to \gamma\gamma$ is left out of the reconstructed signal B. The following sections examine each of these backgrounds in turn.

7.3.1 $B^0 \rightarrow K^{*0}\gamma$

Considering the similarities in the physics of the decays $B^0 \to K^{*0}\gamma$ and $B^0 \to \rho^0\gamma$ it is unsurprising $B^0 \to K^{*0}\gamma$ proves a formidable background. Most of this background component is suppressed by the PID requirements placed on the final state charged pion candidates. However, with a branching fraction approximately 40 times larger than the upper limit for $B^0 \to \rho^0\gamma$, even a misidentification rate of a few percent can result is a significant background component.

As Figure 7.4 shows, even the helicity structure of the background is signal-like. The only thing that allows discrimination of this background is the negative shift in its ΔE^* distribution, which results from misassigning a pion mass to a true kaon. ΔE^* is the key to measuring this background component in the final likelihood fit.

7.3.2 $B^0 \to \rho^0(\eta/\pi^0)$

These background components are strongly suppressed by the π^0 and η vetoes for the high energy photon. However, there are instances in which the second π^0 or η daughter photon is not reconstructed or falls beneath the minimum energy threshold. As Figure 7.5 indicates, these backgrounds are more easily identified than $B^0 \to K^{*0}\gamma$. They have a significantly different helicity structure. To suppress these background components, the helicity angle is required to satisfy $|\cos \Theta_H| < 0.75$. The remaining contribution is estimated in the likelihood fit by including a component representing the η/π^0 background. The key variables in the fit are the Fisher discriminant, which inherits the difference in the helicity distribution, and ΔE^* , which has a large negative-side tail from the missing energy of the lost η/π^0 daughter photon.

The primary difficulty with this particular background is that the branching fractions for the decays $B^0 \rightarrow \rho^0(\eta/\pi^0)$ are not well measured. The branching fractions and efficiencies



Figure 7.4: The m_{ES} , ΔE^* , $\cos(\theta_B)$, and $\cos(\theta_H)$ distributions of the $B^0 \to \rho^0 \gamma$ signal MC, $B \to K^* \gamma$ background and combinatoric *B* background. All quality cuts have been applied and all distributions are normalized to the same area.



Figure 7.5: The m_{ES} , ΔE^* , $\cos(\theta_B)$, and $\cos(\theta_H)$ distributions for the $B^0 \to \rho^0 \gamma$ signal and $\rho(\pi^0/\eta)$ backgrounds. All quality cuts have been applied and all distributions are normalized to the same area.

are too small to allow independent measurement of the yield of this class of background in the likelihood fit. Thus, the uncertainty in the branching fractions must be incorporated into the systematic uncertainty of the measurement after the fit is performed.

7.3.3 Strategy

Separate components representing the distributions of the combinatoric $B, B^0 \to K^{*0}\gamma$, and $B^0 \to \rho^0(\eta/\pi^0)$ backgrounds are included in the likelihood fit to data. See Chapter 9.

		fit region	signal region		
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})	
$B^+ \to \rho^+ \gamma$	15102	$47.33 {\pm} 0.38$	12581	$39.43 {\pm} 0.35$	
$B^0 \overline{B}^0$ combinatoric	226	$316{\pm}21$	32	44.7 ± 7.9	
B^+B^- combinatoric	370	469 ± 24	101	128 ± 13	
$B^+ \to K^{*+} \gamma$	242	$12.6 {\pm} 0.8$	106	$5.52 {\pm} 0.54$	
$B \to K^{*+} \pi^0$	72	$0.09 {\pm} 0.01$	17	$0.022{\pm}0.005$	
$B \to K^{*+} \eta$	76	$0.76 {\pm} 0.09$	14	$0.14{\pm}0.04$	
$B \to \rho^+ \pi^0$	2739	$28.50 {\pm} 0.54$	1580	$16.44{\pm}0.41$	
$B \to \rho^+ \eta$	2174	$23.64{\pm}0.51$	1090	$11.85 {\pm} 0.36$	
$B \to \pi^+ \eta$	669	$3.9{\pm}0.2$	81	$0.47{\pm}0.05$	
$B \to K^+ \eta$	22	$0.08{\pm}0.02$	3	$0.01 {\pm} 0.01$	
$B \to K_S \eta$	293	$0.71 {\pm} 0.04$	69	$0.16{\pm}0.02$	
$B \to \pi^+ \pi^0$	148	$2.46{\pm}0.20$	17	$0.28 {\pm} 0.06$	
$B \to K^+ \pi^0$	25	25 0.31 ± 0.06		$0.05 {\pm} 0.02$	
$B \to K_S \pi^0(^*)$	15	$0.33{\pm}0.08$	5	$0.11{\pm}0.05$	

Table 7.5: The yields of the peaking background components for the $B^+ \to \rho^+ \gamma$ decay mode at 200 fb⁻¹. All yields are normalized using the branching fractions assumed in Table 7.3. The error on the yield is the statistical error only and does not include the uncertainty in the branching fractions.

7.4 Mode $B^+ \to \rho^+ \gamma (\rho^+ \to \pi^+ \pi^0)$

The sources of B backgrounds for the $B^+ \to \rho^+ \gamma$ (Table 7.5) are analogous to those of $B^0 \to \rho^0 \gamma$. Random combinations of candidates form the numerically dominant background. Dangerous backgrounds that peak in the signal region in m_{ES} and ΔE^* come from $B^+ \to K^{*+} \gamma$, $B^+ \to \rho^+ \eta$, and $B^+ \to \rho^+ \eta$, which are covered in more detail below.

7.4.1 $B^+ \rightarrow K^{*+}\gamma$

The relationship between the decays $B^+ \to \rho^+ \gamma$ and $B^+ \to K^{*+} \gamma$ is analogous to that between $B^0 \to \rho^0 \gamma$ and $B^0 \to K^{*0} \gamma$. Misidentification of a final state K^+ as π^+ leads to a false candidate from a $B^+ \to K^{*+} \gamma$ decay that mimics many of the distributions of true signal candidates (Figure 7.6). The key to discerning this background in the final fit is the negative shift in the ΔE^* distribution due to the incorrect mass hypothesis assigned to the misidentified kaon.

7.4.2 $B^+ \to \rho^+(\eta/\pi^0)$

A below-threshold or missing photon from $\eta/\pi^0 \to \gamma\gamma$ decay can lead to a $B^+ \to \rho^+(\eta/\pi^0)$ decay that closely resembles a $B^+ \to \rho^+\gamma$ signal decay. Again, these backgrounds can be distinguished by their helicity distribution (Figure 7.7) and the long negativeside tail on their ΔE^* distributions due to the missing photon. Candidates are required to satisfy $|\cos \Theta_H| < 0.7$ to suppress most candidates from these backgrounds.

7.4.3 Strategy

Separate components representing the distributions of the combinatoric B, $B^+ \to K^{*+}\gamma$, and $B^+ \to \rho^+(\eta/\pi^0)$ backgrounds are included in the likelihood fit to data. See



Figure 7.6: The m_{ES} , ΔE^* , $\cos(\theta_B)$, and $\cos(\theta_H)$ distributions for the $B^+ \rightarrow \rho^+ \gamma$ signal, $B \rightarrow K^* \gamma$ background, and combinatoric *B* background. All quality cuts have been applied and all distributions are normalized to the same area.



Figure 7.7: The m_{ES} , ΔE^* , $\cos(\theta_B)$, and $\cos(\theta_H)$ distributions of the $B^0 \to \rho^0 \gamma$ signal and $\rho(\pi/\eta)$ backgrounds. All quality cuts have been applied and all distributions are normalized to the same area.

	fit region		signal region	
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})
$B^0 ightarrow \omega \gamma$	10888	$17.06 {\pm} 0.16$	9438	$14.79 {\pm} 0.15$
$B^0 \overline{B}^0$ combinatoric	37	$51.7 {\pm} 8.5$	8	11.17 ± 3.95
B^+B^- combinatoric	47	$59.5 {\pm} 8.7$	9	$11.40{\pm}3.80$
$B^0 \to K^{*0} \gamma, K^{*0} \to K^+ \pi^-$	84	$5.31{\pm}0.58$	18	$1.14{\pm}0.27$
$B^0 \to K^{*0} \gamma, K^{*0} \to K^0_s \pi^0$	18	$1.14{\pm}0.27$	6	$0.38 {\pm} 0.16$
$B^+ \to K^{*+} \gamma, K^{*+} \to K^+ \pi^0$	26	$1.35 {\pm} 0.27$	8	$0.42{\pm}0.15$
$B^+ \to K^{*+} \gamma, K^{*+} \to K^0_s \pi^+$	38	$1.98{\pm}0.32$	11	$0.57 {\pm} 0.17$
$B \to K^{*0} \eta$	33	$0.48 {\pm} 0.084$	9	$0.131 {\pm} 0.043$
$B \to K^{*+} \eta$	4	$0.040 {\pm} 0.020$	2	$0.020{\pm}0.014$
$B o ho^0 \pi^0$	56	$0.11 {\pm} 0.014$	12	$0.023 {\pm} 0.006$
$B o ho^0 \eta$	56	$0.068 {\pm} 0.009$	11	$0.013 {\pm} 0.004$
$B \to \rho^+ \pi^0$	33	$0.34{\pm}0.06$	10	$0.10{\pm}0.03$
$B \to \rho^+ \eta$	40	$0.43 {\pm} 0.07$	11	$0.12{\pm}0.04$
$B ightarrow \omega \eta$	791	$8.45 {\pm} 0.30$	476	$5.09{\pm}0.23$
$B \to \omega \pi^0$	711	$1.90{\pm}0.07$	554	$1.48 {\pm} 0.06$
$B \to \pi^+ \eta$	6	$0.035 {\pm} 0.014$	0	0 ± 0
$B \to K^+ \eta$	0	0 ± 0	0	0 ± 0
$B \to K_S \eta$	16	$0.038 {\pm} 0.009$	5	$0.012{\pm}0.005$
$B \to \pi^+ \pi^0$	0	$0{\pm}0$	0	$0{\pm}0$
$B \to K^+ \pi^0$	1	$0.01 {\pm} 0.01$	1	$0.01 {\pm} 0.01$
$B \to K_S \pi^0(^*)$	0	0 ± 0	0	0 ± 0

Table 7.6: The yields of the peaking background components for the $B^0 \to \omega \gamma$ decay mode at 200 fb⁻¹. All yields are normalized using the branching fractions assumed in Table 7.3. The error on the yield is the statistical error only and does not include the uncertainty in the branching fractions.

Chapter 9.

7.5 Mode $B^0 \to \omega \gamma(\omega \to \pi^+ \pi^- \pi^0)$

The *B* backgrounds besetting the $B^0 \to \omega \gamma$ analysis are enumerated in Table 7.6, and are analogous to those plaguing the $B \to \rho \gamma$ modes. Again, $B \to K^* \gamma$ decays and $B^0 \to \omega(\eta/\pi^0)$ decays form the most dangerous peaking *B* backgrounds while the combinatoric background dominates numerically.

7.5.1 $B \rightarrow K^* \gamma$

The most dangerous source of background from $B \to K^* \gamma$ decays are decays in the mode $B^0 \to K^{*0} \gamma$, $K^{*0} \to K_s^0 \pi^0$, $K_s^0 \to \pi^+ \pi^-$, which has the same set of final state particles as the signal $B^0 \to \omega \gamma$. This background is reduced to very low levels by the vertex flight distance cut detailed in Chapter 4.5.3. The remaining combined $B \to K^* \gamma$ represents a minor background. Instead of including a component in the likelihood fit representing the $B \to K^* \gamma$ background, the small bias induced on the signal yield is subtracted from the results of the likelihood fit. The bias is small enough that the systematic uncertainty introduced by the subtraction is relatively small and comparable to the systematic uncertainty necessarily accompanying a component representing $B \to K^* \gamma$ in the likelihood fit.

7.5.2 $B^0 \rightarrow \omega(\eta/\pi^0)$

Like similar decays in the $B \to \rho \gamma$ analyses, events of the decays $B^0 \to \omega \eta$ and $B^0 \to \omega \pi^0$, where one of the final state η/π^0 daughter photons is not reconstructed, peak in the signal region in m_{ES} and ΔE^* . And like the analogous decays in $B \to \rho \gamma$ modes, this background is distinguishable from the signal decay by its helicity distribution (Figure 7.8). To suppress these backgrounds, candidates are required to satisfy $|\cos \Theta_H| < 0.7$.

As indicated in Table 7.3, the branching fraction $\mathcal{B}(B^0 \to \omega \pi^0)$ has not been well measured. This uncertainty contributes to the systematic uncertainty of the measurement as described in Chapter 10.

7.5.3 Strategy

A component representing the distribution of the $B^0 \to \omega(\eta/\pi^0)$ background is included in the likelihood fit to data. See Chapter 9.



Figure 7.8: The m_{ES} , ΔE^* , $\cos(\theta_B)$, $\cos(\theta_D)$ and $\cos(\theta_H)$ distributions for the $B^0 \to \omega \gamma$ signal MC and $\omega(\eta/\pi^0)$ background. All quality cuts have been applied and all distributions are normalized to the same area.

Chapter 8

Monte Carlo efficiencies of selection criteria

This chapter contains summary tables of the cumulative efficiencies of the selection criteria described in Chapter 4-7.

8.1 $B^0 \rightarrow K^{*0}\gamma, K^{*0} \rightarrow K^+\pi^-$

Table 8.1 displays the cumulative efficiency of each of the selection criteria when applied to fully simulated $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) Monte Carlo data. Table 8.2 lists the expected yields of various backgrounds under the complete set of $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) selection criteria based on Monte Carlo studies.

Cut	Passed	Rel. ϵ %	Cum. ϵ %
$B^0 \to K^{*0} \gamma (K^{*0} \to K^+ \pi^-)$ Events Generated	72000		
Beta Skim/Candidate Finding/Truth Match	50681	70.4	70.4
High-Energy Photon Quality 4.2			
No problematic channels	49883	98.4	69.3
$N_{xtal} > 4$	49883	100.0	69.3
Second Moment < 0.002	48894	98.0	67.9
$-0.74 < \cos \theta_{\gamma} < 0.93$	48038	98.2	66.7
25 cm Isolation from Neutral bumps	44441	92.5	61.7
$25~\mathrm{cm}$ Isolation from Charged bumps	43349	97.5	60.2
π^0 Veto	41516	95.8	57.7
η^0 Veto	39742	95.7	55.2
Charged-Track Quality 4.3			
K^{\pm} GoodTracksLoose	37498	94.4	52.1
π^{\pm} GoodTracksLoose	35422	94.5	49.2
K^{\pm} SMSTight	28873	81.5	40.1
π^{\pm} Not SMSTight	28547	98.9	39.6
Vertex fit χ^2	26679	93.5	37.1
$\sigma_{\Delta z}$ Cut	25360	95.1	35.2
$0.8 < m_{K\pi} < 1.0 \text{GeV}/c^2$	22224	87.6	30.9
Fit Region 4.6			
$-0.3 < \Delta E < 0.3$ GeV	21474	96.6	29.8
$5.2 < m_{ES,rescaled} < 5.3 \mathrm{GeV}/c^2$	21474	100.0	29.8
NN > 0.425719	18368	85.5	25.5
Best Candidate Selection 4.7			
Best Candidate Selection	18308	99.7	25.4
Overall efficiency	18525		25.7
Expected signal yield in 81.9 fb^{-1}	590.0		

Table 8.1: Efficiency of the applied cuts for $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ signal. Relevant section numbers are listed for each of the categories of cuts.

	Lumi.	Raw Yields		in 81.9 ${\rm fb}^{-1}$		
	$({\rm fb}^{-1})$	$K^+\pi^-$	$K^-\pi^+$	$K^+\pi^-$	$K^{-}\pi^{+}$	Total
Generic B^+B^-	286.53	21	17	6.0	4.9	10.9
Generic $B^0 \bar{B}^0$	295.79	27	31	7.5	8.6	16.1
$B^0 \to X^0_s \gamma$	291.00	120	136	33.8	38.3	72.0
$B^+ \to X_s^+ \gamma$	280.42	120	126	35.0	36.8	71.8
Generic $B^0 \to K^{*0} \gamma$	2571.43	0	0	0.0	0.0	0.0
Generic $B^+ \to K^{*+} \gamma$	2714.29	181	166	5.5	5.0	10.5
Sum of B bkgnds.				87.8	93.5	181.3
uds continuum	39.90	391	385	802.6	790.3	1592.8
$c\bar{c}$ continuum	42.09	261	257	507.9	500.1	1007.9
$\tau^+\tau^-$	45.70	7	9	12.5	16.1	28.7
$uds + c\bar{c} + \tau^+\tau^-$				1323.0	1306.5	2629.5
Off-Resonance Data	9.49	173	144	1493.0	1242.7	2735.8

Table 8.2: $K^{*0} \to K^+ \pi^-$ Monte Carlo Yields by CP Mode
8.2 $B^0 \rightarrow \rho^0 \gamma$

Table 8.3 displays the cumulative efficiency of each of the selection criteria when applied to fully simulated $B^0 \rightarrow \rho^0 \gamma$ Monte Carlo data. Table 8.4 lists the expected yields of various backgrounds under the complete set of $B^0 \rightarrow \rho^0 \gamma$ selection criteria based on Monte Carlo studies.

Description	Events	$eff.(\pm \text{ stat.})$	cul. eff.(\pm stat.)
MC Events	115000		
BtoXGamma criteria			
$1.5 \text{ GeV} < E_{\text{max}} < 3.5 \text{ GeV}$			
$n_{\rm GTL} > 2$			
$R_2 < 0.9$	81836	$0.712{\pm}0.002$	$0.712 {\pm} 0.002$
high-energy photon			
$-0.74 < \cos(\theta) < 0.93$	79645	$0.973 {\pm} 0.003$	$0.693{\pm}0.002$
No problem crystal	79645	$1.0{\pm}0.0$	$0.693{\pm}0.002$
$n_{\rm crystals} > 4$	79645	$1.0{\pm}0.0$	$0.693{\pm}0.002$
Second moment < 0.002	78802	$0.989{\pm}0.004$	$0.685 {\pm} 0.002$
$> 25 \mathrm{cm}$ isolation	70976	$0.901{\pm}0.003$	$0.617 {\pm} 0.002$
π^0/η veto	60780	$0.856{\pm}0.003$	$0.529 {\pm} 0.002$
Charged Tracks			
GoodTracksLoose	58591	$0.964{\pm}0.004$	$0.509{\pm}0.002$
π^{\pm} selector	38948	$0.665 {\pm} 0.003$	$0.339{\pm}0.002$
ρ^0 Meson			
$\overline{m_{\pi\pi}}$	30123	$0.773 {\pm} 0.004$	$0.262{\pm}0.002$
$\chi^2 \text{ prob}$	28942	$0.961{\pm}0.006$	$0.252{\pm}0.001$
$ \cos(\theta_H) < 0.75$	25252	$0.873 {\pm} 0.005$	$0.220{\pm}0.001$
Vertexing			
$ \Delta z < 0.4\mathrm{cm}, \Delta z_{\mathrm{err}} < 0.04\mathrm{cm}$	24147	$0.956{\pm}0.006$	$0.210{\pm}0.001$
NN output			
NN> 0.6	19435	0.805 ± 0.006	0.169 ± 0.001
fit region			
total	17480		0.152 ± 0.001

Table 8.3: The efficiencies of different cuts for the $B^0 \rightarrow \rho^0 \gamma$ decay mode with CM2-based signal Monte Carlo. Truth-matching is not required for the individual cut efficiencies, only for the last line-item. The efficiency of the corresponding cuts in SP5 signal Monte Carlo is shown in column 2.

		fit region		ignal region
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})
$B^0 o ho^0 \gamma$	18382	$16.8 {\pm} 0.1$	17035	$15.6 {\pm} 0.2$
$B^0 \to K^{*0} \gamma, \ K^{*0} \to K^+ \pi^-$	320	$11.4{\pm}0.6$	260	$9.2{\pm}0.6$
$B \to \rho^0 \pi^0$	4471	$1.60 {\pm} 0.02$	3328	$1.20{\pm}0.02$
$B o ho^0 \eta$	1052	$0.99{\pm}0.03$	615	$0.58{\pm}0.02$
Sum		$2.59{\pm}0.04$		$1.77 {\pm} 0.03$
other $B^0 \overline{B^0}$ backgrounds	58	$33.3 {\pm} 4.4$	15	$8.6{\pm}2.2$
other B^+B^- backgrounds	37	43.2 ± 7.1	3	$3.5{\pm}2.0$
Sum		76.5 ± 8.3		12.1±3.0
off-peak	322	4025 ± 224	24	$300{\pm}61$

Table 8.4: The expected yield for the $B^0 \rightarrow \rho^0 \gamma$ decay mode at 200 fb⁻¹ after applying all cuts. The yields are obtained using the assumed branching fraction or cross sections shown in Table 7.3.

8.3 $B^+ \rightarrow \rho^+ \gamma$

Table 8.5 displays the cumulative efficiency of each of the selection criteria when applied to fully simulated $B^+ \rightarrow \rho^+ \gamma$ Monte Carlo data. Table 8.6 lists the expected yields of various backgrounds under the complete set of $B^+ \rightarrow \rho^+ \gamma$ selection criteria based on Monte Carlo studies.

Description	Events	eff.(\pm stat.)	cul. eff.(\pm stat.)
MC Events	118000		
BtoXGamma criteria			
$1.5 \text{ GeV} < E_{\max} < 3.5 \text{ GeV}$			
$n_{\rm GTL} > 2$			
$R_2 < 0.9$	63753	$0.540{\pm}0.002$	$0.540{\pm}0.002$
high-energy photon			
$-0.74 < \cos(\theta) < 0.93$	62439	$0.979 {\pm} 0.004$	$0.529{\pm}0.002$
No problem crystal	62439	$1.0{\pm}0.0$	$0.529{\pm}0.002$
$n_{\rm crystals} > 4$	62439	$1.0{\pm}0.0$	$0.529{\pm}0.002$
Second moment < 0.002	61758	$0.989{\pm}0.004$	$0.523{\pm}0.002$
$> 25\mathrm{cm}$ isolation	55560	$0.900{\pm}0.004$	$0.471 {\pm} 0.002$
π^0/η veto	47410	$0.853 {\pm} 0.004$	$0.402{\pm}0.002$
π^0 selections			
m_{π^0}	45322	$0.956{\pm}0.004$	$0.384{\pm}0.002$
$\cos(\theta) < 0.6$	43077	$0.950{\pm}0.005$	$0.365{\pm}0.002$
Charged Tracks			
GoodTracksLoose	42409	$0.984{\pm}0.005$	$0.359{\pm}0.002$
π^{\pm} selector	34233	$0.807 {\pm} 0.004$	$0.290{\pm}0.002$
ρ^0 Meson			
$m_{\pi\pi}$	27217	$0.795{\pm}0.005$	$0.231{\pm}0.001$
$ \cos(\theta_H) < 0.7$	23157	$0.851 {\pm} 0.006$	$0.196{\pm}0.001$
NN output			
NN> 0.6	17674	$0.763 {\pm} 0.006$	$0.150 {\pm} 0.001$
fit region			
total	12447		0.105 ± 0.001

Table 8.5: The efficiencies of different cuts for the $B^+ \rightarrow \rho^+ \gamma$ decay mode with CM2-based signal Monte Carlo. Truth-matching is not required for the individual cut efficiencies, only for the last line-item. The efficiency of the corresponding cuts in SP5 signal Monte Carlo is shown in column 2.

		fit region	si	gnal region
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})
$B^+ \to \rho^+ \gamma$	16132	$28.7{\pm}0.2$	14232	$25.3 {\pm} 0.2$
$B^+ \to K^{*+} \gamma, \ K^{*+} \to K^+ \pi^0$	151	$4.4{\pm}0.4$	102	$3.0{\pm}0.3$
$B \to \rho^+ \pi^0$	3897	$7.0{\pm}0.1$	2609	$4.7{\pm}0.1$
$B \to \rho^+ \eta$	769	$6.5{\pm}0.2$	413	$3.5{\pm}0.2$
Sum		13.5 ± 0.3		$8.2{\pm}0.2$
other $B^0 \overline{B^0}$ backgrounds	130	$74.7 {\pm} 6.5$	29	$16.6 {\pm} 3.0$
other B^+B^- backgrounds	82	$95.6{\pm}10.5$	17	$19.8 {\pm} 4.8$
Sum		170.3 ± 12.4		36.5 ± 5.7
off-peak	564	$7050{\pm}296$	32	400 ± 70

Table 8.6: The expected yield for the $B^+ \to \rho^+ \gamma$ decay mode at 200 fb⁻¹ after applying all cuts. The yields are obtained using the assumed branching fraction or cross sections shown in Table 7.3.

8.4 $B^0 \rightarrow \omega \gamma$

Table 8.7 displays the cumulative efficiency of each of the selection criteria when applied to fully simulated $B^0 \to \omega \gamma$ Monte Carlo data. Table 8.8 lists the expected yields of various backgrounds under the complete set of $B^0 \to \omega \gamma$ selection criteria based on Monte Carlo studies.

Description	Events	eff.(\pm stat.)	cul. eff.(\pm stat.)
MC Events	118000		
BtoXGamma criteria			
$1.5 \text{ GeV} < E_{\text{max}} < 3.5 \text{ GeV}$			
$n_{\rm GTL} > 2$			
$R_2 < 0.9$	58953	$0.500{\pm}0.002$	$0.500{\pm}0.002$
high-energy photon			
$-0.74 < \cos(\theta) < 0.93$	57705	$0.979 {\pm} 0.004$	$0.489{\pm}0.002$
No problem crystal	57705	$1.0{\pm}0.0$	$0.489{\pm}0.002$
$n_{\rm crystals} > 4$	57705	$1.0{\pm}0.0$	$0.489{\pm}0.002$
Second moment < 0.002	57096	$0.989{\pm}0.004$	$0.484{\pm}0.002$
$> 25 \mathrm{cm}$ isolation	51377	$0.900{\pm}0.004$	$0.435{\pm}0.002$
π^0/η veto	43856	$0.854{\pm}0.004$	$0.372{\pm}0.002$
Charged Tracks			
GoodTracksLoose	40524	$0.924{\pm}0.005$	$0.343{\pm}0.002$
π^{\pm} selector	36028	$0.889 {\pm} 0.005$	$0.305{\pm}0.002$
ω Meson			
χ^2 prob	34566	$0.959{\pm}0.005$	$0.293{\pm}0.002$
K_S flight distance veto	33435	$0.967{\pm}0.005$	$0.283{\pm}0.002$
$m_{\pi\pi\pi}$	18411	$0.551{\pm}0.004$	$0.156{\pm}0.001$
$\cos\theta_H < 0.7$	16175	$0.879 {\pm} 0.007$	$0.137{\pm}0.001$
Vertexing			
$ \Delta z < 0.4\mathrm{cm}, \Delta z_{\mathrm{err}} < 0.04\mathrm{cm}$	15090	$0.933 {\pm} 0.008$	$0.128 {\pm} 0.001$
NN output			
NN> 0.65	11486	$0.761 {\pm} 0.007$	$0.097 {\pm} 0.001$
fit region			
total	8297		$0.070 {\pm} 0.001$

Table 8.7: The efficiencies of different cuts for the $B^0 \to \omega \gamma$ decay mode with CM2-based signal Monte Carlo. Truth-matching is not required for the individual cut efficiencies, only for the last line-item. The efficiency of the corresponding cuts in SP5 signal Monte Carlo is shown in column 2.

	fit region		S	signal region
Description	raw	(200 fb^{-1})	raw	(200 fb^{-1})
$B^0 ightarrow \omega \gamma$	10662	$9.5{\pm}0.1$	9591	$8.5 {\pm} 0.1$
$B \to \omega \pi^0$	243	$0.50{\pm}0.03$	171	$0.35 {\pm} 0.03$
$B \to \omega \eta$	267	$2.2{\pm}0.1$	162	$1.3 {\pm} 0.1$
Sum		$2.7{\pm}0.1$		$1.7{\pm}0.1$
$B^0 \to K^{*0} \gamma$	76	$2.7{\pm}0.3$	21	$0.7{\pm}0.2$
$B^+ \to K^{*+} \gamma$	78	$2.3 {\pm} 0.3$	28	$0.8{\pm}0.2$
Sum		$5.0 {\pm} 0.4$		$1.6 {\pm} 0.2$
other $B^0 \overline{B^0}$ backgrounds	20	11.5 ± 2.6	3	$1.7{\pm}1.0$
other B^+B^- backgrounds	12	$14.0{\pm}4.0$	1	$1.2{\pm}1.2$
Sum		25.5 ± 4.7		2.8 ± 1.5
off-peak	141	1763 ± 148	14	$175 {\pm} 46$

Table 8.8: The expected yield for the $B^0 \to \omega \gamma$ decay mode at 200 fb⁻¹ after applying all cuts. The yields are obtained using the assumed branching fraction or cross sections shown in Table 7.3.

Chapter 9

Signal extraction for individual modes

This chapter presents the likelihood fit that determines the signal yield of each signal mode individually. The fit estimates the raw number of signal events surviving the selection criteria. Converting this measurement into an estimation of the branching fractions requires an estimated efficiency of the selection criteria, which will be described in Chapter 10 along with a treatment of the systematic uncertainties of the measurement and the estimates for the individual branching fraction.

