

Leptonic Decays of the Charged B Meson

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Luke A. Corwin, B.S.

* * * * *

The Ohio State University

2008

Dissertation Committee:

Professor Klaus Honscheid, Adviser

Professor Eric Braaten

Professor Richard Kass

Professor Mike Lisa

Professor Hayrani Oz

Approved by

Adviser
Graduate Program in
Physics

ABSTRACT

We present a search for the decay $B^+ \rightarrow \ell^+ \nu_\ell$ ($\ell = \tau, \mu, \text{ or } e$) in $(458.9 \pm 5.1) \times 10^6$ $\Upsilon(4S)$ decays recorded with the *BABAR* detector at the SLAC PEP-II *B*-Factory. A sample of events with one reconstructed exclusive semi-leptonic *B* decay ($B^- \rightarrow D^0 \ell^- \bar{\nu} X$) is selected, and in the recoil a search for $B^+ \rightarrow \ell^+ \nu_\ell$ signal is performed. The τ is identified in the following channels: $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$, $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$, $\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$, and $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$. The analysis strategy and the statistical procedure is set up for branching fraction extraction or upper limit determination. We determine from the dataset a preliminary measurement of $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) = (1.8 \pm 0.8 \pm 0.1) \times 10^{-4}$, which excludes zero at 2.4σ , and $f_B = 255 \pm 58$ MeV. Combination with the hadronically tagged measurement yields $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) = (1.8 \pm 0.6) \times 10^{-4}$. We also set preliminary limits on the branching fractions at $\mathcal{B}(B^+ \rightarrow e^+ \nu_e) < 7.7 \times 10^{-6}$ (90% C.L.), $\mathcal{B}(B^+ \rightarrow \mu^+ \nu_\mu) < 11 \times 10^{-6}$ (90% C.L.), and $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) < 3.2 \times 10^{-4}$ (90% C.L.).

“Great are the works of the LORD, studied by all who delight in them.”

- Pslam 111:2 (NRSV)

ACKNOWLEDGMENTS

For most of my graduate career, I have had the verse on the previous page above my desk. It reminds me to acknowledge and thank God for creating a Universe that obeys fixed laws and giving me the opportunity to joyfully study them first. My parents Anita and Ron Abts have always supported and encouraged me in my academic and scientific career, even though they did not always understand what I was studying. They sacrificed more than I can understand to provide the best home and education that they could.

At Independence, WI Public Schools, I was blessed with several teachers who encouraged my intellectual development inside and outside of science. In particular, I acknowledge my English teacher, Ms. Susan Solli, who taught me the value and beauty of words. Ms. Solli and Ms. Heidi Bragger were coaches on the forensics team; they were critical in developing my presentation skills. Mrs. Pamela Lehmeier was the director of the high school and grade school bands; she brought out my musical talent. I continued playing the tuba through graduate school and thus learned a great deal about the connections between music and physics.

As an undergraduate at the University of Minnesota, I had the opportunity to work with Professors Ron Poling and Jon Urheim on my first particle physics research project, which was on the CLEO experiment. Prof. Urheim was my supervisor, and under his direction, I obtained a preliminary result. This experience confirmed that I wanted to stay in particle physics as a graduate student. I also thank my lab-mates Valery Frolov, Alexander Scott and Alexander Smith for all of their help. Also in Minnesota, I had several pastors and friends on campus and in church who helped me accept that scientific work was a legitimate calling and taught me how to better

integrate my faith and my vocation. In particular, I thank Dave & Cheri Burkum and Steve Knight at Christian Student Fellowship, Steve Treichler of Hope Community Church, and Pat Khanke at St. Paul Fellowship.

At Ohio State, I am grateful to Prof. Honscheid, who originally requited me for this research group and has been an excellent adviser since then. The entire *BABAR* group here has been a source of help, friendship, and pizza. I specifically acknowledge Dr. Joseph Regensburger, Dr. Quincy Wong, and Don Burdette. James Morris has also been an invaluable, reliable, and patient purveyor of computer support. My spiritual journey has been guided by members of the Christian Graduate Student Alliance (CGSA), Summit United Methodist Church, and Continuum Church. At CGSA, I thank Bob Trube, Gary Nielson, Markus & Stephanie Dickinson, and Paul Rimmer. At Summit, I thank Pastor Linda Wallick, Wil Ranney, and Risë Straight. I also thank Kelly, whose introduction into my life and heart has been the most pleasant surprise of the past five years, for her patience, understanding, compassion, humor, honesty, and intellect.

At SLAC, several members of the Leptonic bottom and charmed Analysis Working Group made valuable contributions to this analysis. This analysis would not have been possible without the assistance of Dr. Steve Sekula and Dr. Paul “Jack” Jackson, the postdoctoral researchers with whom I worked most closely. I also acknowledge the members of All Nations Christian Fellowship and the InterVarsity Graduate Christian Fellowship and Stanford, including Pete Sommer and Diane Schouten. I owe a very personal thaks to Peter Kockelman.

This work is supported by the U.S. Department of Energy.

VITA

- December 20, 1980 Born - Sheboygan, WI, USA
- May 1999 Valedictorian, Independence Public Schools, WI, USA
- May 2003 B.S. Physics, *summa cum laude*,
University of Minnesota - Twin Cities, Minneapolis, MN, USA
- June 2003 - Present Ph.D. Candidate, The Ohio State University, Columbus, OH, USA

PUBLICATIONS

1. B. Aubert *et al.*, “A search for $B^+ \rightarrow \tau^+ \nu_\tau$,” Phys. Rev. D **76**, 052002 (2007).
2. L. A. Corwin, “The search for $B^+ \rightarrow \tau^+ \nu_\tau$ at BaBar,” Nucl. Phys. Proc. Suppl. **169**, 70 (2007).

FIELDS OF STUDY

Major Field: Experimental High Energy Physics.

TABLE OF CONTENTS

	Page
Abstract	ii
Dedication	iii
Chapters:	
1. Introduction	1
1.1 Particle Physics	1
1.2 The Standard Model of Particle Physics	3
1.2.1 Fundamental Particles	4
1.2.2 Fundamental Interactions	7
1.2.3 CKM Matrix	13
1.2.4 Unanswered Questions	14
1.3 Notes To Reader	15
2. Fully Leptonic Decays of the Charged B Meson	17
2.1 Experimental Challenges	18
2.1.1 Neutrinos and Tagging	18

2.1.2	Helicity Suppression	19
2.2	Theoretical Predictions	19
2.3	Previous Experimental Searches	22
2.4	Experimental Procedure	22
3.	The <i>B</i> Factory at SLAC	26
3.1	The Linear Accelerator	26
3.2	Positron Electron Project II	28
3.3	The <i>BABAR</i> Detector	33
3.3.1	Silicon Vertex Tracker	34
3.3.2	Drift Chamber	36
3.3.3	Detector of Internally Reflected Cerenkov light	39
3.3.4	Electromagnetic Calorimeter	42
3.3.5	Instrumented Flux Return	45
3.4	Particle Identification	49
3.4.1	Identification of Individual Particles	49
3.4.2	Composite Particle Lists	51
3.5	Blinding and Monte Carlo Simulation of the <i>BABAR</i> Detector	52
3.6	Data Quality Monitoring	54
3.7	Run Divisions	55
4.	Experimental Details and Data Sets	56
4.1	Event Selection Outline	56
4.2	Data	57
4.3	Monte Carlo Simulations	58

4.3.1	Generic Background Monte Carlo Samples	58
4.3.2	Signal Monte Carlo Samples	58
4.4	Skim Selection	59
4.5	Tag Selection	60
4.5.1	Seeding Method	60
4.5.2	Accepted Signal Modes	61
4.6	Tag Selection for $B^+ \rightarrow \ell^+ \nu_\ell$	62
4.6.1	Optimization Procedure	63
4.6.2	Optimization Results	65
4.7	Tag B Efficiency and Yield	66
4.8	Fitting vs. Cuts	66
5.	Experimental Search for $B^+ \rightarrow \ell^+ \nu_\ell$	72
5.1	Separating $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$ from $B^+ \rightarrow \tau^+ \nu_\tau$	73
5.2	Extra Energy	79
5.3	Likelihood Ratio	80
5.3.1	Variables Included in LHR	82
5.3.2	Final PDF Selection	89
5.4	Optimization Procedure	92
5.4.1	Photon Pair Conversion	94
5.4.2	Signal and Overall Efficiency	95
6.	Control Samples and Systematic Corrections	98
6.1	Control Samples	98
6.1.1	D^0 Mass Sideband Sample	98

6.1.2	E_{extra} Sideband Sample	99
6.1.3	LHR Sideband Samples	99
6.1.4	$p'_{\text{sig } \ell}$ Sideband Sample	99
6.1.5	Double Tag Sample	100
6.2	Background Prediction	102
6.2.1	Background Prediction using 2D Side Bands	102
6.2.2	Background Prediction using 1D Side Bands	103
6.2.3	Choice of Background Prediction Method	107
7.	The Feldman-Cousins Method	110
7.1	The Feldman Cousins Method	110
7.2	Application of Systematic Errors	115
8.	Systematics	116
8.1	Systematic Error from Background Prediction (N_{BG})	116
8.2	Systematic Error from B-Counting (N_{BB})	117
8.3	Systematic Error from Tagging Efficiency (ε_{tag})	117
8.3.1	$D^0 \rightarrow K^- \pi^+$ tags	118
8.3.2	$D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ tags	119
8.4	Systematic Error from Signal Efficiency (ε_{sig})	119
8.4.1	Modeling of E_{extra} shape	119
8.4.2	Tracking Efficiency for Signal Side	120
8.4.3	π^0 Selection	121
8.4.4	Particle Identification	121
8.5	Summary of Systematic Corrections and Uncertainties	122

9. Results	124
9.1 $B^+ \rightarrow e^+ \nu_e$	124
9.2 $B^+ \rightarrow \mu^+ \nu_\mu$	124
9.3 $B^+ \rightarrow \tau^+ \nu_\tau$	126
9.3.1 Exclusion of the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ mode	127
9.3.2 Excess seen in $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ mode	127
9.3.3 Results for $B^+ \rightarrow \tau^+ \nu_\tau$	133
9.3.4 $f_B V_{ub} $	135
9.3.5 Combined Result	136
9.3.6 Constraints on the Type II Two Higgs Doublet Model	137
9.4 Summary	140

LIST OF TABLES

Table		Page
1.1	A summary of the known fundamental constituents of matter [1] is shown. The names of the particles and the symbols used to represent them are given. Also listed are their electric charge relative to the electric charge of the proton. Their mass, in units of MeV/c^2 , is given. Notice all 12 particles are grouped into three generations of increasing mass and decreasing lifetime.	4
1.2	Properties of the composite particles relevant to this paper [1] are shown. The same units are used in this table as in Table 1.1. The n is the neutron and \bar{n} is the anti-neutron. The quark content of the particles is given in parentheses next to the particle symbol. Charge is given in units of the charge of the proton.	5

1.3	Mediators of the fundamental interactions. Charge (Q) is again in units of the charge of the proton. Spin is given in units of \hbar . Note that the weak force is unusual in having three mediators. The relative strengths of the interactions are given for interacting particles at two example distances.	7
2.1	Existing and predicted upper limits (at 90% C.L.) and measurements of the branching fractions of the $B^+ \rightarrow \ell^+ \nu_\ell$ decay. The inputs to the Naive SM predictions are from [1], except $ V_{ub} $ and f_B , for which we have included the latest values. CKM Fitter and UT Fitter are groups that collect and summarize the most current particle physics results and use them to predict unmeasured quantities, such as those in the table. The semileptonic tag measurements from <i>BABAR</i> are the results found in this analysis.	21
3.1	Production cross-sections in the PEP-II e^+e^- collider at 10.58 GeV center-of-mass.	30
3.2	Particle Identification Selectors. The two <code>Tracks</code> lists do not use particle identification. $ d_0 $ is the distance, transverse to the beam axis, from the interaction point to the point of closest approach for the track. $ z_0 $ is the distance, along the beam axis, from the interaction point to the point of closest approach for the track. \mathcal{L} has the same meaning as in Equation 3.9.	51
3.3	Data samples used in the analysis.	55

4.1	MC samples used in the analysis. Luminosity of each MC sample are obtained by dividing the number of events with the appropriate cross-sections. The cross-sections for all species except $B\bar{B}$ are taken from <i>The BaBar Physics Book</i> [2]. The $B\bar{B}$ cross-section is calculated by dividing the number of $\Upsilon(4S)$ in the on-peak data, assuming that the production of B^+B^- and $B^0\bar{B}^0$ are equal.	68
4.2	Signal MC samples used in the analysis. Luminosity of each MC sample are obtained by dividing the number of events with the appropriate cross-sections.	69
4.3	Branching fraction of the decays involved in $B^- \rightarrow D^{*0}\ell^-\bar{\nu}_\ell$ tag reconstruction [1].	70
4.4	Branching fraction of the τ decay modes used for $B^+ \rightarrow \tau^+\nu_\tau$ signal search [1].	70
4.5	Number of events in each training and validation MC sample.	70
4.6	Comparison of Tag Efficiencies for different cuts. All of the yields are raw, unweighted yields for the signal MC sample for the mode given, representing Runs 1-6.	71
5.1	Comparison of two different minimum neutral cluster energy requirements for E_{extra} . After our data sample has passed the tag selection requirements and the signal mode in question has been reconstructed, we optimize a requirement on <i>only</i> E_{extra} . Comparing the resulting Figures of Merit (same as used in §5.4) indicates which definition has the most discriminating power. The number in the superscript indicates the minimum neutral cluster energy requirement in MeV.	80