The use of the likelihood fit is described in Chapter 3.3. The likelihood function $\mathcal{L}(\vec{n}, \vec{\alpha})$ over a set of independent variables \vec{x} is constructed from a set of PDFs $\mathcal{P}_i(\vec{x}_j; \vec{\alpha}_i)$. for the various candidate hypotheses *i* (signal, continuum background, combinatoric *B* backgrounds, etc.). For these analyses, the $\mathcal{P}_i(\vec{x}_j; \vec{\alpha}_i)$ are constructed as if the variables \vec{x} are uncorrelated for each candidate hypothesis. This assumption leads to separable PDF's of the form

$$\mathcal{P}_i(\vec{x}_j; \vec{\alpha}_i) \equiv \prod_k \mathcal{P}_{ik}(x_k; \vec{\alpha}_{ik})$$

Construction of such PDFs becomes simply a matter of fitting a parameterized shape to each distribution of each variable for each candidate hypothesis. However, the assumption also leads

to a source of systematic error that must be examined carefully. The detailed BABAR Monte Carlo is used to determine the shapes of the PDFs.

The likelihood fits are implemented in the RooFit [46] package for the Root [47] data analysis framework.

This chapter begins with a recapitulation of the independent variables upon which the fits are based, followed by a description of the construction, verification, and application to real data of the likelihood fit for each signal mode.

9.1 Likelihood fit variables

Chapter 3.3 describes the variables m_{ES} and ΔE^* that form the foundation of discrimination against the remaining combinatoric backgrounds, both from continuum and B decays. m_{ES} and ΔE^* are included as independent variable in the fit for each of the four analysis modes. In these variables, properly reconstructed signal candidates portray strongly peaking distributions, while combinatoric backgrounds show flatter distributions.

For the $B^0 \to K^{*0}\gamma$ mode, m_{ES} and ΔE^* are the only independent variables in the fit. The signal $B^0 \to K^{*0}\gamma$ is prominent in these variables and no further variables are required to discriminate signal from background.

For the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ decays m_{ES} and ΔE^* are not enough. The branching fractions for these modes are so small as to make their peaks appear as mere statistical fluctuations of the large backgrounds. m_{ES} and ΔE^* alone do not give enough discrimination against B backgrounds that peak in the signal region. To enhance the power of the fit to discriminate signal from background, the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ fits include two additional independent variables: the output of the neural network o_{nn} , and the Fisher discriminant F. The neural network provides another variable in which the signal distribution peaks strongly while the continuum distribution is smooth. The Fisher discriminant aids in discrimination against certain dangerous B backgrounds described in Chapter 7.

9.2 Mode
$$B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$$

9.2.1 Fit components

Three candidate hypotheses are included as components in the fit: signal candidates, continuum background candidates, and B background candidates.

When the likelihood fit is applied to real data, several parameters defining the shapes of distributions are allowed to 'float'-that is they become variables in the parameter space over which the likelihood function in maximized. It is desirable to float as many of these shape parameters as possible to avoid systematic uncertainties associated with subtle differences that may exist between the simulated detector response in the Monte Carlo and the actual detector response to real events. In general, it is not possible to float every parameter. If the likelihood function is too general, the maximization fit may fail to converge or converge to pathological local minima. Table 9.1 indicates which of the parameters of the component PDFs are fixed and which of them are allowed to float in the likelihood fit.

9.2.1.1 Signal distributions

Studies on the BABAR detailed Monte Carlo events show that the m_{ES} distribution of signal decays is well represented by a simple Gaussian distribution (Appendix A.1). Both of the Gaussian parameters $\langle m_{ES} \rangle$ and $\sigma_{m_{ES}}$ are allowed to float in the fit.

The ΔE^* distribution has a significant tail on the negative side due to shower losses in the calorimetric energy measurements. This asymmetric peak distribution is well described



Figure 9.1: $K^{*0} \to K^+\pi^-$ Component Fit to signal $K^{*0} \to K^+\pi^-$ Monte Carlo of Gaussian $(m_{ES}) \times \text{Crystal Ball } (\Delta E^*).$

by a Crystal Ball function (Appendix A.3):

$$\mathcal{F}_{CB}(\Delta E^*) = C_{CB} \cdot \left\{ \begin{array}{ll} e^{-\frac{(\Delta E^* - \langle \Delta E^* \rangle)^2}{2\sigma \Delta E^{*2}}} & \text{for } \Delta E^* > \langle \Delta E^* \rangle - \alpha_{\Delta E^*} \sigma_{\Delta E^*} \\ \frac{\left(\frac{\alpha_{\Delta E^*}}{\alpha_{\Delta E^*}}\right)^{n_{\Delta E^*}} \cdot e^{-\frac{\alpha_{\Delta E^*}^2}{2}}}{\left(\frac{\langle \Delta E^* \rangle - \Delta E^*}{\sigma_{\Delta E^*}} + \frac{n_{\Delta E^*}}{\alpha_{\Delta E^*}} - \alpha_{\Delta E^*}\right)^{n_{\Delta E^*}}} & \text{for } \Delta E^* \le \langle \Delta E^* \rangle - \alpha_{\Delta E^*} \sigma_{\Delta E^*}} \end{array} \right\},$$

where C_{CB} is a normalization constant, $\langle \Delta E^* \rangle$ and $\sigma_{\Delta E^*}$ are the mean and standard deviation of the Gaussian part of the function, $\alpha_{\Delta E^*}$ controls the location of the connection between the Gaussian part and exponential tail, and $n_{\Delta E^*}$ controls the shape of the exponential tail. Figure 9.1 shows fits of these distributions to the BABARMonte Carlo signal events.

Studies with Monte Carlo events indicate that allowing the Crystal Ball parameters $\alpha_{\Delta E^*}$ and $n_{\Delta E^*}$ to float in the fit frequently leads to convergence failures. When the likelihood fit is applied to data, these two parameters are fixed to the values estimated from the fit to Monte Carlo depicted in Figure 9.1 and listed in Table 9.1. The other two Crystal Ball parameters defining the location and width of the peak ($< \Delta E^* >$ and $\sigma_{\Delta E^*}$ are allowed to float. The systematic uncertainties associated with deviations of the true signal distributions of m_{ES} and ΔE^* from these functional shapes used in the likelihood fit are estimated in Chapter 10.8.1.



Figure 9.2: $K^{*0} \to K^+\pi^-$ Component Fit to continuum Monte Carlo of Argus $(m_{ES}) \times 1^{st}$ order polynomial (ΔE^*) .

9.2.1.2 Continuum distributions

The m_{ES} distribution of continuum background conforms to the well-known Argus function (Appendix A.4)

$$\mathcal{F}_{Argus}(m_{ES}) = C_{Argus} \cdot \frac{m_{ES}}{E_{BEAM}} \cdot \sqrt{1 - \frac{m_{ES}^2}{E_{BEAM}^2}} \cdot e^{-\xi_{m_{ES}} \left(1 - \frac{m_{ES}^2}{E_{BEAM}^2}\right)} ,$$

where C_{Argus} is a normalization constant. The Argus function is a threshold or cutoff function derived from phase-space considerations in random candidate reconstruction. The cutoff parameter E_{BEAM} is fixed by the known CM energy of the interaction and is not allowed to float in the fit. The Argus shape parameter $\xi_{m_{ES}}$ is allowed to float.

The ΔE^* distribution of the continuum background can be represented by a simple first order polynomial. The slope of the polynomial is allowed to float in the fit. A first order polynomial makes a one parameter PDF-the constant of the polynomial is fixed by the normalization condition of the PDF. Figure 9.2 and Table 9.1 show the results of fitting these shape functions to Monte Carlo continuum events.

9.2.1.3 *B* background distributions

The component PDF for the B background component must represent a proportional sum of the various types of B background described in Chapter 7.1, namely

- Down-feed from $B \to X_s \gamma$ decays,
- Cross-feed from $B^+ \to K^{*+}\gamma, K^{*+} \to K^+\pi^0$,
- Combinatoric coincidences.

As Chapter 7.1 indicates, the proportion of $B \to X_s \gamma$ decays relative to other B decay processes was not well measured at the time this analysis was performed. Nor is the yield of any of the B background components large enough to reliable estimate the relative proportions. The PDF included in the fit represents the shape of the distributions of a sum of these components in proportions based on the best estimates available. The systematic uncertainty associated with the uncertain proportion of $B \to X_s \gamma$ is investigated in Chapter 10.8.1 by varying the shape of the B background PDF to represent different proportions of the background components.

Several functional representations of the m_{ES} distribution of the best-estimate sum *B* background were tested. The best representation was obtained with a Novosibirsk function (Appendix A.2) as depicted in Figure 9.3:

$$\mathcal{F}_{Novo}(m_{ES}) = C_{Novo} \cdot \exp(-\frac{1}{2} \left(\frac{\ln(1 + \tau_{m_{ES}} \cdot (m_{ES} - \langle m_{ES} \rangle) \cdot \frac{\sinh(\tau_{m_{ES}}\sqrt{\ln 4})}{\sigma_{m_{ES}}\tau m_{ES}\sqrt{\ln 4}}}{\tau_{m_{ES}}}\right)^2 + \tau_{m_{ES}}^2) ,$$

where C_{Novo} is a normalization constant, $\langle m_{ES} \rangle$ and $\sigma_{m_{ES}}$ control the location of the maximum and width of the distribution similar to Gaussian parameters, and $\tau_{m_{ES}}$ controls the shape of the tail. The ΔE^* distribution follows an exponential shape well. Monte Carlo studies indicate that allowing any of the parameters of these shapes to float leads to convergence problems in the fit to data Table 9.1. All of the parameters are fixed in the fit to data to the values obtained in a



Figure 9.3: $K^{*0} \to K^+\pi^-$ Component Fit to non-signal B Monte Carlo of Novosibirsk (m_{ES}) × exponential (ΔE^*).

fit to the Monte Carlo distributions of the B background. The systematic uncertainty associated with fixing this shape in the fit is investigated in Chapter 10.8.1.

9.2.1.4 Direct CP asymmetry

Another goal of the fit is to measure the direct CP asymmetry a_{CP} of the signal in addition to yields of the three candidate hypotheses. To do this, the candidates over which the fit is applied are divided into two disjoint populations depending on the charge of the charged final state candidate tagged as the kaon. Candidates with a positively charged kaon are placed in the population X_+ representing $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ decays, while candidates with a negative kaon candidate are assigned to the population X_- representing the conjugate decay $\overline{B}^0 \to \overline{K}^{*0}\gamma$, $\overline{K}^{*0} \to K^-\pi^+$.

Each of these populations has a PDF representing each of the three candidate hypotheses. This doubles the number of the number of PDFs as each component PDF is reproduced for each candidate hypothesis–a total of six components in the fit. With each of the six components is associated a yield, e.g. $n_{sig,+}$ represents the estimated number of signal candidates in population X_+ . With each candidate hypothesis is associated a direct CP asymmetry $a_{CP,i} = \frac{n_{i,-} - n_{i,+}}{n_{i,-} + n_{i,+}}$.

The shapes of the distributions for the candidate hypotheses of population X_+ are expected to be identical to those of population X_- . Thus the values of the floated parameters α_+ in each distribution over population $X_+ \mathcal{P}_{i,+}(\vec{x}; \vec{\alpha}_{i,+})$ are constrained to the same values as their corresponding parameters α_- in population X_- .

Monte Carlo studies indicate that the *B* background yield is too small for reliable independent estimation of $a_{CP,BBkg}$ (Chapter 9.2.2). Thus in the likelihood fit to data it is constrained $a_{CP,BBkg} = a_{CP,ContBkg} \equiv a_{CP,Bkg}$. With the population specific yields $n_{i,\pm}$ related to the total hypothesis yield n_i and CP asymmetry $a_{CP,i}$ by $n_{i,\pm} = n_i(1 \mp a_{CP,i})$, the likelihood function becomes:

$$\mathcal{L}(\vec{n},\vec{\alpha}) = \exp\left(-\sum_{i\in H_{ev}} n_i\right)$$

$$+ \left[\prod_{j\in X_+} \left(\sum_{i\in H_{ev}} n_i(1-a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right) + \prod_{j\in X_-} \left(\sum_{i\in H_{ev}} n_i(1+a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right)\right],$$

$$+ \left[\prod_{j\in X_+} \left(\sum_{i\in H_{ev}} n_i(1-a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right) + \prod_{j\in X_-} \left(\sum_{i\in H_{ev}} n_i(1+a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right)\right],$$

$$+ \left[\prod_{j\in X_+} \left(\sum_{i\in H_{ev}} n_i(1-a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right) + \prod_{j\in X_-} \left(\sum_{i\in H_{ev}} n_i(1+a_{CP,i})\mathcal{P}_i(\vec{x}_j;\vec{\alpha}_i)\right)\right],$$

where the H_{ev} sums are over event hypotheses $H_{ev} = \{sig, BBkg, ContBkg\}$.

The parameters over which the likelihood function are maximized are the shape parameters of the individual PDFs listed in Table 9.1, the total yields in each of the candidate hypotheses n_i , and the two CP asymmetry parameters $a_{CP,sig}$ and $a_{CP,Bkg}$ -a total 11 parameters.

9.2.2 Toy Monte Carlo studies

Toy Monte Carlo studies are a tool for examining the robustness of the likelihood fit and for estimating some of the systematic uncertainties associated with the fitting procedure. In a Toy Monte Carlo study, the likelihood fit is applied to many (usually ≥ 100) prototype data sets that are generated via Monte Carlo methods. Each data set is referred to as a toy data set or experiment and represents the content of real data under a well-defined scenario. The collection of experiments generated under the same scenario is usually termed an ensemble. The

Data Set/Parameter		float/fit	Value±Error
Signal MC			
$Gaussian(m_{ES})$	$< m_{ES} > (\text{GeV})$	float	$5.27946{\pm}0.00011$
	$\sigma_{m_{ES}}$ (GeV)	float	$0.002628{\pm}0.000077$
Crystal Ball(ΔE^*)	$<\Delta E^* > (\text{GeV})$	float	$0.01802{\pm}0.00367$
	$\sigma_{\Delta E^*}$ (GeV)	float	$0.04273 {\pm} 0.00257$
	$lpha_{\Delta E^*}$	fix	$0.700083 {\pm} 0.136$
	$n_{\Delta E^*}$	fix	$5.19 {\pm} 3.85$
Continuum background I	MC		
$\operatorname{Argus}(m_{ES})$	$\xi_{m_{ES}}$	float	$-11.00{\pm}2.18$
	E_{BEAM} (GeV)	fix	5.29
$1^{\rm st}$ order Poly (ΔE^*)	$P0_{\Delta E^*} \; (\text{GeV}^{-1})$	float	$-0.697 {\pm} 0.111$
Off resonance data			
$\operatorname{Argus}(m_{ES})$	$\xi_{m_{ES}}$	float	$-17.68{\pm}2.04$
	E_{BEAM} (GeV)	fix	5.27
$1^{\rm st}$ order Poly (ΔE^*)	$P0_{\Delta E^*} \; (\text{GeV}^{-1})$	float	$-1.037 {\pm} 0.105$
$B\bar{B}$			
Novosibirsk (m_{ES})	$< m_{ES} >_{Bbkg} (GeV)$	fix	$5.28107{\pm}0.00145$
	$\sigma_{m_{ES},Bbkg}$ (GeV)	fix	$0.0117 {\pm} 0.0013$
	$ au_{Bbkg}$	fix	-1.061 ± 0.129
Exponential(ΔE^*)	$c_{Bbkg} \; (\mathrm{GeV}^{-1})$	fix	-6.640 ± 0.593

Table 9.1: $K^{*0} \to K^+\pi^-$ PDF parameters fit to fully simulated Monte Carlo and off resonance data



Figure 9.4: $K^{*0} \to K^+\pi^-$ Full Fit to SP4 Monte Carlo: The data are a weighted sum of all components. The PDF is the combined PDF for the branching fraction fit.

ensemble averages and distributions of the fit values of the floating parameters and of maximum likelihood value itself convey information about the robustness and fidelity of the fit function.

These qualitative statements can be clarified using the A_{CP} measurement in $B^0 \to K^{*0}\gamma$ as an example. The PDFs used in the likelihood function itself create an useful first scenario to test the intrinsic robustness of the fit. If the experiments are generated in accordance with the likelihood PDFs–with the assumptions that m_{ES} and ΔE^* are uncorrelated for each candidate hypothesis and the distributions conform exactly to the empirical function fit to the fully simulated BABAR Monte Carlo-then the likelihood fit applied to the experiments should reproduce the values of the floating parameters used to generate the experiments. A toy ensemble generated in this way, directly from the PDF used to fit it, is termed a 'pure' toy ensemble. Of course, there will be statistical fluctuations of the fit parameters, but the ensemble distributions of the fit values of the parameters should follow a Gaussian (A.1) shape with a mean at the generator value. A more commonly used figure of merit is termed the 'pull' p(a) of a floating parameter a. For each toy data set, the likelihood fit estimates a value a_{est} and an error $\sigma(a)$ for each of the floating parameters. The pull p(a) is defined in terms of the value of the parameter used in generating the data set a_{gen} as $p(a_{est}) = (a_{est} - a_{gen})/\sigma(a)$. The distribution of the pull p(a) over an ensemble should be Gaussian with mean 0 and σ 1. Significant deviations from these values in the pull distributions are indicative of a possible problem with the likelihood fit.

Studies on the fully simulated BABAR Monte Carlo give estimates for the yields N_i of each of the six candidate hypotheses (2 CP states × {signal, continuum, B background}) expected in real data (Table 8.2). Each toy data set contains a randomly generated number of candidates n_i of each type *i*. The n_i are generated from a Poisson distribution with mean N_i . Hence, the ensemble distribution of n_i is Poisson with mean at the expected value N_i . The n_i candidates are generated from the appropriate PDF described in Chapter 9.2.1 via the common rejection method. For example, the two dimension PDF for signal candidates is a Gaussian in m_{ES} times a Crystal Ball in ΔE^* normalized to unit volume over the fit region:

$$\mathcal{P}_{sig}(m_{ES}, \Delta E^*; \vec{\alpha}) \equiv \mathcal{F}_{Gauss}(m_{ES}) \times \mathcal{F}_{CB}(\Delta E^*)$$

Events are ordered pairs of $(m_{ESj}, \Delta E_j^*)$ values generated randomly and uniformly over the fit region. Along with an event is a randomly generated selection value s_j . If

 $s_j < \mathcal{P}_{sig}(m_{ESj}, \Delta E_j^*; \vec{\alpha})$ the event is included in the toy data set. This procedure is repeated until n_i events have been included in the data set for each candidate hypothesis. In the toy study to be described next, 751 independent data sets were generated in this way to form the ensemble.

Applying a pure toy Monte Carlo study to the $B^0 \to K^{*0}\gamma$ likelihood function disqualifies $a_{CP,BBkg}$ as a free parameter as follows. The data sets for this study are generated from the PDFs described in Chapter 9.2.1. The numbers of events for each of the six hypotheses are Poisson-distributed about the mean numbers expected in the real 81.9 fb⁻¹ data set according to studies with the detailed BABAR Monte Carlo (Table 8.2) with the assumption that there is no CP asymmetry for any of the three candidate types (signal, continuum, *B* background) $(a_{CP,i} = 0)$. The ensemble contains 751 toy data sets generated with these assumptions. The likelihood function described in Chapter 9.2.1 is maximized over each of the toy data sets in the ensemble. The list of floated parameters is that listed in Table 9.1 and the total yields and CP asymmetries for each of the three candidate hypotheses. Note that this is 12 total parameters. The constraint $a_{CP,BBkg} = a_{CP,ContBkg}$ is not enforced, $a_{CP,BBkg}$ and $a_{CP,ContBkg}$ are varied separately.

Table 9.2 shows the resulting mean and RMS values for the floating fit parameter distributions over the ensemble. With an RMS value of 0.260, it is clear that $a_{CP,BBkg}$ cannot be measured well by the likelihood fit even over these most ideal of data set. Several toy data sets produce estimates for $a_{CP,BBkg}$ with very large errors. These large errors are commonly

	Float A	CP, BBkg	$A_{CP,BBkg}$	$= A_{CP,ContBkg}$
Variable	Mean	RMS	Mean	RMS
nBBkg	170.1	51.6	170.1	51.8
nContBkg	2232.4	67.0	2232.4	67.1
nSig	581.4	28.5	581.35	28.54
$a_{CP,BBkg}$	-0.000	0.260	N/A	
$a_{CP,ContBkg}$	-0.000	0.026	-0.00	0.026
$a_{CP,Sig}$	0.0020	0.047	0.00	0.047
$P0_{\Delta E}$	-0.715	0.154	-0.715	0.155
ξ	-11.65	3.14	-11.65	3.14
$<\Delta E>$	0.0179	0.0030	0.018	0.003
$\sigma_{\Delta E}$	0.0428	0.0026	0.0427	0.0026
$< m_{ES} >$	5.27948	0.00013	5.27948	0.00013
$\sigma_{m_{ES}}$	0.00258	0.00012	0.00258	0.00011

Table 9.2: $K^{*0} \to K^+ \pi^-$ Values of Fit Parameters for CP Asymmetry fit from pure toy Monte Carlo Study.

accompanied by very low values of the estimated error on N_{BBkg} . These appear as bins with small but nonzero populations far removed from the main peak in the center two plots of Figure 9.5A and expanded in Figure 9.6, which show the estimated error distributions of N_{BBkg} and $a_{CP,BBkg}$. Such unusual errors are commonly a symptom of convergence problems with the likelihood fit. Not only is $a_{CP,BBkg}$ poorly measured, its inclusion may be leading the fit into regions of parameter space that make fit convergence difficult.

This toy Monte Carlo study was repeated, this time including the constraint $a_{CP,BBkg} = a_{CP,ContBkg}$. The same ensemble of data sets was used, only the likelihood function was changed to reflect the constraint and reduce the number of free parameters from 12 to the 11 proposed for the final fit to data. The results of this study are also presented in Table 9.2 and Figure 9.5. Note from Figures 9.5 and 9.6 that removing $a_{CP,BBkg}$ as a free parameter eliminates the instances of unusual error in the N_{BBkg} .

A final pure toy Monte Carlo study was conducted to verify the fit's sensitivity to an actual $a_{CP,sig}$ in data. According to the RMS of the $a_{CP,sig}$ distribution in Table 9.2



B) $A_{CP,BBkg} = A_{CP,ContBkg}$

Figure 9.5: $K^{*0} \to K^+\pi^-$ Pure toy Monte Carlo : Distributions of N_{BBkg} (first row) and $a_{CP,ContBkg}$ (second row) for two toy studies: A) in which $a_{CP,BBkg}$ is free to float in the fit and B) in which $a_{CP,BBkg}$ is fixed equal to $a_{CP,ContBkg}$. For both $a_{CP,generated} = 0$. The first column shows the distributions of the estimated parameters N_{BBkg} and $a_{CP,ContBkg}$. The second column shows the distributions of the estimated uncertainties $\sigma(N_{BBkg})$ and $\sigma(a_{CP,BBkg})$. The third column shows the distributions of the pulls $N_{BBkg}/\sigma(N_{BBkg})$ and $a_{CP,ContBkg}/\sigma(a_{CP,BBkg})$. Enlarged versions of the plots of $\sigma(N_{BBkg})$ and $\sigma(a_{CP,BBkg})$ for the freely floating $a_{CP,BBkg}$ study (A) appear in Figure 9.6, which more clearly shows the anomalous values.



Figure 9.6: Enlarged versions of the plots of distributions $\sigma(N_{BBkg})$ and $\sigma(a_{CP,ContBkg})$ from Figure 9.5 A. The distributions are from pure toy Monte Carlo study of $K^{*0} \to K^+\pi^-$ in which $a_{CP,BBkg}$ is freely estimated by the likelihood fit. Note the occassional occurrence of estimated errors far from the central peak. This is often indicative of a failure in the likelihood fit procedure.

for the last $a_{CP,sig} = 0$ pure toy, an asymmetry of $a_{CP,sig} = 0.1$ should be unambiguously measurable. An ensemble of 751 toy data sets was generated from the PDFs described in Chapter 9.2.1 in the same way as for the previous toy Monte Carlo study with $a_{CP,sig} =$ 0.1. That is, the mean number of representative signal candidates included in the population X_+ is no longer equal to the mean number of candidates included in X_- for the data set. Rather, they have means set by the value of $a_{CP,sig}$: $N_{+,sig} = N_{sig}(1 - a_{CP,sig})/2 = 0.45N_{sig}$, $N_{-,sig} = N_{sig}(1 + a_{CP,sig})/2 = 0.55N_{sig}$. The summary of mean and RMS values for the estimated values of the floating fit parameters shown in Table 9.3 indicates that, indeed, the fit is sensitive to values of $a_{CP,sig} \approx 0.1$.

A final toy Monte Carlo study is conducted to examine the effects of correlations between m_{ES} and ΔE^* in data on the likelihood fit, which assumes that the variable are uncorrelated. Table 9.4 shows that in the fully simulated *BABAR* Monte Carlo data. m_{ES} and ΔE^* are not strongly correlated for any of the three event types for $K^{*0} \to K^+\pi^-$. However, the stronger correlations in other $B \to K^*\gamma$ modes led to a further investigation, which was also applied to $K^{*0} \to K^+\pi^-$ for uniformity.

In order to introduce the correlations in m_{ES} and ΔE^* into the data, the data sets for the toy Monte Carlo ensemble include events from the fully simulated *BABAR* Monte Carlo, which simulate these correlations correctly. The continuum and *B* background components of the data set are generated from the appropriate PDFs described in Chapter 9.2.1. All of the signal events included in the data sets are drawn from the fully simulated *BABAR* Monte Carlo. The process of mixing fully simulated data with PDF generated data is known as 'embedding' by *BABAR* analysts, and a toy Monte Carlo study that uses this technique to construct the ensemble data sets is called an embedded toy Monte Carlo study.

Embedded toy Monte Carlo studies represent a necessary compromise. Ideally, every data set in a toy Monte Carlo study would draw all of its events from the fully simulated Monte

Carlo. However, usually only enough continuum and generic B Monte Carlo is generated in the BABAR simulation production to create a single data set of appropriate size. The simulation is time and compute intensive and it is impossible to generate more than one independent set of fully simulated Monte Carlo, so independent sets of continuum and B background data are generated from the PDFs-the implicit assumption that their m_{ES} and ΔE^* correlations are negligible is supported by the fully simulated Monte Carlo. Specific signal modes, like the decays $B^0 \to K^{*0}\gamma, K^{*0} \to K^+\pi^-$, are generated in relatively large quantities because of their need in the development for analysis. The set of fully generated $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^$ events used in the development of this analysis contains as many events as are expected in 2571.4 fb⁻¹-enough to create 31 independent data sets equivalent to 81.9 fb⁻¹. This is still not enough to create the desired ensemble of > 100 independent toy data sets, so a random selection of these events are included in each toy data set with the recognition that the data sets will not be completely independent. The statistics of the toy Monte Carlo study are interpreted as if there were just 31 independent data sets in the ensemble. If all of the N toy data sets were independent, the statistical error $\sigma(x)$ on the ensemble average of the fit parameter x would be computed as the ensemble standard deviation $\sigma(x) \approx \frac{1}{N-1} \sum (x - \langle x \rangle)^2$. Since the N data sets of the ensemble are not all independent, the number of independent data sets that could be constructed $N_{ind} = 31$ is used instead: $\sigma(x) \approx \frac{1}{N_{ind} - 1} \sum (x - \langle x \rangle)^2$.

Table 9.3 displays the means of the fit parameters over the ensemble of 751 signalembedded toy data sets, and Table 9.5 tabulates their pulls. These toys were generated with an average 590 embedded signal decays equally distributed between the two modes. It is clear that the likelihood fit underestimates the number of signal events. To determine the necessary correction factor, the distribution of the ratios of the number fit yields of each candidate hypothesis $N_{i,fit}$ to the number of generated or included events of each type $N_{i,gen}$, $N_{i,fit}/N_{i,gen}$, is fit with a Gaussian distribution. The fit mean of the Gaussian is used as a correction factor in the final

result. The fit $\sigma/\sqrt{N_{ind}}$ of the Gaussian is incorporated into the estimated systematic error (Chapter 10.8.1). These efficiency distributions are displayed in Figure 9.7 and summarized in Table 9.6.

	Pure Toy	Pure Toy Monte Carlo		ded Toy	
	$(A_{CP,S})$	$S_{ig} = 0.1)$	$(A_{CP,Si})$	$(A_{CP,Sig} = 0.0)$	
Variable	Mean	RMS	Mean	RMS	
nBBkg	175.6	52.3	190.4	54.6	
nContBkg	2223.2	65.9	2221.9	70.11	
nSig	581.1	28.7	573.8	28.7	
$a_{CP,ContBkg}$	-0.000	0.021	-0.002	0.021	
$a_{CP,Sig}$	0.102	0.045	0.00	0.047	
$P0_{\Delta E}$	-0.705	0.151	-0.692	0.158	
ξ	-11.46	3.31	-11.32	3.37	
$<\Delta E>$	0.0181	0.0028	0.020	0.003	
$\sigma_{\Delta E}$	0.0427	0.0026	0.039	0.0023	
$< m_{ES} >$	5.27948	0.00013	5.27952	0.00012	
$\sigma_{m_{ES}}$	0.00259	0.00011	0.00256	0.00011	

Table 9.3: $K^{*0} \to K^+ \pi^-$ Values of Fit Parameters for CP Asymmetry fit from pure toy Monte Carlo Study ($A_{CP,Sig} = 0.1$) and Embedded Toy study.