5.2	This is a list of all variables used as signal PDFs for the various decay modes. Variables in bold are used for both Continuum and $B\bar{B}$ background. Variables in plain roman font are used for Continuum only. Variables in <i>italics</i> are used for $B\bar{B}$ only.	91
5.3	Optimized ranges from which we accept signal candidates, which are mostly given by our optimization procedure. The exceptions are the $p'_{\text{sig } \ell}$ ranges for the two leptonic τ decay modes, which were chosen as described in §5.1, and the upper limits of the $p'_{\text{sig } \ell}$ ranges for the other two leptonic B decays, which were chosen to incorporate almost all of the signal MC samples.	94
5.4	Signal predictions using optimized cuts. BG predictions from the E_{extra} sideband, as described in 6.2.2. The figure of merit (FOM) is calculated from the second and third columns.	95
5.5	Overall efficiency ($\varepsilon \equiv \varepsilon_{\text{sig}} \times \varepsilon_{\text{tag}}$) of optimized signal selection for all modes in Runs 1-6. These are unnormalized numbers of events for the relevant signal MC Test samples. The Yield is the number of events passing the optimized selection. The third column is the total number of events generated for the signal MC sample divided by 3; the division counterbalances the reduction in yield caused by using only the Test MC samples. The fourth column is the quotient of the Yield over the Number Generated, which is, by definition the same as the product of the tag and signal efficiencies.	96
6.1	BG predictions using the D^0 mass sideband and peaking generic MC.	104

6.2	BG Predictions from E_{extra} sideband. R_{MC} is the ratio of events in the sideband to events in the signal region of E_{extra} in the background MC. $N_{\text{data,SideB}}$ is the number of events in the E_{extra} sideband in data. $N_{\text{MC,Sig}}$ is the number of normalized events in the E_{extra} signal region of the background MC samples. This is the background prediction taken solely from the MC samples. $N_{\text{exp,Sig}}$ is the product of R_{MC} and $N_{\text{data,SideB}}$; it is the background prediction extrapolated from the data sideband using the MC samples.	106
6.3	BG Predictions from $LHR_{\text{cont.}}$ sideband. All variables are as described in the caption of Table 6.2.	106
6.4	BG Predictions from $LHR_{B\bar{B}}$ sideband. All variables are as described in the caption of Table 6.2.	107
6.5	BG Predictions from $p'_{\text{sig } \ell}$ sideband. All variables are as described in the caption of Table 6.2.	107
6.6	Comparison of BG predictions from various sidebands.	109
8.1	Single and double tag yields in the full data and MC samples of $D^0 \rightarrow K^-\pi^+$ events.	119
8.2	Single and double tag yields in full data and MC samples of $D^0 \rightarrow K^-\pi^+\pi^-\pi^+$ events.	119
8.3	Summary of systematic corrections, uncertainties, and fractional uncertainties	122
8.4	The corrected tag and signal efficiencies. Two errors are quoted: the first is the MC statistical uncertainty, and the second is the systematic error computed from the sources in this section.	123

9.1	Each branching fraction is calculated for each of the four τ decay modes separately.	129
9.2	Comparison of background predictions (from E_{extra} sideband) and observed yields in on-peak data for the two tag leptons. The fourth column is given by $(N_{\text{obs}} - N_{\text{BG}})/\sigma_{N_{\text{BG}}}$. Note that these studies were performed <i>before</i> the $m_{\ell\ell}$ requirement was applied.	132
9.3	Comparison of background predictions (from E_{extra} sideband) and observed yields in on-peak data for the two subsets of the signal region of E_{extra} . The fourth column is given by $(N_{\text{obs}} - N_{\text{BG}})/\sigma_{N_{\text{BG}}}$. Note that these studies were performed <i>before</i> the $m_{\ell\ell}$ requirement was applied.	132
9.4	The observed number of on-resonance data events in the signal region are shown, together with the number of expected background events and the number of expected signal events, which is taken from our signal MC samples. The FOM for the τ modes is $N_{\text{sig}}/\sqrt{N_{\text{obs}}}$ and $N_{\text{sig}}/(1.5 + \sqrt{N_{\text{BG}}})$ for the other two modes. Note that N_{sig} and hence FOM is negative for some modes.	139
9.5	$B^+ \rightarrow \tau^+ \nu_\tau$ with the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ channel.	141
9.6	$B^+ \rightarrow \tau^+ \nu_\tau$ without the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ channel.	142
9.7	$B^+ \rightarrow \mu^+ \nu_\mu$	143
9.8	$B^+ \rightarrow e^+ \nu_e$	143

LIST OF FIGURES

Figure	Page
1.1 The strong interaction mediates the decay of a $b\bar{b}$ meson with sufficient mass to produce two charged B mesons. The helix represents a gluon that mediates the strong interaction. The interactions also causes a $u\bar{u}$ pair to form as the b and \bar{b} grow farther apart. This decay is of importance to many analyses at <i>BABAR</i> , including this one, because it is the primary decay mode of the $\mathcal{T}(4S)$	8
1.2 Two examples of the electromagnetic interaction are shown. At left is an example where an electron and positron collide to produce a photon, represented by the wavy lines, that then generates a $\mu^+\mu^-$ pair. This is one example of matter-antimatter annihilation producing new particles. At right is the Feynman diagram of the decay $\pi^0 \rightarrow \gamma\gamma$.	10
1.3 Two examples of the weak interaction are shown. At left is the decay studied in this analysis, the decay of a charged B meson via the weak interaction into a charged and neutral lepton. At right is the decay of a neutron, which is also governed by the weak interaction.	11

1.4	This two-photon fusion event is an example of a Feynman diagram that has both electromagnetic and strong interactions. Virtual photons emitted from the initial electron and positron produce a u and \bar{u} . The strong interaction produces another quark-antiquark pair to satisfy color confinement. The strong interaction also binds the quarks into two mesons. This process can produce background events when the hadrons produced by the fused photons are misreconstructed as a D^0 , one of the electrons is misidentified as a tag lepton and the other as the signal lepton.	13
2.1	Feynman diagram for the purely leptonic B decay $B^+ \rightarrow \ell^+ \nu_\ell$	18
2.2	The reconstructed invariant mass of each pair of charged K and π mesons in each event in Runs 1-5 is shown. Notice the peak at the mass of the D^0	24
3.1	Overview of SLAC, including PEP-II and <i>BABAR</i>	27
3.2	Cross-section of the <i>BABAR</i> detector, viewing in the direction of the positron beam. All distance measurements are in mm.	31
3.3	Cross-section of the <i>BABAR</i> detector, side view. All distance measurements are in mm.	32
3.4	A schematic cross-section of the SVT, looking down the beam line, is shown.	35
3.5	Schematic cross-section of the SVT, viewed from the side	35
3.6	A schematic cross-section of the Drift Chamber is shown.	36

3.7	Performance of the DCH separating particles of different types using dE/dx as a function of track momentum. Protons are represented as p and deuterons, which are a heavy hydrogen nucleus consisting of a proton and neutron, are denoted d	37
3.8	A schematic view of the wires and cells within a small section of the DCH is shown.	38
3.9	Schematic view of the DIRC, viewed from the side.	40
3.10	Schematic view of a single quartz bar and readout within the DIRC.	41
3.11	Schematic view of the upper half of the EMC, viewed from the side.	43
3.12	Schematic view of a single crystal and readout within the EMC.	44
3.13	Schematic view of the IFR, including barrel and end caps.	45
3.14	Schematic view of a signal planar resistive plate chamber (RPC), which was part of the original IFR.	46
3.15	Efficiency history for 12 months starting in June 1999 for RPC modules showing different performance: a) highly efficient and stable; b) continuous slow decrease in efficiency; c) more recent, faster decrease in efficiency.	47
3.16	The cross section of a limited streamer tube is shown.	47

5.1	In this representation of semileptonic B tagging, the dashed white vector is $\vec{p}_{D^0\ell}^*$. The red vectors, which are constrained to lie on the surface of the blue cone, represent possible directions for the signal B momentum. The angle $\theta_{B-D^0\ell}$ defines the cone with respect to $\vec{p}_{D^0\ell}^*$. The gray spot is the B vertex, and the red cone represents the possible paths of the signal B projected back through the vertex. The green ellipse represents the beam spot. This graphic is <i>not</i> to scale.	74
5.2	This plot is of three different calculations of the signal lepton momentum in the $B^+ \rightarrow e^+\nu_e$ (left) and $B^+ \rightarrow \mu^+\nu_\mu$ (right) modes. The black solid line shows p_{sig}^* ; the red solid line shows $p'_{\text{sig } \ell}$ calculated using the Beam Spot method; the green dashed line shows $p'_{\text{sig } \ell}$ calculated using the Y-Average Method. These distributions were made using the signal MC samples representing Runs 1-6. Tag selection is applied, but no signal selection is applied beyond requiring that the selected signal mode is reconstructed.	77
5.3	Background and signal distributions for $p'_{\text{sig } \ell}$ from both the Beam Spot and Y-Average methods. No significant difference is seen in background distributions for the two methods. All MC and data samples normalized to Run 1-6 data luminosity.	78
5.4	Shown are the $p'_{\text{sig } \ell}$ distributions for the regions of overlap between the $B^+ \rightarrow \tau^+\nu_\tau$, $B^+ \rightarrow e^+\nu_e$ (left), and $B^+ \rightarrow \mu^+\nu_\mu$ (right) signal MC samples. All distributions are scaled to Runs 1-6 data luminosity, except $B^+ \rightarrow e^+\nu_e$ which has its branching fraction scaled up to 1×10^{-7}	79

5.5	At left, we have the momentum spectra, in the center of mass frame, for the signal leptons in the $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ and $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ modes. We can see that muon identification becomes less efficient than electron identification below about 1.2 GeV. At right, we have the same momentum spectrum with distributions of the different modes reconstructed as pions. This shows that, at low momenta many of the true muons and electrons are reconstructed as pions. Also, we see a sizable contribution from real $\rho^+ \rightarrow \pi^+ \pi^0$ where we have missed the π^0 . All distributions are taken from the signal MC sample and scaled to Run 1-6 data luminosity. The dashed lines are the true distributions for the indicated particle that pass the tag selection. The solid lines show the distributions that pass the tag cuts and are reconstructed as the given mode.	97
6.1	Distribution of E_{extra} in double tags with different sets of selection requirements. Data and MC are normalized to Run 1-6 luminosity in the top left and top right graphs. In the bottom graph the same cuts are applied as in the top right graph, but Data and MC are normalized to unity.	101
6.2	Proposed construction for 2D background prediction	102
7.1	The fits to N_{sig} for several values of N_{true}	111
7.2	The distribution of N_{true} vs. N_{sig} . For each integer value of N_{true} , we generated 100,000 random numbers and fit the resulting N_{sig} distribution to the sum of the Gaussian functions.	112
7.3	An example of the distribution of R	113

9.1	<p>Total extra energy is plotted after all cuts have been applied to the $B^+ \rightarrow e^+\nu_e$ mode. Off-resonance data and MC have been normalized to the on-resonance luminosity. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC samples have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2 and summed together. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow e^+\nu_e$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.</p>	125
9.2	<p>Total extra energy is plotted after all cuts have been applied to the $B^+ \rightarrow \mu^+\nu_\mu$ mode. Off-resonance data and MC have been normalized to the on-resonance luminosity. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC samples have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2 and summed together. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \mu^+\nu_\mu$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.</p>	126

- 9.3 $m_{\ell\ell}$ is plotted after all cuts have been applied for the $\tau^+ \rightarrow e^+\nu_e\bar{\nu}_\tau$ mode. Events in this distribution are required to pass all selection criteria, except the $m_{\ell\ell}$ requirement. The background MC sample has been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5. . . . 128
- 9.4 The branching fractions calculated from each τ decay channel are plotted. 1 = $\tau^+ \rightarrow e^+\nu_e\bar{\nu}_\tau$, 2 = $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$, 3 = $\tau^+ \rightarrow \pi^+\bar{\nu}_\tau$, 4 = $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$. The horizontal line is a fit to a constant. 129
- 9.5 Total extra energy is plotted for the mode $B^+ \rightarrow \tau^+\nu_\tau(\tau^+ \rightarrow e^+\nu_e\bar{\nu}_\tau)$. This distribution is taken from the sidebands of the D^0 mass distribution (§6.1.1). All other optimized signal selection criteria, have been applied. Comparing this to Figure 9.6 shows that the excess of signal events in this mode occurs in the peak of the D^0 mass distribution, which suggests that it is not due to two-photon fusion. The generic MC samples are scaled to data luminosity. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5. . . . 130

9.6	<p>Total extra energy is plotted after all cuts have been applied in the modes (left) $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and (right) $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$. The background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.</p>	133
9.7	<p>Total extra energy is plotted after all cuts have been applied in the mode (left) $\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$ and (right) $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$. The background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.</p>	134

9.8	Total extra energy is plotted after all cuts have been applied with all $B^+ \rightarrow \tau^+ \nu_\tau$ modes combined. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5. . . .	135
9.9	Total extra energy is plotted after all cuts have been applied for $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and with no Bremsstrahlung recovery. Events in this distribution are required to pass all selection criteria. The MC is not scaled. .	136
9.10	The confidence band produced by the Feldman Cousins method. The central band represents the 1σ confidence level; the next band out represents the 90% confidence level. The vertical red line represents the value of N_{sig} based on the number of observed events. The top left plot is for $B^+ \rightarrow \tau^+ \nu_\tau$; top right is for $B^+ \rightarrow \mu^+ \nu_\mu$, and the bottom plot is for $B^+ \rightarrow e^+ \nu_e$	137

9.11 Shown are the regions in phase space ($\tan \beta$ vs. m_{H^\pm}) of the Two Higgs Doublet Model that are excluded by all relevant measurements and searches, including this one. The colored regions are excluded at the 95% confidence level [3]. The two green wedges are the region excluded by the world average of $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$ before my measurement was made public. Since my measurement is consistent with the average, these wedges will not change greatly when it is included. 138

CHAPTER 1

INTRODUCTION

1.1 Particle Physics

Particle physics studies the fundamental laws and constituents of the Universe. As such, it strives to answer questions humans have asked for all of recorded history. Many answers have been proposed over the millennia. Greek, Chinese, Japanese, and Hindu cultures all shared the idea that a few basic “elements” or “phases” constitute the world. In Western thought, the most famous of these are the four elements (Earth, Air, Fire, and Water) of ancient Greece.