Data set	Correlation
Continuum MC	-0.045
Off-peak Data	0.027
Generic B MC	-0.084
Signal MC	0.020

Table 9.4: The correlations between m_{ES} and ΔE^* for continuum background Monte Carlo, off-peak data, generic B Monte Carlo and signal Monte Carlo for reconstructed $B^0 \to K^{*0}\gamma$ $(K^{*0} \to K^+\pi^-)$ candidates. The events in the fitting region are selected and all the cuts are applied.

9.2.3 Fit results on data

After the likelihood fit procedure has been vetted by toy Monte Carlo studies the fit can be applied to the real data. The results are depicted in Figure 9.8 with the fit values of

	Pure Toy MC		Embedded 7	Toy MC
Description	Mean of Pull	σ of Pull	Mean of Pull	σ of Pull
$A_{CP,Sig}$	0.042	1.002	-0.014	1.013
$A_{CP,Bkg}$	-0.012	1.006	-0.110	1.004
Sig. Yield	-0.318	0.997	-0.589	1.019
Cont. Bkg. Yield	0.243	0.992	0.086	1.040
B Bkg. Yield	-0.185	0.995	0.207	1.046
$< m_{ES} >$	0.017	1.004	0.336	0.950
$<\Delta E^*>$	-0.003	1.011	0.636	1.097
$\sigma_{m_{ES}}$	-0.133	1.076	-0.310	1.014
$\sigma_{\Delta E^*}$	-0.050	1.055	-1.596	1.036
ξ	-0.209	0.957	-0.114	1.023
$P0_{\Delta E^*}$	-0.136	1.022	0.013	1.051

Table 9.5: $K^{*0} \to K^+\pi^-$ Toy Monte Carlo Pulls, A_{CP} fit: The mean and σ of the pull distributions for both Pure and Signal-Embedded Toy studies for each of the floating parameters.

Yield Parameter	Pure Toy	Signal Embedded Toy
ϵ_{nSig}	$0.99958{\pm}0.00082$	$0.97386{\pm}0.00085$
$\sigma(\epsilon_{nSig})$	$0.02592{\pm}0.00059$	$0.02746{\pm}0.00062$
$\epsilon_{nContBkg}$	$1.00046 {\pm} 0.00070$	$1.00155 {\pm} 0.00070$
$\sigma(\epsilon_{nContBkg})$	$0.02216{\pm}0.00050$	$0.02263 {\pm} 0.00050$
ϵ_{nBBkg}	$0.9960{\pm}0.0090$	$1.0658{\pm}0.0092$
$\sigma(\epsilon_{nBBkq})$	$0.2833 {\pm} 0.0064$	$0.2983{\pm}0.0066$

Table 9.6: $K^{*0} \to K^+ \pi^-$ Toy Monte Carlo yield efficiencies



Figure 9.7: $K^{*0} \rightarrow K^+\pi^-$ Embedded Toy MC study of fits with randomly selected simulated signal data. Top row: distributions for the estimated component yields. Bottom row: distributions of the ratio of estimated to generated component yield with Gaussian fit.

the parameters summarized in Table 9.7. There are no surprises in the fit. The fit values of the parameters are used to estimate the branching fraction and CP asymmetry of $B^0 \to K^{*0}\gamma$ in Chapter 10.8.1.

A final toy Monte Carlo study is used to validate the fit maximum value of the likelihood function. Assume that the true distribution of real data is represented by the parameters and yields produced by the likelihood fit. How likely is the value of the likelihood function under this assumption? A signal embedded toy Monte Carlo study is used to answer this question. The data sets in the study ensemble are composed of embedded events from the fully simulated signal Monte Carlo and generated events from the continuum and B background PDFs. The shape parameters of the continuum background PDF used in generation are fixed to the values produced by the likelihood fit to data. The B background shape was fixed for the likelihood fit to data and remained fixed in the same shape for generation of the toy events. The numbers of events of each type included in data set are the values estimated by the likelihood fit to data.



Figure 9.8: $K^{*0} \to K^+\pi^-$: CP fit to onpeak RunI + RunII data. The fit regions are projected onto m_{ES} (right) and ΔE^* (left). The points with errors represent the binned data. The blue lines represent the complete fit PDF. The black, red, and green lines represent the signal, continuum, and *B* background components.

In the case of the signal component, the number of events in each data set is the fit yield scaled up by the reciprocal of the fit efficiency determined in Chapter 9.2.2. The likelihood fit was applied to each data set in the ensemble exactly as it was applied to the real data. The resulting distribution of the negative log likelihood (NLL) value appears in Figure 9.9. The NLL value produced by the fit lies well within the toy Monte Carlo distribution signifying a good fit.

Modes	$K^{*0} \to K^+ \pi^-$			
Param	BF fit	A_{CP} fit		
Signal	582.6	583 ± 30		
$A_{CP,sig}$	N/A	-0.069 ± 0.046		
Cont. Bkg.	2579.8	$2579{\pm}70$		
B Bkg.	141.5	142 ± 54		
$A_{CP,bkg}$	N/A	$0.006 {\pm} 0.020$		
$< m_{ES} >$	5.27980	$5.27980 {\pm} 0.00013$		
$\sigma_{m_{ES}}$	0.00254	$0.00255{\pm}0.00012$		
$<\Delta E^*>$	-0.0109	-0.0109 ± 0.0031		
$\sigma_{\Delta E^*}$	0.0450	$0.0450{\pm}0.0027$		
ξ	-12.6	-12.58 ± 3.0		
$P0_{\Delta E^*}$	-0.492	-0.490 ± 0.14		

Table 9.7: Comparison of results from branching fraction and A_{CP} fits to data



Figure 9.9: A_{CP} Goodness of fit test. Negative maximum likelihood distributions of a signal embedded toy studies for $K^{*0} \to K^+\pi^-$. The yields and parameters for PDF are fixed from unblinded results, the negative maximum likelihoods from the unblinded results are marked by red lines in the same figures.

9.3 Modes $B \to \rho \gamma$ and $B^0 \to \omega \gamma$

The likelihood fit functions for the decays $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$, and $B^0 \to \omega \gamma$ are more complicated than that for $(B^0 \to K^{*0} \gamma, K^{*0} \to K^+ \pi^-)$. The relative backgrounds yields are expected to be much larger than for $B^0 \to K^{*0} \gamma$, and as such must be treated more carefully and completely. Clear discrimination must be established against the peaking backgrounds, which have concentrations of events in the signal region in m_{ES} and ΔE^* . To enhance separation between signal and the various backgrounds each candidate hypothesis is represented by a four dimensional PDF over m_{ES} , ΔE^* , the neural network output o_{nn} , and the Fisher discriminant F.

The likelihood fit for each mode has PDF components representing signal decays, continuum background decays, and peaking $B \to (\rho/\omega)(\eta/\pi^0)$ decays. In addition, the $B \to \rho\gamma$ decay modes have PDF components representing peaking $B \to K^*\gamma$ backgrounds and combinatoric *B* backgrounds. The $B \to K^*\gamma$ and combinatoric *B* backgrounds are much less significant for the $B^0 \to \omega\gamma$ mode and do not need independent PDFs (Chapter 7.5).

The shapes of these PDFs for each mode are fixed to distributions of the variables in the fully simulated *BABAR* Monte Carlo and in data sidebands. Table 9.9 summarizes the shapes used for each PDF. The determination of each of these shapes will be described later in this chapter.

As in the $B^0 \to K^{*0}\gamma$ analysis, the likelihood functions are vetted using toy Monte Carlo studies and finally applied to data.

9.3.1 Signal components

The PDFs used to represent the signal components of each of the $B \rightarrow (\rho/\omega)\gamma$ modes are depicted in Figures 9.10-9.12.

Component	Yield	m_{ES}	ΔE^*	Neural Net Shape	Fisher Shape
Signal	Fit	MC	MC	MC, val: CS	MC
peaking B background	MC	MC	MC	same as signal	MC
combinatoric ${\cal B}$ background	Fit	MC	MC	same as signal	MC
continuum background	Fit	Fit	Fit	side band on-peak	side band on-peak

Table 9.8: The treatment of the normalization and shape parameters of the individual PDF components. The combinatoric *B* background is negligible in the $B^0 \rightarrow \omega \gamma$ decay mode and is not treated in the likelihood fit. The neural net signal shape is fixed by Monte Carlo but validated by a control sample. The neural net shape for peaking *B* background and combinatoric *B* background is fixed to be the same as the shape in signal.

Mode	Description	m_{ES}	ΔE^*	neural net	Fisher
$B^0 o ho^0 \gamma$	signal	Crystal-Ball	Crystal-Ball	Crystal-Ball	KEYS
	$B \to K^* \gamma$ bkg.	Crystal-Ball	Crystal-Ball	same as signal	KEYS
	$B \to \rho(\pi/\eta)$ bkg	Crystal-Ball	Crystal-Ball	same as signal	KEYS
	Combinatoric B bkg	KEYS	KEYS	same as Signal	KEYS
	Continuum bkg	ARGUS	2nd poly	Histogram PDF	KEYS
$B^+ \to \rho^+ \gamma$	signal	Crystal-Ball	Crystal-Ball	Crystal-Ball	KEYS
	$B \to K^* \gamma$ bkg	Crystal-Ball	Crystal-Ball	same as Signal	KEYS
	$B \to \rho(\pi/\eta)$ bkg	Crystal-Ball	Crystal-Ball	same as Signal	KEYS
	Combinatoric B bkg	KEYS	KEYS	same as signal	KEYS
	Continuum bkg	ARGUS	2nd poly	Histogram PDF	KEYS
$B^0 o \omega \gamma$	signal	Crystal-Ball	Crystal-Ball	KEYS	KEYS
	$B \to \omega(\pi/\eta)$ bkg	Crystal-Ball	Crystal-Ball	same as Signal	KEYS
	Continuum bkg	ARGUS	2nd poly	Histogram PDF	KEYS

Table 9.9: The standard functions used to build the final PDF for $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$, and $B^0 \to \omega \gamma$ decay modes. Descriptions of the functions can be found in [48] (ARGUS), [49] (Crystal-Ball), and [50] (KEYS).

Studies on the BABAR detailed signal Monte Carlo indicate that the m_{ES} and ΔE^* distributions for each of the three $B \to (\rho/\omega)\gamma$ signal modes are well represented by Crystal Ball functions (Appendix A.3). The ΔE^* distributions display the same significant negative-side tail also observed in $B^0 \to K^{*0}\gamma$ (Chapter 9.2.1). This tail is largely due to shower losses in the measurement of the high energy photon momentum. These losses also contribute to a low-side tail in m_{ES} more significantly than observed in $B^0 \to K^{*0}\gamma$.

The m_{ES} and ΔE^* distributions shapes have a theoretic foundation. This is not the case for the distributions of the neural network (NN) outputs o_{nn} and the Fisher discriminants F. More empirical representational functions must be selected for these variables. The NN distributions for the $B^0 \rightarrow \rho^0 \gamma$ and $B^+ \rightarrow \rho^+ \gamma$ modes fit Crystal Ball functions well, as seen in Figures 9.10 and 9.11. The $B^0 \rightarrow \omega \gamma$ NN distribution does not fit any relatively simple and common function and must be represented by a one-dimensional KEYS PDF (Appendix A.5). A KEYS function is a non-parametric shape that constructs a smooth distribution from a finite set of discrete points.

The Fisher discriminant F distributions for each of the signal modes also must be empirically represented by a KEYS PDF.

In the *BABAR* data set, the signal yields of the signal modes are not expected to be large enough to estimate any of the distribution shape parameters. The shapes of the signal distributions are fixed to those fit to fully simulated signal Monte Carlo. The effect of fixing the signal shape will be examined after the application of the fit to data for determination of a systematic error associated therewith.

9.3.2 Continuum background components

The PDFs used to represent the continuum background components of each of the $B \rightarrow (\rho/\omega)\gamma$ modes are depicted in Figures 9.13-9.15.



Figure 9.10: The PDFs for the signal in $B^0 \rightarrow \rho^0 \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated signal Monte Carlo from which the PDF shapes are derived.



Figure 9.11: The PDFs for the signal in $B^+ \rightarrow \rho^+ \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated signal Monte Carlo from which the PDF shapes are derived.



Figure 9.12: The PDFs for the signal in $B^0 \to \omega \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated signal Monte Carlo from which the PDF shapes are derived.
The m_{ES} and ΔE^* distributions of the continuum background events for the

 $B \to (\rho/\omega)\gamma$ modes are similar to those for $B^0 \to K^{*0}\gamma$. The m_{ES} distributions are well represented by an Argus shape (Appendix A.4). The ΔE^* distributions conform well to a simple second order polynomial.

Like the signal components, the continuum components have NN and F distributions that must be modeled empirically. Rather than the Monte Carlo events, the wealth of real events in the sidebands outside of fit region are used to create the continuum NN and F PDF models. If judiciously selected these real events, which are primarily from continuum processes with small contributions from combinatoric B candidates, represent the distributions of the continuum candidates in the fit region better than simulated Monte Carlo events. Each mode uses the union of two sidebands for developing the continuum background distributions: an upper ΔE^* sideband, and a mass-veto sideband that inverts the analysis cut on the ρ or ω mass. Explicitly defining these sidebands:

- Upper ΔE^* sideband (all modes): candidates passing all of the analysis cuts save the fit region cut in the $m_{ES} \times \Delta E^*$ region [5.2, 5.3]GeV × [0.3, 0.6]GeV.
- Mass veto sideband (B → ργ modes): candidates passing all of the analysis cuts save the fit region and m_{ππ} cut in the m_{ππ} × m_{ES} × ΔE* region [0.5, 0.63] U[0.94, 1.10]GeV × [5.2, 5.26]GeV × [-0.35, 0.6]GeV.
- Mass veto sideband (B⁰ → ωγ modes): candidates passing all of the analysis cuts save the fit region and m_{ππ} cut in the m_{ππ} × m_{ES} × ΔE* region [0.7, 0.764] U[0.795, 0.85]GeV × [5.2, 5.26]GeV × [-0.35, 0.6]GeV.

The NN PDFs used for each of the modes in the likelihood fits are the histograms of the neural network output for the union of the sidebands. That is, the PDF shape is the



Figure 9.13: The PDFs for the continuum background in $B^0 \to \rho^0 \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the sideband in *BABAR* data from which the PDF shapes are derived. In the case of the neural network (NN) distribution, the PDF is itself the sideband histogram so the two exactly coincide.

normalized step function corresponding to a histogram of the NN values of candidates in the sidebands.

The Fisher F PDFs are non-parametric KEYS PDFs (Appendix A.5) generated by the sideband events for each of the signal modes.

Continuum processes are the single largest source of candidates. As such, the shape parameters for the Argus and polynomial functions in m_{ES} and ΔE^* can be easily determined in the likelihood fit. For each mode, the Argus parameter ξ , the first and second polynomial coefficients P_1 and P_2 , and the total continuum yield are independently estimated in the likelihood fit. The constant terms of the secon order polynomial PDFs is fixed by normalization conditions on the PDFs.



Figure 9.14: The PDFs for the continuum background in $B^+ \rightarrow \rho^+ \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the sideband in *BABAR* data from which the PDF shapes are derived. In the case of the neural network (NN) distribution, the PDF is itself the sideband histogram so the two exactly coincide.



Figure 9.15: The PDFs for the continuum background in $B^0 \rightarrow \omega \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the sideband in *BABAR* data from which the PDF shapes are derived. In the case of the neural network (NN) distribution, the PDF is itself the sideband histogram so the two exactly coincide.

9.3.3 $B \rightarrow K^* \gamma$ background components

The PDFs used to represent the $B \to K^* \gamma$ background components for the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ modes are depicted in Figures 9.16 and 9.17.

Table 8.8 shows that the analysis cuts eliminate almost all of the background candidates from $B \to K^* \gamma$ decays in the $B^0 \to \omega \gamma$ mode. No component PDF is included to represent this background in the $B^0 \to \omega \gamma$ mode.

 $B \to K^* \gamma$ processes do create a significant peaking background for the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ modes (Tables 8.4 and 8.6).

The distributions of the four variables for $B \to K^* \gamma$ processes are qualitatively similar to those of the signal. In fact, the neural network output distributions, which is entirely dependent on the rest of the event (ROE), are almost identical to those for the respective signal modes. The neural network distributions of the $B \to K^* \gamma$ backgrounds for the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ modes are exactly those Crystal Ball functions used for the respective signal decays.

The m_{ES} and ΔE^* PDFs of $B \to K^* \gamma$ for both of the $B \to \rho \gamma$ decay modes are Crystal Ball functions fit to the fully simulated *BABAR* Monte Carlo data. Although the signal decays are also represented by Crystal Ball functions, the fit values of the Crystal Ball parameters for the background are significantly different and lead to a distinct PDF shape (see Figures 9.10 and 9.11 for signal shapes). In particular, the peak in the ΔE^* distribution for $B \to K^* \gamma$ backgrounds lies significantly beneath that for the signal, making the ΔE^* distribution the key discriminating variable in the distribution.

The Fisher discriminant F distributions for the $B \to K^* \gamma$ backgrounds are represented by KEYS PDFs generated from the fully simulated Monte Carlo events.

The yields of the $B \to K^* \gamma$ are too small to permit independent measurement of either their shape parameter or even their yields. Instead of treating $B \to K^* \gamma$ backgrounds as



Figure 9.16: The PDFs for the $B \to K^* \gamma$ background in $B^0 \to \rho^0 \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated $B \to K^* \gamma$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the $B \to K^* \gamma$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.

biases that must be subtracted from the final signal yields, fixed PDF components representing $B \to K^* \gamma$ are included in the likelihood functions. This accomplishes a distributed background subtraction in the fits that reduces the contribution from $B \to K^* \gamma$ backgrounds to the systematic uncertainties. The shape parameters of the PDFs are fixed to the values determined from the Monte Carlo fits. The yields of the $B \to K^* \gamma$ are fixed to values estimated from the Monte Carlo efficiency of the selection applied to $B \to K^* \gamma$ decays and the measured $B \to K^* \gamma$ branching fractions. Fixing the shape and yield of this background component contributes systematic uncertainties to the final fit yield.



Figure 9.17: The PDFs for the $B \to K^* \gamma$ background in $B^+ \to \rho^+ \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated $B \to K^* \gamma$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the $B \to K^* \gamma$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.

9.3.4 π^0/η background components

The PDFs used to represent the $B \to (\rho/\omega)(\eta/\pi^0)$ background components of each of the $B \to (\rho/\omega)\gamma$ modes are depicted in Figures 9.18-9.20.

Backgrounds from the processes $B \to \rho \eta$, $B \to \rho \pi^0$, $B \to \omega \eta$, and $B \to \omega \pi^0$, referred to collectively as $B \to (\rho/\omega)(\eta/\pi^0)$, represent a significant signal-like background in each of the $B \to (\rho/\omega)\gamma$ signal modes. Each likelihood fit has a component representing this background.

Like the $B \to K^* \gamma$ backgrounds, the $B \to (\rho/\omega)(\eta/\pi^0)$ backgrounds have neural network distributions that are almost identical to those for the signal and m_{ES} and ΔE^* distributions that are qualitatively similar to those for signal. The neural network PDFs of the $B \to \rho(\eta/\pi^0)$ backgrounds for the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ modes are exactly those Crystal Ball functions used for the respective signal decays. The neural network PDF for the $B \to \omega(\eta/\pi^0)$ background component is exactly the KEYS PDF used to represent the signal neural network distribution.

The m_{ES} and ΔE^* PDFs for each mode are Crystal Ball functions fit the appropriate fully simulated Monte Carlo distributions. Although the signal PDFs are also Crystal Ball functions, their parameters differ from those of the π^0/η backgrounds. The Fisher discriminant PDFs are KEYS PDFs generated from the fully simulated Monte Carlo data.

The key variables discriminating $B \to (\rho/\omega)(\eta/\pi^0)$ events from signal events in the fit are ΔE^* and the Fisher discriminant F (see Chapters 7.3.2, 7.4.2, and 7.5.2, and compare Figure 9.10 to Figure 9.18, Figure 9.11 to Figure 9.19 and Figure 9.12 to Figure 9.20).

Also like those of the $B \to K^* \gamma$ backgrounds, the yields of the $B \to (\rho/\omega)(\eta/\pi^0)$ are too small for either their distribution shapes or their total yields to be measured independently. Their shapes in the likelihood fit are fixed to those fit to the fully simulated BABAR Monte Carlo events. Their yields are fixed to values estimated from the Monte Carlo efficiency of the selection



Figure 9.18: The PDFs for the $\rho(\pi/\eta)$ background in $B^0 \to \rho^0 \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated $\rho(\pi/\eta)$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the $\rho(\pi/\eta)$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.

applied to $B \to (\rho/\omega)(\eta/\pi^0)$ decays and the best estimates of the $B \to (\rho/\omega)(\eta/\pi^0)$ branching fractions. As noted in Chapter 7 and Table 7.3, the branching fractions of these processes are not all well measured. This adds an additional uncertainty to the final measurement.

9.3.5 Combinatoric *B* background components

The PDFs used to represent the combinatoric B background components of the

 $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ modes are depicted in Figures 9.21 and 9.22.

The analysis cuts reduce the combinatoric B background to ignorable levels in the $B^0 \rightarrow \omega \gamma$ (Table 8.8). No PDF component is included to represent combinatoric B backgrounds for this mode.



Figure 9.19: The PDFs for the $\rho(\pi/\eta)$ background in $B^+ \to \rho^+ \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated $\rho(\pi/\eta)$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the $\rho(\pi/\eta)$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.



Figure 9.20: The PDFs for the $\omega(\pi/\eta)$ background in $B^0 \to \omega \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated $\rho(\pi/\eta)$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the $\rho(\pi/\eta)$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.

The combinatoric B backgrounds in the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ analyses are dominated by $B \to X_s \gamma$ processes. Since the fully simulated Monte Carlo sample of $B \to X_s \gamma$ events represents the number of events in a much larger sample of equivalent real events than the generic B samples, the shapes of the combinatoric B background m_{ES} , ΔE^* and Fisher discriminant FPDFs are fixed to the distributions of the $B \to X_s \gamma$ Monte Carlo. The higher statistics of the sample will lead to a much smaller systematic uncertainty in the PDF shapes.

As for $B \to K^* \gamma$ and $B \to (\rho/\omega)(\eta/\pi^0)$, the neural network PDFs for the combinatoric *B* backgrounds are fixed to exactly the distributions used for the signal. The neural network output distributions for $B \to X_s \gamma$ do not conform to the signal NN shape as well as those for the other *B* backgrounds. However, a toy Monte Carlo study shows that the deviation is negligible as a source of uncertainty.

The m_{ES} , ΔE^* , and Fisher distributions of the combinatoric *B* background component all resist representation by well known parametric functions. The PDFs for all three variables for both the $B^0 \rightarrow \rho^0 \gamma$ and $B^+ \rightarrow \rho^+ \gamma$ analyses are represented by KEYS PDFs generated from the $B \rightarrow X_s \gamma$ Monte Carlo.

The KEYS PDFs have no parameters that can be floated in the likelihood fit. There are no shape parameters for the combinatoric B background floating in the likelihood fit. However, the total yields of the combinatoric B background is large enough in the $B^0 \rightarrow \rho^0 \gamma$ and $B^+ \rightarrow \rho^+ \gamma$ analyses that they should be independently estimable (see Tables 8.4, 8.6 and 8.8).

9.3.6 Summary of likelihood functions

Table 9.9 (p. 161) summarizes the PDF shapes used for each component of each of the modes.

Tables 9.10 to 9.12 detail the parameters that are allowed to float in the likelihood fits and the values of the PDF shape parameters that are fixed in the likelihood fits. Table 9.8



Figure 9.21: The PDFs for the combinatoric B background in $B^0 \to \rho^0 \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated generic $B\overline{B}$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the generic $B\overline{B}$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.



Figure 9.22: The PDFs for the combinatoric *B* background in $B^+ \to \rho^+ \gamma$ decay mode. The lines represent the functional form of the PDFs used in the likelihood fit. The points are a histogram of the fully simulated generic $B\overline{B}$ Monte Carlo. The neural network (NN) distribution is the same as that for the signal component and not derived from the generic $B\overline{B}$ Monte Carlo. The remaining components are derived from the Monte Carlo data sets against which they're plotted.

indicates how the values of each of the parameters are fixed.

The likelihood functions for each of the two $B \to \rho \gamma$ modes are four dimensional $(m_{ES}, \Delta E^*, o_{nn}, \text{ and } F)$ functions with five candidate hypotheses (signal, continuum, $B \to K^* \gamma$ background, η/π^0 background, and combinatoric B background) and a total of six floating parameters: three continuum PDF shape parameters $\xi_{m_{ES}}$, $P_{1,\Delta E^*}$, $P_{w,\Delta E^*}$ and three yields for signal, continuum, and combinatoric B backgrounds.

The likelihood function for the $B^0 \to \omega \gamma$ mode is a four dimensional $(m_{ES}, \Delta E^*, o_{nn},$ and F) function with three candidate hypotheses (signal, continuum, and η/π^0 background) and a total of five floating parameters: three continuum PDF shape parameters $\xi_{m_{ES}}, P_{1,\Delta E^*},$ $P_{w,\Delta E^*}$ and two yields for signal, continuum background.

9.3.7 Toy Monte Carlo studies

As for the $B^0 \to K^{*0}\gamma$ analysis (Chapter 9.2.2), the likelihood fits for the $B \to (\rho/\omega)\gamma$ modes are validated with toy Monte Carlo studies.

First, pure toy Monte Carlo studies are performed to test the robustness of the prescribed fitting procedure and to obtain an initial estimate of the fit's sensitivity to expected signal levels. Studies on the fully simulated *BABAR* Monte Carlo data, combined with the best available estimates and measurements of the branching fractions of various data types, lead to a projected composition of the real 191 fb⁻¹ data set for each of the modes summarized in Table 9.13. These yields and the PDFs described in Chapters 9.3.1 through 9.3.5 are used to generate ensembles of toy data sets for each mode. The parameters designated to float in the likelihood fit are fixed to the values estimated from the development fits to sideband data for generating the data sets. The likelihood fits are applied to each of the respective toy data sets and the resulting statistics compiled into Tables 9.14 and 9.15. Table 9.14 shows the pulls of the floating parameters over the pure toy ensembles. The pulls for the signal yields are depicted in

	Signal	Combinato	ric backgrounds
Domain	$B^0 o ho^0 \gamma$	cont. bkg.	$B\overline{B}$ bkg.
Yield (events)	float	float	float
$m_{ES}({ m ~GeV})$	$< m_{ES} >= 5.27978 \pm 0.00003$ $\sigma_{m_{ES}} = 0.00293 \pm 0.00002$ $\alpha_{m_{ES}} = 1.24 \pm 0.03$ $n_{m_{ES}} = 5.4 \pm 0.6$	ξ (float)	KEYS
$\Delta E^*(\text{ GeV})$	$<\Delta E^*>=0.005\pm0.002$ $\sigma_{\Delta E^*}=0.0368\pm0.0005$ $\alpha_{\Delta E^*}=0.71\pm0.03$	P_1 (float) P_2 (float)	KEYS
NN	$n_{\Delta E^*} = 3.6 \pm 0.4$ $< NN >= 0.9784 \pm 0.0002$ $\sigma_{NN} = 0.0071 \pm 0.0002$ $\alpha_{NN} = 0.229 \pm 0.009$ $n_{NN} = 1.45 \pm 0.04$	Histogram PDF	signal NN
Fisher	KEYS	KEYS	KEYS
	$B^0 \to K^{*0} \gamma$	ing B backgrounds $B \rightarrow$	$ ho^0(\pi^0/\eta)$
Yield (events)	$15.2{\pm}6.1$	$2.5^{+3.3}_{-2.5}$	
$m_{ES}(\text{ GeV})$	$< m_{ES} >= 5.2799 \pm 0.0003$ $\sigma_{m_{ES}} = 0.0029 \pm 0.0002$ $\alpha_{m_{ES}} = 0.87 \pm 0.14$ $n_{m_{ES}} = 10 \pm 0$	$< m_{ES} >=5.28$ $\sigma_{m_{ES}} =0.0035 \pm$ $\alpha_{m_{ES}} =0.58 \pm 0.$ $n_{m_{ES}} =10 \pm 0$	03±0.0002 ±0.0001 03
$\Delta E^* ($ GeV $)$	GeV) $\begin{cases} < \Delta E^* > = -0.081 \pm 0.005 & < \Delta E^* \\ \sigma_{\Delta E^*} = 0.049 \pm 0.004 & \sigma_{\Delta E^*} \\ \alpha_{\Delta E^*} = 0.74 \pm 0.12 & \alpha_{\Delta E^*} \\ n_{\Delta E^*} = 10 \pm 0 & n_{\Delta E^*} \end{cases}$		49±0.006 0.004 03
NN	signal NN	signal NN	
Fisher	KEYS	KEYS	

Table 9.10: All fixed or floating parameters in the final likelihood fit for $B^0 \rightarrow \rho^0 \gamma$ decay mode. The correction of the peak of ΔE^* has already applied on signal MC and peaking *B* backgrounds.