During the past 400 years, the tools of science have developed and allowed us to provide rigorously testable answers to these ancient questions. In 1909, the famous gold foil experiment demonstrated that atoms contain a dense nucleus that is extremely small relative to the cloud of electrons surrounding it [4]. Later discoveries showed that nuclei are made of protons and neutrons bound together. However, in 1964 it was suggested by Feynman, Gell-Mann and Zweig that protons and neutrons are not fundamental either and are made of quarks [5]. So far all evidence suggests that quarks are fundamental. Electrons, which are another constituent of an atom, are also fundamental. They belong to the lepton family. Both quarks and leptons are presently thought of as the fundamental building blocks of nature.

We have also discovered that the Universe contains additional types of quarks and leptons beyond those present in ordinary matter. These particles interact via four fundamental interactions. The full mathematical description of these particles and forces, excluding gravity, is called the Standard Model.

In particle physics, particles are conventionally represented by symbols as in Table 1.1. Particle reactions and decays are written in a manner similar to that of chemistry and nuclear physics, with symbols on either side of an arrow (\rightarrow); the particles on the left side of the arrow are present before the decay or reaction; the particles on the right side are present afterward. For instance, $n \rightarrow pe^-\bar{\nu}_e$ represents a neutron decaying into a proton, electron, and anti-neutrino; this process is known as β decay.

To understand the symbolism and notation used for energy and momentum, we must recall that Einstein's Special Theory of Relativity relates rest mass (m), energy (E), and momentum (\vec{p}) through

$$E^2 = (|\vec{p}|c)^2 + (mc^2)^2 \Rightarrow m = \frac{1}{c^2} \sqrt{E^2 - (|\vec{p}|c)^2}. \quad (1.1)$$

If a particle is at rest ($|\vec{p}| = 0$), Equation 1.1 reduces to $E = mc^2 \Rightarrow m = \frac{E}{c^2}$. Note that rest mass is the mass of a particle in its own reference frame or any frame where it is at rest.

The standard unit of energy in particle physics is the electron volt (eV), which is the energy required to move an electron through an electric potential difference of 1 volt [6]. Mass is written in units of MeV/ c^2 or GeV/ c^2 , where 1 GeV/ c^2 = 1000 MeV/ c^2 = 10^9 eV/ c^2 , and momentum in units of MeV/ c or GeV/ c .

For every particle, an anti-particle exists with equal mass but opposite additive quantum numbers, such as electric charge. Anti-particles are denoted with either a

bar over the particle symbol (e.g. $\bar{\nu}_e$) or the opposite charge in a superscript (e.g. e^+). When a particle and anti-particle collide, they undergo a process called annihilation, in which both disappear and an equal amount of mass and energy appear in their place in accordance with the laws of quantum mechanics.

If the particle and anti-particle carry kinetic energy into the collision, that energy can also be transformed into other forms of mass and energy. In other words, colliding particles can transform energy into matter. This is the primary technique used in particle physics. High-energy collisions allow physicists to produce particles that are very unstable; they can only be studied if they are produced within or near a detector. Most are so unstable that they can only travel microscopic distances from appearance to decay and so cannot reach active detector elements. Information about these particles must be reconstructed using special relativity from measurements of those decay products that have long enough lifetimes to reach active detector elements.

1.2 The Standard Model of Particle Physics

The Standard Model (SM) is the theory of particle physics. It uses Noether's theorem, which is one of the most useful mathematical tools in physics. According to Noether's theorem, invariance with regard to a transformation leads to conservation laws. For example, in classical mechanics, invariance under translation through space leads to conservation of momentum.

In the SM, we assume that the Lagrangian describing particle behavior is invariant under local phase transformations. Invariance under these transformations is known as gauge invariance or gauge symmetry. The simplest set of such transformations are of the form $e^{i\alpha(x)}$, where $\alpha(x)$ is a real number that can vary arbitrarily over space

and time. This group of transformation is known as the gauge group $U(1)$. A more complicated gauge group ($SU(3)$) is also used in the standard model; it consists of a 3×3 matrix of local phase transformations.

The SM Lagrangian was derived by imposing gauge symmetry under these groups on previously extant effective theoretical descriptions of individual particles and interactions [7]. In the SM, gauge invariance leads to several conservation and other laws that are discussed below.

1.2.1 Fundamental Particles

Leptons			Quarks		
Particle	Q	Mass (MeV/ c^2)	Particle	Q	Mass (MeV/ c^2)
electron (e^-)	-1	0.511	up (u)	$+\frac{2}{3}$	1.5 - 3.0
e neutrino (ν_e)	0	$< 2 \times 10^{-6}$	down (d)	$-\frac{1}{3}$	3 - 7
muon (μ^-)	-1	105.7	charm (c)	$+\frac{2}{3}$	$(1.25 \pm 0.09) \times 10^3$
μ neutrino (ν_μ)	0	$< 2 \times 10^{-6}$	strange (s)	$-\frac{1}{3}$	95 ± 25
tau lepton (τ^-)	-1	1776.84 ± 0.17	top (t)	$+\frac{2}{3}$	$(174.2 \pm 3.3) \times 10^3$
τ neutrino (ν_τ)	0	$< 2 \times 10^{-6}$	bottom (b)	$-\frac{1}{3}$	$(4.70 \pm 0.07) \times 10^3$

Table 1.1: A summary of the known fundamental constituents of matter [1] is shown. The names of the particles and the symbols used to represent them are given. Also listed are their electric charge relative to the electric charge of the proton. Their mass, in units of MeV/ c^2 , is given. Notice all 12 particles are grouped into three generations of increasing mass and decreasing lifetime.

The SM contains six quarks and six leptons, which are the fundamental particles constituting ordinary matter and are summarized in Table 1.1. Both are found in

Particle	Charge	Mass (MeV/c ²)	Lifetime (sec.)
n (udd)	0	939.56536 ± 0.00008	885.7 ± 0.8
\bar{n} ($\bar{u}\bar{d}\bar{d}$)	0	939.56536 ± 0.00008	885.7 ± 0.8
$\Upsilon(4S)$ ($b\bar{b}$)	0	$(10.5794 \pm 0.0012) \times 10^3$	$\sim 10^{-23}$
B^+ ($\bar{b}u$)	+1	5297.0 ± 0.6	$(1.638 \pm 0.011) \times 10^{-12}$
B^0 ($\bar{b}d$)	0	5279.4 ± 0.5	$(1.530 \pm 0.009) \times 10^{-12}$
D^0 ($c\bar{u}$)	0	1864.5 ± 0.4	$(410.1 \pm 1.5) \times 10^{-15}$
K^+ ($u\bar{s}$)	+1	493.677 ± 0.016	$(1.2385 \pm 0.0024) \times 10^{-8}$
π^0 ($u\bar{u}/d\bar{d}$)	0	134.9766 ± 0.0006	$(8.4 \pm 0.6) \times 10^{-17}$
π^+ ($u\bar{d}$)	+1	139.57018 ± 0.00035	$(2.6033 \pm 0.0005) \times 10^{-8}$

Table 1.2: Properties of the composite particles relevant to this paper [1] are shown. The same units are used in this table as in Table 1.1. The n is the neutron and \bar{n} is the anti-neutron. The quark content of the particles is given in parentheses next to the particle symbol. Charge is given in units of the charge of the proton.

three “generations” (or “families”). All twelve particles have intrinsic angular momentum (spin) of $\frac{1}{2}\hbar$ and are therefore known as fermions.

Each lepton generation has one charged lepton and one neutral lepton, known as a neutrino. Experimental evidence has consistently shown that lepton number is conserved. All leptons have lepton number $L = 1$, and anti-leptons have $L = -1$. Lepton number conservation means that L cannot change during any reaction or decay. For instance, if no leptons are present at the beginning of a reaction, equal numbers of leptons and anti-leptons must be present at the end.

Each generation of quarks contains one quark with charge $+\frac{2}{3}$ and one with charge $-\frac{1}{3}$, in units of the charge of the proton. The number of possible quark combinations produces a veritable zoo of hadrons, a few of which are listed in Table 1.2.

We are made of the first generation of matter; our atoms are made of electrons, protons, and neutrons. Protons and neutrons are bound states of (uud) and (udd)

quark combinations, respectively. The other two generations are more massive and unstable, but they can be produced via the collisions of other particles at sufficiently high energy.

Such collisions occur naturally. Particles, mostly protons, are ejected into the cosmos in a very wide spectrum of energies from various sources; these are called cosmic rays. Collisions between cosmic rays and other particles, such as atomic nuclei in Earth's atmosphere, occur at energies spanning several orders of magnitude. Most cosmic rays carry a few GeV of energy or less; their flux decreases according to a power law with increasing energy.

The most energetic cosmic rays collide with the atmosphere with energies orders of magnitude higher than the energies produced in any device ever built by humans. That is one reason scientists are confident that the Large Hadron Collider will not produce destructive black holes. Though such energetic collisions are rare, cosmic rays collide with nuclei in Earth's atmosphere and other celestial bodies with sufficient frequency and energy that such black holes would have been produced naturally if they could be produced at the LHC [8].

To study particles produced in high energy collisions, we must surround the collision point with varied detectors and recording instruments. Cosmic ray collisions do not occur frequently enough to make this practical. Particle accelerators are built to produce collisions at sufficient frequencies and under sufficiently controlled conditions to allow for precision measurements.

Interaction	Mediator	Q	Spin	Mass (GeV/ c^2)	Relative Strength	
					10^{-18} m	3×10^{-17} m
Strong	gluon (g)	0	1	0	25	60
Electromagnetic	photon(γ)	0	1	0	1	1
Weak	W^\pm	± 1	1	80.40 ± 0.03	0.8	10^{-4}
	Z^0	0	1	91.188 ± 0.002		

Table 1.3: Mediators of the fundamental interactions. Charge (Q) is again in units of the charge of the proton. Spin is given in units of \hbar . Note that the weak force is unusual in having three mediators. The relative strengths of the interactions are given for interacting particles at two example distances.

1.2.2 Fundamental Interactions

Four fundamental interactions have been found in nature, through which all physical objects detect and interact with one another. Two have macroscopic ranges and we feel them in everyday life; they are electromagnetism and gravity. The other two, the weak and strong interactions, are effective only at subatomic distance scales and so have only been discovered and described in the twentieth century.

The SM accounts for three of the four fundamental interactions: weak, strong, and electromagnetic. Gravity is not included, but it is sufficiently weak on subatomic scales that it can be ignored in most situations. The relative strengths of the other three forces are given by their coupling constants in Table 1.3.

Each of these three interaction is mediated by a specific particle or set of particles. All mediators, also known as force carriers, are gauge bosons. Bosons have integer spin in units of \hbar . Mathematically, the gauge invariant groups in the standard model predict the existence of these mediators, which is why they are called *gauge* bosons. Particles interact by exchanging these mediators, which are listed in Table 1.3. The

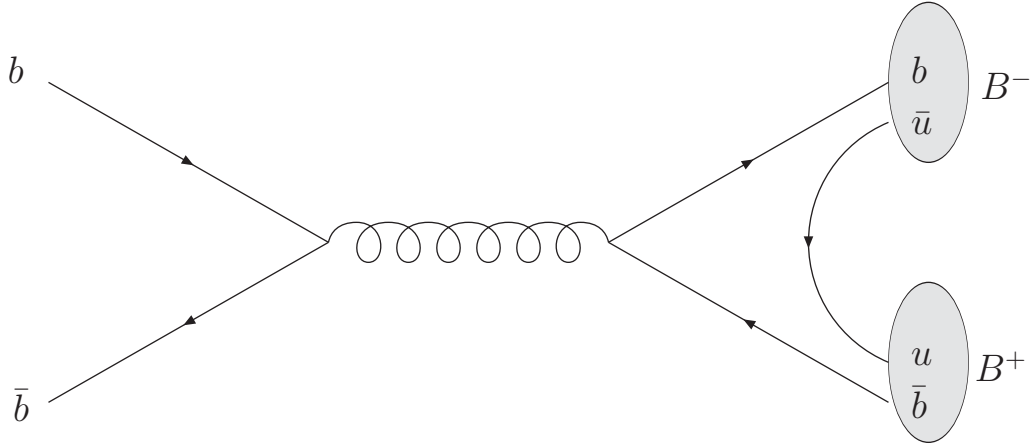


Figure 1.1: The strong interaction mediates the decay of a $b\bar{b}$ meson with sufficient mass to produce two charged B mesons. The helix represents a gluon that mediates the strong interaction. The interactions also causes a $u\bar{u}$ pair to form as the b and \bar{b} grow farther apart. This decay is of importance to many analyses at $BABAR$, including this one, because it is the primary decay mode of the $\Upsilon(4S)$.

force carriers are fundamental particles, so a complete list of all fundamental particles is the union of the force carriers, the quarks, and the leptons.

Interactions between particles are represented using Feynman diagrams. These are used to visually represent and calculate probabilities for various processes within the standard model.

Strong Interaction

The strong interaction is described by quantum chromodynamics (QCD). It is effective only over short ranges on the order of nuclear size, but at those ranges, it is the strongest of the four interactions.

The charge of the strong interaction is color. Any particle that interacts strongly has red (R), green (G), blue (B), anti-red (\bar{R}), anti-green (\bar{G}), or anti-blue (\bar{B}) color

charge. Some particles simultaneously carry more than one color. The term “color” is purely mnemonic; it has no relation to actual wavelengths of visible light.

The appropriately named gluon mediates the strong force. The gauge invariant SU(3) group reveals that we have eight massless gluons that carry two color charges and can thus interact with each other. Quarks interact by exchanging gluons. When a quark emits or absorbs a gluon, it changes color; the gluon carries the colors of the initial and final quarks. When quarks or gluons are pulled apart, an arbitrary number of gluon-gluon interactions can occur, making a “net” of gluons between them. This phenomenon is non-perturbative, which means that small changes in the system have significant consequences. This renders calculations very difficult.

This also means that the strength of the strong interaction *increases* with distance. So, when two quarks or gluons are pulled apart, the total energy stored in the gluon field increases until it is sufficient to form new quarks. Thus, the SM predicts that quarks will never be found in isolation. No confirmed experimental evidence of free quarks has been found.

Similarly, no color charged object can be found in isolation; therefore, any observable particle must be color neutral. This property is known as color confinement. This can be achieved by having one quark of each color (the baryons) or a quark and an antiquark that have a color and its anti-color (the mesons). Residual effects of the strong interaction also binds baryons together to form atomic nuclei. Particles containing five quarks (pentaquarks) have been postulated, but no conclusive evidence of their existence has been found.