	Signal	Combinato	ric backgrounds
Domain	$B^+ \to \rho^+ \gamma$	cont. bkg.	$B\overline{B}$ bkg.
Yield (events)	float	float	float
$m_{ES}({ m ~GeV})$	$< m_{ES} >= 5.27951 \pm 0.00003$ $\sigma_{m_{ES}} = 0.00309 \pm 0.00003$ $\alpha_{m_{ES}} = 1.09 \pm 0.03$ $n_{m_{ES}} = 6.5 \pm 0.9$	ξ (float)	KEYS
$\Delta E^*($ GeV $)$	$<\Delta E^* >=-0.009\pm 0.002$ $\sigma_{\Delta E^*} = 0.0490\pm 0.0007$ $\alpha_{\Delta E^*} = 0.60\pm 0.03$ $n_{\Delta E^*} = 5.8\pm 1.6$	P_1 (float) P_2 (float)	KEYS
NN	$< NN >= 0.9660 \pm 0.0005$ $\sigma_{NN} = 0.0241 \pm 0.0007$ $\alpha_{NN} = 0.64 \pm 0.03$ $n_{NN} = 1.12 \pm 0.05$	Histogram PDF	signal NN
Fisher	KEYS	KEYS	KEYS
	$B^+ \to K^{*+} \gamma$	ng B backgrounds $B \rightarrow $	$ ho^+(\pi^0/\eta)$
Yield (events)	$5.4{\pm}1.7$	$12.9 {\pm} 4.0$	
$m_{ES}(\text{ GeV})$	$< m_{ES} >= 5.280 \pm 0.0006$ $\sigma_{m_{ES}} = 0.0033 \pm 0.0005$ $\alpha_{m_{ES}} = 0.7 \pm 0.2$ $n_{m_{ES}} = 10 \pm 0$	$< m_{ES} >= 5.27$ $\sigma_{m_{ES}} = 0.0042 \pm$ $\alpha_{m_{ES}} = 0.73 \pm 0.$ $n_{m_{ES}} = 10 \pm 0$	90±0.0002 ±0.0002 05
$\Delta E^*($ GeV)	$<\Delta E^{*} >=-0.087\pm0.02$ $\sigma_{\Delta E^{*}} =0.073\pm0.012$ $\alpha_{\Delta E^{*}} =0.6\pm0.3$ $n_{\Delta E^{*}} =10\pm0$	$<\Delta E^{*} >=-0.08 \\ \sigma_{\Delta E^{*}} ==0.090 \pm 0 \\ \alpha_{\Delta E^{*}} ==0.28 \pm 0. \\ n_{\Delta E^{*}} ==10 \pm 0$	80±0.006).005 05
NN	signal NN	signal NN	
Fisher	KEYS	KEYS	

Table 9.11: All fixed or floating parameters in the final likelihood fit for $B^+ \rightarrow \rho^+ \gamma$ decay mode. The correction of the peak of ΔE^* has already applied on signal MC and peaking *B* backgrounds.

	Signal	Backgrounds		
Domain	$B^0 ightarrow \omega \gamma$	$B ightarrow \omega(\pi^0/\eta)$	cont. bkg.	
Yield (events)	float	$2.6^{+0.8}_{-1.2}$	float	
	$< m_{ES} > = 5.27982 \pm 0.00004$	$< m_{ES} >= 5.2794 \pm 0.0004$		
$m_{EG}(\text{GeV})$	$\sigma_{m_{ES}}=\!\!0.00305{\pm}0.00003$	$\sigma_{m_{ES}} = 0.0043 \pm 0.0003$	۶ (float)	
mes(dev)	$\alpha_{m_{ES}} = 1.20 \pm 0.04$	$\alpha_{m_{ES}} = 0.83 \pm 0.10$	ζ (moat)	
	$n_{m_{ES}} = 4.3 \pm 0.6$	$n_{m_{ES}} = 10 \pm 0$		
	$<\Delta E^{*}>=-0.004\pm0.002$	$<\Delta E^*>$	P_1 (float)	
$\Delta E^*(\text{GeV})$	$\sigma_{\Delta E^*} = 0.0440 \pm 0.0007$	$= -0.180 \pm 0.012$	1 [(110au)	
	$\alpha_{\Delta E^*} = 0.62 \pm 0.03$	$\sigma_{A,E*} = 0.112 \pm 0.009$	$P_{\rm a}$ (float)	
	$n_{\Delta E^*} = 4.9 \pm 1.0$	$0 \Delta E^* = 0.112 \pm 0.003$	1 2 (Hoat)	
NN	KEYS	signal NN	Histogram PDF	
Fisher	KEYS	KEYS	KEYS	

Table 9.12: All fixed or floating parameters in the final likelihood fit for $B^0 \to \omega \gamma$ decay mode. The correction of the peak of ΔE^* has already applied on signal MC and peaking B backgrounds.

Mode	Signal	$B \to K^* \gamma$ bkg.	$(\rho/\omega)(\pi/\eta)$ bkg.	Cont. bkg.	Comb. B bkg.
$B^0 ightarrow ho^0 \gamma$	17	12	3	4000	70
$B^+ \to \rho^+ \gamma$	29	4	14	7000	170
$B^0 \to \omega \gamma$	10	n/a	3	1700	n/a

Table 9.13: The expected fit region yields of each component for each decay mode based on Monte Carlo efficiencies and the assumed branching fractions from Table 3.3. These yields are used in the toy Monte Carlo studies. Refer to Tables 8.4, 8.6, and 8.8 for more detailed yield figures.

mean/σ	$N_{ m sig}$	$N_{ m cont}$	N_{BBkg}	ξ	P_1	P_2
$B^0 \to \rho^0 \gamma$	-0.05/1.02	0.02/1.01	-0.09/0.98	-0.04/0.97	-0.01/0.96	-0.04/0.97
$B^+ \to \rho^+ \gamma$	-0.02/1.00	0.18/1.01	-0.45/1.00	-0.08/1.00	-0.15/1.00	0.02/0.97
$B^0 \to \omega \gamma$	-0.10/0.99	-0.02/0.99	n/a	-0.05/1.05	-0.05/0.97	-0.05/1.00

Table 9.14: The mean and sigma of the pull distributions for all floating parameters in the likelihood fit. The listed σ s are the widths of pull distributions, not the error in the estimated mean. These are expected to be 1.0 based on the definition of the pull.

Mode	Toy Input	Mean Fit Yield	Mean Fit Error	$\operatorname{Error}(\%)$
$B^0 \to \rho^0 \gamma$	17	17.1	9.3	54.4
$B^+ \to \rho^+ \gamma$	29	29.8	13.4	45.0
$B^0 \to \omega \gamma$	10	9.7	6.1	62.9

Table 9.15: The estimated sensitivity from the likelihood fit for each decay mode with 200 fb^{-1} data.

Figure 9.23. The pulls of the signal yields indicate that the fit is an unbiased estimator of the signal yield under the assumption that the PDFs of the fit correspond to the true PDFs. The largest of the signal yield pulls belonging to the $B^0 \rightarrow \omega \gamma$ mode indicates a possible systematic deviation from the input (true) value of the sigma yield of approximately 1/10 the statistical error of a single fit, an insignificant deviation. Table 9.15 indicates that the fit can estimate signal levels based on the assumed branching fractions in Table 3.3 at a 2σ level, still insufficient to claim observation of these individual modes in *BABAR* data.

As for the $B^0 \to K^{*0}\gamma$ analysis, the more important toy Monte Carlo study is the signal-embedded toy study. These use toy data sets with background data generated from the likelihood fit PDFs incorporated with signal data selected from the fully simulated *BABAR* Monte Carlo in the manner previously described (Chapter 9.2.2). The results of this set of toy studies are summarized in Tables 9.16 through 9.18. The pull distributions of the signal yield estimate for each mode is displayed Figure 9.24. Figure 9.25 shows the distributions of the significance of the yield estimates $N/\sigma(N)$ for each of the decay modes. The signal embedded toy Monte Carlo



Figure 9.23: The pull distribution of the signal yield for each decay mode, pure toy.

Par	Gen	Raw E	Ensemble Par	Gaussian Fit	to Pull dist	
	value	$Mean \pm RMS$	Err±RMS	Pull±RMS	Gauss Mean	Gauss σ
N _{sig}	17	17 ± 10	$9.3{\pm}1.1$	$-0.08{\pm}1.08$	-0.03 ± 0.03	$1.00{\pm}0.03$
N _{cont}	4000	4003 ± 76	$72.5{\pm}0.8$	$0.04{\pm}1.05$	$0.06{\pm}0.03$	$1.03{\pm}0.02$
N_{BBkg}	70	68 ± 38	$37.6{\pm}1.8$	$-0.10{\pm}1.02$	$-0.08 {\pm} 0.03$	$0.98{\pm}0.02$
ξ	-12.8	$-12.8{\pm}1.9$	$1.88 {\pm} 0.03$	$0.03{\pm}1.02$	$-0.02{\pm}0.03$	$0.97{\pm}0.02$
P_1	-1.29	$-1.29{\pm}0.11$	$0.11{\pm}0.00$	$0.00{\pm}0.99$	$-0.01{\pm}0.03$	$1.01{\pm}0.03$
P_2	1.43	$1.44{\pm}0.65$	$0.65{\pm}0.03$	$-0.02{\pm}1.01$	$0.01{\pm}0.03$	$0.99{\pm}0.02$

Table 9.16: $B^0 \rightarrow \rho^0 \gamma$: The raw statistical mean and RMS for the values of the fit parameters, error of the fit parameters, and pulls of the fit parameters for the signal-embedded toy Monte Carlo study.

Par	Gen	Raw E	Ensemble Par	Gaussian Fit	to Pull dist	
	value	$Mean \pm RMS$	$\mathrm{Err}\pm\mathrm{RMS}$	Pull±RMS	Gauss Mean	Gauss σ
Nsig	29	32 ± 14	$13.4{\pm}1.3$	$0.11{\pm}1.06$	$0.20{\pm}0.03$	$1.02{\pm}0.03$
$N_{\rm cont}$	7000	7014 ± 92	$90.5{\pm}0.7$	$0.15{\pm}1.02$	$0.16{\pm}0.03$	$1.02{\pm}0.02$
N_{BBkg}	170	152 ± 39	$38.7{\pm}1.8$	$-0.51{\pm}1.02$	$-0.51{\pm}0.03$	$1.02{\pm}0.02$
ξ	-6.58	$-6.6{\pm}1.4$	$1.43 {\pm} 0.01$	$-0.01{\pm}0.99$	$-0.01{\pm}0.03$	$0.99{\pm}0.02$
P_1	-1.036	$-1.05{\pm}0.07$	$0.07 {\pm} 0.00$	$-0.17{\pm}1.01$	$-0.16{\pm}0.03$	$1.03 {\pm} 0.02$
P_2	-0.302	-0.27 ± 0.44	$0.44{\pm}0.01$	$0.04{\pm}1.01$	$0.04{\pm}0.03$	$0.99{\pm}0.02$

Table 9.17: $B^+ \rightarrow \rho^+ \gamma$: The raw statistical mean and RMS for the values of the fit parameters, error of the fit parameters, and pulls of the fit parameters for the signal-embedded toy Monte Carlo study.

study verifies the key conclusions of the pure toy study: that the fit is robust and that it is an unbiased estimator of the signal yield. The complexity of the fit will required more toy Monte Carlo studies described in Chapter 10 to determine the efficiency of the fit procedures and the systematic uncertainties associated with the assumptions that go into the likelihood fits.

9.3.8 Signal yields

At last, the likelihood fits are applied to the data. The individual fit yields from the likelihood fit procedures on data are summarized in Table 9.19. The fits are displayed in projection plots superposed on histograms of the data in Figures 9.26 to 9.28. These projection

Par	Gen	Raw E	Ensemble Par	Gaussian Fit	to Pull dist	
	value	$Mean \pm RMS$	${\rm Er}\pm{\rm RMS}$	Pull±RMS	Gauss Mean	Gauss σ
Nsig	10	$10.3{\pm}6.5$	$6.0{\pm}1.0$	-0.11 ± 1.14	-0.06 ± 0.04	$0.99{\pm}0.03$
$N_{\rm cont}$	1700	$1698 {\pm} 41$	41 ± 3	$-0.06{\pm}0.98$	$-0.04{\pm}0.03$	$0.95{\pm}0.02$
ξ	-5.6	$-5.7{\pm}2.7$	$2.8{\pm}0.2$	$-0.05 {\pm} 0.95$	$-0.10{\pm}0.03$	$0.91{\pm}0.02$
P_1	-1.267	$-1.26 {\pm} 0.14$	$0.14{\pm}0.01$	$0.04{\pm}0.97$	$0.05{\pm}0.03$	$0.96{\pm}0.02$
P_2	-0.331	$-0.31 {\pm} 0.88$	$0.87{\pm}0.08$	$-0.04{\pm}1.01$	$-0.04{\pm}0.03$	$0.98{\pm}0.02$

Table 9.18: $B^0 \rightarrow \omega \gamma$: The raw statistical mean and RMS for the values of the fit parameters, error of the fit parameters, and pulls of the fit parameters for the signal-embedded toy Monte Carlo study.



Figure 9.24: The pull distribution of the signal yield for the signal-embedded toy study for each decay mode.



Figure 9.25: The significance distribution of the signal yield for the signal-embedded toy study for each decay mode.

plots do not show the complete data set in the fit region. Cuts are placed on the undisplayed variables in each plot to enhance the signal peak. As a result, the signal peaks in the projection plots can appear more prominent than would plots of the complete data set. The PDFs are scaled appropriately to accurately reflect the composition of the displayed portion of the data.

As for the $B^0 \to K^{*0}\gamma$ mode, the maximum value of the likelihood function determined by the fit to data is validated by comparison with distributions from a toy Monte Carlo study. Unlike the $B^0 \to K^{*0}\gamma$ mode, the toy Monte Carlo studies used as a goodness of fit test for the $B \to (\rho/\omega)\gamma$ modes are pure toy Monte Carlo studies rather than signal embedded toy Monte Carlo studies. The purpose of this validation study is to determine whether the likelihood value resulting from the fit to data is consistent with the assumption that the distribution of the data is exactly that produced by the fit. The pure Monte Carlo study addresses this in a pure way. In the $B^0 \to K^{*0}\gamma$ study, embedded data sets were used to acknowledge that the PDFs naively represented the signal m_{ES} and ΔE^* distributions as uncorrelated. The assumptions for the embedded study are that the yields and background distributions are exactly those produced by the fit and that the signal distributions are exactly those of the fully simulated Monte Carlo. For the $B \rightarrow (\rho/\omega)\gamma$ analyses, the researchers decided that additional assumptions involved in the embedded toy study were unnecessary complications on a test that should be a simple validation. It should be noted that the likelihood value is dominated by the vast continuum background, which is generated from a PDF in both kinds of toy ensemble. The numeric differences in the likelihood between using purely PDF generated signal data and using embedded fully simulated signal data are insignificant.

For each mode, an ensemble of 950 pure toy data sets is generated from the respective PDFs for each mode. The shape parameters that float in the likelihood fit are fixed to their values produced by the likelihood fit to data for generation of the toy data set. The likelihood fits are then applied to each of their toy data sets, and the maximum likelihood values recorded

	$B^0 \to \rho^0 \gamma \text{ mode}$		$B^+ \to \rho^+ \gamma \text{ mode}$		$B^0 \to \omega \gamma \text{ mode}$	
Yield	fit result	expected	fit result	expected	fit result	expected
$N_{\rm sig}$	$0.3^{+7.2}_{-5.4}$	$16.1 {\pm} 8.9$	$25.6^{+15.3}_{-13.9}$	27.5 ± 12.8	$8.3^{+5.7}_{-4.5}$	$9.1{\pm}5.8$
$N_{\rm cont. \ bkg.}$	$4269{\pm}73$	3852 ± 214	6850 ± 90	$6747 {\pm} 283$	1378 ± 37	1687 ± 142
$N_{B \text{ bkg.}}$	$80{\pm}36$	$73.2 {\pm} 7.9$	$175 {\pm} 40$	163 ± 12	n/a	n/a

Table 9.19: The fitted yields for the Run1-4 data set and the expected yields. The expected yields are estimated from Monte Carlo with the exception on the continuum background, where off-peak data is used. The quoted uncertainties on the expectations are the expected statistical fit uncertainties based on toy Monte Carlo trial fits, not a theoretic uncertainty of the quoted expected values.

in the plots of Figure 9.29. For each of the modes, the likelihood value produced by the fit to data agrees very well with those produced by the toy studies. The conclusion is that, when the likelihood fit is applied to data that truly originate from the PDFs produced by the fit to data, the resulting likelihood value is consistent with that produced by the likelihood fit to data. This reinforces the hypothesis that the likelihood fit produces an accurate representation of the data.

The generally accepted threshold for evidence of a particular decay is 3σ significance of the yield. As Table 9.19 indicates, this analysis yields no significant evidence for any of the decay modes $B^0 \rightarrow \rho^0 \gamma$, $B^+ \rightarrow \rho^+ \gamma$, or $B^0 \rightarrow \omega \gamma$. We must wait for more data for the first evidence of these decays in BABAR data.



Figure 9.26: The projection plots of the unblinded results for $B^0 \rightarrow \rho^0 \gamma$ decay mode, where a), b), c), and d) are the projection onto m_{ES} , ΔE^* , neural net, and Fisher, respectively. The red solid line is the total fit and the dashed blue line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286 \text{ GeV}/c^2$, $-0.10 < \Delta E^* < 0.05 \text{ GeV}$ and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 50%, 59%, 78%, and 49% for m_{ES} , ΔE^* , neural net, and Fisher projections.



Figure 9.27: The projection plots of the unblinded results for $B^+ \rightarrow \rho^+ \gamma$ decay mode, where a), b), c), and d) are the projection onto m_{ES} , ΔE^* , neural net, and Fisher, respectively. The red solid line is the total fit and the dashed blue line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286 \text{ GeV}/c^2$, $-0.10 < \Delta E^* < 0.05 \text{ GeV}$ and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 42%, 56%, 66%, and 41% for m_{ES} , ΔE^* , neural net, and Fisher projections.



Figure 9.28: The projection plots of the unblinded results for $B^0 \to \omega \gamma$ decay mode, where a), b), c), and d) are the projection onto m_{ES} , ΔE^* , neural net, and Fisher, respectively. The red solid line is the total fit and the dashed blue line is the total background. The plots are projections with the following cuts applied if that variable is not being plotted: $5.272 < m_{ES} < 5.286 \text{ GeV}/c^2$, $-0.10 < \Delta E^* < 0.05 \text{ GeV}$ and $\mathcal{N} > 0.9$. The selection efficiency for signal events is 46%, 58%, 70%, and 45% for m_{ES} , ΔE^* , neural net, and Fisher projections.



Figure 9.29: The $-\log L$ value of the unblinded fit compared to the $-\log L$ distribution from toy Monte Carlo for each decay mode. The $-\log L$ values of the unblinded fit are labeled by red arrows.

Chapter 10

Systematic uncertainties and individual measurements

The estimated number of real signal events passing the selection criteria is just one of the components of the branching fractions measurements. Completing the measurement requires an estimate of the total number of $\Upsilon(4S)$ events and an estimate of the efficiency of the selection criteria for signal events. The measurement of the total number of $\Upsilon(4S)$ decays in the data set is discussed in Chapter 10.1. The selection efficiency is estimated mostly from studies with the fully simulated *BABAR* Monte Carlo. The tables in Chapter 8 summarize the efficiency of the selection criteria on Monte Carlo events. Differences between the simulated event content and detector response and the actual events and detector response lead to systematic differences between the efficiencies on Monte Carlo events and actual events. Although a great deal of work has gone into making the *BABAR* Monte Carlo an accurate and complete description of the experiment, measurable differences between the Monte Carlo and real events still remain. Each of these systematic deviations of the Monte Carlo from the true data leads to a source of error in the measurement. These are systematic errors tending to shift the mean of the measurement away from the true value in a definite direction, rather than stochastic errors representative of statistical fluctuations of the measurement about the true value of the measured branching fractions. Many such sources of systematic error in the estimated efficiency are considered in this chapter. Chapter 10.11 summarizes these errors and the necessary adjustments to the Monte Carlo selection efficiency. With all the components finally in place, the branching fraction measurements and upper limits are computed in Chapter 10.12.

10.1 *B* counting

The determination of the number of $\Upsilon(4S) \to B\overline{B}$ in BABAR data is described fully in [51]. Briefly, the strategy is to compare the ratios of the number of hadronic events $N_{\rm MH}$ to the number of $e^-e^+ \to \mu^-\mu^+$ event $N_{\mu\mu}$ for data taken with the collision CM energy at the $\Upsilon(4S)$ resonance and for data taken off resonance. Assuming that any increase in the relative number of hadronic events is due to $\Upsilon(4S)$ production, the number of $\Upsilon(4S)$ events N_{Υ} passing the hadronic selection criteria is given by:

$$N_{\Upsilon} = N_{\rm MH}^{on} - N_{\mu\mu}^{on} \cdot \frac{N_{\rm MH}^{off}}{N_{\mu\mu}^{off}} \cdot \kappa ,$$

where κ is factor close to 1 that accounts for small differences between on resonance data and off resonance data in the hadronic cross section, muonic cross section and selection efficiencies. The total number of $\Upsilon(4S)$ events in the data N_{Υ}^0 is then calculated from N_{Υ} and the efficiency of the hadronic selection criteria ϵ_{Υ} by

$$N_{\Upsilon}^0 = N_{\Upsilon} / \epsilon_{\Upsilon}$$

The efficiency ϵ_{Υ} is determined with Monte Carlo simulated events.

This procedure gives $N_{\Upsilon}^0 = (88.21 \pm 0.97) \times 10^6$ for the data set used in the $B^0 \to K^{*0}\gamma$ measurement, and $N_{\Upsilon}^0 = (210.9 \pm 2.3) \times 10^6$ for the data set used in the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ analyses. The 1.1% uncertainty associated measured with N_{Υ}^0 for each data set reflects the uncertainty with which ϵ_{Υ} is known. The efficiency is determined from fully simulated Monte Carlo, which is known to deviate somewhat from actual data. Each of the selection criteria for events included in the determination of N_{Υ}^0 was examined in an independent study. The total uncertainty in ϵ_{Υ} is an accumulation of the systematic differences between data and Monte Carlo in these studies. This 1.1% uncertainty must be included as a multiplicative systematic uncertainty for each branching fraction measurement.

10.2 Tracking efficiency

The analyses described in this thesis use the BABAR GoodTracksLoose selection criteria for all charged candidates used in reconstruction. BABAR maintains a tracking group dedicated to studying the efficiencies of the standard selection criteria. Their studies indicate that the BABAR Monte Carlo overestimates the detection efficiency for charged particles. The result is a necessary downward correction of the selection efficiency from the value determined in Monte Carlo studies.

The methods recommended by the tracking group for determining the efficiency corrections for the GoodTracksLoose selection are described in [52]. The efficiency of the Good-TracksLoose selection on both data and Monte Carlo is measured by comparing tracks detected in the SVT and DCH. The efficiency of the DCH can be estimated by determining the number of charged candidates reconstructed in the SVT that are also reconstructed by the DCH. Since the GoodTracksLoose selection criteria depend almost entirely on the DCH, the GoodTracksLoose efficiency is essentially the DCH efficiency for tracks passing the GoodTracksLoose criteria.

The analyses described in this thesis all employ the standard *BABAR* software tools based on these methods to determine their efficiency corrections. At a procedural level, these tools integrate the ratio of the real data and Monte Carlo efficiencies as a function of candidate charge, transverse momentum, polar angle, azimuthal angle, and the event's charged track multiplicity over the distribution of charged particles in the signal decay to produce a total correction to the Monte Carlo signal selection efficiency.

This method indicates that the signal selection efficiency in Monte Carlo signal data must be corrected by the following factors for each of the signal modes:

- $B^0 \rightarrow K^{*0} \gamma$: 0.986
- $B^0 \rightarrow \rho^0 \gamma$: 0.9878
- $B^+ \rightarrow \rho^+ \gamma$: 0.9961
- $B^0 \rightarrow \omega \gamma$: 0.9860

This efficiency correction carries a systematic uncertainty with it associated largely with the details of track reconstruction. The tracking group estimates a systematic uncertainty of 0.8% per final state charged track. This means a 0.8% uncertainty for the $B^+ \rightarrow \rho^+ \gamma$ efficiency correction, and a 1.6% uncertainty for the other three modes.

An estimate of the systematic uncertainty associated with tracking efficiency asymmetries between positively and negatively charged particles is obtained by applying the prescription separately to samples of $B^0 \to K^{*0}\gamma$ and $\overline{B}^0 \to \overline{K}^{*0}\gamma$. No significant asymmetry is observed, but the 0.35% discrepancy is adopted as a systematic uncertainty on the measurement of $A_{CP}(B^0 \to K^{*0}\gamma)$.

10.3 PID efficiency

The standard BABAR particle identification selectors are also maintained and studied by a dedicated analysis group [53]. As with the tracking, there are small but measurable differences in the efficiencies of the selectors on real and Monte Carlo simulated data. They also maintain a standard set of tools for determining the necessary corrections and systematic uncertainties.

The key identification problem in each of the analysis modes is the discrimination of kaons and pions. The performance of each of the selectors is determined by analyzing samples of π^{\pm} and K^{\pm} from reconstructed decays of D^* mesons. These decays can be reconstructed from kinematic signatures independently of the variables used by the PID selectors.

In general, the Monte Carlo efficiencies are higher than those for real data. A correction factor is determined in almost the same way as the charged particle tracking efficiency correction. The ratio of data to Monte Carlo efficiency as a function of candidate charge, total momentum, and polar angle are integrated over the momentum spectra of the charged candidates. For the $B^0 \rightarrow K^{*0}\gamma$ measurement, this procedure gives a correction factor to the Monte Carlo signal efficiency of 0.9822. A 1% systematic uncertainty is associated with this correction, mostly due to the differences in the momenta spectra of the candidates used to measure the efficiencies and the those of the candidates used in the signal reconstruction.

Applying this procedure separately to $B^0 \to K^{*0}\gamma$ and $\overline{B}^0 \to \overline{K}^{*0}\gamma$ samples indicates no significant bias on $A_{CP}(B^0 \to K^{*0}\gamma)$. The 1.0% discrepancy in PID efficiency on the two samples is adopted as a systematic uncertainty for the measurement of $A_{CP}(B^0 \to K^{*0}\gamma)$.

The PID selection efficiencies in $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ for charged pion selection show no significant difference between data and Monte Carlo. A conservatively estimated systematic uncertainty of 2% is assigned to each mode associated with the pion selection. The $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ selection use a non-standard PID selector developed specificially for this analysis. It is less thoroughly studied than the standard PID selectors. The 2% systematic uncertainty could be reduced with further analysis of the selector; however, such a reduction would be insignificant relative to the total systematic uncertainty of the measurements.

For the $B \to \rho \gamma$ modes, the rate at which K^{\pm} are misidentified as π^{\pm} must also be

considered. $B \to K^* \gamma$ decays represent an important background for these modes, and the likelihood fit uses a fixed yield to effectively subtract this background from the signal yield. PID studies indicate that the Monte Carlo data underestimate the misidentification rate. The $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ background yield in the $B^0 \to \rho^0\gamma$ analysis must be increased by a factor of 1.32 for the fit to data. This correction has a systematic bias uncertainty of ± 0.32 candidate on the $B^0 \to K^{*0}\gamma$ yield. The $B^+ \to K^{*+}\gamma(K^{*+} \to K^+\pi^0)$ background yield in the $B^+ \to \rho^+\gamma$ analysis must be increased by a factor of 1.31 with a ± 0.31 candidate systematic bias uncertainty. The likelihood fits applied to the data in Chapter 9.3 have incorporated these corrections.

10.4 π^0 and single photon efficiency

The BABAR Neutral Reconstruction Analysis Working Group is a team of BABAR physicists that studies issues and efficiencies of the reconstruction neutral final state candidates and π^0 s. They have produced an analysis and a set of tools for correcting the efficiencies of reconstructed decays that incorporate reconstructed π^0 s from one the standard sets of reconstruction criteria, i.e. the pi0DefaultMass criteria (Chapter 4.4) used in the $B^+ \to \rho^+ \gamma$ and $B^0 \to \omega \gamma$ analyses.

The data and Monte Carlo efficiencies of π^0 s and photons are measured independently of these analyses by analyzing a sample of π^0 s from reconstructed τ^{\pm} decays. This analysis is fully described in [54]. In general, the Monte Carlo overestimates the efficiency with which π^0 s are reconstructed.

A correction factor for the Monte Carlo signal efficiency is determined by the nowfamiliar method of integrating the measured ratio of data to Monte Carlo π^0 efficiency as a function of π^0 momentum over the pi^0 spectrum for the $B^+ \to \rho^+ \gamma$ and $B^0 \to \omega \gamma$ modes. This results in correction factors of 0.971 for the $B^+ \to \rho^+ \gamma$ efficiency and 0.966 for the $B^0 \to \omega \gamma$ efficiency. There is a 3% systematic uncertainty associated with each correction factor.

The efficiencies for the high energy B daughter photons do not differ significantly between data and Monte Carlo. No correction to the efficiency is indicated. The BABAR Neutral Reconstruction Analysis Working Group's photon efficiency study implies an associated systematic uncertainty of 2.5% in the $B^0 \to K^{*0}\gamma$ measurement and a 3.0% systematic uncertainty in each of the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ measurements. In the $B^+ \to \rho^+\gamma$ and $B^0 \to \omega\gamma$ analyses, this error must be added linearly with the 3% π^0 systematic uncertainty because the sources of uncertainty for the two efficiencies are highly correlated.

10.5 Systematic uncertainty for photon quality selection

The *B* daughter photon selection criteria described in Chapter 4.2 have been used exclusive radiative penguin measurements largely without change since the first iteration of the $B \to K^* \gamma$ measurements described in [32] and [55] and published in [13]. The systematic uncertainty associated with these selection criteria was well measured in that analysis and the results have been adopted for the analyses described in this thesis.

The efficiencies of π^0/η veto and of the photon candidate isolation criteria are fundamentally dependent on of the rest of the event. A random calorimeter cluster generated by noise, machine backgrounds, or the decay of the non-signal B can accidentally combine with a true signal photon causing a failure of the π^0/η veto, or fall near enough to the signal photon to cause a failure of the isolation criteria. After the calorimeter acceptance, these are the largest sources of inefficiency for the photon quality selection. To compare the data and Monte Carlo rates of inefficiency for these cuts, the analysts for [13] embedded Monte Carlo simulated signal photon candidates in on-resonance data and Monte Carlo simulated $B^0\overline{B}^0$ decays. They then
determined the rate with which these embedded photon candidates were accidentally vetoed by the selection criteria. The results indicate a 1% discrepancy between data and Monte Carlo efficiencies for the π^0/η veto, and a 2% discrepancy for the isolation criteria. These differences are adopted as systematic uncertainties for each of the decay modes.

10.6 Neural network systematic

Differences between the neural network output distributions for data and signal Monte Carlo are evaluated with an independent sample of $B \to D\pi$ events. Most of the neural network input variables are dependent only on the decay of the non-signal B (the ROE) and on the momentum of the high energy photon. The spectrum of the momentum of the B daughter pion in $B \to D\pi$ decays is very similar to that of the photon in $B \to K^*\gamma$ and $B \to (\rho/\omega)\gamma$ exclusive decays. Since the non-signal B decays independently of the signal B, the substitution of the B daughter pion in $B \to D\pi$ events for the photon in the exclusive signal modes yields distributions for most of the neural network input variables that are almost identical to those of the signal, complete with appropriate correlations among them. In fact, for the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ modes the neural network is designed so that it is dependent only on variables with this character, so that the output distributions of the exact neural networks used in the analyses, when applied to $B \to D\pi$ events under the substitution $\pi_B \to \gamma_B$, are almost identical to the neural network output distributions for $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ signal events.

The neural network for the $B^0 \to K^{*0}\gamma$ measurement uses two inputs that rely on information from the reconstructed signal B: the B decay angle $\cos \Theta_B$ and the K^* helicity angle $\cos \Theta_H$. The distribution of $\cos \Theta_B$ should be the same for any properly fully reconstructed Bmeson. With respect to this variable the $B \to D\pi$ sample is a good representation of $B^0 \to K^{*0}\gamma$ signal. The $B \to D\pi$ distribution of $\cos \Theta_H$ is not the same as that of $B^0 \to K^{*0}\gamma$. However, the signal $\cos \Theta_H$ distribution results from conservation of angular momentum in the decay of the K^* . The physical content of $\cos \Theta_H$ is unambiguous and modeled very well by the Monte Carlo simulated events. It is not a significant source of discrepancy between signal data and Monte Carlo simulated events. Further, $\cos \Theta_H$ is largely uncorrelated with the rest of the neural network input variables. In the study of the neural network efficiency with the $B \to D\pi$ sample, the value for $\cos \Theta_H$ calculated for the reconstructed $B \to D\pi$ candidates is substituted with a random value generated from the $\cos \Theta_H$ distribution of signal $B^0 \to K^{*0}\gamma$ Monte Carlo. This modification makes the $B \to D\pi$ a very good independent sample upon which the neural network output distribution discrepancies between real *BABAR* events and *BABAR* signal Monte Carlo can be evaluated. These differences are adopted as the systematic uncertainties associated with the neural network efficiencies in Tables 10.6 and 10.8.

10.7 Fit efficiency systematic

The efficiencies of the likelihood fits are determined with embedded toy Monte Carlo studies described in Chapter 9.2.2 and Chapter 9.3.7. The use of the embedded toys accounts for differences between the idealized PDFs and the fully simulated Monte Carlo distributions, which more accurately represent the true distributions. The statistical error of the efficiency for these toy Monte Carlo studies are adopted as systematic errors associated with these efficiencies. The result is a 0.9% error for $B^0 \to K^{*0}\gamma$, 10.2% for $B^0 \to \rho^0\gamma$, 8.4% for $B^+ \to \rho^+\gamma$, and 5.4% for $B^0 \to \omega\gamma$.

10.8 Fit systematics

The sources of systematic uncertainty described to this point have all been associated with the signal efficiency. They are all 'multiplicative' uncertainties that scale with the signal yield and that are expressed as a fraction or percent of that yield. The systematic uncertainties described in this section, which are associated with the likelihood fit, do not scale directly with the signal yield. They arise from possible systematic deviations between the idealized distributions (PDFs) in the likelihood fits and the true distributions in the data. As a result, their scaling properties are usually much more complicated, with terms that also scale with the various background levels. The uncertainties in the signal yield associated with these effects are evaluated after the likelihood fit has been applied to data as signal yield biases, or 'additive' uncertainties, rather than multiplicative efficiency uncertainties.

The evaluation of additive fit systematics follow a common pattern: a feature of the PDFs that may differ from the true distributions is identified. This feature of the PDF is varied within limits representing the degree of uncertainty in the feature. The effect on the signal yield is determined by applying the varied PDFs either in additional likelihood fits to the data or in toy Monte Carlo studies. The resulting deviation from the measured fit yield is adopted as an additive systematic uncertainty on the fit yield.

The additive systematics in the $B^0 \to K^{*0}\gamma$ analysis have been treated superficially as multiplicative uncertainties by expressing the signal yield bias uncertainties as fractions of the signal yield. When expressed in this way, the additive fit systematic uncertainties are small and comparable to the efficiency systematics. This allows for a uniform treatment of the fit systematics and the efficiency systematics, and leads numerically to the same result as explicitly treating the additive systematics separately. The separation between signal yield bias (additive) and signal yield efficiency (multiplicative) errors is rigorously maintained in the $B \to \rho\gamma$ and $B^0 \to \omega \gamma$ analyses. This avoids the alarm and possible confusion of, for example, expressing the +1.7 event signal yield uncertainty for the $B^0 \to \rho^0 \gamma$ mode (Table 10.4) as a 567% uncertainty on the $0.3^{+7.2}_{5.4}$ signal yield (Table 9.19).

10.8.1 $B^0 \rightarrow K^{*0} \gamma$

Three significant sources systematic uncertainty associated with the fit PDF are examined, each with a toy Monte Carlo study:

- 1. The accuracy of the uncorrelated parameterized PDF used to represent the $B\overline{B}$ background shape,
- 2. The relative proportion of $B \to X_s \gamma$ in the $B\overline{B}$ background and
- 3. The effect of fixing the ΔE^* tail parameter $\alpha_{\Delta E^*}$ for the signal PDF.

The PDF representing the $B\overline{B}$ background component in the likelihood fit is the best representation in simple, available functions consistent with the fully simulated BABAR Monte Carlo. However, it is just a simple parameterized function of a complicated sum of various individual decay component distributions, and it carries the assumption that m_{ES} and ΔE^* are uncorrelated. The effect of these assumptions on signal yield is investigated using a toy Monte Carlo study that embeds fully simulated BABAR $B\overline{B}$ background events in toy data sets of signal and continuum that are generated from the PDFs. A Pure toy Monte Carlo study (Chapter 9.2.2) indicates that, if the PDFs are accurate and complete representations of the data, then the fit is unbiased. The PDF-generated $B\overline{B}$ candidates are exchanged for fully simulated $B\overline{B}$ candidates and the resulting bias in the signal efficiency of the fit is taken as a systematic uncertainty associated with the simplicity of the $B\overline{B}$ background PDF. There is sufficient BABAR $B\overline{B}$ Monte Carlo to create 3.4 independent *B*-background embedded toy data sets. The toy ensemble consisted of 1000 embedded toy data sets with randomly selected *B*-background candidates. The signal efficiency of this embedded toy study is 0.9836 ± 0.0049 , a -1.64% bias on the signal estimate, which is taken as the associated systematic uncertainty.

Chapter 9.2.1 points out that there is a relatively large uncertainty attached to current measurements of branching fraction $\mathcal{B}(B \to X_s \gamma)$, which is the dominant source of B background candidates. Toy Monte Carlo studies were conducted after the likelihood fit was applied to data to estimate the effect on the signal yield of both overestimating and underestimating the size of the $B \to X_s \gamma$ component relative to the remainder of the B backgrounds. Because the B background yield is independently estimated in the likelihood fit, this uncertainty enters only through the B background PDF shape. Three pure toy Monte Carlo studies were conducted to investigate this source of uncertainty. The first was a control using the B background PDF used in the likelihood fit to data and based on the best available estimate of the relative size of the $B \rightarrow X_s \gamma$ component to generate the toy data sets. The second investigated the effect of underestimating the relative size of the $B \to X_s \gamma$ component. For this study the B background candidates in the toy data set are generated from a PDF fit to exclusively $B \to X_s \gamma$ data, representing a case in which $B \to X_s \gamma$ overwhelmingly dominates other sources of B background. The third study examined the effect on the signal yield of overestimating the relative size of the $B \to X_s \gamma$ component. In this study, the PDF used to generate the B background component is a fit to BABAR Monte Carlo with the relative size of the $B \rightarrow X_s \gamma$ reduced to 0.8 of its size in the control study. In each of the studies, signal and continuum background candidates are generated from PDFs with their parameters fixed to the output values of the likelihood fit to data.

The results of these three studies are summarized in Table 10.1. The signal efficiency of second study is 0.41% less than that for the baseline control study. The third study has a signal efficiency 0.31% greater than that of the control study. The maximum difference of 0.41%

	Toy 1		То	Toy 2		Toy 3	
	Base	eline	Underestin	Underestimate $b \to s\gamma$		Overestimate $b \to s\gamma$	
Variable	Mean	RMS	Mean	RMS	Mean	RMS	
nSig	574.66	29.24	571.34	28.80	576.26	29.51	
$\frac{nSig_{fit}}{nSig_{gen}}$	0.98613	0.02812	0.98208	0.02820	0.98926	0.02908	
nContBkg	2590.38	74.31	2581.18	72.63	2596.47	70.29	
$\frac{nContBkg_{fit}}{nConBkg_{gen}}$	1.00491	0.01966	1.00185	0.02032	1.00536	0.01920	
nBBkg	137.37	54.74	147.67	57.38	135.01	54.48	
$\frac{nBBkg_{fit}}{nBBkg_{gen}}$	0.9689	0.3846	1.0397	0.3974	0.9474	0.3784	
$< m_{ES} >$	5.2798019	0.0001270	5.2797917	0.0001356	5.2797990	0.0001335	
$\sigma_{m_{ES}}$	0.0025395	0.0001089	0.0025244	0.0001105	0.0025374	0.0001141	
$<\Delta E>$	-0.010763	0.003235	-0.010696	0.003245	-0.011202	0.003462	
$\sigma_{\Delta E}$	0.044642	0.002761	0.044753	0.002778	0.045023	0.003015	
ξ	-13.141	2.921	-12.840	3.085	-13.062	2.965	
$P0_{\Delta E}$	-0.5017	0.1412	-0.5025	0.1372	-0.5012	0.1446	

is adopted as a systematic uncertainty associated with the uncertain B background content.

Table 10.1: $K^{*0} \to K^+\pi^-$ Branching Fraction $B\bar{B}$ Systematics: Toy MC fit result summary

In addition to the possible correlations in m_{ES} and ΔE^* considered in Chapter 9.2.2, the PDF used in the likelihood fit may differ in shape from the actual data. The functional forms used to model the signal distributions are well established. The most significant possible difference is the uncertain value of the fixed tail parameter $\alpha_{\Delta E^*}$. The large background in the negative ΔE^* makes $\alpha_{\Delta E^*}$ impractical to estimate independently in the likelihood fit. Toy Monte Carlo studies indicate that attempting to float $\alpha_{\Delta E^*}$ in the likelihood fit frequently leads to intolerable fit failures. The effect of fixing the value of $\alpha_{\Delta E^*}$ must be considered a source of systematic uncertainty.

Two additional toy Monte Carlo studies were performed after the likelihood fit to data to estimate the size of this uncertainty. A fit to the fully simulated *BABAR* Monte Carlo gives $\alpha_{\Delta E^*} = 0.700 \pm 0.136$, which is the fixed value used in the likelihood fit to data. The data sets for the toy Monte Carlo ensembles were generated from signal PDFs with $\alpha_{\Delta E^*}$ fixed at $\pm 1\sigma$, that is, at 0.564 and 0.836 for the two studies respectively. The other parameters in the generating PDFs were fixed to the results from the likelihood fit to data. The control study for the *B* background shape systematic also serves as a baseline for this pair of studies.

The results of these toy Monte Carlo studies are summarized in Table 10.2. The fit efficiency for the $\alpha_{\Delta E^*} = 0.700 - 0.136$ case is 1.3% lower than that of the baseline study. The corresponding efficiency deviation for the $\alpha_{\Delta E^*} = 0.700 + 0.136$ case is 1.3% higher than that of the baseline. The systematic error associated with fixing the parameter $\alpha_{\Delta E^*}$ in the likelihood fit is estimated to be 1.3%.

	Toy 1		То	Toy 2		Toy 3	
	Baseline		$\alpha_{\Delta E^*}$	$\alpha_{\Delta E^*} - \sigma_{\alpha}$		$+ \sigma_{\alpha}$	
Variable	Mean	RMS	Mean	RMS	Mean	RMS	
nSig	574.66	29.24	567.44	28.38	581.65	28.84	
$\frac{nSig_{fit}}{nSig_{gen}}$	0.98613	0.02812	0.97304	0.03068	0.99920	0.02798	
nContBkg	2590.38	74.31	2571.10	71.59	2612.00	69.30	
$\frac{nContBkg_{fit}}{nConBkg_{gen}}$	1.00491	0.01966	0.99626	0.02000	1.01239	0.01944	
nBBkg	137.37	54.74	166.73	57.08	109.9285	54.06	
$\frac{nBBkg_{fit}}{nBBkg_{gen}}$	0.9689	0.3846	1.1820	0.4022	0.7744	0.3732	
$< m_{ES} >$	5.2798019	0.0001270	5.2798035	0.0001287	5.2798095	0.0001326	
$\sigma_{m_{ES}}$	0.0025395	0.0001089	0.0025174	0.0001143	0.0025448	0.0001070	
$<\Delta E>$	-0.010763	0.003235	-0.015636	0.004052	-0.007948	0.002931	
$\sigma_{\Delta E}$	0.044642	0.002761	0.048479	0.003465	0.042167	0.002595	
ξ	-13.141	2.921	-12.446	2.902	-13.716	2.866	
$P0_{\Delta E}$	-0.5017	0.1412	-0.4745	0.1436	-0.5325	0.1425	

Table 10.2: $K^{*0} \to K^+\pi^-$ Branching Fraction $\alpha_{\Delta E^*}$ Systematics: Toy MC fit result summary

Adding these three systematic errors in quadrature gives the 2.2% fitting function error in the systematics summary Table 10.6.

A final set of toy Monte Carlo studies was performed to examine the effect on the measurement of $A_{CP}(B^0 \to K^{*0}\gamma)$ of fixing the *B* background CP asymmetry to that of continuum background. *BABAR* has recently published ([56]) a measurement of the CP asymmetry in $B \to X_s \gamma$ processes, which dominate the *B* background of $B^0 \to K^{*0} \gamma$. Pure toy Monte Carlo ensembles were generated with the CP asymmetry of the *B* background at the extremes of the 1σ interval measured in [56]. The maximum ensemble mean deviation from the fit $A_{CP}(sig) = 0.0$ among these studies was 0.25%. This is adopted as a systematic uncertainty associated with fixing the CP content of the *B* background.

10.8.2 $B \rightarrow \rho \gamma$ and $B^0 \rightarrow \omega \gamma$ modes

The more complicated likelihood functions for the $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ modes combined with the smaller yield relative to backgrounds make the systematic uncertainties associated with these likelihood fits much more complicated. There are many more ways in which the likelihood function can differ from the actual shape of the data distributions, and these differences may lead to effects that are significant relative to the small signal yields. Many sources of possible systematic uncertainty have been studied in detail using toy Monte Carlo ensembles. The general form of these toy Monte Carlo studies are the same: one feature of the likelihood function is selected for evaluation. Ensembles of toy data sets are then generated from PDFs with their parameter values fixed to those resulting from or fixed in the likelihood fit to data save for the feature or parameter under study. Usually two ensembles are generated per parameter, each representing a 1 σ deviation in the parameter or feature from its fixed value, where σ is an error estimate from the component fits in Chapter 9.3.1. The resulting ensemble mean deviations in the signal yields from their estimated values on the real data are adopted as 1σ estimates of the systematic errors associated with the likelihood parameter under study. Table 10.3 summarizes the results of these tests as biases on the final signal yields from the likelihood fit to data. The details of these systematic error estimates follow.

For each mode, the shapes of the signal m_{ES} and ΔE^* distributions are fixed to the component fits to fully simulated Monte Carlo data described in Chapter 9.3.1. The systematic

	$B^0 o ho^0 \gamma$		$B^+ \to \rho^+ \gamma$		$B^0 \rightarrow$	$\omega\gamma$
Description	Sig. bias	error	Sig. bias	error	Sig. bias	error
Fixing of sig. PDFs	0.0	$\substack{+0.22\\-0.02}$	0.0	$\substack{+1.54\\-0.99}$	0.0	$^{+0.15}_{-0.16}$
Binning of histogram PDFs	0.0	$\substack{+0.2\\-0.7}$	0.0	$^{+0.0}_{-0.7}$	0.0	$^{+0.5}_{-0.0}$
Statistics of histogram PDFs	0.0	$\substack{+0.4\\-0.4}$	0.0	$\substack{+0.4\\-0.4}$	0.0	$\substack{+0.4\\-0.4}$
$B \to K^* \gamma$ bkg.	0.0	$^{+1.59}_{-1.29}$	0.0	$^{+0.92}_{-0.81}$	-1.05	$^{+1.05}_{-1.05}$
$\overline{(ho/\omega)(\pi/\eta)}$ bkg.	0.0	$\substack{+0.56\\-0.64}$	0.0	$^{+1.31}_{-1.20}$	0.0	$^{+0.26}_{-0.19}$
B bkg.	n/a	ı	n/a	a	0.0	$\substack{+0.0\\-1.5}$
Best Candidate Selection	n/a	ı	0.0	$\substack{+0.0\\-1.2}$	n/a	ì
Total	0.0	$^{+1.7}_{-1.6}$	0.0	$^{+2.2}_{-2.2}$	-1.05	$^{+1.3}_{-1.9}$

Table 10.3: The summary of all systematic uncertainties associated with the likelihood fit for each decay mode. For each type of systematic uncertainty, a bias on the signal yield and the error(events) on the signal bias are calculated. Lines in **bold** have been recomputed using the unblinded data set. All other numbers are extracted from the toy Monte Carlo studies described in Section 10.

uncertainties associated with fixing these shapes is estimated with a battery of toy Monte Carlo studies as described above. An independent toy study was conducted for each of the 8 parameters in the m_{ES} and ΔE^* distributions. The 8 positive and 8 negative bias estimates from these studies are added in quadrature in Table 10.4 and summarized as the 'Fixing of sig. PDFs' line item of Table 10.3.

The continuum m_{ES} and ΔE^* shapes parameters are estimated independently in the likelihood fit and are not expected to be a significant source of error. Systematic uncertainties associated with the fixed PDFs for the neural network and Fisher distributions must be considered. As a reminder: the neural network and Fisher PDFs for continuum candidates are fixed to shapes obtained from sidebands of the BABAR data. There are no Monte Carlo–data differences to consider, but use of a step function to represent the neural network distribution must be considered as a source of systematic uncertainty. Two effects associated with the use of the histogram/step function were considered: the somewhat arbitrarily chosen number of bins or steps and possible fluctuations in the step levels due to the finite statistics of the sideband

Description	$B^0 \to \rho^0 \gamma$	$B^+ \to \rho^+ \gamma$	$B^0 \to \omega \gamma$			
	Fixing of signal	parameters				
$< m_{ES} >$	$0.00\substack{+0.10\\-0.01}$	$0.00\substack{+0.19 \\ -0.08}$	$0.00\substack{+0.01\\-0.02}$			
$\sigma_{m_{ES}}$	$0.00^{+0.11}_{-0.02}$	$0.00\substack{+0.27\\-0.16}$	$0.00\substack{+0.04\\-0.05}$			
$\alpha_{m_{ES}}$	$0.00\substack{+0.07\\-0.00}$	$0.00\substack{+0.15\\-0.05}$	$0.00\substack{+0.06\\-0.07}$			
$n_{m_{ES}}$	$0.00\substack{+0.08\\-0.00}$	$0.00\substack{+0.13\\-0.04}$	$0.00^{+0.12}_{-0.11}$			
$<\Delta E^{*}>$	$0.00\substack{+0.06\\-0.00}$	$0.00\substack{+0.37\\-0.29}$	$0.00\substack{+0.05\\-0.07}$			
$\sigma_{\Delta E^*}$	$0.00\substack{+0.08\\-0.00}$	$0.00^{+0.20}_{-0.10}$	$0.00\substack{+0.00\\-0.02}$			
$\alpha_{\Delta E^*}$	$0.00\substack{+0.07\\-0.00}$	$0.00\substack{+0.92\\-0.73}$	$0.00\substack{+0.02\\-0.04}$			
$n_{\Delta E^*}$	$0.00\substack{+0.05\\-0.00}$	$0.00^{+1.11}_{-0.56}$	$0.00\substack{+0.01\\-0.02}$			
Total	$0.00^{+0.22}_{-0.02}$	$0.00^{+1.54}_{-0.99}$	$0.00^{+0.15}_{-0.16}$			
	Binning of h	istogram				
25(30)	-0.72	-0.30	+0.15			
35(45)	-0.38	-0.40	+0.48			
45(60)	0.00	0.00	0.00			
55(75)	+0.20	-0.70	+0.52			
65(90)	-0.49	-0.20	+0.44			
Error.	$0.00\substack{+0.20\\-0.70}$	$0.00\substack{+0.00\\-0.70}$	$0.00\substack{+0.50\\-0.00}$			
Fixing of peaking B background yield						
$B \to K^* \gamma$ background	$0.00^{+1.58}_{-1.29}$	$0.00\substack{+0.92\\-0.81}$	n/a			
$(ho/\omega)/(\pi/\eta)$	$0.00\substack{+0.56\\-0.64}$	$0.00^{+1.31}_{-1.20}$	$0.00\substack{+0.26\\-0.19}$			

Table 10.4: The summary of all systematic uncertainties associated with the likelihood fit for each decay mode. When changing the binning of the continuum histogram PDF, the default bins are 45, 60, and 45 for $B^0 \rightarrow \rho^0 \gamma$, $B^+ \rightarrow \rho^+ \gamma$ and $B^0 \rightarrow \omega \gamma$ decay respectively. The variation of bins are the same between $B^0 \rightarrow \rho^0 \gamma$ and $B^0 \rightarrow \omega \gamma$ decay mode, shown in the first column, and the variation for $B^+ \rightarrow \rho^+ \gamma$ is shown in the bracket.

sample. The bin/step width was selected so that the statistical error of each bin/step level was less than 10%. Uncertainty due to this arbitrary binning was examined with several pure toy Monte Carlo studies in which the ensemble generated with the default step function PDF and fit with a likelihood function containing a continuum neural network PDF with a different bin/step width. A reasonably wide range of step sizes were examined, and the maximum mean estimated signal yield deviations from the generated values were adopted as systematic uncertainties. The various step widths tested and the resulting deviations appear in Table 10.4. The adopted systematic uncertainties appear as the 'Binning of histogram PDFs' line item in Table 10.3.

The toy Monte Carlo ensembles used to estimate the effects of statistical step level fluctuations were generated using smooth 5th order polynomial functions to represent the continuum neural network distributions. These 5th order polynomials come from fits to the sideband distribution. The deviations in signal yield were adopted as the systematic uncertainties appearing as the 'Statistics of histogram PDFs' line item in Table 10.3.

As detailed in Chapter 9.3.3 and 9.3.4, the PDF shapes and yields for the background sources $B \to K^*\gamma$ and $B \to (\rho/\omega)(\eta/\pi^0)$ are fixed in the likelihood fit. The uncertainties of the fixed yields are the dominant sources of systematic uncertainty associated with these backgrounds. Monte Carlo studies indicate that uncertainties in background shapes have a negligible impact on the signal yield. The magnitude of these uncertainties are estimated with embedded toy Monte Carlo studies. Candidates from the fully simulated BABAR Monte Carlo samples of these backgrounds are embedded into toy Monte Carlo data sets with all other components generated from the PDFs fit to data. To study the systematic uncertainty associated with the fixed $B \to K^*\gamma$ yields, three toy ensembles are generated with embedded $B \to K^*\gamma$ background events. One ensemble is a control in which the mean expected number of $B \to K^*\gamma$ events are embedded in each data set. In one of the other two, the number of embedded $B \to K^*\gamma$ events is 1σ greater than the mean expected; in the other 1σ fewer $B \to K^*\gamma$ events are embedded. σ is the estimated statistical error (\sqrt{N}) in the $B \to K^*\gamma$ yield based on the Monte Carlo efficiency of the signal selection criteria on the $B \to K^*\gamma$ background. The ensemble mean deviations are adopted as systematic uncertainties. The results appear as the ' $B \to K^*\gamma$ bkg' line item in Table 10.3. Note that for the $B^0 \to \omega\gamma$ analysis this uncertainty appears in addition to a bias due to the omission of a component representing the $B \to K^*\gamma$ background in the likelihood fit.

The systematic uncertainties associated with the fixed yields of the $B \to (\rho/\omega)(\eta/\pi^0)$ backgrounds are estimated with the same process as that used for the $B \to K^* \gamma$ backgrounds. The results appear as the ' $(\rho/\omega)(\pi/\eta)$ bkg' line item in Table 10.3.

This procedure is also repeated for the combinatoric B background yield in the $B^0 \to \omega \gamma$ analysis. The level of the combinatoric B background is expected to be negligible in the $B^0 \to \omega \gamma$ mode, so no bias is assigned, but a systematic uncertainty is assigned based on an embedded toy Monte Carlo study. Because the combinatoric B background yield is an independently estimated parameter in the $B^0 \to \rho^0 \gamma$ and $B^+ \to \rho^+ \gamma$ analyses, no systematic uncertainty is associated with their yields. The B bkg line item in Table 10.3 summarizes these uncertainties.

10.9 Systematic bias of the best candidate selection

When multiple reconstructed B candidates in an event pass all of the selection criteria, the one with the smallest value of $|\Delta E^*|$ is selected for inclusion in the likelihood fit sample; the others are discarded. This method can bias the ΔE^* distribution for continuum events, creating an artificial peak around 0. This hypothetical peak is not modeled in the likelihood fit PDF for the continuum ΔE^* distribution; Its possible existence in data may bias the estimated signal yield.

Mode	signal Monte Carlo	on-peak data
$B^0 \to \rho^0 \gamma$	1.03	1.02
$B^+ \to \rho^+ \gamma$	1.15	1.08
$B^0 \to \omega \gamma$	1.14	1.10

Table 10.5: The average candidate multiplicity for each decay mode. All cuts have been applied to the events and candidates. Only on-peak data collected in Run 2 is shown in the region [5.20, 5.29] GeV/ $c^2 \times [-0.6, 0.6]$ GeV in the $m_{ES} - \Delta E^*$ plane.

The multiplicity of candidates surviving the selection criteria for a subset of the BABAR data (Run2) appears in Table 10.5. The candidate multiplicity for the $B^0 \rightarrow \rho^0 \gamma$ mode is negligibly different from 1. Additional study is afforded the possible bias in the $B^+ \rightarrow \rho^+ \gamma$ and $B^0 \rightarrow \omega \gamma$ analyses.

The possibility of an artificial ΔE^* peaking structure in continuum backgrounds is studied in a subset of real *BABAR* data set used in the likelihood fit: the set of all on-peak events passing all of the selection criteria in the ΔE^* fit range ([-0.3, 0.3] GeV) with $m_{ES} < 5.26$ GeV. Figure 10.1 shows the ΔE^* distribution of the best $B^+ \rightarrow \rho^+ \gamma$ candidate in the each of the events of this subset with multiple $B^+ \rightarrow \rho^+ \gamma$ candidates. The distribution displays a clear peak consistent with selection bias. A small Gaussian distribution is added to the continuum's polynomial ΔE^* distribution to model the effect of this bias. The center and width of the Gaussian are determined with a fit to the best-of-multiple-candidate ΔE^* distributions for the $B^+ \rightarrow \rho^+ \gamma$ and $B^0 \rightarrow \omega \gamma$ modes (shown in Figure 10.1 for the $B^+ \rightarrow \rho^+ \gamma$ mode). In this fit, the polynomial component is fixed to that estimated in the full likelihood fit to determine signal yield; the normalizations of the Gaussian and polynomial components (numbers of candidates belonging to each) are allowed to float.