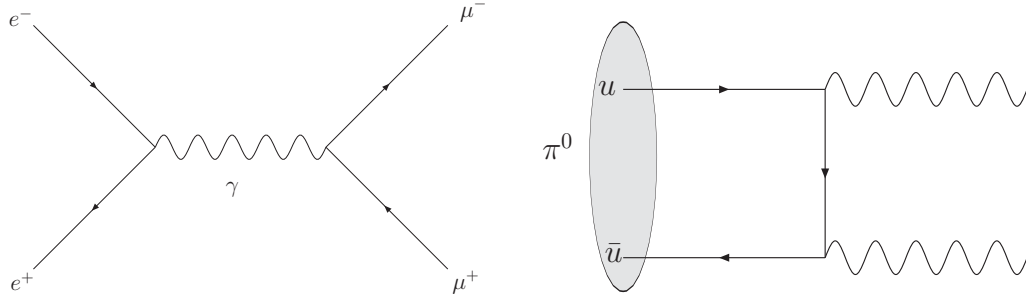


Figure 1.2: Two examples of the electromagnetic interaction are shown. At left is an example where an electron and positron collide to produce a photon, represented by the wavy lines, that then generates a $\mu^+\mu^-$ pair. This is one example of matter-antimatter annihilation producing new particles. At right is the Feynman diagram of the decay $\pi^0 \rightarrow \gamma\gamma$.

The strong interaction preserves the flavor of the quark; for example, a t quark cannot decay into a b or u quark via the strong interaction. An example of the strong interaction is shown in the Feynman diagram in Figure 1.1.

Electromagnetic Interaction

The theory describing electromagnetism is called quantum electrodynamics (QED); it employs the $U(1)$ gauge group. Of the three forces described by the SM, it is the only one with an infinite range. Consequently, it is the only one that we are conscious of in everyday life. Electromagnetism affects all electrically charged particles and electromagnetic radiation; in everyday contexts, it determines the chemical properties of atoms by governing the interactions of electrons with protons and each other.

Since it has a charge with only two values (positive and negative) and the mediator does not interact with itself, calculations involving the electromagnetic interaction are simpler than for the strong interaction. Charged particles interact by exchanging

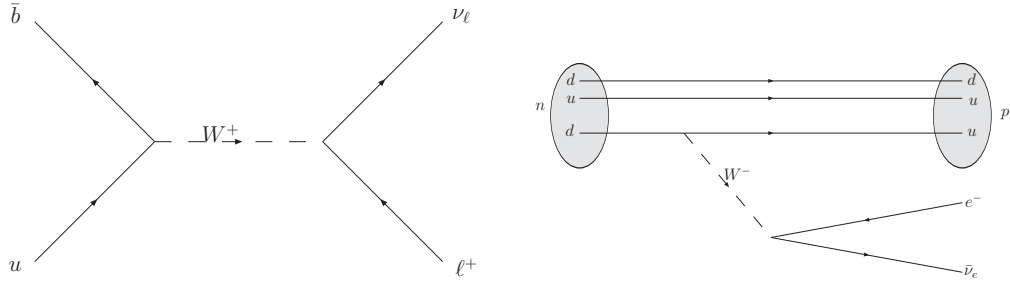


Figure 1.3: Two examples of the weak interaction are shown. At left is the decay studied in this analysis, the decay of a charged B meson via the weak interaction into a charged and a neutral lepton. At right is the decay of a neutron, which is also governed by the weak interaction.

photons. Since photons do not interact with each other, they can be free particles. As such, they constitute visible light and all other forms of electromagnetic radiation. Two examples of Feynman diagrams for the electromagnetic interaction are shown in Figure 1.2.

Weak Interaction

The weak interaction affects all known particles; however, it has intrinsic helicity preferences. Helicity is the dot product of linear momentum and spin. Negative helicity particles have anti-parallel linear momentum and spin. Positive helicity particles have parallel linear momentum and spin. The weak interaction interacts exclusively with particles that have negative helicity or anti-particles that have positive helicity. If a particle has one helicity in a given reference frame and it has mass, other reference frames exist in which it has the other helicity. One can imagine overtaking a particle so that its momentum in the observer's changes direction, and the helicity is

thus reversed. Massless particles travel at the speed of light, so they have the same helicity in all reference frames.

In addition to its helicity preference, the weak interaction is also unique because its force carriers are massive, unlike the photon or gluon. The Lagrangian that describes the weak interaction in the SM is invariant under the $SU(2)$ gauge group, which consists of two dimensional complex vectors of local phase transformations. The parameters of the Lagrangian are chosen such that they allow a process called spontaneous symmetry breaking to reveal the existence of massive W^\pm and Z^0 .

As its name suggests and as is quantified in Table 1.3, the weak interaction is weak compared to the other interactions. This relative weakness is not predicted by the SM; it is an empirically determined quantity. The exact masses of the force carriers and fundamental particles are also empirically determined and not predicted by the SM.

Its effects are so small that they are only noticeable when the strong and electromagnetic interactions cannot act. Neutrinos are neutral and not affected by the strong force, so they only interact weakly and via gravity. Therefore, any process involving neutrinos on the subatomic scale is sensitive to the weak interaction. Unfortunately, neutrinos react with matter so weakly and rarely that they are undetectable by *BABAR*. This makes my analysis challenging since the neutrinos are such an important physical component of this analysis and many analyses at *BABAR*.

Unlike the strong interaction, the weak interaction can change quarks from one flavor to another. The mechanism for this change is discussed in Section 1.2.3. The three decays for which I search in this analysis are all mediated by the weak interaction, as seen in Figure 1.3.

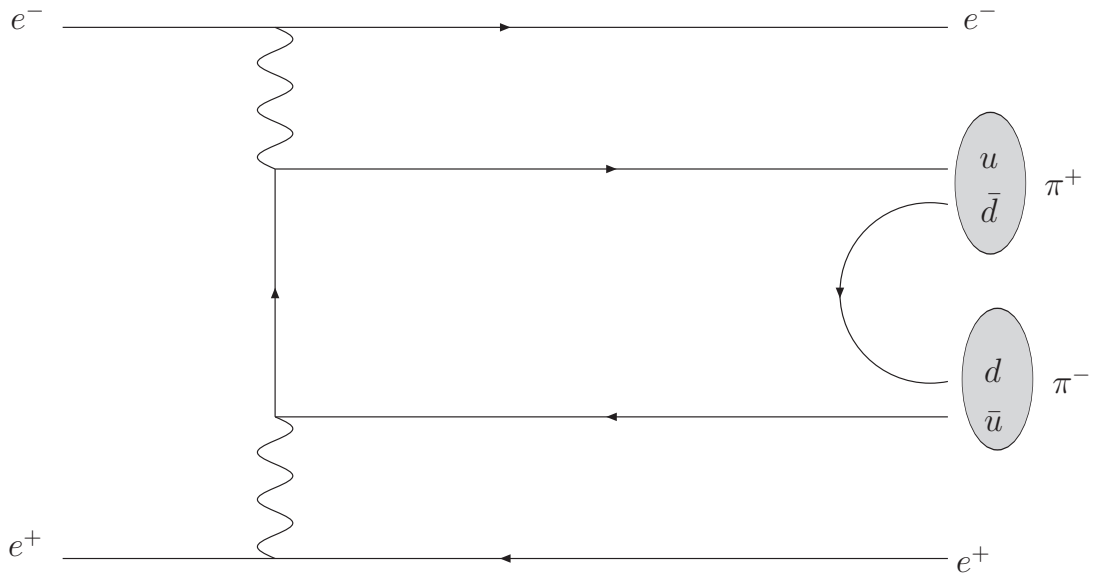


Figure 1.4: This two-photon fusion event is an example of a Feynman diagram that has both electromagnetic and strong interactions. Virtual photons emitted from the initial electron and positron produce a u and \bar{u} . The strong interaction produces another quark-antiquark pair to satisfy color confinement. The strong interaction also binds the quarks into two mesons. This process can produce background events when the hadrons produced by the fused photons are misreconstructed as a D^0 , one of the electrons is misidentified as a tag lepton and the other as the signal lepton.

Note that many decays and processes involve more than one interaction. For instance, the two photon fusion process shown in Figure 1.4 is mediated by the strong and electromagnetic interactions.

1.2.3 CKM Matrix

The weak interaction achieves generational change on the quarks because it couples to weak eigenstates d' , s' , and b' rather than mass eigenstates d , s , and b . In other words, the electromagnetic and strong interactions see as a d quark, the weak interaction sees as a linear combination of d' , s' , and b' . Conversely, what the weak

interaction sees as a d' , the other interactions see as a linear combination of $d, s,$ and b .

These linear combinations are summarized in the Cabibbo-Kobayashi-Maskawa (CKM) the matrix, which is shown in Equation 1.2.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.2)$$

$|V_{ij}|^2$ is the probability of the quark q_i being transformed via the weak interaction into the quark q_j . Assuming that only three generations of quarks exist, the CKM matrix must be unitary, otherwise the probabilities of quarks being in some states could be greater than 1. Mathematically, this means means $U^\dagger U = I$, where U^\dagger is the complex conjugate of the transpose of U and I is the identity matrix.

The CKM matrix has four free parameters: three real numbers and a complex phase. The current experimental ranges established for the absolute values of the entries in the matrix are shown in Equation 1.3 [1].

$$|V_{CKM}| = \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{pmatrix} \quad (1.3)$$

1.2.4 Unanswered Questions

During the past thirty years, the SM has become one of the most thoroughly experimentally verified theories in science. However, we know it is not complete because it does not solve certain problems. The SM cannot account for gravity. It does not predict the masses of the fundamental particles, the values of the CKM matrix elements, the relative strengths of the fundamental interactions, or several other parameters; those must be experimentally determined. It does not account for

Dark Matter and Dark Energy, which constitute approximately 96% of the Universe. Other questions also remain unanswered.

Many hypotheses have been proposed beyond the SM to solve some or all of these problems. Any such hypothesis must make predictions that are consistent with the vast array of experimental and observational results confirming the SM. Therefore, each new result that confirms the SM places constraints on any hypothesis beyond the SM. Constraining, eliminating, or supporting hypotheses beyond the SM is a necessary step toward achieving a more complete understanding of the most fundamental laws governing physical reality.

One particle predicted by SM, the Higgs boson (H^0), has not yet been found. Its discovery is the primary goal of the Large Hadron Collider. This analysis is not sensitive to any aspects of the SM Higgs; however, some hypotheses beyond the standard model postulate a charged Higgs (H^+) [9]. The possible existence of H^+ can be constrained by this measurement.

1.3 Notes To Reader

Throughout this analysis I will be taking charge conjugation as implicit. Charge conjugation refers to a particle's anti-particle analog. For example when I refer to studies of the decay $B^+ \rightarrow \tau^+ \nu_\tau$, this is meant to include the charge conjugate decay $B^- \rightarrow \tau^- \bar{\nu}_\tau$.

Also, to reduce clutter in equations, particle physics uses a set of units in which $\hbar = c = 1$. This means that speed becomes a unit-less fraction of the speed of light and mass and momentum are measured in unites of energy (e.g. GeV). Distances

are still measured in meters, and time is measured in units of distance.

$$c = 1 \Rightarrow 3 \times 10^8 \text{m/s} = 1 \Rightarrow 1\text{s} = 3.3 \times 10^{-9}\text{m}. \quad (1.4)$$

Unless otherwise noted, these units will be used for the remainder of this document.

CHAPTER 2

FULLY LEPTONIC DECAYS OF THE CHARGED B MESON

The purely leptonic decays $B^+ \rightarrow \ell^+ \nu$ ($\ell = e, \mu$ or τ) provide sensitivity to the Standard Model (SM) parameter V_{ub} , the CKM matrix element, and f_B , the meson decay constant that describes the overlap of the quark wave functions within the meson. In the SM, the decay proceeds via quark annihilation into a W^+ boson (Figure 2.1), and the branching fraction is given by

$$\mathcal{B}(B^+ \rightarrow \ell^+ \nu) = \frac{G_F^2 m_B m_\ell^2}{8\pi} \left(1 - \frac{m_\ell^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B, \quad (2.1)$$

where G_F is the Fermi constant, m_B and m_ℓ are the B meson and lepton masses, τ_B is the B meson lifetime. The branching fraction has a simple dependence on $|V_{ub}| \cdot f_B$.

In the SM context, observation of $B^+ \rightarrow \ell^+ \nu$ provides a direct measurement of f_B , as $|V_{ub}|$ is measured from semileptonic B -meson decays. Currently, it is the only good experimental environment in which f_B can be measured.

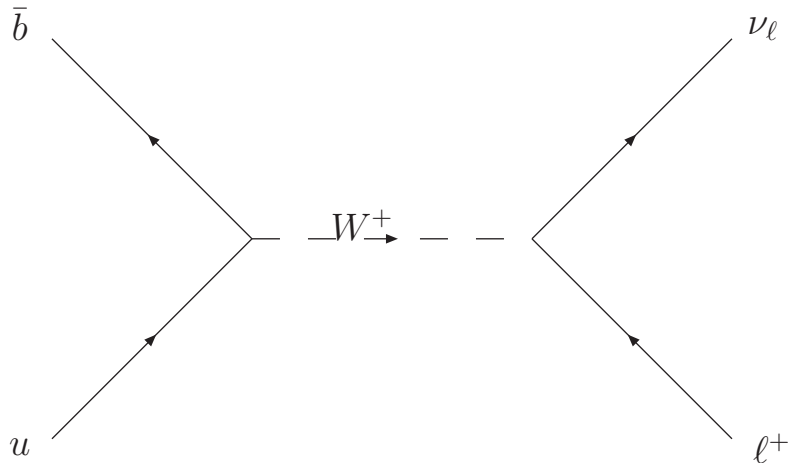


Figure 2.1: Feynman diagram for the purely leptonic B decay $B^+ \rightarrow \ell^+ \nu_\ell$.