The proportion of the total continuum background that peaks in this way is estimated by applying the polynomial+Gaussian to the entire $|\Delta E^*| < 0.3$ GeV, $m_{ES} < 5.26$ GeV subset of real data. The polynomial parameters are fixed to the values produced by the likelihood fit



Figure 10.1: The on-peak events in the fit region with $m_{ES} < 5.26 \text{ GeV}/c^2$ with multiple candidates are fitted with a Gaussian plus a second order polynomial. The parameters of the second order polynomial are fixed to the same values as for the unblinded fit. The parameters of the Gaussian and the relative normalization of the Gaussian and polynomial distributions are estimated in the fit. $B^+ \rightarrow \rho^+ \gamma$ decay mode.

to data, and the Gaussian shape parameters are fixed to the values from the fit to the bestof-multiple-candidate subset as determined above. The noramlizations of the two components are allowed to float in the fit yielding the results depicted in Figure 10.2. This sideband study indicates $(1.3 \pm 1.1)\%$ of the continuum candidates for the $B^+ \rightarrow \rho^+ \gamma$, and $(0.6 \pm 2.5)\%$ for the $B^0 \rightarrow \omega \gamma$ mode, appear in this Gaussian peak. For $B^0 \rightarrow \omega \gamma$, the peaking yield is negligible and not considered a significant source of systematic uncertainty.

For the $B^+ \rightarrow \rho^+ \gamma$ mode, the bias uncertainty associated with the best candidate selection is estimated by performing a second likelihood fit to the data. For this second fit, another candidate type hypothesis is added to the likelihood fit to represent this possible peaking continuum background. This hypothesis is represented by the same m_{ES} , neural network, and Fisher distributions as the default non-peaking continuum background, with the Gaussian fit to the sideband representing the ΔE^* distribution. The yield of this new peaking continuum background component is fixed to be 1.3% of the total continuum yield in the original fit to data. The resulting shift in the signal yield estimate is adopted as a systematic uncertainty. This appears as the 'Best Candidate Selection' line item in Table 10.3.



Figure 10.2: The on-peak events in the fit region with $m_{ES} < 5.26 \text{ GeV}/c^2$ are fitted with a Gaussian plus a second order polynomial. The parameters of the second order polynomial are fixed to the same values as for the unblinded fit. The parameters of the Gaussian are fixed to those determined from the fit to the multiple candidate events in the $m_{ES} < 5.26 \text{ GeV}/c^2$ region. The relative normalization of the Gaussian and polynomial distributions are estimated in the fit. Left: $B^+ \to \rho^+ \gamma$ decay mode. Right: $B^0 \to \omega \gamma$ decay mode.

10.10 Nuclear interaction effects in $A_{CP}(B^0 \to K^{*0}\gamma)$

Charged particles undergo interactions with the detector before entering the tracking volume of the DCH. The cross sections for these interactions may be CP asymmetric–the cross sections for particles may differ from the corresponding cross sections for anti-particles. Because the tracking asymmetries are only measured in the DCH, such asymmetries are not accounted for in the analysis of the tracking asymmetry and may induce an artificial observed CP asymmetry in the measurement.

To estimate the magnitude of such a possible effect, this analysis follows the method used by [57]. Using a simplified model for the material between the interaction point and the DCH's inner edge, the momentum dependent interaction cross sections tabulated by [17] are integrated over the momentum distributions of the K^{\pm} and π^{\pm} from signal decays. The resulting 0.20% difference between $K^{+}\pi^{-}$ and $K^{-}\pi^{+}$ is adopted as a systematic uncertainty in the measurement of $A_{CP}(B^{0} \to K^{*0}\gamma)$.

	$K^{*0} \to K^+ \pi^-$		
Monte Carlo Efficiency	0.25	57	
Descriptions	Factor	σ	
B Counting	1.000	1.1%	
Tracking Eff	0.986	1.6%	
PID	0.982	1.0%	
Neutrals Efficiency	1.000	2.5%	
Distance cut	1.000	2.0%	
$\pi^0(\eta)$ veto	1.000	1.0%	
Neural network	1.001	3.0%	
Cross- and down-feed	1.000	2.2%	
Signal eff.	0.978	0.9%	
Total Correction	0.949	5.3%	
Corrected Efficiency	0.244	5.6%	
Sub-Mode BR	0.66	37	

Table 10.6: This table lists the factors needed to correct the Monte Carlo efficiency and their systematic errors for $B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$. The final estimated efficiency is also given.

10.11 Summary of systematics and efficiency

Table 10.6 collects the efficiency corrections and sources of systematic uncertainty for the $B^0 \to K^{*0}\gamma$ branching fraction analysis. Table 10.8 duplicates this summary for the $B \to \rho\gamma$ and $B^0 \to \omega\gamma$ modes. Table 10.7 shows the systematic uncertainties associated with the measurement of $A_{CP}(B^0 \to K^{*0}\gamma)$

10.12 Individual mode measurements

With the conclusion of the analysis of sources of systematic uncertainty, measurements for the branching fractions of the individual decay modes can be assembled.

Systematic errors on A_{CP} (%)	
Description	$K^+\pi^-$
Tracking efficiency	0.35
Charged particle identification	1.00
Nuclear interaction asymmetry	0.20
<i>B</i> -background asymmetry	0.25
Total	1.1

Table 10.7: Fractional systematic uncertainties on the branching fraction $\mathcal{B}(B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-))$ and absolute systematic uncertainties on CP asymmetry $A_{CP}(B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-))$.

	B^0 -	$B^0 \to \rho^0 \gamma$ $B^+ \to \rho^+ \gamma$		B^0	$\rightarrow \omega \gamma$			
Description	factor	error	factor	error	factor	error		
Multiplicative Efficiencies								
Number of B events	1.000	0.011	1.000	0.011	1.000	0.011		
Tracking eff.	0.988	0.016	0.996	0.008	0.986	0.016		
PID	1.000	0.020	1.000	0.020	1.000	0.020		
Photon (π^0/γ) eff.	1.000	0.030	0.971	0.060	0.966	0.060		
π^0, η veto	1.000	0.010	1.000	0.010	1.000	0.010		
Photon isolation cut	1.000	0.020	1.000	0.020	1.000	0.020		
$m_{\pi\pi(\pi)}$ mass cut	1.000	0.000	1.000	0.000	1.000	0.000		
Neural net	1.000	0.046	1.000	0.018	1.000	0.046		
MC statistics/bias	1.000	0.102	1.000	0.081	1.000	0.054		
and B bkg modeling								
Total multiplicative	0.988	0.121	0.967	0.108	0.952	0.100		
	Additive Corrections to Signal Yield							
Description	bias	error	bias	error	bias	error		
fit bias/systematic	0.0	+1.7 -1.6	0.0	+2.2 -2.2	-1.05	+1.3 -1.9		

Table 10.8: The summary of all systematic uncertainties for each decay mode. For each type of systematic uncertainty, a correction factor on the signal efficiency and the relative error on the signal efficiency are calculated. For additive corrections, all numbers are in events.

10.12.1 $B^0 \rightarrow K^{*0}\gamma$ branching fraction

The likelihood fit to data yields $N_{sig} = 583 \pm 30 \ B^0 \to K^{*0}\gamma$, $K^{*0} \to K^+\pi^-$ events passing the selection criteria (Table 9.7). Table 10.6 gives the cumulative efficiency of the selection criteria and likelihood fit as $\epsilon = 0.244 \times (1.000 \pm 0.055)$. The fraction of $B^0 \to K^{*0}\gamma$ decays in which the K^{*0} decays by $K^{*0} \to K^+\pi^-$ is f = 2/3 by isospin arguments. The number of $B\overline{B}$ events in the data set is estimated in Chapter 10.1 to be $N_{B\overline{B}} = 88.206 \times (1.000 \pm 0.011) \times$ 10^6 . Assuming that the fraction of $\Upsilon(4S) \to B^0\overline{B}^0$ decays equals the number of $\Upsilon(4S) \to B^+B^$ decays, the total number of B^0 or \overline{B}^0 decays occurring in the data is $N_{B^0} = N_{B\overline{B}}$. Putting all of these factors together gives the measurement for the branching fraction

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = \frac{N_{sig}}{\epsilon f N_{B\overline{B}}} = (3.92 \pm 0.20(\text{stat}) \pm 0.22(\text{sys})) \times 10^{-5}$$

. According to the common convention in high energy experimental measurements, the measurement uncertainty has been quoted in two parts: the statistical uncertainty (stat) associated with the fit estimate of N_{sig} , and the experimental systematic uncertainty (sys) associated with ϵ and $N_{B\overline{B}}$. The total measurement uncertainty is understood to be the sum in quadrature of these two components.

10.12.2 $B^0 \rightarrow K^{*0}\gamma$ CP asymmetry

The CP asymmetry $A_{CP}(B^0 \to K^{*0}\gamma)$ requires no efficiency scaling or bias correction applied to its measured value from the likelihood fit. Adding the total systematic uncertainty from Table 10.7 gives:

$$A_{CP}(B^0 \to K^{*0}\gamma) = -0.069 \pm 0.046(\text{stat}) \pm 0.011(\text{sys}).$$

10.12.3 $B \rightarrow (\rho/\omega)\gamma$ branching fraction limits

Measurements of the branching fractions for the modes $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$, and $B^0 \to \omega \gamma$ are calculated in the same way. The fit yields for each of the modes N_{sig} is given in Table 9.19. Biases on the estimated N_{sig} , efficiencies ϵ , and sub-mode decay fractions f appear in Table 10.8. Chapter 10.1 has the total number of $B\overline{B}$ events. The measured branching fractions appear in Table 10.9. The biases are subtracted from the fit yields to get an effective signal yields which are used in branching fraction calcualations. Note that for none of these three modes does the central value of the estimated branching fraction exceed 3σ , where σ is the total uncertainty on the branching fraction. In fact, for each case the number of signal candidates estimated by the likelihood fit to data is consistent with 0. In such cases, it is customary, and more useful, to quote a 90% confidence level (CL) upper limit on the branching fraction \mathcal{B}_{90} implied by the measurement. \mathcal{B}_{90} is a value of the branching fraction for which, were the measurement repeated independently several times, in 90% of those experiments the true value of the branching fraction \mathcal{B}_0 would be less than the estimated \mathcal{B}_{90} .

The basic procedure adopted for setting upper limits is that prescribed by the BABAR Statistics Working Group [58]. Given a measured parameter x and a likelihood function $\mathcal{L}(x)$, the 90% CL upper limit is the value x_{90} that satisfies:

$$0.9 = \frac{\int_{-\infty}^{x_{90}} \mathcal{L}(x) dx}{\int_{-\infty}^{+\infty} \mathcal{L}(x) dx}.$$
(10.1)

The maximum likelihood fit procedure for determining the signal yield provides a ready-made likelihood function for N_{sig} . However, this likelihood function does not include the systematic uncertainties of the measurement.

Reference [59] provides a formalism for incorporating systematic uncertainties into an experimental upper limit determination. Let x be a parameter to be measured by an experiment with an uncertain sensitivity S. Given a likelihood function $\mathcal{L}(xS)$ for the measured product xS

and the probability density W(S) for S, the likelihood function for x can be found by integrating $\mathcal{L}(xS)$ over the density W(S):

$$\mathcal{L}(x) = \int \mathcal{L}(xS)W(S)dS \tag{10.2}$$

In the present case, the efficiency ϵ of the selection criteria from, e.g., Table 8.3 is the 'sensitivity' of the measurement. It is assumed that the true efficiency is normally distributed about the mean values $\hat{\epsilon}_{sig}$ from Tables 8.3, 8.5, and 8.7 with σ_{ϵ} given by the systematic error on the efficiency from Table 10.8:

$$W(\epsilon; \hat{\epsilon}, \sigma_{\epsilon}) = \frac{1}{\sqrt{2\pi\sigma_{\epsilon}^2}} e^{-(\epsilon-\hat{\epsilon})^2/2\sigma_{\epsilon}^2}.$$

If N_{eff} is the effective total number of signal decays in the data set, the measured signal yield $N_{sig} = \epsilon N_{eff}$, hence the situation is precisely that described by Equation 10.2.

When evaluated as a function of N_{sig} with the other floated parameters fixed to their values at the maximum, the likelihood function used in the measurement closely conforms to a normal distribution (to within a normalization factor) centered at the maximum likelihood value \hat{N}_{sig} with a standard deviation given by the estimated error on the yield σ_N^2

$$\mathcal{L}(N_{sig}) \propto e^{-(N_{sig} - \hat{N}_{sig})^2 / 2\sigma_N^2} = e^{-(\epsilon N_{eff} - \hat{N}_{sig})^2 / 2\sigma_N^2}.$$

The integral required is:

$$\mathcal{L}(N_{eff}) \propto \int e^{-(\epsilon N_{eff} - \hat{N}_{sig})^2 / 2\sigma_N^2} e^{-(\epsilon - \hat{\epsilon})^2 / 2\sigma_\epsilon^2} d\epsilon = \frac{2\pi\sigma_\epsilon\sigma_N}{\hat{\epsilon}\sqrt{2\pi(\frac{N_{eff}^2\sigma_\epsilon^2}{\hat{\epsilon}^2} + \frac{\sigma_N^2}{\hat{\epsilon}^2})}} e^{-\frac{(N_{eff} - \frac{N_{sig}}{\hat{\epsilon}})^2}{2(\frac{eff}{\hat{\epsilon}^2} + \frac{\sigma_N^2}{\hat{\epsilon}^2})}}.$$
 (10.3)

This is the likelihood that is to be integrated to determine the confidence level in Equation 10.1.

Equation 10.3 has several noteworthy features. First, the form of $\mathcal{L}(N_{eff})$ is similar to a Gaussian distribution with mean at the expected estimate of the total number of signal decays in the complete data set $\hat{N}_{eff} = \hat{N}_{sig}/\hat{\epsilon}$, and an N_{eff} -dependent standard deviation $\sigma_{N_{eff}}^2 = \frac{N_{eff}^2 \sigma_{\epsilon}^2}{\hat{\epsilon}^2} + \frac{\sigma_N^2}{\hat{\epsilon}^2}$. In fact, in the absence of systematic uncertainties ($\sigma_{\epsilon} = 0$), the original



Figure 10.3: Plot showing Equation 10.3 for several values of σ_{ϵ} . $\epsilon = 1.0$ has been assumed. Plot taken from [60].

normal distribution is exactly recovered. Further, the terms $\frac{N_{eff}\sigma_{\epsilon}}{\hat{\epsilon}}$ and $\frac{\sigma_N}{\hat{\epsilon}}$ are easily recognized respectively as the statistical and systematic contributions to the uncertainty $\sigma_{N_{eff}}$.

The magnitude of the systematic uncertainty $\sigma_{\epsilon}/\hat{\epsilon}$ determines by how much $\mathcal{L}(N_{eff})$ deviates from a Gaussian likelihood. As suggested by Figure 10.3, the Gaussian approximation is good for systematic uncertainties up to about 20%; the 90% CL differs negligibly from the Gaussian. In the present case, the systematic uncertainties are about 10%; the Gaussian 90% CL upper limits (N_{eff}^{90}) have been used in lieu of evaluating the integrals:

$$N_{eff}^{90} = \hat{N}_{eff} + 1.28 \times \sigma_{N_{eff}}.$$

The resulting 90% CL upper limits for the three decay modes are listed in the last column of Table 10.9.

				$\mathcal{B}(\pm \mathrm{stat.} \pm \mathrm{sys.})$	90%C.L. U.L.
Description	$N_{ m sig}$	fit sys.	ϵ	$(\times 10^{-6})$	$(\times 10^{-6})$
$B^0 \to K^{*0} \gamma$	583 ± 30	N/A	$0.224{\pm}0.056$	$39.2 \pm 2.0 \pm 2.2$	N/A
$B^0 \to \rho^0 \gamma$	$0.3^{+7.2}_{-5.4}$	$^{+1.7}_{-1.6}$	$0.158{\pm}0.019$	$0.01^{+0.22+0.05}_{-0.16-0.05}$	0.36
$B^+ \to \rho^+ \gamma$	$25.6^{+15.3}_{-13.9}$	$^{+2.2}_{-2.2}$	$0.132{\pm}0.014$	$0.92\substack{+0.55+0.13\\-0.50-0.13}$	1.76
$B^0 ightarrow \omega \gamma$	$8.3^{+5.7}_{-4.5}$	$^{+1.3}_{-1.9}$	$0.086{\pm}0.009$	$0.46^{+0.31+0.08}_{-0.25-0.12}$	0.97
combined $B \to \rho \gamma$	269^{+126}_{-120}	$^{+40.2}_{-44.5}$	n/a	$0.64_{-0.28-0.10}^{+0.32+0.10}$	1.16

Table 10.9: Summary of the fit yields, systematic errors, signal efficiency, branching fraction, and the 90% upper limit (U.L.) for $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$, $B^0 \to \omega \gamma$, and the combined result.

Chapter 11

Combined mode measurements

Although the measurements presented in Chapter 10.12 are independently relevant, when considered in various combinations these measurements acquire even more power to test details of the Standard Model.

The measurement of \mathcal{B} and A_{CP} in the decays $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ described in this thesis was carried out in conjunction with similar analyses for the closely related decays $B^0 \to K^{*0}\gamma(K^{*0} \to K^0_s\pi^0), B^+ \to K^{*+}\gamma(K^{*+} \to K^+\pi^0), \text{ and } B^+ \to K^{*+}\gamma(K^{*+} \to K^0_s\pi^+).$ Of these, $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ and $B^0 \to K^{*0}\gamma(K^{*0} \to K^0_s\pi^0)$ provide largely independent measurements of $\mathcal{B}(B^0 \to K^{*0}\gamma)$ and can be combined for an improved measurement of this quantity. These measurements can be further combined with the $\mathcal{B}(B^+ \to K^{*+}\gamma)$ from the two charged modes for an improved average $\mathcal{B}(B \to K^*\gamma)$, which is theoretically related to the $b \to s\gamma$ transition rate. The measurements of A_{CP} can be combined in a similar way.

The $B \to K^* \gamma$ measurements can also be combined into the isospin violation parameter Δ_{0-} :

$$\Delta_{0-} = \frac{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) - \Gamma(B^- \to K^{*-} \gamma)}{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) + \Gamma(B^- \to K^{*-} \gamma)}$$
(11.1)

The decay widths $\Gamma(B^0 \to \rho^0 \gamma)$, $\Gamma(B^+ \to \rho^+ \gamma)$, and $\Gamma(B^0 \to \omega \gamma)$ are related by the

isospin relation Equation 1.3 (p. 5) repeated below:

$$\Gamma(B^0 \to \rho^0 \gamma) = \Gamma(B^0 \to \omega \gamma) = 0.5 \Gamma(B^+ \to \rho^+ \gamma).$$
(11.2)

Although insufficient data exist to date to test this relation, a refined upper limit on the theoretical parameters (related to $b \rightarrow d\gamma$) underlying each of these transition widths can be obtained by combining the measurements for the three modes assuming the truth of this relation.

Finally, the $B \to K^* \gamma$ and $B \to (\rho/\omega) \gamma$ measurements can be combined to estimate the ratio of CKM elements $|V_{td}/V_{ts}|$ (see Chapter 1.2):

$$\frac{\mathcal{B}(B \to (\rho/\omega)\gamma)}{\mathcal{B}(B \to K^*\gamma)} = \left|\frac{V_{td}}{V_{ts}}\right|^2 \left(\frac{1 - m_\rho^2/M_B^2}{1 - m_{K^*}^2/M_B^2}\right)^3 \zeta^2 |1 + \Delta R| , \qquad (11.3)$$

where m_{ρ} , m_{K^*} , and M_B are the masses of the ρ , K^* , and B mesons respectively, ζ arises from flavor-SU(3) symmetry breaking between ρ/ω and K^* , and the correction ΔR accounts for annihilation diagrams.

11.1 $B \to K^* \gamma$ modes

The analysis of $B^0 \to K^{*0}\gamma(K^{*0} \to K^+\pi^-)$ presented in this thesis was carried out simultaneously and with close collaboration to similar analyses of the $B^0 \to K^{*0}\gamma(K^{*0} \to K_s^0\pi^0)$, $B^+ \to K^{*+}\gamma(K^{*+} \to K^+\pi^0)$, and $B^+ \to K^{*+}\gamma(K^{*+} \to K_s^0\pi^+)$ decay modes. Tables 11.1 and 11.2 summarize the results of these four analyses. Note that direct CP violating asymmetry was not measured for the $B^0 \to K^{*0}\gamma(K^{*0} \to K_s^0\pi^0)$ mode. The $K^{*0} \to K_s^0\pi^0$ final state does not unambiguously tag the originating B as B^0 or \overline{B}^0 , unlike the final states of the other three modes. The CP analysis of this mode must employ time-dependent methods and has been carried out on *BABAR* data in the independent analysis [61]. Table 11.3 details the sources of systematic uncertainty associated with each measurement. Many of these systematic uncertainties for the various modes are based on the same studies and samples, e.g. the tracking

efficiency systematics all come from the same study performed by the *BABAR* Tracking Efficiency Task Force. Such uncertainties are correlated among the four decay modes. Before interesting combinations of these results can be constructed, a formalism for dealing with the correlations among the systematic uncertainties is required.

				Isospin	BF \pm stat \pm sys
Description	Unblind Yield	Efficiency	(σ_{sys})	factor	$(\times 10^{-5})$
$K^{*0} \to K^+ \pi^-$	582.6 ± 29.7	0.244	(5.8%)	0.667	$4.06 \pm 0.21 \pm 0.23$
$K^{*0} \to K^0_s \pi^0$	61.8 ± 15.3	0.153	(12.3%)	0.114	$4.02 \pm 0.99 \pm 0.49$
$K^{*+} \to K^+ \pi^0$	250.9 ± 22.6	0.174	(9.4%)	0.333	$4.89 \pm 0.45 \pm 0.45$
$K^{*+} \to K^0_s \pi^+$	156.9 ± 15.7	0.221	(6.3%)	0.229	$3.52 \pm 0.36 \pm 0.21$

Table 11.1: Unblinded signal yield and the branching fraction for each decay mode on Run1 and Run2 data.

Description	$A_{CP} \pm \text{stat} \pm \text{sys} (\times 10^{-5})$
$K^{*0} \to K^+ \pi^-$	$-0.069 \pm 0.046 \pm 0.011$
$K^{*+} \to K^+ \pi^0$	$0.084 \pm 0.075 \pm 0.007$
$K^{*+} \to K^0_s \pi^+$	$0.061 \pm 0.092 \pm 0.007$

Table 11.2: Unblinded A_{CP} for each decay mode on Run1 and Run2 data.

11.1.1 Statistical formalism

The method employed by this analysis for combining the various $B \to K^* \gamma$ measurements is a specialization of a more general method for combining measurements with correlated uncertainties described in [62]. The formalism is essentially a minimum χ^2 method similar to the standard linear least squares fit. The summary of the method provided in this section will use the estimate of an 'average' branching fraction $\mathcal{B}(B \to K^* \gamma)$ from the four modes as a concrete example to give the method context.

Let \vec{y} be a vector of measurements. In the present case, \vec{y} contains the four branching

	Systematic errors on $\mathcal{B}(\%)$				
Description	$K^+\pi^-$	$K_s^0 \pi^0$	$K^+\pi^0$	$K_s^0 \pi^+$	
Number of B events	1.1	1.1	1.1	1.1	
$R^{+/0}$	2.4	2.4	2.4	2.4	
Tracking efficiency	1.6		0.8	0.8	
Charged particle identification	1.0		1.0	1.0	
Photon efficiency	2.5	7.6	7.6	2.5	
Photon isolation cut	2.0	2.0	2.0	2.0	
π^0, η veto	1.0	1.0	1.0	1.0	
K_s efficiency		3.0		3.0	
Neural network	3.0	3.5	2.7	2.8	
PDF parameterization	2.2	7.3	2.7	1.4	
MC statistics/fit bias	0.9	3.2	2.4	1.6	
Total	5.8	12.3	9.4	6.3	
		Systematic erro	ors on A_{CP} (%)		
Tracking efficiency	0.35		0.25	0.25	
Charged particle identification	1.00		0.55	0.53	
Nuclear interaction asymmetry	0.20		0.35	0.15	
<i>B</i> -background asymmetry	0.25		0.25	0.25	
Total	1.1		0.7	0.7	

Table 11.3: Fractional systematic uncertainties on the branching fractions $\mathcal{B}(B \to K^* \gamma)$ and absolute systematic uncertainties on CP asymmetry $A_{CP}(B \to K^* \gamma)$.

fraction measurements

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \mathcal{B}(B \to K^*\gamma)[K^{*0} \to K^+\pi^-] \\ \mathcal{B}(B \to K^*\gamma)[K^{*0} \to K_s^0\pi^0] \\ \mathcal{B}(B \to K^*\gamma)[K^{*+} \to K^+\pi^0] \\ \mathcal{B}(B \to K^*\gamma)[K^{*+} \to K_s^0\pi^+] \end{pmatrix}$$
(11.4)

Let V be the (symmetric, square) covariance matrix of the measurements

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{12}^2 & \sigma_2^2 & \cdots & \sigma_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n}^2 & \sigma_{1n}^2 & \cdots & \sigma_n^2 \end{pmatrix}$$
(11.5)

For the current example, the diagonal elements σ_i^2 are the total sum squared uncertainties (statistical plus systematic) for each mode. The off-diagonal elements σ_{ij}^2 contain the correlated parts of the systematic uncertainties. Finally, let $\vec{\theta}$ a vector of parameters in a model that is to be fit to the measurements. Let $\vec{\theta}$ be related to the model's 'predicted' values \vec{f} of the parameters measured in the vector \vec{y} by the linear transform A

$$\vec{f} = A\vec{\theta} \tag{11.6}$$

In the example case, $\vec{\theta}$ is a single parameter, the branching fraction $\mathcal{B}(B \to K^* \gamma)$, and A is a column vector of ones; each mode gives a measured value for $\mathcal{B}(B \to K^* \gamma)$.

The goal is to find the values for $\vec{\theta}$ that best account for the measurements \vec{y} . The χ^2 difference between the model and measurements is given by

$$\chi^2 \equiv (\vec{y} - A\vec{\theta})^T V^{-1} (\vec{y} - A\vec{\theta}) \tag{11.7}$$

Using the minimum χ^2 criteria for the best–fit values of $\vec{\theta}$ gives

$$\vec{\theta} = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y}$$
(11.8)

The covariance matrix W of the fit values of the parameters $\vec{\theta}$ is

$$W = (A^T V^{-1} A)^{-1} (11.9)$$

Note that for a diagonal V and a one component $\vec{\theta}$, this method reduces to the standard method for computing a weighted average of independent measurements with uncertainties. Instead of the independent weights $\sigma_i^{-2} / \sum_k \sigma_k^{-2}$, the correlated weight vector \vec{w} is

$$\vec{w} = (A^T V^{-1} A)^{-1} A^T V^{-1}.$$
(11.10)

It is customary to resolve the variance in derived measurements $\vec{\theta}$ into statistical and systematic parts when quoting them in publication. Two methods have been used in previous analyses for separating the quoting uncertainties. The first involves repeating the calculation of W_{stat} with a diagonal V_{stat} incorporating only the (independent) statistical uncertainties on \vec{y} . The square root of the resulting variances are quoted as the statistical component of the uncertainty $\sigma_{stat}(\theta_i) = \sqrt{W_{stat,ii}}$. The systematic component completes the total variance W_{ii} when added in quadrature with $\sigma_{stat}(\theta_i)$, $\sigma_{syst}(\theta_i) = \sqrt{W_{ii} - W_{stat,ii}}$. This seems to be the most often used method. However, it's rigorously problematic. Using V_{stat} instead of the full Vwill produce different values for the estimated $\vec{\theta}$. The resulting covariance matrix W_{stat} solves a different problem and may not represent the statistical contribution to W. For concision, this method will be subsequently referred to as the *statistical-remainder* method of resolving the uncertainties.

A second method involves performing standard Gaussian error propagation on the weighted formula for $\vec{\theta}$ separately for the statistical and systematic parts of the covariance matrix V. Specializing the discussion to a single component θ for simplicity, the weight vector \vec{w} is calculated as in Equation 11.10 above using the complete covariance matrix V, with $\theta = \vec{w} \cdot \vec{y}$ by Equation 11.8. The weights \vec{w} are then assumed constants in a linear equation defining θ , and the uncertainty in θ calculated by a standard Gaussian propagation of errors. The covariance

matrix V is resolved into a diagonal matrix of statistical uncertainties $V_{stat,ii} = \sigma_{stat}(y_i)^2$ and V_{syst} containing the correlated systematic uncertainties:

$$V = V_{stat} + V_{syst}$$

The error in θ is then computed from V and the partial derivatives $d_i \equiv \partial \theta / \partial y_i = w_i$

$$\begin{aligned} \sigma_{\theta}^2 &= \vec{d}V\vec{d}^T \\ &= \vec{d}V_{stat}\vec{d}^T + \vec{d}V_{syst}\vec{d}^T \\ &= \sigma_{\theta,stat}^2 + \sigma_{\theta,syst}^2 \end{aligned}$$

For concision, this method will be subsequently referred to as the *propagation-of-errors* method of resolving the uncertainties.

Both methods have been applied to the current analysis to check consistency. The errors quoted in the published results ([1]) come from the second method.

11.1.2 Correlation of systematic branching fraction uncertainties

It remains to identify which sources of systematic uncertainty are correlated among the $B \to K^* \gamma$ modes and to construct the covariance matrices for the combined branching fraction calculations. Refer to Table 11.3 for a list of sources of systematic uncertainties, and Chapter 10.8 for a description of these uncertainties.

The four modes each incur systematic uncertainties related to the high energy photon. For three of these sources of systematic uncertainty, each mode uses an uncertainty based on the same study and data set. Thus the uncertainties associated with photon detection efficiency, the photon isolation cuts, and the π^0/η veto must be considered correlated between each pair of modes. The systematic uncertainties associated with B counting come from a standard BABAR analysis, and must be considered correlated between each pair of mode measurements.

The three modes in which the K^* decays to a K^{\pm} or π^{\pm} each derives its uncertainty associated with charged candidate tracking efficiency from the same Tracking Efficiency Task Force study, and hence must be considered correlated between any pair of them.

The uncertainty associated with K_S^0 efficiency in the $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K_s^0\pi^0$) and $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K_s^0\pi^+$) are derived from the same Neutrals Working Group study and must be considered correlated.

And finally, the uncertainty associated with the charged kaon candidate identification in the $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) and $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) come from the same PID group study, and are correlated. The sources of correlated uncertainties among the four modes is summarized in Table 11.4.

The contribution to an off-diagonal element of the covariance matrix is the product of the uncertainties from the given source for the two modes involved. This corresponds to a degree of correlation equal to the geometric mean of the two uncertainties. For example, the tracking efficiency gives a 1.6% uncertainty to $B^0 \to K^{*0}\gamma \ (K^{*0} \to K^+\pi^-)$ and a 0.8% uncertainty to $B^+ \to K^{*+}\gamma \ (K^{*+} \to K^+\pi^0)$. The contribution to the off-diagonal element is the product

$$\sigma_{\text{trk},ij}^2 = (0.016\mathcal{B}(B^0 \to K^{*0}\gamma(K^+\pi^-))) \cdot (0.008\mathcal{B}(B^+ \to K^{*+}\gamma(K^+\pi^0)))$$

= $(0.016 \cdot 3.92 \times 10^{-5}) \cdot (0.008 \cdot 4.90 \times 10^{-5})$
= 2.46×10^{-13}

Tables 11.5 and 11.6 show the total covariance matrix V and the systematic covariance matrix V_{syst} thus calculated for the four $B \to K^* \gamma$ analyses.

Mode	$K^{*0} \to K^+ \pi^-$	$K^{*0} \to K^0_s \pi^0$	$K^{*+} \to K^+ \pi^0$	$K^{*+} \to K^0_s \pi^+$
$K^{*0} \to K^+ \pi^-$	-	Common 4	Common 4	Common 4
			Tracking	Tracking
			Particle ID	
$K^{*0} \to K^0_s \pi^0$	-	-	Common 4	Common 4
				K_s^0
$K^{*+} \to K^+ \pi^0$	-	-	-	Common 4
				Tracking

Table 11.4: Correlated systematics of the four $B \to K^* \gamma$ modes. "Common 4" refers to the five systematics that are common to all modes (B Counting, Neutrals Efficiency, Distance Cut, and π^0/η veto. Diagonal entries are simply the total systematic for each mode. The matrix is symmetric, so all entries below the diagonal are omitted

	「	$K^{*0} \to K^+ \pi^-$	$K^{*0} \to K^0_s \pi^0$	$K^{*+} \to K^+ \pi^0$	$K^{*+} \to K^0_s \pi^+$
	$K^{*0} \to K^+ \pi^-$	0.9435×10^{-11}	0.4115×10^{-11}	0.5458×10^{-11}	$0.1963 imes 10^{-11}$
V =	$K^{*0} \to K^0_s \pi^0$	0.4115×10^{-11}	12.29×10^{-11}	1.256×10^{-11}	0.4833×10^{-11}
	$K^{*+} \to K^+ \pi^0$	0.5458×10^{-11}	1.256×10^{-11}	4.074×10^{-11}	0.4443×10^{-11}
	$K^{*+} \to K^0_s \pi^+$	0.1963×10^{-11}	0.4833×10^{-11}	0.4443×10^{-11}	1.723×10^{-11}

Table 11.5: Covariance matrix of the $B \to K^* \gamma$ measurements

	-	$K^{*0} \to K^+ \pi^-$	$K^{*0} \to K^0_s \pi^0$	$K^{*+} \to K^+ \pi^0$	$K^{*+} \to K^0_s \pi^+$
	$K^{*0} \to K^+ \pi^-$	0.5064×10^{-11}	0.4115×10^{-11}	0.5458×10^{-11}	0.1963×10^{-11}
$V_{syst} =$	$K^{*0} \to K^0_s \pi^0$	0.4115×10^{-11}	2.4×10^{-11}	1.256×10^{-11}	0.4833×10^{-11}
	$K^{*+} \to K^+ \pi^0$	0.5458×10^{-11}	1.256×10^{-11}	2.052×10^{-11}	0.4443×10^{-11}
	$K^{*+} \to K^0_s \pi^+$	0.1963×10^{-11}	0.4833×10^{-11}	0.4443×10^{-11}	0.4386×10^{-11}

Table 11.6: Systematic covariance matrix of the $B \to K^* \gamma$ measurements

11.1.3 Combined branching fractions

All is now prepared for combining the $B \to K^* \gamma$ branching fractions. This section presents the computation of three 'average' branching fractions from a combination of the four individual branching fraction measurements:

•
$$\mathcal{B}(B^0 \to K^{*0}\gamma)$$
 from $B^0 \to K^{*0}\gamma \ (K^{*0} \to K^+\pi^-)$ and $B^0 \to K^{*0}\gamma \ (K^{*0} \to K^0_s\pi^0)$,

•
$$\mathcal{B}(B^+ \to K^{*+}\gamma)$$
 from $B^+ \to K^{*+}\gamma \ (K^{*+} \to K^+\pi^0)$ and $B^+ \to K^{*+}\gamma \ (K^{*+} \to K^0_s\pi^+)$,

• $\mathcal{B}(B \to K^* \gamma)$ from all four modes.

The computation of $\mathcal{B}(B^0 \to K^{*0}\gamma)$ from $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) and $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K_s^0\pi^0$) involves only the 2 × 2 neutral sector in the upper left-hand corner of the covariance matrices in Tables 11.5 and 11.6. Applying the statistical formalism to this sector gives $\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.061 \pm 0.303) \times 10^{-5}$, where the uncertainty is the total uncertainty. The relative weights for the two neutral modes assigned by the procedure (Equation 11.10):

$$w_0 \equiv w(K^{*0} \to K^+ \pi^-) = 0.9571$$
(11.11)
$$w_1 \equiv w(K^{*0} \to K_s^0 \pi^0) = 0.0429$$
(11.12)

The more precise measurement of $\mathcal{B}(B^0 \to K^{*0}\gamma \ (K^{*0} \to K^+\pi^-))$ dominates. Using the statistical-remainder method for resolving the total uncertainty into statistical and systematic parts gives

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.061 \pm 0.205(\text{stat}) \pm 0.224(\text{syst})) \times 10^{-5}$$
.

Applying the propagation-of-errors method gives

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.061 \pm 0.205(\text{stat}) \pm 0.224(\text{syst})) \times 10^{-5}$$

In this case, the dominance of one of the modes leads to almost identical results for the two methods.

Repeating this set of procedures for the $B^+ \to K^{*+}\gamma$ sector in the lower right hand corner of the covariance matrices gives $\mathcal{B}(B^+ \to K^{*+}\gamma) = (3.874 \pm 0.373) \times 10^{-5}$ and relative weights

$$w_2 \equiv w(K^{*+} \to K^+ \pi^0) = 0.2606$$
(11.13)
$$w_3 \equiv w(K^{*+} \to K_s^0 \pi^+) = 0.7394.$$

(11.14)

The statistical-remainder method resolves the uncertainty as

$$0.373(tot) = 0.280(stat) \pm 0.246(syst).$$

The propagation-of-errors breakdown gives

$$\mathcal{B}(B^+ \to K^{*+}\gamma) = (3.874 \pm 0.290(\text{stat}) \pm 0.235(\text{syst})) \times 10^{-5}$$

In this case, using only the statistical uncertainties in the statistical-remainder would method lead to a significantly different weighting of the two modes—(0.3886/0.6114) instead of (0.2606/0.7394)—and a significantly different estimate for $\mathcal{B}(B^+ \to K^{*+}\gamma)$.

Finally, the total average $\mathcal{B}(B \to K^* \gamma)$ using all four modes is

$$\mathcal{B}(B \to K^* \gamma) = (3.930 \pm 0.174 (\text{stat}) \pm 0.196 (\text{syst})) \times 10^{-5}$$
,

where the quoted uncertainties come from the propagation-of-errors method. The statisticalremainder method gives $0.165(\text{stat}) \pm 0.203(\text{syst})$, which agrees reasonably well.

11.1.4 Isospin asymmetry

The isospin violation parameter most often used in literature is

$$\Delta_{0-} \equiv \frac{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) - \Gamma(B^- \to K^{*-} \gamma)}{\Gamma(\overline{B}^0 \to \overline{K}^{*0} \gamma) + \Gamma(B^- \to K^{*-} \gamma)} .$$
(11.15)

It will be calculated in two steps to simplify the treatment of systematic uncertainties. The first step isolates the terms that rely only the measured branching fractions in an isospin ratio I.

$$I \equiv \frac{\mathcal{B}(B^0 \to K^{*0}\gamma)}{\mathcal{B}(B^+ \to K^{*+}\gamma)}$$
(11.16)

The average values for $\mathcal{B}(B^0 \to K^{*0}\gamma)$ and $\mathcal{B}(B^+ \to K^{*+}\gamma)$ from Chapter 11.1.3 are used in the calculation of I. In terms of the individual mode branching fraction measurements and their relative weights w_i in the average branching fraction measurements (Equations 11.12 and 11.14):

$$I = \frac{w_0 \mathcal{B}(B^0 \to \gamma K^{*0}(K^+\pi^-)) + w_1 \mathcal{B}(B^0 \to \gamma K^{*0}(K_S^0\pi^0))}{w_2 \mathcal{B}(B^+ \to \gamma K^{*+}(K^+\pi^0)) + w_3 \mathcal{B}(B^0 \to \gamma K^{*+}(K_S^0\pi^+))}$$
(11.17)

This relation is taken as defining I, and its uncertainty is calculated from the covariance matrices in Tables 11.5 and 11.6 by the propagation-of-errors method described in Chapter 11.1.1:

$$I = 1.048 \pm 0.093 \text{stat} \pm 0.055 \text{syst}$$
(11.18)

The second step relates I to the isospin violation parameter Δ_{0-} :

$$\Delta_{0-} = \frac{IR^{+/0}\frac{\tau^+}{\tau^0} - 1}{IR^{+/0}\frac{\tau^+}{\tau^0} + 1} , \qquad (11.19)$$

where τ^+ and τ^0 are the B^+ and B0 lifetimes respectively, and $R^{+/0}$ is the B production ratio defined as

$$R^{+/0} \equiv \frac{\Gamma(\Upsilon(4S) \to B^+B^-)}{\Gamma(\Upsilon(4S) \to B^0\overline{B}^0)}$$

The lifetime ratio from [17] is $\tau^+/\tau^0 = 1.083 \pm 0.017$. The uncertainty on this value is classified as a systematic uncertainty and combined with the systematic uncertainty from *I* in the quoted result for Δ_{0-} . The charge production ratio measured at *BABAR* ([63]) $R^{+/0} = 1.006 \pm 0.048$ is not significantly different from 1. The uncertainty contributed to Δ_{0-} from $R^{+/0}$ is quoted as a separate systematic uncertainty. The end result:

$$\Delta_{0-} = 0.063 \pm 0.044 (\text{stat}) \pm 0.028 (\text{syst}) \pm 0.024 (R^{+/0})$$

This corresponds to a 90% confidence interval of

$$-0.031 < \Delta_{0-} < 0.157$$

11.1.5 CP asymmetry

An average A_{CP} is computed from the measured A_{CP} for the $B^0 \to K^{*0}\gamma$ $(K^{*0} \to K^+\pi^-), B^+ \to K^{*+}\gamma \ (K^{*+} \to K^+\pi^0)$, and $B^+ \to K^{*+}\gamma \ (K^{*+} \to K_s^0\pi^+)$ analyses in the same way as the combined branching fractions.

The sources of systematic uncertainties in the measurement of A_{CP} are listed in Table 11.3. The systematic uncertainties associated with the tracking efficiency asymmetry are estimated from the same study for each of the 3 modes, and must be considered correlated for any pair. The uncertainties associated with nuclear interaction asymmetries before the tracking volume must be considered correlated for modes that share a final state particle. Thus the $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) is correlated to the $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) because they share a K^{\pm} , and with $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K_s^0\pi^+$) because of the common π^{\pm} . The $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) and $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K_s^0\pi^+$) nuclear interaction uncertainties are uncorrelated. The uncertainty associated with charged kaon identification asymmetries in the $B^0 \to K^{*0}\gamma$ ($K^{*0} \to K^+\pi^-$) and $B^+ \to K^{*+}\gamma$ ($K^{*+} \to K^+\pi^0$) modes come from the same study and must be considered correlated. The uncertainties associated with B-background asymmetries are determined by independent toy Monte Carlo studies for each mode, and are not correlated. These correlations are summarized in Table 11.7. The resulting total and systematic covariance matrices appear in Tables 11.7 and 11.9.

Mode	$K^{*0} \to K^+ \pi^-$	$K^{*+} \to K^+ \pi^0$	$K^{*+} \to K^0_s \pi^+$
$K^{*0} \to K^+ \pi^-$		Tracking	Tracking
	-	Nuclear interactions	Nuclear interactions
		Particle ID	
$K^{*+} \to K^+ \pi^0$	-	-	Tracking

Table 11.7: Correlated systematics of the three A_{CP} modes. Diagonal entries are simply the total systematic for each mode. The matrix is symmetric, so all entries below the diagonal are omitted

$$V = \begin{bmatrix} K^{*0} \to K^{+}\pi^{-} & K^{*+} \to K^{+}\pi^{0} & K^{*+} \to K_{s}^{0}\pi^{+} \\ K^{*0} \to K^{+}\pi^{-} & 2.238 \times 10^{-3} & 7.075 \times 10^{-5} & 1.175 \times 10^{-5} \\ K^{*+} \to K^{+}\pi^{0} & 7.075 \times 10^{-5} & 5.68 \times 10^{-3} & 0.625 \times 10^{-5} \\ K^{*+} \to K_{s}^{0}\pi^{+} & 1.175 \times 10^{-5} & 0.625 \times 10^{-5} & 8.507 \times 10^{-3} \end{bmatrix}$$

Table 11.8: Covariance matrix of the ${\cal A}_{CP}$ measurements

$$V = \begin{bmatrix} K^{*0} \to K^{+}\pi^{-} & K^{*+} \to K^{+}\pi^{0} & K^{*+} \to K_{s}^{0}\pi^{+} \\ \hline K^{*0} \to K^{+}\pi^{-} & 12.25 \times 10^{-5} & 7.075 \times 10^{-5} & 1.175 \times 10^{-5} \\ \hline K^{*+} \to K^{+}\pi^{0} & 7.075 \times 10^{-5} & 5.5 \times 10^{-5} & 0.625 \times 10^{-5} \\ \hline K^{*+} \to K_{s}^{0}\pi^{+} & 1.175 \times 10^{-5} & 0.625 \times 10^{-5} & 4.284 \times 10^{-5} \end{bmatrix}$$

Table 11.9: Systematic covariance matrix of the ${\cal A}_{CP}$ measurements
The relative weights for the three measurements produced by the least- squares proce-

dure:

$$w(K^{*0} \to K^+ \pi^-) = 0.6054$$

 $w(K^{*+} \to K^+ \pi^0) = 0.2341$
 $w(K^{*+} \to K_s^0 \pi^+) = 0.1605$

and the combined average A_{CP} :

$$A_{CP}(B \to K^* \gamma) = -0.012 \pm 0.036 (\text{stat}) \pm 0.008 (\text{syst})$$

The quoted uncertainties are the result of the propagation-of-errors method for resolving the uncertainty into statistical and systematic components. The statistical-remainder method gives an almost identical segmentation. This corresponds to a 90% confidence interval of

$$-0.073 < A_{CP}(B \to K^* \gamma) < 0.048$$

11.2 $B \rightarrow \rho \gamma, B^0 \rightarrow \omega \gamma$ modes

The $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ branching fraction measurements are combined at the level of the likelihood fit, rather than at the *post hoc* level of the $B \to K^* \gamma$ analyses. A single likelihood function is fit to candidates reconstructed in all three modes $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$, and $B^0 \to \omega \gamma$. The mechanics of the simultaneous fit to the three populations of candidates is the same as that applied to the two CP-conjugate populations in the $B^0 \to K^{*0} \gamma$ as described in Chapter 9.2.1. This likelihood fit produces an estimate of the sum of the number of signal decays in the three modes—a total signal yield, which is expected to have a smaller relative statistical error than any of the individual mode measurements. The sizes of the systematic uncertainties of this combined yield are estimated with toy Monte Carlo studies after the combined fit.

11.2.1 Combined likelihood fit

The likelihood function fit used in the combined fit is the product of the individual mode likelihood functions defined in Chapter 9.3

$$\mathcal{L}_{comb} = \mathcal{L}(B^0 \to \rho^0 \gamma) \cdot \mathcal{L}(B^+ \to \rho^+ \gamma) \cdot \mathcal{L}(B^0 \to \omega \gamma)$$

with an additional constraint on the signal yields. This constraint derives from the theoretical relationship among the decay widths of the three modes.

$$\Gamma(B^0 \to \rho^0 \gamma) = \Gamma(B^0 \to \omega \gamma) = \frac{1}{2} \Gamma(B^+ \to \rho^+ \gamma).$$

In terms of the branching fractions $\mathcal{B}(\text{mode})$ and the *B* lifetime ratio τ_{B^+}/τ_{B^0} this relationship is

$$\mathcal{B}(B^0 \to \rho^0 \gamma) = \mathcal{B}(B^0 \to \omega \gamma) = \frac{1}{2} \frac{\tau_{B^0}}{\tau_{B^+}} \mathcal{B}(B^+ \to \rho^+ \gamma)$$

The mean number of signal events \bar{N}_{sig} (mode) expected to survive the selection criteria with efficiencies ϵ (mode) follow:

$$\frac{\bar{N}_{sig}(B^0 \to \rho^0 \gamma)}{\epsilon(B^0 \to \rho^0 \gamma)} = \frac{\bar{N}_{sig}(B^0 \to \omega \gamma)}{\epsilon(B^0 \to \rho^0 \gamma)} = \frac{1}{2} \frac{\tau_{B^0}}{\tau_{B^+}} \frac{\bar{N}_{sig}(B^0 \to \omega \gamma)}{\epsilon(B^0 \to \rho^0 \gamma)}$$

To encode this relationship as a constraint in the likelihood fit, the formerly independent mode signal yields N_{sig} are defined in terms of a single total 'effective signal yield' N_{eff} , which is to be estimated (floated) in the likelihood fit

$$N_{\rm eff} = \frac{N_{sig}(B^+ \to \rho^+ \gamma)}{\epsilon(B^+ \to \rho^+ \gamma)} + \frac{\tau_{B^+}}{\tau_{B^0}} [\frac{N_{sig}(B^0 \to \rho^0 \gamma)}{\epsilon(B^0 \to \rho^0 \gamma)} + \frac{N_{sig}(B^0 \to \omega \gamma)}{\epsilon(B^0 \to \omega \gamma)}]$$

The N_{sig} are related to N_{eff} by

$$N_{sig}(B^+ \to \rho^+ \gamma) \equiv \frac{1}{2} N_{\text{eff}} \epsilon (B^+ \to \rho^+ \gamma)$$
$$N_{sig}(B^0 \to \rho^0 \gamma) \equiv \frac{1}{4} \frac{\tau_{B^0}}{\tau_{B^+}} N_{\text{eff}} \epsilon (B^0 \to \rho^0 \gamma)$$
$$N_{sig}(B^0 \to \omega \gamma) \equiv \frac{1}{4} \frac{\tau_{B^0}}{\tau_{B^+}} N_{\text{eff}} \epsilon (B^0 \to \omega \gamma)$$

	B^0	$B^0 \to \rho^0 \gamma$ $B^+ \to \rho^+ \gamma$		$\rightarrow \rho^+ \gamma$	$B^0 \to \omega \gamma$
Description	$N_{ m cont}$	$N_{B \rm \ bkg.}$	$N_{ m cont}$	N_B bkg.	$N_{ m cont}$
Individual fit	4269 ± 73	$80{\pm}36$	$6850 {\pm} 90$	$175 {\pm} 40$	1378 ± 37
Combined fit	4270 ± 73	73 ± 35	6850 ± 95	181 ± 39	1380 ± 37

Table 11.10: Summary of the continuum background and combinatoric B background yield from individual fits and the combined fit.

The end result of the procedure is quoted as an 'average' combined branching fraction

$$\mathcal{B}(B \to (\rho, \omega)\gamma)$$

$$\mathcal{B}(B \to (\rho, \omega)\gamma) \equiv \frac{1}{2} \cdot \left[\mathcal{B}(B^+ \to \rho^+ \gamma) + \frac{\tau_{B^+}}{\tau_{B^0}} \cdot \left(\mathcal{B}(B^0 \to \rho^0 \gamma) + \mathcal{B}(B^0 \to \omega\gamma)\right)\right]$$
(11.20)

that is related to the fit effective signal yield $N_{\rm eff}$ and the number of $B\overline{B}$ events in the data set $N_{B\overline{B}}$ by $N_{\rm eff} = 2N_{B\overline{B}}\mathcal{B}(B \to (\rho, \omega)\gamma).$

11.2.2 Fit results

The result of applying the combined likelihood fit described in Chapter 11.2.1 to the data is an estimated total effective signal yield of $N_{\text{eff}} = 269^{+126}_{-120}$. Figure 11.1 depicts the fit functional shapes on the data. The independently estimated background yields are listed in Table 11.10. As one would expect, the background estimates are in excellent agreement with their values from the individual likelihood fits.

11.2.3 Systematic uncertainties

Each source of systematic uncertainty afflicting each of the individual mode measurements will also affect the estimation of N_{eff} (Tables 10.8). The sizes of the resulting uncertainties in N_{eff} are estimated with toy Monte Carlo studies in the same way that the systematic uncertainties associated with the individual fits were determined (Chapter 10.8). First, a 'control' toy ensemble is generated from best-fit PDFs from the combined likelihood fit. Then the effect



Figure 11.1: Projections of the combined fit to $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ in the four discriminating variables: (a) m_{ES} , (b) ΔE^* , (c) \mathcal{N} , and (d) \mathcal{F} . The points are data, the solid line is the total PDF and the dashed line is the background only PDF. The selections applied, unless the variable is projected, are: $5.272 < m_{ES} < 5.286$ GeV/ c^2 , $-0.10 < \Delta E^* < 0.05$ GeV and $\mathcal{N} > 0.9$; the selection efficiencies for signal events are 45%, 57%, 70%, and 44% for the m_{ES} , ΔE^* , \mathcal{N} and \mathcal{F} projections, respectively.

of a systematic uncertainty is determined by generating a toy ensemble from a PDF with its relevant parameter varied by $\pm 1\sigma$. The resulting deviations in the mean value of N_{eff} from its mean value in the control study are adopted as the systematic uncertainty on N_{eff} associated with the given input systematic uncertainty.

There is no need to perform a toy study for each source of systematic uncertainty. Uncertainties that all effect the same parameter in the likelihood fit can be aggregated into a single study. The sources of systematic uncertainty fall into three general categories

- 1. uncertainties/biases on the individual fit yields due to the PDFs (Table 10.3),
- 2. uncertainties on the individual mode reconstruction efficiencies ϵ (mode) that are uncorrelated among the three signal modes (neural network efficiency and MC statistics in Table 10.8)
- 3. uncertainties on ϵ (mode) that are correlated among the three modes (tracking efficiency, PID systematic, π^0/γ efficiency, π^0/η veto, and photon isolation cut in Table 10.8).

The first two categories can be treated in a straightforward manner with a pair of toy studies $(+1\sigma, -1\sigma)$ per mode per category. To treat the third category of correlated systematics properly, the affected parameters (selection efficiencies ϵ (mode)) must be varied in a correlated way.

The systematic uncertainties for category 1 on N_{eff} associated with each of the modes are determined by generating toy Monte Carlo ensembles with the number of signal events in the mode under study varied by $\pm 1\sigma$ from its value determined in the combined likelihood fit. The resulting deviations are adopted as a 1σ uncertainty on N_{eff} for the given mode. The results of this procedure are listed in Table 11.12.

Table 11.11 shows the total correlated and uncorrelated uncertainties associated with ϵ (mode) expressed as a fraction of ϵ (mode). The systematic uncertainty on N_{eff} associated with

Description	$B^0 o ho^0 \gamma$	$B^+ \to \rho^+ \gamma$	$B^0 \to \omega \gamma$
Correlated error.	0.047	0.068	0.070
Uncorrelated error.	0.112	0.083	0.071

Table 11.11: Break down of the per mode systematic uncertainties on the signal efficiency from Table 10.8 into contributions correlated and uncorrelated among the signal modes.

the uncorrelated part of the systematic uncertainty on ϵ (mode) for a given mode is determined by generating toy Monte Carlo ensembles varying ϵ (mode) by $\pm 1\sigma$. The resulting mean deviations N_{eff} from the control mean are adopted as 1σ systematic uncertainties on N_{eff} . The results are listed in Table 11.12.

The toy ensembles for category 3 are generated with all three ϵ (mode) simultaneously varied in the same direction (up or down) by their associated 1 σ uncertainty. By simultaneously varying the ϵ (mode), the correlations in the uncertainties are preserved. The results are adopted as 1 σ uncertainties on N_{eff} and are also listed in Table 11.12.

Summing the systematic uncertainties in quadrature gives a total uncertainty of $^{+40}_{-45}$ on $N_{\rm eff}.$

11.2.4 Combined limit

The total effective signal yield is

$$N_{\rm eff} = 269^{+126}_{-120} ({\rm stat})^{+40}_{-45} ({\rm syst}).$$

The resulting combined branching fraction is 2.1σ from 0

$$\mathcal{B}(B \to (\rho, \omega)\gamma) = (0.64^{+0.32}_{-0.28}(\text{stat})^{+0.10}_{-0.10}(\text{syst})) \times 10^{-6}$$

Applying the methods for determining upper limits on branching fractions described in Chapter 10.12.3, the upper limit on $\mathcal{B}(B \to (\rho, \omega)\gamma)$ at the 90% confidence level is

$$\mathcal{B}(B \to (\rho, \omega)\gamma) < 1.16 \times 10^{-6}$$

	Input variation (2	$N_{\rm eff}$ output		
Description	$B^0 \to \rho^0 \gamma \text{ mode}$	$B^+ \to \rho^+ \gamma \text{ mode}$	$B^0 \to \omega \gamma \text{ mode}$	err.(events)
<u> </u>	$^{+1.7}_{-1.6}$	_	_	$^{+10.9}_{-9.8}$
Signal bias	-	$^{+2.2}_{-2.2}$	-	$^{+12.9}_{-12.6}$
(category r)	-	-	$^{+1.3}_{-1.9}$	$^{+27.5}_{-34.6}$
Uncompleted oppor	$0.176 \\ 0.140$	-	-	$^{+17.2}_{-16.7}$
(category 2)	-	$\substack{0.143\\0.121}$	-	$^{+4.6}_{-5.5}$
(0000901) 2)	-	-	$0.092 \\ 0.080$	$^{+4.0}_{-4.6}$
Correlated error	0.165	0.141	0.092	+15.6
(category 3)	0.151	0.123	0.080	-14.3
Total				$^{+40.2}_{-44.5}$

Table 11.12: Summary of all the systematic uncertainties for the combined effective signal yield. Each contribution to the systematic uncertainty in the last column is based on a toy Monte Carlo study. The toy ensemble is generated from the fit PDF with either the number of events in a given signal mode N_{mode} (for category 1) or the efficiency of the mode ϵ_{mode} (for categories 2 and 3) varied as indicated in the table. The deviation of the toy ensemble mean fit result N_{eff} from its data fit value is taken as the 1σ contribution to the systematic uncertainty. For example, the first line of the table indicates that a toy ensemble was generated with $N_{B^0 \rightarrow \rho^0 \gamma}$ varied by +1.7 and with $N_{B^+ \rightarrow \rho^+ \gamma}$ and $N_{B^0 \rightarrow \omega \gamma}$ unchanged producing a change in the fit N_{eff} by +10.9. The first line of the "Correlated error" row indicates that a toy ensemble was generated with $\epsilon_{B^0 \rightarrow \rho^0 \gamma} = 0.165$, $\epsilon_{B^+ \rightarrow \rho^+ \gamma} = 0.141$ and $\epsilon_{B^0 \rightarrow \omega \gamma}$ resulting in a +15.6 deviation in N_{eff} .

11.3 $|V_{td}/V_{ts}|$

Equation 11.3 (repeated below) relates the branching fractions $\mathcal{B}(B \to (\rho/\omega)\gamma)$ and $\mathcal{B}(B \to K^*\gamma)$ to the ratio of CKM elements $|V_{td}/V_{ts}|$.