2.1 Experimental Challenges

2.1.1 Neutrinos and Tagging

Conservation of charge and lepton number force the charged B decay to produce a neutrino when it produces a charged lepton. In the case of $B^+ \rightarrow \tau^+ \nu_\tau$, the τ lifetime is so short that it will decay while still inside of the detector, producing one or two additional neutrinos. Neutrinos are neutral leptons; therefore, they interact only via the weak interaction. This means that they have an extremely small probability of interacting with other matter. We can only detect particles that interact with the material in the detector, and the probability of neutrino interactions are so small that they are undetectable by *BABAR*. This means that information is invariably lost from these decays, which is a major experimental challenge. Specifically, we lose the energy and momentum of the neutrino, which prevents us from reconstructing the mass of the B^+ or τ . More creative methods are needed.

The technique used in this analysis for addressing this challenge takes advantage of the decay of the $\Upsilon(4S)$ system into two charged B mesons. We can reconstruct one of them from a decay mode that is well-understood and from which we can detect most or all of the daughter particles. This is called the tag B and will be denoted as such or as B^- for the remainder of this document.

If we have properly reconstructed the tag B , we know that what remains in the event must be from the other B , which we call the signal B . We search for our desired decays in the signal B .

2.1.2 Helicity Suppression

The decaying B^+ has a zero spin; conservation of angular momentum forces the spins of its two daughters to be aligned and anti-parallel. Conservation of momentum forces the linear momenta of the two daughters to also be in opposite directions in the B rest frame. The neutrino is constrained by its nearly zero mass and the weak interaction to always have a negative helicity.

So the positively charged lepton is forced to have positive helicity. Since positively charged leptons are antiparticles, positive helicity is more difficult to achieve for lighter particles, as discussed in §1.2.2. The dependence on lepton mass (m_ℓ^2) in Equation 2.1 arises from helicity suppression and heavily suppresses the rate for lighter leptons.

2.2 Theoretical Predictions

The Standard Model estimate of all three $B^+ \rightarrow \ell^+ \nu_\ell$ branching fractions are given in Table 2.1. The “Naive SM Prediction” is obtained by inserting $|V_{ub}| = (3.96 \pm 0.15_{-0.23}^{+0.20}) \times 10^{-3}$ [10]. The remaining values, except for f_B , are taken from the Particle Data Group (PDG) [1].

The latest values of f_B are from other measurements or limits on this branching fraction, and using those could bias our prediction. Therefore, we wish to use the best theoretical predictions for f_B . Similar overlap parameters exist for many mesons, such as the π^+ . Some early calculations predicted that these parameters would be proportional to the sum of the masses of the quarks within the meson [11, 12]. This would predict $f_B \approx 4.2 \text{ GeV}$. The best current predictions are obtained using lattice QCD calculations. I chose the value used by the PDG, which was calculated by a collaboration including Prof. Junko Shigemitsu of the Ohio State Physics Department, $f_B = 0.216 \pm 0.022 \text{ GeV}$ [13].

		$B^+ \rightarrow e^+ \nu_e$	$B^+ \rightarrow \mu^+ \nu_\mu$	$B^+ \rightarrow \tau^+ \nu_\tau$
Prediction	Naive SM	$(1.4 \pm 0.3) \times 10^{-11}$	$(5.8 \pm 1.3) \times 10^{-7}$	$(1.3 \pm 0.3) \times 10^{-4}$
	CKM Fitter [14]	$0.89^{+0.12}_{-0.09} \times 10^{-11}$	$(3.8^{+0.5}_{-0.4}) \times 10^{-7}$	$(0.93^{+0.12}_{-0.09}) \times 10^{-4}$
	UT Fitter [15]	-	-	$(0.86 \pm 0.16) \times 10^{-4}$
PDG Values [1]		$< 9.8 \times 10^{-7}$	$< 1.7 \times 10^{-6}$	$(1.4 \pm 0.4) \times 10^{-4}$
Inclusive Meas.	BABAR [16]	-	$< 1.3 \times 10^{-6}$	N/A
	Belle [17]	$< 9.8 \times 10^{-7}$	$< 1.7 \times 10^{-6}$	N/A
Hadronic Tag Meas.	BABAR	$< 5.2 \times 10^{-6}$ [18]	$< 5.6 \times 10^{-6}$ [18]	$(1.8^{+1.0}_{-0.9}) \times 10^{-4}$ [19]
	Belle	-	-	$(1.8 \pm 0.7) \times 10^{-4}$ [20]
Semilep. Tag Meas.	BABAR (my analysis)	$< 7.7 \times 10^{-6}$	$< 11 \times 10^{-6}$	$(1.8 \pm 0.8 \pm 0.1) \times 10^{-4}$
	Belle [21]	-	-	$(1.65^{+0.38+0.35}_{-0.37-0.37}) \times 10^{-4}$

Table 2.1: Existing and predicted upper limits (at 90% C.L.) and measurements of the branching fractions of the $B^+ \rightarrow \ell^+ \nu_\ell$ decay. The inputs to the Naive SM predictions are from [1], except $|V_{ub}|$ and f_B , for which we have included the latest values. CKM Fitter and UT Fitter are groups that collect and summarize the most current particle physics results and use them to predict unmeasured quantities, such as those in the table. The semileptonic tag measurements from BABAR are the results found in this analysis.

2.3 Previous Experimental Searches

Searches for the three $B^+ \rightarrow \ell^+ \nu_\ell$ modes have been ongoing since the early 1990's. The ALEPH collaboration set an upper limit, at the 90% confidence level, of $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) < 1.8 \times 10^{-3}$ in 1994 [22]. The CLEO II collaboration set the first upper limits of $\mathcal{B}(B^+ \rightarrow e^+ \nu_e) < 1.5 \times 10^{-5}$ and $\mathcal{B}(B^+ \rightarrow \mu^+ \nu_\mu) < 2.1 \times 10^{-5}$ in 1995 [23].

Two different tag B decay modes are used in *BABAR*, and both have been used to search for $B^+ \rightarrow \tau^+ \nu_\tau$. In previous *BABAR* analyses [18, 19], a fully reconstructed hadronic tag has been used to search for $B^+ \rightarrow \ell^+ \nu_\ell$ decay. For the hadronic tag, the reconstruction begins by reconstructing a D^0 or D^{*0} . Charged and neutral hadrons are added in an attempt to reconstruct the tag B^- . If a tag B^- can be reconstructed the remainder of each event is then searched for the signal mode.

This analysis is the first attempt at a semileptonic tagged measurement of $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$. Two previous searches for $B^+ \rightarrow \tau^+ \nu_\tau$ have been conducted at *BABAR* using the semileptonic tags. The first was conducted using $(231.8 \pm 2.6) \times 10^6$ $\Upsilon(4S)$ decays; it set an upper limit at 90% confidence level of 2.6×10^{-4} [24]. The second was conducted using $(383 \pm 4) \times 10^6$ $\Upsilon(4S)$ decays; it set an upper limit at 90% confidence level of 1.7×10^{-4} [25].

My analysis is an updated and expanded version of the analysis on the first 347 fb^{-1} of data. The current best upper limits and branching fractions are shown in Table 2.1.

2.4 Experimental Procedure

In any particle physics experiment, we can only detect a limited set of particles directly. The rest decay before they reach active detector elements. Detectable

particles include charged pions (π^\pm), kaons (K^\pm), electrons (e^\pm), muons (μ^\pm), and photons (γ). We are only able to measure a limited number of observables for these particles. These include charge, momentum (p), energy loss (dE/dx), velocity, and position over time. The velocity is calculated using the Cerenkov angle (θ_c). When a charged particle passes through transparent material with sufficient velocity, it emits Cerenkov light at an angle

$$\cos \theta_c = 1/n\beta, \tag{2.2}$$

where β is the fraction of the speed of light at which the particle is traveling.

We face a fundamental challenge of translating these observables into the physical quantities we want to measure. We need at least two of these quantities to identify a particle.

The information provided by the detector allows us to reconstruct the four vector of the particles we can directly detect. We can sum their four vectors and calculate the invariant mass of the sum, as defined in

$$P_i^\mu \equiv (E_i, \vec{p}_i) \tag{2.3}$$

$$M(12)^2 = (P_1^\mu + P_2^\mu) \cdot (P_1^\mu + P_2^\mu) \tag{2.4}$$

$$\Rightarrow M(12)^2 = m_1^2 + m_2^2 + 2(E_1 E_2 - \vec{p}_1 \cdot \vec{p}_2) \tag{2.5}$$

For this invariant mass to be accurate, we must correctly identify the detected particles. Each particle has a unique mass, so if the two particles in question were the only particles produced by a common source, the invariant mass uniquely identifies

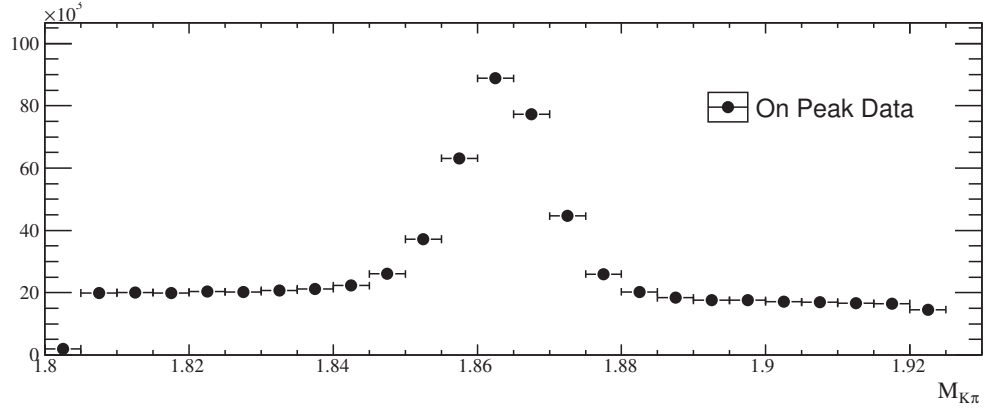


Figure 2.2: The reconstructed invariant mass of each pair of charged K and π mesons in each event in Runs 1-5 is shown. Notice the peak at the mass of the D^0 .

that source. For example, consider $D^0 \rightarrow K^-\pi^+$. We add the four vectors for all combinations of kaons and pions in the event and calculate $M_{K\pi}^2$ for each pair. If a given pair really came from a D^0 the invariant mass will be at or near the D^0 mass. If not, it will be at a random value. The resulting distribution will have a peak around the D^0 mass if real $D^0 \rightarrow K^-\pi^+$ are present in the data sample. Figure 2.2 shows such a peak in $M_{K\pi}$.

This procedure works well, in principle, for identifying various particles and decays; however, several types of events can result in fake signals. These are our backgrounds.

The simplest background is from combinatoric events. When combined, two unrelated tracks can, purely by chance, result in values of variables, such as mass, that are at or near the values where we expect our signal events to occur. These events usually have a flat or otherwise easily modeled background shape, such as the flat

background in the $M_{K\pi}$ in Figure 2.2. Such backgrounds can usually be predicted with precision using Monte Carlo simulation.

Another background source is incorrectly identified particles. Leptons can be misidentified as kaons or pions, pions can be misidentified as kaons, etc. Only a few percent of particles are misidentified, but when searching for rare decays, accounting for this is important.

The most pernicious background for this analysis are events in which some particles are not detected at all. The *BABAR* detector does not cover the entire region surrounding the interaction point. Detector elements are not present between the barrel and end cap or along the beam pipe. The decay $B^+ \rightarrow \overline{D^0} \ell^+ \overline{\nu}_\ell$ can be misreconstructed as $B^+ \rightarrow \ell^+ \nu_\ell$ if the components of the D^0 are lost down the beam pipe. This does not occur often, but since this decay occurs orders of magnitude more often than our target decays, it is our largest source of background.

CHAPTER 3

THE B FACTORY AT SLAC

The study of rare decays¹ of the B mesons is facilitated by B factories, which produce large numbers of B mesons. The factory at the Stanford Linear Accelerator Center² (SLAC) consists of the Linear Accelerator (“linac”), Positron Electron Project II (PEP-II) [26], and the *BABAR* detector [27]. A similar and competing B factory in Tsukuba, Japan consists of the KEK accelerator and the Belle detector [28].

3.1 The Linear Accelerator

The linac is housed in a straight tunnel that is two miles long. At the western end of the linac, electrons are released from a semiconductor using a polarized laser and organized into bunches. The electrons are accelerated to an energy of 10 MeV and circulated through damping rings that damp their motion perpendicular to the beam direction. Next, the electrons flow into the longest section of the linac, which is powered by klystrons above the ground. The klystrons generated microwave radiation that accelerates the electrons and positrons (e^+) to their final energies.

¹Multiple different physical phenomena and particles are studied at *BABAR*, but the search for decays that are rare or difficult to detect is the most important for this analysis.