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

The parameter ζ is the ratio of transition form factors that involves non-perturbative physics; ΔR parameterizes the calculable small dynamical differences between the $B \to K^* \gamma$ and $B \to \rho \gamma$ transitions. The current theoretical determinations of both of these parameters are [10]

$$\zeta = 0.85 \pm 0.10$$

 $\Delta R = 0.10 \pm 0.10.$

Of these, the non-perturbative form factor ratio ζ is considered least well known.

The ratio of the central values of the branching fractions from the analyses presented in this thesis is

$$\frac{\mathcal{B}(B \to (\rho/\omega)\gamma)}{\mathcal{B}(B \to K^*\gamma)} = \frac{(0.64 \pm 0.34) \times 10^{-6}}{(3.930 \pm 0.262) \times 10^{-5}} = 0.0163 \pm 0.0087 \;,$$

where the quoted uncertainty is obtained by standard propagation-of-errors. The 90% confidence level upper limit is

$$\frac{\mathcal{B}(B \to (\rho/\omega)\gamma)}{\mathcal{B}(B \to K^*\gamma)} < 0.027.$$

Using the measured central value for the branching fraction ratio and the quoted theoretical parameter values the CKM ratio $|V_{td}/V_{ts}|$ is estimated to be

$$|V_{td}/V_{ts}| = 0.142 \pm 0.038(\exp) \pm 0.017(\zeta) \pm 0.006(\Delta R)$$
,

where the contributions to the uncertainty for each of the parameters has been explicitly retained. The total uncertainty is the square root of the quadrature sum of the components $\sigma(|V_{td}/V_{ts}|) = 0.042$. The 90% confidence level upper limit is

$$|V_{td}/V_{ts}| < 0.195.$$

Ignoring the theoretic uncertainties, as is sometimes done in experimental publications, the 90% confidence level upper limit is $|V_{td}/V_{ts}| < 0.190$.

Chapter 12

Summary and conclusions

Table 12.1 summarizes the experimental measurements produced by the analysis described in this thesis. For ease of comparison, the theoretical calcualations of the measured values from Table 1.1 have been reprinted in Table 12.2.

The branching fractions $\mathcal{B}(B \to K^*\gamma)$ are the most precisely measured values for any radiative penguin decay of B mesons. Although they provide an experimental standard among such measurements, they are not precise tests of the Standard Model. As indicated in Chapter 1.3 and Table 12.2, theoretical calculations of the branching fractions have uncertainties that are over five times those of experimental determinations. Naively, it may appear significant that the theoretical values are almost twice the experimental values, but these differences are actually between $1.1\sigma_{\text{theory}}$ and $1.4\sigma_{\text{theory}}$ and should more properly be interpreted as constraining the uncertain parameter space of the theoretical Standard Model calculations.

The CP $A_{CP}(B \to K^*\gamma)$ and isospin $\Delta_{0-}(B \to K^*\gamma)$ asymmetries show more promise as precision tests of the Standard Model. New physics interactions, most importantly those supersymmetric models, which may have only a very minor effect on the branching fractions, can have large asymmetry effects. The experimental values for $A_{CP}(B \to K^*\gamma)$ and $\Delta_{0-}(B \to K^*\gamma)$

Value	
$\mathcal{B}(B^0 \to K^{*0}\gamma)$	$(3.92 \pm 0.20 \pm 0.24) \times 10^{-5}$
$A_{CP}(B \to K^* \gamma)$	$-0.013 \pm 0.036 \pm 0.010$
$\Delta_{0-}(B \to K^* \gamma)$	$0.050 \pm 0.045 \pm 0.037$
$\mathcal{B}(B^0 \to \rho^0 \gamma)$	$< 0.4 \times 10^{-6}$
$\mathcal{B}(B^+ \to \rho^+ \gamma)$	$<1.8\times10^{-6}$
$\mathcal{B}(B^0 o \omega \gamma)$	$< 1.0 \times 10^{-6}$
$\mathcal{B}(B \to (\rho/\omega)\gamma)$	$< 1.2 \times 10^{-6}$
$\left V_{td}/V_{ts} ight $	< 0.190

Table 12.1: Collected table of measured results presented in this thesis. Values for $B \to K^* \gamma$ are quoted as (central value)±(statistical uncertainty)±(systematic uncertainty). The upper limits for $\mathcal{B}(B \to \rho \gamma)$ and $\mathcal{B}(B^0 \to \omega \gamma)$ are quoted at 90% confidence level.

Calculations			
$\mathcal{B}(B^0 \to K^{*0} \gamma) (\times 10^{-5})$	$7.09^{+2.47}_{-2.27}$ [5]	$7.6^{+3.5}_{-3.0}$ [6]	7.0 ± 2.7 [7]
$A_{CP}(B \to K^* \gamma)$	$< 1\% \ [5, 8]$		
$\Delta_{0-}(B \to K^* \gamma)$	$(8.0^{+2.1}_{-3.2})\% \times (0.3/T_1^{B \to K^*})$ [9]		
$\mathcal{B}(B^0 \to \rho^0 \gamma) (\times 10^{-6})$	$0.49 \pm 0.18 (\text{th}) \pm 0.04 (\text{ex})$ [7]	$0.66 \pm 0.20 [10]$	$0.76^{+0.26}_{-0.23}$ [5]
$\mathcal{B}(B^+ \to \rho^+ \gamma) (\times 10^{-6})$	$0.90 \pm 0.33 (\text{th}) \pm 0.10 (\text{ex})$ [7]	1.35 ± 0.42 [10]	$1.58^{+0.53}_{-0.46}$ [5]

Table 12.2: Reproduction of Table 1.1. Next to leading order theoretical predictions of parameters for the $B \to K^* \gamma$ and $B \to \rho \gamma$ decays. Columns represent different published calculations for each observable. The source of each value is referenced in square brackets [] beside it. The value of $\mathcal{B}(B^0 \to \omega \gamma)$ is theoretically predicted by isospin symmetry to be the same as $\mathcal{B}(B^0 \to \rho^0 \gamma)$. The factor $T_1^{B \to K^*}$ in the prediction for Δ_{0-} is a form factor which is the dominant source of the theoretical error. Determinations of $T_1^{B \to K^*}$ include $0.32^{+0.04}_{-0.02}$ ([11]), 0.38 ± 0.06 ([12]), and 0.27 ± 0.04 ([6]).

presented in this analysis are consistent with the Standard Model. However, the uncertainties on these measurements have not yet ruled out the possibility for discovery of new physics effects. As *BABAR* continues to accumulate data, these measurements will become more precise and remain interesting tests of the Standard Model.

While evidence for the flavor changing neutral current transition $b \to s\gamma$ is now wellestablished—indeed spawning precision measurements—evidence for the similar current $b \to d\gamma$ remains elusive. Neither in the analysis presented here, nor in the corresponding analysis by the Belle collaboration [16], is evidence for $b \to d\gamma$ observed in any of the channels $B^0 \to \rho^0 \gamma$, $B^+ \to \rho^+ \gamma$ or $B^0 \to \omega \gamma$ individually, nor is evidence observed in the sum of these three channels. The 90% confidence level upper limits presented in this analysis are solidly within the range of theoretically calculated values (Table 12.2) and are cause, perhaps, for disappointment, but not for concern about the verity of the Standard Model.

Although the lack of evidence for $B \to \rho \gamma$ means that there is still no direct measurement for $|V_{td}/V_{ts}|$, and hence the length of the leg of unitarity triangle \overline{AB} (Chapter 1.2, Table 1.2), the upper limit of this value is beginning to provide an interesting constraint on the CKM parameters. Figure 12.1 demonstrates the constraint in the $(\overline{\rho}, \overline{\eta})$ plane of the 90% confidence level upper limit on $|V_{td}/V_{ts}|$ presented in Chapter 11.3. The constraints presented in the figure use only the experimental uncertainty, ignoring the uncertainty in the theoretical parameters ζ and ΔR (see Equation 11.3, p. 11.3). The $|V_{td}/V_{ts}|$ limit is consistent with the global CKM fit. Further, with improved theoretical calculations, it can even add a new constraint to the allowed $(\overline{\rho}, \overline{\eta})$ region. The uncertainty in the theoretical parameter ζ significantly affects the CKM constraint, but it is hoped that lattice calculations soon will lead to a significant reduction in this uncertainty.

With BABAR and Belle continuing to take data, it is hoped that soon the decays $B \to \rho \gamma$ will begin yielding precision measurements, just as the $B \to K^* \gamma$ decays have, providing

additional tests of the Standard Model through $A_{CP}(B \to \rho \gamma)$ and $\Delta_{0-}(B \to \rho \gamma)$ and fulfilling their promise of an independent measurement of $|V_{td}/V_{ts}|$.



Figure 12.1: Limit in the $(\overline{\rho}, \overline{\eta})$ plane of the BABAR limit on $|V_{td}/V_{ts}|$ superposed on the winter 2004 results of the CKMfitter group's global CKM fit of relevant experimental results to the unitarity triangle (Figure 1.3). The purple arcs represent two determinations of the $|V_{td}/V_{ts}|$ 90% confidence level upper limit calculated with only the experimental uncertainty. The two arcs use different values of the theoretical parameters $(\zeta, \Delta R)$: $(\zeta, \Delta R) = (0.85, 0.10)$ for the inner arc and $(\zeta, \Delta R) = (0.75, 0.10)$ for the outer arc. The allowed region of the limits lie inside the circles, that is, between the point (1, 0) and the purple circles.

Appendix A

Functions

A.1 Gaussian function

Most introductory books on statistics, including [62], include a treatment of the Gaussian, or normal, distribution. Its functional form is

$$\mathcal{F}_{\text{Gauss}}(x; \langle x \rangle, \sigma) = C_{\text{Gauss}} \cdot \exp(-\frac{(x - \langle x \rangle)^2}{2\sigma^2}) , \qquad (A.1)$$

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

A.2 Novosibirsk function

The Novosibirsk function is a three parameter function describing an asymmetric peak:

$$\mathcal{F}_{\text{Novo}}(x; \langle x \rangle, \sigma, \tau) = C_{\text{Novo}} \cdot \exp\left(-\frac{1}{2}\left(\frac{\ln(1 + \Lambda(\sigma, \tau) \cdot \tau \cdot (x - \langle x \rangle))}{\tau}\right)^2 + \tau^2\right), \quad (A.2)$$

where $\Lambda(\sigma, \tau) = \sinh(\tau \sqrt{\ln 4}) / \sigma \tau \sqrt{\ln 4}$ and C_{Novo} is a normalization constant. Qualitatively, $\mathcal{F}_{\text{Novo}}$ is similar to a Gaussian function with an enhanced tail on one side of the peak,

the side of the peak determined by the sign of τ . The parameters $\langle x \rangle$ and σ represent the mode and width of the peak, and $|\tau|$ parameterizes the significance of the tail. In the limit $\tau \to 0$, the Novosibirsk function becomes a Gaussian:

$$\lim_{\tau \to 0} \mathcal{F}_{\text{Novo}}(x; \langle x \rangle, \sigma, \tau) = \mathcal{F}_{\text{Gauss}}(x; \langle x \rangle, \sigma)$$

A.3 Crystal Ball function

The Crystal Ball function [64] is another function meant to represent a Gaussian distribution with an enhanced tail. This intention is made explicit by defining a cutoff above which the Crystal Ball function is exactly Gaussian and below which it is an inverse power law:

1

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

The parameters $\langle x \rangle$ and σ are the mode and width of the distribution and equivalent to the Gaussian mean and standard deviation in the region $x > \langle x \rangle - \alpha \sigma$. The enhanced tail is defined by a parameter α defining the position of the cutoff and the power n. The parameters α and n are both constrained to be ≥ 0 . $C_{\rm CB}$ is a normalization constant. The function is defined to be continuous and smooth at the cutoff.

A.4 Argus function

The Argus function describes a smooth cutoff behavior determined by kinematic constraints of a system. It was first used to describe a continuum background in m_{ES} by the ARGUS collaboration [65].

$$\mathcal{F}_{\text{Argus}}(x;\xi, E_{\text{BEAM}}) = C_{\text{Argus}} \cdot \frac{x}{E_{\text{BEAM}}} \cdot \sqrt{1 - \frac{x^2}{E_{\text{BEAM}}^2}} \cdot e^{-\xi \left(1 - \frac{x^2}{E_{\text{BEAM}}^2}\right)}$$
(A.4)

The parameter E_{BEAM} is the function's endpoint and it is uniformly 0 for $x > E_{\text{BEAM}}$. The parameter ξ affects the shape of the distribution and is usually called the 'Argus shape parameter', or simply the 'shape parameter'. In the analyses described in this thesis, the continuum endpoint E_{BEAM} in m_{ES} is half the total CM energy of the e^+e^- collision $(\frac{\sqrt{s}}{2})$.

A.5 KEYS non-parametric distribution

KEYS (an acronym for Kernal Estimating Your Shapes) is an implementation of the method of Kernal Estimation to obtain an unbinned, non-parametric, empirical estimation of a probability distribution function (PDF). Kernal Estimation as it is used in particle physics is fully described in [50].

Like a histogram, Kernal Estimation is an empirical estimation of an underlying PDF based on a sample of n data points $\{t_i\}$ randomly drawn from that PDF. The sample defines a discrete *empirical probability density function* $epdf(x) = \sum_i \delta(x - t_i)$. A histogram is a step function estimate of the underlying PDF obtained by integrating epdf(x) over defined intervals $(x_j, x_{j+1}]$. Kernal Estimation, instead, arrives at a smooth, continuous estimate $\hat{f}_0(x)$ of the underlying PDF f(x) by convolving epdf(y) with a smooth, continuous kernal function K(x-y):

$$\hat{f}_0(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-t_i}{h}\right)$$

where h is the smoothing parameter that is assigned based on the granularity of the sample. Many choices for the kernal function K(x) exist.

Where KEYS PDFs are used in this analysis, they uniformly employ a simple Gaussian kernal function:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Further, the analysis uniformly has used a software implementation of the KEYS included in the RooFit [46] package. Kernal Estimation has advantages over the simple histogram as an estimation of a PDF in that it is smooth and free from systematic effects introduced by the arbitrary choice of bin intervals end points. However, it requires more computation than a simple histogram and can significantly increase the time required to perform likelihood fits if implemented inefficiently.

Appendix B

Acronyms

- CKM matrix Cabibbo–Kobayashi–Maskawa quark–mixing matrix
- **CL** Confidence level (p. 220)
- CM frame Center of momentum frame (p. 13)
- CMS Center of momentum system. Same as CM frame.
- **CP symmetry** Discrete symmetry under a combination of charge conjugation (C) and parity (P) transformations.
- **DCH** Drift Chamber (Chapter 2.4)
- **DIRC** Detector of Internally Reflected Cherenkov radiation (Chapter 2.6)
- DOCA Distance of closest approach to the IP (p. 56)
- **EMC** Electromagnetic Calorimeter (Chapter 2.7)
- FW moment Fox–Wolfram moment (p. 51)
- GTL Good Tracks Loose charged candidate selection criteria (p. 50)

- IFR Instrumented Flux Return (p. 18)
- **IP** Interaction point (p. 18)
- **ISR** Initial state radiation (p. 67)
- **KEYS** Kernal estimating your shapes (Appendix A.5)
- LH Likelihood (p. 57)
- MC Monte Carlo, sometimes Monte Carlo data (p. 33)
- MSE Mean squared error (p. 100)
- **NN** Neural network (Chapter 6.1)
- NLL Negative log likelihood (p. 158)
- **NLO** Next to leading order (p. 7)
- **OPE** Operator product expansion (p. 9)
- **PDF** Probability density function (p. 42)
- **PEP–II** Positron–Electron Project II (Chapter 2.1)
- **PID** Particle identification (Chapter 4.3.2)
- **PMT** Photomultiplier tube (p. 27)
- **POCA** Point of closest approach to the IP (p. 25)
- QCD Quantum chromodynamics
- **RMS** Root mean squared
- **ROE** Rest of the event (p. 66)

SLAC Stanford Linear Accelerator Center (p. 13)

- \mathbf{SM} Standard Model (p. 1)
- SNNS Stuttgart Neural Network Simulator (ref. [38])
- **SP** Simulation production (p. 36)
- ${\bf SVT}~$ Silicon Vertex Tracker (Chapter 2.3)

Bibliography

- [1] B. Aubert et al. Measurement of branching fractions, and CP and isospin asymmetries, for $B \rightarrow K^* \gamma$. Phys. Rev., D70:112006, 2004.
- [2] B. Aubert et al. Search for the radiative penguin decays $B^+ \to \rho^+ \gamma$, $B^0 \to \rho^0 \gamma$, and $B^0 \to \omega \gamma$. *Phys. Rev. Lett.*, 94:011801, 2005.
- [3] Paolo Gambino and Mikolaj Misiak. Quark mass effects in $\bar{B} \to X_s \gamma$. Nucl. Phys., B611:338–366, 2001.
- [4] B. Aubert et al. Determination of the branching fraction for inclusive decays $B \to X_s \gamma$. 2002. Contributed to 31st International Conference on High Energy Physics (ICHEP 2002).
- [5] Stefan W. Bosch and Gerhard Buchalla. The radiative decays $B \to V\gamma$ at next-to-leading order in QCD. Nucl. Phys., B621:459–478, 2002.
- [6] M. Beneke, T. Feldmann, and D. Seidel. Systematic approach to exclusive $B \rightarrow V l^+ l^-, V \gamma$ decays. Nucl. Phys., B612:25–58, 2001.
- [7] Ahmed Ali and A. Y. Parkhomenko. Branching ratios for $B \to \rho \gamma$ decays in next-to-leading order in α_s including hard spectator corrections. *Eur. Phys. J.*, C23:89–112, 2002.
- [8] Alexander L. Kagan and Matthias Neubert. Direct CP violation in $B \to X_s \gamma$ decays as a signature of new physics. *Phys. Rev.*, D58:094012, 1998.

- [9] Alexander L. Kagan and Matthias Neubert. Isospin breaking in $B \to K^* \gamma$ decays. *Phys.* Lett., B539:227–234, 2002.
- [10] Ahmed Ali, Enrico Lunghi, and Alexander Ya. Parkhomenko. Implication of the $B \rightarrow (\rho, \omega)\gamma$ branching ratios for the CKM phenomenology. Submitted to Phys. Lett. B, 2004.
- [11] Luigi Del Debbio, Jonathan M. Flynn, Laurent Lellouch, and Juan Nieves. Latticeconstrained parametrizations of form factors for semileptonic and rare radiative B decays. Phys. Lett., B416:392–401, 1998.
- [12] Patricia Ball and Vladimir M. Braun. Exclusive semileptonic and rare B meson decays in QCD. Phys. Rev., D58:094016, 1998.
- [13] B. Aubert et al. Measurement of $B \to K^* \gamma$ branching fractions and charge asymmetries. *Phys. Rev. Lett.*, 88:101805, 2002. *BABAR* Analysis Document (BAD) 203.
- [14] M. Nakao et al. Measurement of the $B \to K^* \gamma$ branching fractions and asymmetries. *Phys. Rev.*, D69:112001, 2004.
- [15] B. Aubert et al. Search for the radiative decays $B \to \rho \gamma$ and $B^0 \to \omega \gamma$. Phys. Rev. Lett., 92:111801, 2004.
- [16] K. Abe et al. Search for the $b \rightarrow d\gamma$ process. 2004. hep-ex/0408137.
- [17] S. Eidelman et al. Review of particle physics. Physics Letters B, 592:1+, 2004.
- [18] Lincoln Wolfenstein. Parametrization of the Kobayashi–Maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.
- [19] J. Charles et al. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. Eur. Phys. J., C41:1–131, 2005.

- [20] R. Ammar et al. Evidence for penguins: First observation of $B \to K^*(892)\gamma$. Phys. Rev. Lett., 71:674–678, 1993.
- [21] PEP-II: An asymmetric B factory. Conceptual design report. June 1993. SLAC-418.
- [22] B. Aubert et al. The babar detector. Nucl. Instrum. Meth., A479:1-116, 2002.
- [23] ed. Harrison, P. F. and ed. Quinn, Helen R. The babar physics book: Physics at an asymmetric *B* factory. Papers from Workshop on Physics at an Asymmetric *B* Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.
- [24] Glenn F. Knoll. Radiation detection and measurement. John Wiley & Sons, New York, USA, third edition, 2000.
- [25] David J. Lange and Anders Ryd. EvtGen documentation. BABAR Analysis Document (BAD) 522, version 6.
- [26] Torbjorn Sjostrand. PYTHIA 5.7 and JETSET 7.4: Physics and manual. 1995.
- [27] S. Agostinelli et al. GEANT4: A simulation toolkit. Nucl. Instrum. Meth., A506:250–303, 2003.
- [28] Mark R. Convery, S. Dasu, K Köneke, P. Spradlin, P. Tan, and S. Willocq. Improved measurements of branching fractions, direct-CP asymmetry and isopin violation in $B \rightarrow K^*\gamma$. BABAR Analysis Document (BAD) 665, version 5.
- [29] Mark R. Convery, S. Dasu, T. Hadig, K. Köneke, J. Libby, S. Sekula, P. Spradlin, P. Tan, and A. Yarritu. Search for $B \to \rho \gamma$ and $B^0 \to \omega \gamma$ with run 1-4 data. *BABAR* Analysis Document (BAD) 876, version 6.

- [30] Geoffrey C. Fox and Stephen Wolfram. Observables for the analysis of event shapes in $e^+e^$ annihilation and other processes. *Phys. Rev. Lett.*, 41:1581, 1978.
- [31] Dieter S. Best, Colin P. Jessop, and Hirohisa Tanaka. The selection criteria and systematics for high energy photon selection for electroweak penguin analysis. BABAR Analysis Document (BAD) 201, version 1.
- [32] Tilmann Colberg, Alan Eisner, Colin Jessop, Ventzislav Koptchev, Steve Playfer, Teela Pulliam, Klaus Schubert, Bruce Schumm, Hirohisa Tanaka, and Stephane Willoq. Measurement of $BR(B \to K^*\gamma)$ and search for direct CP violation. *BABAR* Analysis Document (BAD) 33, version 16.
- [33] BABAR Micro Database. Candidate lists and maps in the micro database. http://www.slac.stanford.edu/BFROOT/www/doc/ /workbook/nanomicro/v8.8/Micro/CandLists.html.
- [34] Neutral Reconstruction AWG. Neutral Reconstruction AWG. http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals.
- [35] Vertexing and Composition Tools Group. The BABAR vertexing. BABAR Analysis Document (BAD) 102, version 7.
- [36] Roy Aleksan, John Back, Giulia Bellodi, Andre Gaidot, Pierre-Francois Giraud, Paul F. Harrison, Theresa Harrison, Andreas Hoecker, Sandrine Laplace, Francois Le Diberder, Ran Liu, Yibin Pan, Vasilii Shelkov, Jan Stark, Georges H. Vasseur, Jinwei Wu, and Christophe Yeche. Measurement of branching fractions and CP-violating asymmetries in B⁰ → ρ⁻h⁺. BABAR Analysis Document (BAD) 350, version 11.
- [37] Juerg Beringer. BTagger a multivariate tagging algorithm with categories based on the physics of the Btag decay. BABAR Analysis Document (BAD) 317, version 4.

- [38] A. Zell et al. Stuttgart neural network simulator (SNNS). http://www-ra.informatik.uni-tuebingen.de/SNNS/.
- [39] M. S. Srivastava and E. M. Carter. An Introduction to Applied Multivariate Statistics. North-Holland, New York, Amsterdam, Oxford, 1983. ISBN 0444006214.
- [40] B. Aubert et al. B meson decays to η^(')K*, η^(')ρ, η^(')π⁰, ωπ⁰, and φπ⁰. BABAR Analysis Document (BAD) 720, Presented at 39th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions, La Thuile, Italy, 28 Mar - 4 Apr 2004., 2004.
- [41] Daniel A. Bowerman, James Robert Gaillard, Ivo Gough Eschrich, Geoffrey W. Morton, and Michael Borivoje Nikolich. Measurement of the branching fraction and charge asymmetry for the decay $B^{\pm} \rightarrow K^{*\pm}\pi^0$. BABAR Analysis Document (BAD) 814, version 4.
- [42] B. Aubert et al. Measurement of branching fractions and charge asymmetries in $B^{\pm} \rightarrow \rho^{\pm} \pi^{0}$ and $B^{\pm} \rightarrow \rho^{0} \pi^{\pm}$ decays, and search for $B^{0} \rightarrow \rho^{0} \pi^{0}$. BABAR Analysis Document (BAD) 700, submitted to Phys.Rev.Lett., 2003.
- [43] B. Aubert et al. Searches for B⁰ decays to combinations of charmless isoscalar mesons. Based on BABAR Analysis Document (BAD) 705, Presented at 39th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions, La Thuile, Italy, 28 Mar - 4 Apr 2004., 2004.
- [44] B. Aubert et al. Observation of $B^0 \to \omega K^0$, $B^+ \to \eta \pi^+$, and $B^+ \to \eta K^+$ and study of related decays. *Phys. Rev. Lett.*, 92:061801, 2004. *BABAR* Analysis Document (BAD) 753.
- [45] B. Aubert et al. Observation of the decay $B^{\pm} \to \pi^{\pm}\pi^{0}$, study of $B^{\pm} \to K^{\pm}\pi^{0}$, and search for $B^{0} \to \pi^{0}\pi^{0}$. *Phys. Rev. Lett.*, 91:021801, 2003.
- [46] W. Verkerke, D. Kirkby, et al. RooFit toolkit for data modelling. http://roofit.sourceforge.net/.

- [47] R. Brun et al. ROOT: An object-oriented data analaysis framework. http://root.cern.ch/.
- [48] H. Albrecht et al. Reconstruction of B mesons. Phys. Lett., B185:218, 1987.
- [49] Tomasz Skwarnicki. A study of the radiative cascade transitions between the Υ' and Υ resonances. DESY F31-86-02.
- [50] Kyle S. Cranmer. Kernel estimation in high-energy physics. Comput. Phys. Commun., 136:198–207, 2001.
- [51] Christopher Hearty. Measurement of the number of $\Upsilon(4S)$ mesons produced in run 1 (*B* counting). *BABAR* Analysis Document (BAD) 134, version 1.
- [52] Marcella Bona, Gianluca Cavoto, Carlo Dallapiccola, Bill Ford, Sergio Grancagnolo, Natalia Kuznetsova, Alex Olivas, Mirna van Hoek, Erich Varnes, Steve Wagner, and Lei Zhang. Report of the Tracking Efficiency Task Force for 2001. BABAR Analysis Document (BAD) 324, version 2.
- [53] BABAR PID Group. BABAR particle ID. http://www.slac.stanford.edu/BFR00T/www/Physics/Tools/Pid/pid.html.
- [54] Mark Tiller Allen, Mitchell T. Naisbit, and Aaron Roodman. A study of π^0 efficiency. BABAR Analysis Document (BAD) 870.
- [55] Tilmann Colberg. Measurement of the decay fractions of $B^{\pm} \to K^{*\pm}\gamma, (K^{*\pm} \to \pi^0 K^{\pm})$ and $B^0 \to K^{*0}\gamma, (K^{*0} \to \pi^0 K_S^0)$. BABAR Analysis Document (BAD) 133, version 13.
- [56] B. Aubert et al. Measurement of the direct CP asymmetry in $b \rightarrow s\gamma$ decays. *Phys. Rev.* Lett., 93:021804, 2004.

- [57] B. Aubert et al. Study of $B^{\pm} \to J/\psi \ \pi^{\pm}$ and $B^{\pm} \to J/\psi \ K^{\pm}$ decays: Measurement of the ratio of branching fractions and search for direct CP violation. *Phys. Rev. Lett.*, 92:241802, 2004.
- [58] Roger Barlow, Robert Cahn, Glen Cowan, Francesca Di Lodovico, William T. Ford, Gautier Hamel de Monchenault, David G. Hitlin, David Kirkby, Francois Le Diberder, Gerald Lynch, Frank C. Porter, Soeren A. Prell, Art Snyder, Michael D. Sokoloff, and Roland Waldi. Recommended statistical prodedures for BABAR. BABAR Analysis Document (BAD) 318, version 1.
- [59] Robert D. Cousins and Virgil L. Highland. Incorporating systematic uncertainties into an upper limit. Nucl. Instrum. Meth., A320:331–335, 1992.
- [60] B. Aubert et al. Search for the radiative decays $B \rightarrow \rho \gamma$ and $B^0 \rightarrow \omega \gamma$. Phys. Rev. Lett., 92:111801, 2004. BABAR Analysis Document (BAD) 539, based on BABAR Analysis Document (BAD) 354.
- [61] B. Aubert et al. Measurement of time-dependent CP-violating asymmetries in $B^0 \to K^{*0}\gamma$ $(K^{*0} \to K^0_S \pi^0)$ decays. *Phys. Rev. Lett.*, 93:201801, 2004.
- [62] O. Skjeggestad A.G. Frodesen and H. Tøfte. Probability and Statistics in Particle Physics. Universitetsforlaget, Bergen, Norway, 1979.
- [63] B. Aubert et al. Measurement of the B^+/B^0 production ratio from the $\Upsilon(4S)$ meson using $B^+ \to J/\psi \ K^+$ and $B^0 \to J/\psi \ K^0_S$ decays. *Phys. Rev.*, D69:071101, 2004.
- [64] E. D. Bloom and C. W. Peck. Physics with the Crystal Ball detector. Ann. Rev. Nucl. Part. Sci., 33:143, 1983.
- [65] H. Albrecht et al. Exclusive hadronic decays of B mesons. Z. Phys., C48:543-552, 1990.