²As of October 16, 2008, the laboratory was renamed the SLAC National Accelerator Laboratory, which is still commonly abbreviated as SLAC

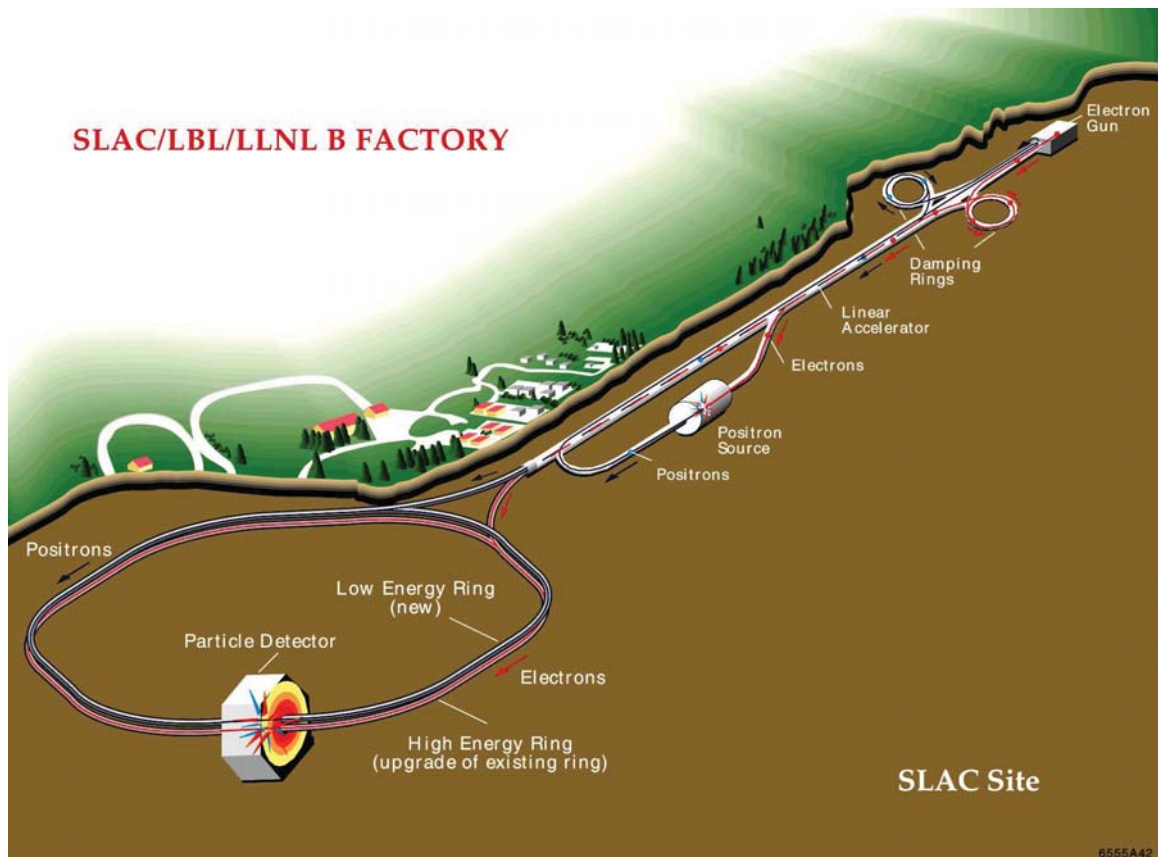


Figure 3.1: Overview of SLAC, including PEP-II and *BABAR*.

Positrons are produced by diverting a portion of the electrons from the linac and colliding them with a tungsten target. These collisions produce e^+e^- pairs; the positrons are directed along a separate line back to the beginning of the linac where they are accelerated by the klystrons to their final energies.

3.2 Positron Electron Project II

PEP-II [26] is primarily composed of two storage rings, with a circumference of 2200m each, at the end of the linac. The two rings are stacked on top of each other in the PEP-II tunnel and circulate the electrons and positrons in opposite directions. The beams are crossed within Interaction Region 2 (IR-2), where the *BABAR* detector is located. The electrons and positrons have energies in the detector rest frame of 8.9 and 3.1 GeV respectively. This produces asymmetric e^+e^- collisions at a center-of-mass (CM) energy equal to the mass of the $\Upsilon(4S)$ resonance, which is 10.58 GeV. The $\Upsilon(4S)$, which is a bound state of a b and \bar{b} quark, is thus given a Lorentz boost of $\beta\gamma = 0.56$ in the direction of the electron beam, where β and γ are defined in Equations 3.1 and 3.2, respectively.

$$\beta = \frac{v}{c} \tag{3.1}$$

$$\gamma = \frac{1}{1 - \beta^2} \tag{3.2}$$

In Equation 3.1, v is the speed of the CM system relative to the lab frame, and c is the speed of light. The asymmetry of the particle beam energies is critical for other analyses performed at *BABAR* but is not necessary for this one.

From Table 1.2, we see that the mass of the $\Upsilon(4S)$ is slightly above twice the mass of the B mesons. It decays into B^+B^- with a branching fraction of $(50.9 \pm 0.7)\%$ and into $B^0\bar{B}^0$ with a branching fraction of $(49.1 \pm 0.7)\%$. For all calculations in this analysis, we assume that the $\Upsilon(4S)$ decays exclusively into equal numbers of $B^0\bar{B}^0$ and B^+B^- pairs.

The e^+e^- collisions can produce multiple final states, including $b\bar{b}$. The production cross-sections are shown in Table 3.1. Since leptons are fundamental particles with no known substructure, they provide an environment with a smaller number of tracks and neutral clusters per collision than a hadron collider, such as the LHC. All of the energy of the electron and positron enters the collision, unlike hadron collisions where only the fraction of the energy carried by the colliding partons enters.

PEP-II was originally designed to deliver an instantaneous luminosity (\mathcal{L}) of $3 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. The luminosity is a measure of the number of collisions per unit time. A more convenient unit for measuring luminosity and other quantities involving areas of subatomic scales is the barn (b), which is the approximate cross sectional area of an atomic nucleus, $1 \text{ b} = 10^{-24} \text{ cm}^2$. The design luminosity can be restated as $\mathcal{L} = 3 \times 10^{-3} \text{ pb}^{-1} \text{ s}^{-1}$.

That goal was achieved in October 2000. The record instantaneous luminosity was $12.07 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$, achieved on August 16, 2007. The total data collected by the experiment in a given period of time is measured in inverse area. This thesis is based on all data (417.6 fb^{-1}) collected by *BABAR* at the $\Upsilon(4S)$ resonance and 42.2 fb^{-1} collected at a CM energy 40 MeV below the resonance. The data set contains $(458.9 \pm 5.1) \times 10^6$ $\Upsilon(4S)$ decays. PEP-II ceased operations on April 7, 2008, so this document is based on the full dataset recored by *BABAR*.

$e^+e^- \rightarrow$	$\sigma(\text{nb})$
$b\bar{b}$	1.10
$c\bar{c}$	1.30 [2]
$u\bar{u} + d\bar{d} + s\bar{s}$	2.09 [2]
$\tau^+\tau^-$	0.92 [30]
$\mu^+\mu^-$	1.15 [30]
e^+e^-	~ 40 [2]

Table 3.1: Production cross-sections in the PEP-II e^+e^- collider at 10.58 GeV center-of-mass.

The probability of any two particles interacting can be expressed in terms of a cross section, which has units of area and is represented by the symbol σ . The units are from the classical mechanical picture of two solid objects physically colliding.

In general, the luminosity and cross section are related by

$$N = \sigma\mathcal{L}, \tag{3.3}$$

where N is the number of occurrences of a given final state. For instance, for the interaction $e^+e^- \rightarrow \mu^+\mu^-$, N is the number of $\mu^+\mu^-$ pairs produced. A finite set of final states are allowed for e^+e^- collisions in PEP-II. Those states and the cross-sections for them, are listed in Table 3.1.

The cross-section for $e^+e^- \rightarrow b\bar{b}$ is determined using Equation 3.3. The number of $\Upsilon(4S)$ decays is obtained from data [29]. We simply divide that number by the total luminosity to obtain the cross-section, assuming all $b\bar{b}$ pairs are produced in the $\Upsilon(4S)$ resonance.

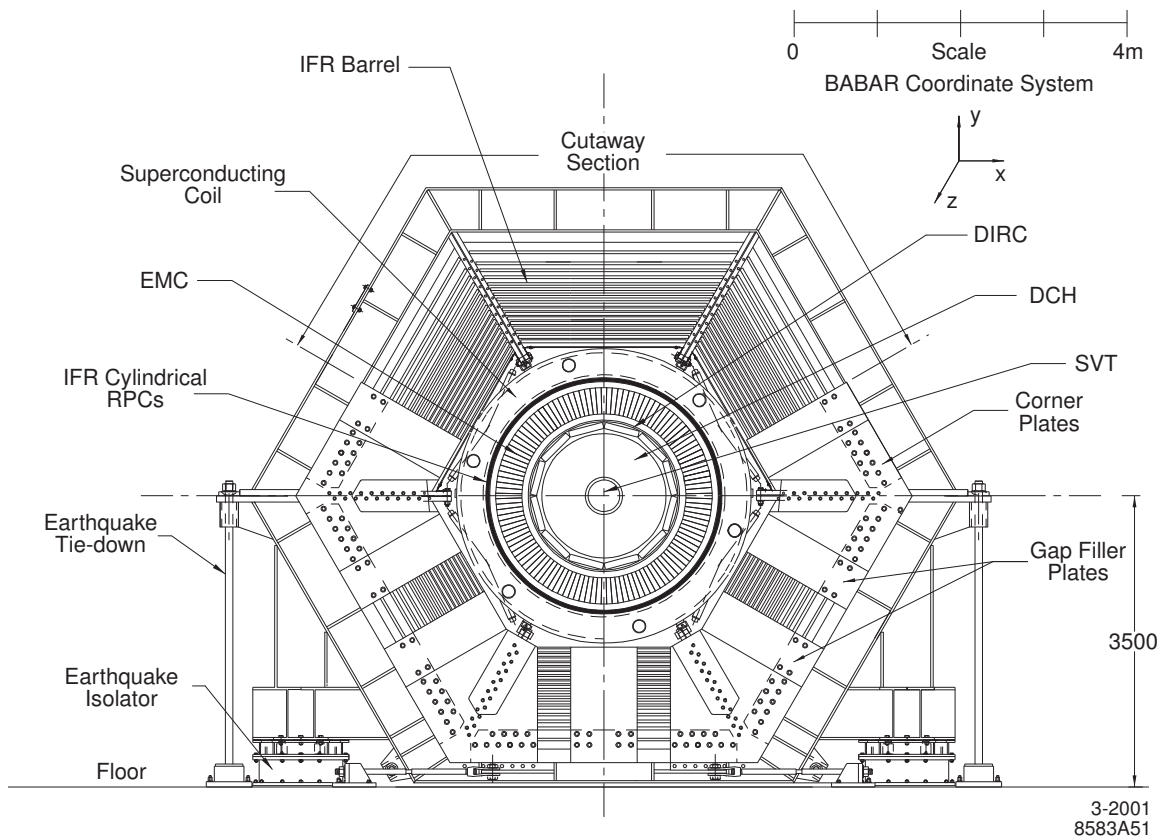


Figure 3.2: Cross-section of the *BABAR* detector, viewing in the direction of the positron beam. All distance measurements are in mm.

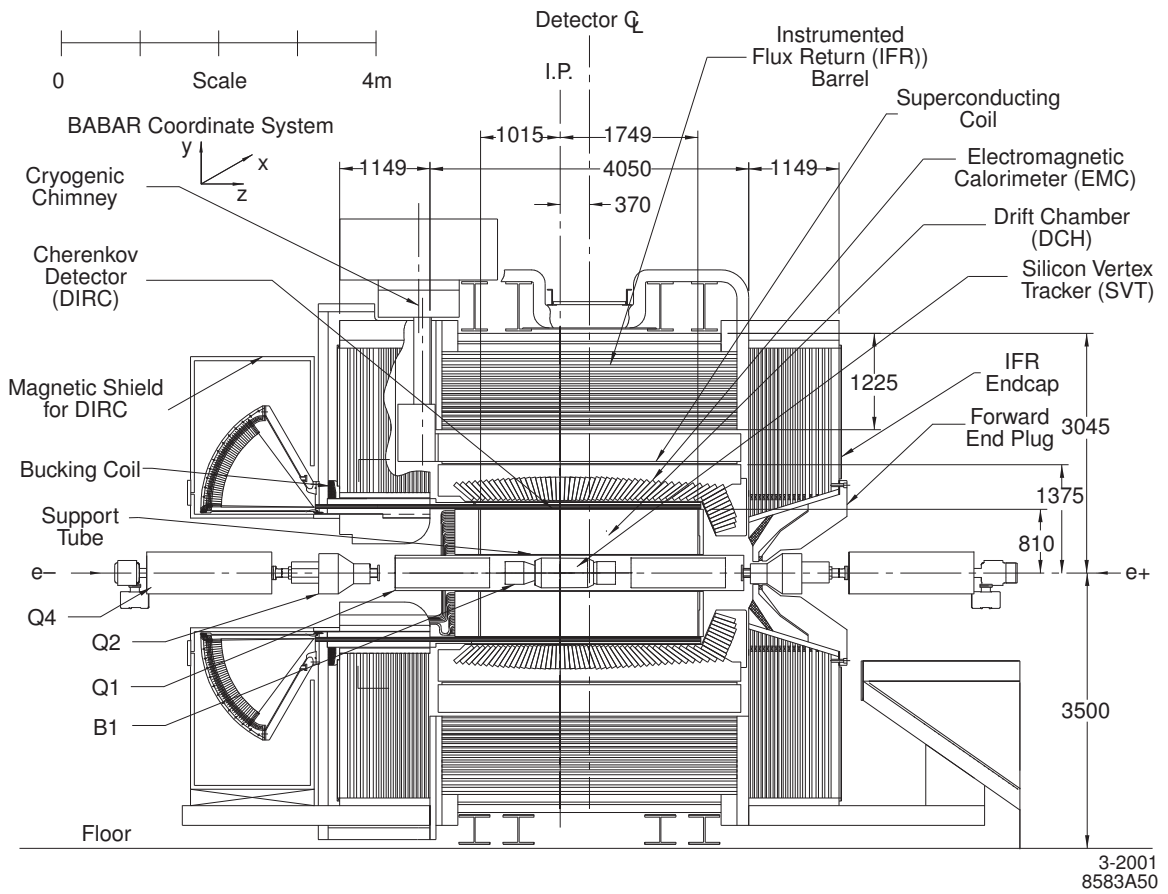


Figure 3.3: Cross-section of the *BABAR* detector, side view. All distance measurements are in mm.

3.3 The *BABAR* Detector

The detector was built to primarily study the properties of the B mesons, which contain either b or \bar{b} quarks [27]. One reads \bar{b} as “bee-bar,” hence the name *BABAR*.

The B meson is very short lived and decays very close to the collision point. To correctly reconstruct a decay, the particle trajectories and momenta need to be precisely measured. Moreover, accurate particle identification needs to be performed to differentiate between various types of charged tracks and neutral clusters. *BABAR* was constructed to satisfy these requirements. It is a typical particle physics detector composed of several components, which can be seen in Figures 3.2 and 3.3.

The Silicon Vertex Tracker (SVT) is located closest to the collision point and provides precision measurement of angles and positions of charged tracks. The Drift Chamber surrounds the SVT; it uses wires at high-voltage in a thin gas to provide particle identification and momentum measurements. The Detector of Internally Reflected Cerenkov light (DIRC) uses the Cerenkov light produced by charged particles to measure their velocity. The Electromagnetic Calorimeter (EMC) uses scintillating crystals to detect the energies of charged and neutral particles, including photons. The Solenoidal Magnet provides a 1.5 T magnetic field, which causes charged particles to curve, allowing measurement of their momentum. The Instrumented Flux Return (IFR) is placed between the steel plates that return the magnetic flux from the solenoid. It detects highly penetrating tracks, which are generated by muons and neutral hadrons.

A standard coordinate system as been established for *BABAR*, with the nominal collision point, also known as the interaction point (IP), as the origin. As seen in Figure 3.3, the z -axis is parallel to the long-axis of the detector, coincident with

the beams. Positive z is the direction of travel for the electrons. The y axis is perpendicular to z , with positive defined as vertical. The x -axis is defined with respect to y and z as in a standard right-handed coordinate system. Positions in the detector are usually described with spherical coordinates. The angle θ is the angle with respect to the z axis, which has the domain 0 to π radians (0 to 180°), corresponding to the positive and negative z directions respectively. The angle ϕ is measured with respect to the y -axis.

3.3.1 Silicon Vertex Tracker

The SVT is shown in Figures 3.4 and 3.5. It consists of the detector elements closest to the IP and surrounds the beam pipe, which has a radius of 27.8 mm. The sensitive elements of the SVT are 300 μm thick double-sided silicon strip detectors, which are arranged on 52 sensors. The sensors come in six different types, varying in size from $43 \times 42 \text{ mm}^2$ to $68 \times 53 \text{ mm}^2$. One side of each sensor consists of z strips that are transverse to the beam axis and measure the z position of charged tracks passing through the SVT. The other side consists of ϕ strips that are orthogonal to the z strips and measure the ϕ position of the tracks.

The detector elements are organized into five layers. The innermost three have radii of 32, 40, and 54 mm; their measurements are crucial to determining if several tracks have a common origin (called a “vertex”). The outer two layers have radii in the range 91-144 mm and provide high-precision tracking information, which links the trajectories to the data from the DCH. The electronic signals from the strips are transmitted to the readout electronics via more than 150,000 channels.

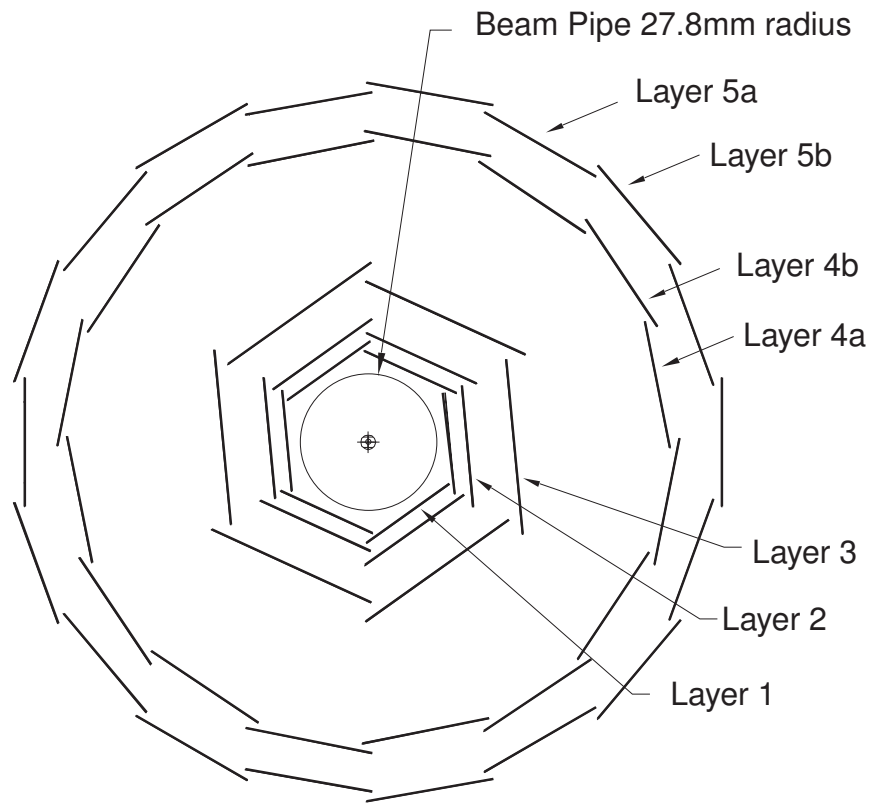


Figure 3.4: A schematic cross-section of the SVT, looking down the beam line, is shown.

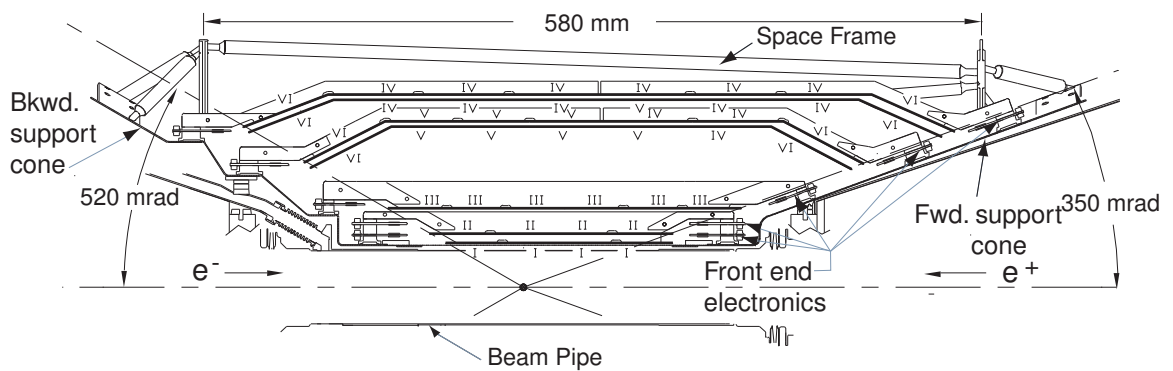


Figure 3.5: Schematic cross-section of the SVT, viewed from the side

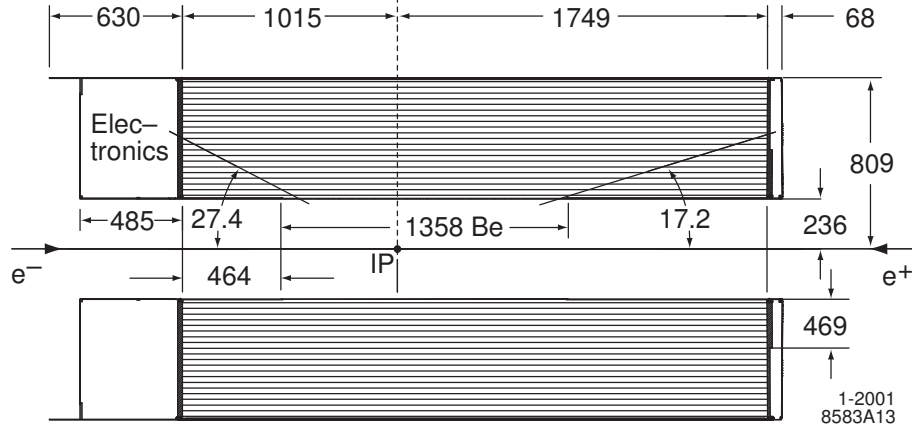


Figure 3.6: A schematic cross-section of the Drift Chamber is shown.

The SVT is sensitive to tracks in the region $150.2^\circ > \theta > 20.1^\circ$. It has a vertex resolution of better than $30 \mu\text{m}$ along the z -axis and $\approx 100 \mu\text{m}$ in the $x - y$ plane. For most tracks, the SVT and DCH are used together to exact tracking and momentum information. However, the DCH cannot detect tracks with low momentum transverse to the beams ($p_T < 120 \text{ MeV}$), so the SVT serves as a standalone detector. Since such tracks are important to the reconstruction of B mesons, the SVT was designed to have good sensitivity into the low p_T region.

Being so close to the interaction point, the SVT was exposed to a large amount of ionizing radiation over its lifetime. To ensure proper operation of the SVT, radiation-hard electronics is used in the detector.

3.3.2 Drift Chamber

The Drift Chamber (DCH) surrounds the SVT, and its primary purpose is to efficiently detect and measure the momentum and other kinematic variables of charged

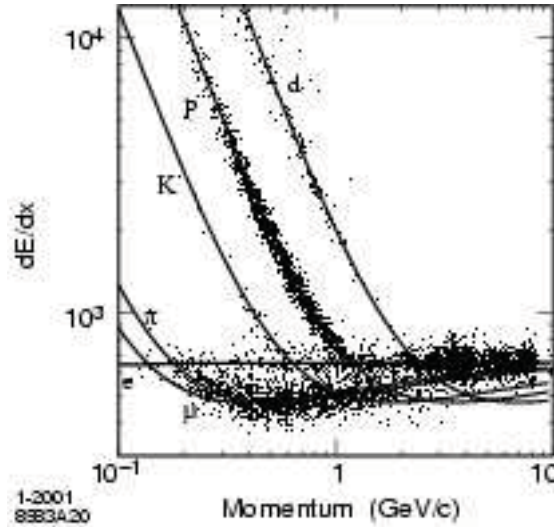
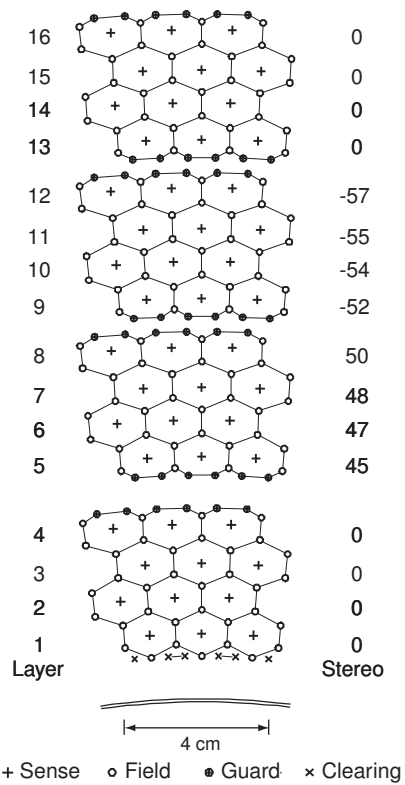


Figure 3.7: Performance of the DCH separating particles of different types using dE/dx as a function of track momentum. Protons are represented as p and deuterons, which are a heavy hydrogen nucleus consisting of a proton and neutron, are denoted d .

particles. Some particles do not decay until they are past the SVT and within the DCH; in these cases, the tracks from the charged daughter particles are measured by the DCH without help from the SVT.

The chamber is physically a 2.8 m long cylinder with an inner radius of 23.6 cm and an outer radius of 80.9 cm. Because of the asymmetric-energy e^+e^- collisions, the center of the DCH is shifted by 37 cm from the interaction point along the electron beam direction. As shown in Figure 3.6, the DCH contains 7104 sense wires that each form the center of a drift cell; these are arranged in 40 layers. Distance information along the z -axis is obtained by placing wires in 24 of the layers at small angles with respect to the z -axis. The arrangement of the wires and cells is shown in Figure 3.8. The sense wires are composed of a tungsten-rhenium alloy and are $20 \mu\text{m}$ in diameter.



1-2001
8583A14

Figure 3.8: A schematic view of the wires and cells within a small section of the DCH is shown.

They have a high voltage (1960 V) applied. The chamber is filled with a gas mixture of 80% helium and 20% isobutane.

When a charged particle passes through the chamber, it can remove electrons from atoms in the gas. Those electrons are attracted to the sense wires. While drifting toward the wire, they knock loose more electrons, creating a cascade, which causes an electronic signal in the sense wire. The time between the e^+e^- collision and the arrival of the signal can be measured with high precision; this localizes the source of the signal to a narrow ring around the wire. These data are used to reconstruct ionization trails through the chamber. Different particles lose different amounts of energy to ionization (dE/dx); this allows the chamber to differentiate between particles up to a momentum of approximately 1 GeV, as seen in Figure 3.7.

A track with $p_T > 180$ MeV that enters the chamber with $17.2^\circ < \theta < 152.6^\circ$ passes through at least half of the layers while curving due to the magnetic field. The measurement of curvature provides a precise measurement of the transverse momentum of the particle. For tracks with sufficiently high p_T , the DCH has a tracking efficiency of approximately 98%. If a particle passes through the SVT and DCH, tracking efficiency is nearly 100%.

3.3.3 Detector of Internally Reflected Cerenkov light

The DIRC surrounds the drift chamber and uses Cerenkov radiation to identify particles. It is composed of a set of fused silica bars around the DCH, a larger water filled region (called the standoff box) at the rear of the detector, and photomultiplier tubes surrounding the standoff box. A schematic of these components is shown in Figure 3.9.

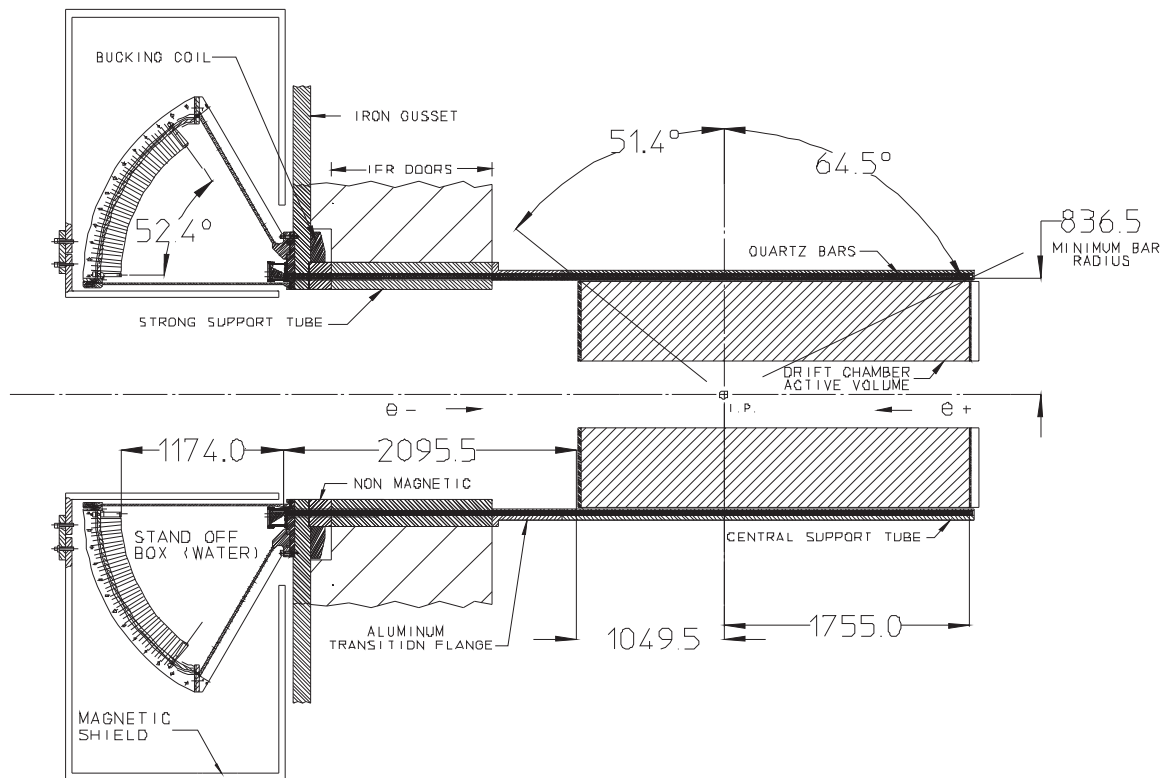


Figure 3.9: Schematic view of the DIRC, viewed from the side.

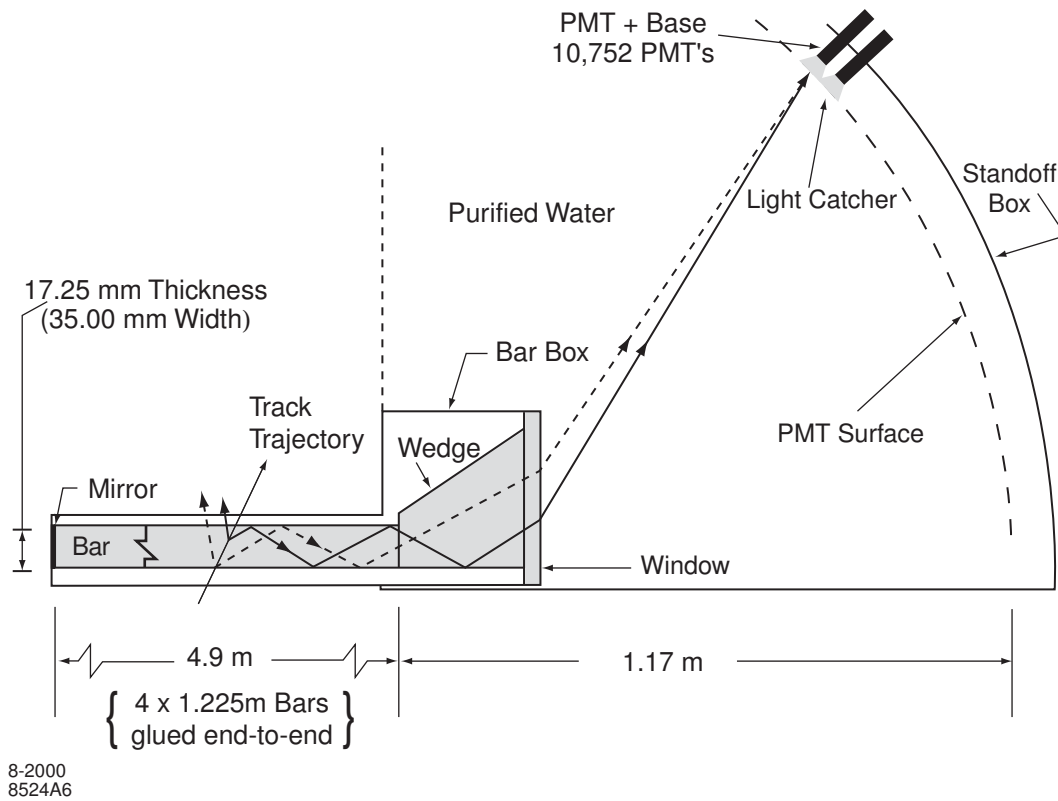


Figure 3.10: Schematic view of a single quartz bar and readout within the DIRC.

The fused silica has index of refraction $n = 1.473$. When a charged particle passes through the bar with sufficient velocity, it emits Cerenkov light at an angle

$$\cos \theta_c = 1/n\beta, \tag{3.4}$$

where β is the fraction of the speed of light at which the particle is traveling. This light is transported down the bar into the standoff box by total internal reflection. The PMTs detect the light and allow the Cerenkov angle and particle velocity to be measured. Once the velocity is known, the mass (and thus the identity) of the particle can be calculated using momentum information from the DCH.

The DIRC is the primary particle identification system for *BABAR* for particles with high enough momentum to render the DCH incapable of effective identification. The DIRC can identify five types of particles: e , μ , π , K , and p .

3.3.4 Electromagnetic Calorimeter

The sub-detectors discussed in previous sections are capable of detecting charged particles. The electromagnetic calorimeter (EMC) can detect charged and neutral particles, including photons, over a wide energy range from 20 MeV to 9 GeV. The active detector elements are 6580 thallium-doped CsI crystals. 5880 are located in a cylindrical barrel and 940 are in the conical forward end cap, as shown in Figure 3.11.

The crystals are scintillators; when a particle passes through them, it deposits energy in the form of an electromagnetic shower. The crystals convert the deposited energy into light, which is internally reflected to photo-diodes for measurement, as shown in Figure 3.12. The crystals yield approximately 50,000 photons per MeV of energy deposited. Particles usually deposit energy in several adjacent crystals,

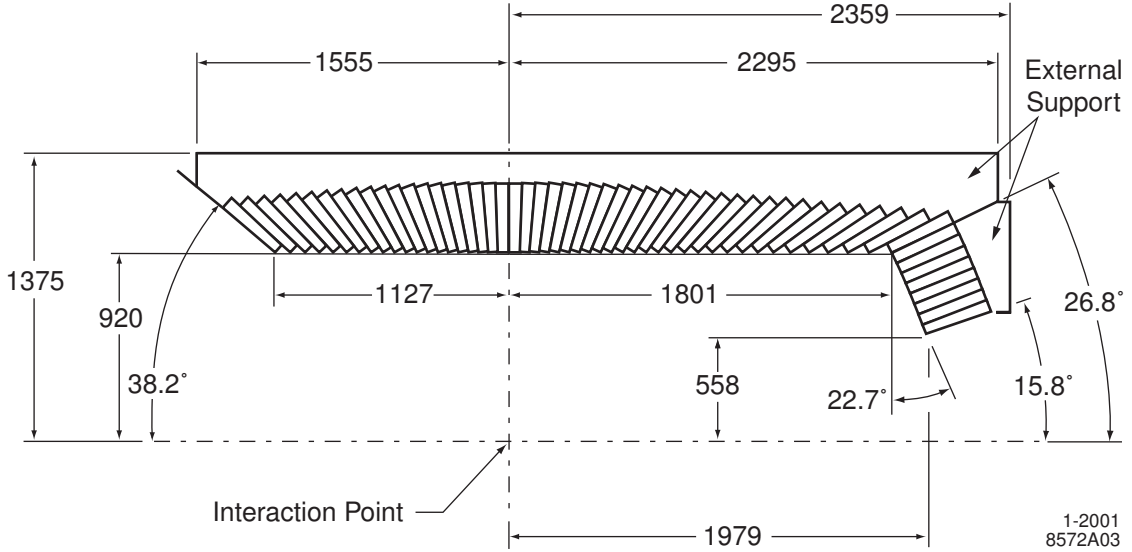


Figure 3.11: Schematic view of the upper half of the EMC, viewed from the side.

forming a cluster. The “primary” crystal in a cluster is required to have at least 10 MeV of energy. Surrounding crystals are considered part of the cluster if their energy exceeds 1 MeV or if they are contiguous (including by corners) with a crystal with an energy deposit of at least 3 MeV.

Reconstruction software attempts to match each cluster in the EMC with a charged track; if the match is successful, the EMC information is incorporated into the putative particle that created the track. If not, the cluster is considered to have come from a neutral particle. An average of 16 clusters are detected per hadronic event, of which 6 are successfully associated with a charged track.

The EMC is very useful for identifying electrons, which is particularly important for the two decay modes that produce electrons in this analysis, *i.e.* $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and $B^+ \rightarrow e^+ \nu_e$. The identification is based on shower energy, lateral shower moments, and track momentum. Since electrons have a much lower mass than any other charged

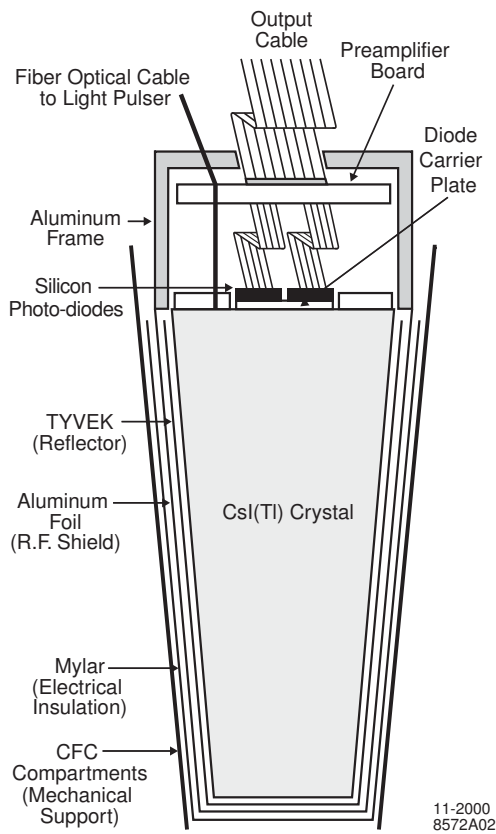


Figure 3.12: Schematic view of a single crystal and readout within the EMC.

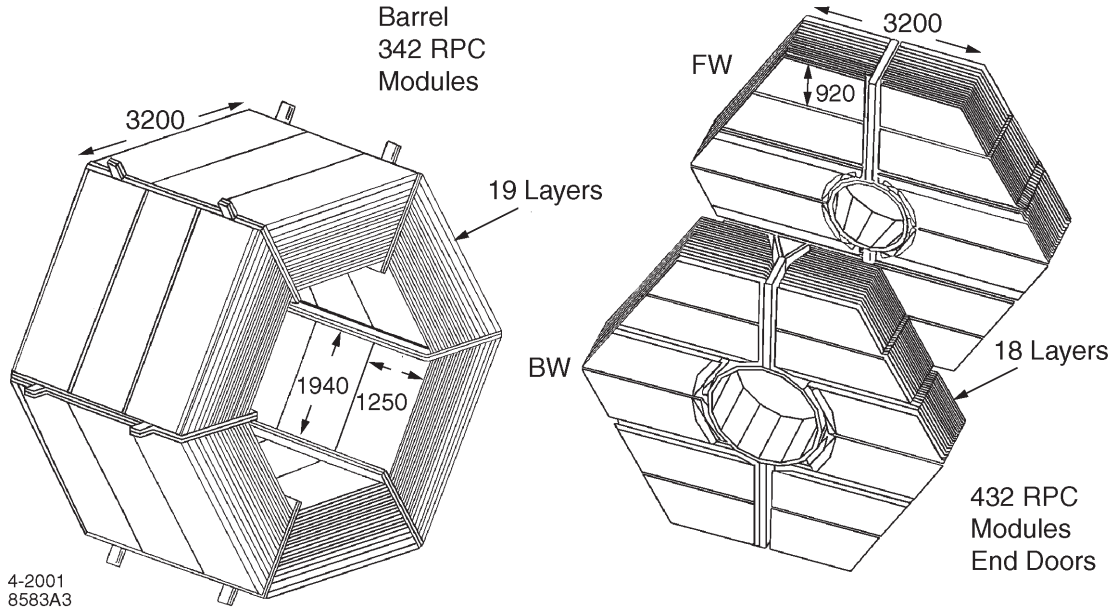


Figure 3.13: Schematic view of the IFR, including barrel and end caps.

particle, the most important variable for identifying them in the EMC is the ratio of energy to track momentum E/p , which is usually close to 1 for electrons and much lower for other particles.

3.3.5 Instrumented Flux Return

The outermost part of the detector is the Instrumented Flux Return (IFR); it covers the polar angle range $15^\circ < \theta < 157^\circ$. It surrounds the solenoidal magnet and is designed to detect particles that penetrate through all other parts of the detector. These particles are primarily muons and neutral hadrons, including neutrons. The IFR is composed of interleaved steel plates and detector elements organized into a hexagonal barrel and two end caps, as shown in Figure 3.13.

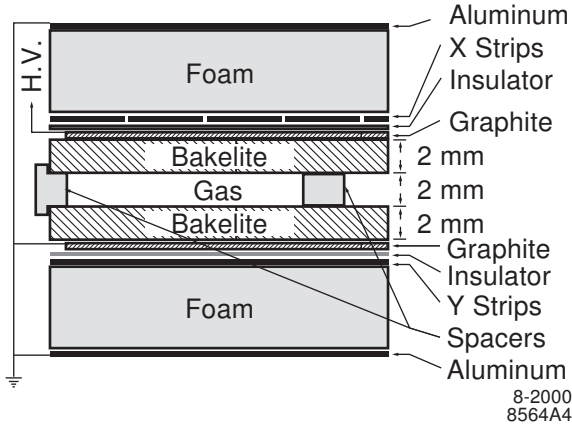


Figure 3.14: Schematic view of a signal planar resistive plate chamber (RPC), which was part of the original IFR.

Until the summer of 2004, the active detecting part of the IFR consisted of 806 resistive plate chambers (RPC), a schematic of which is shown in Figure 3.14, installed in the spaces between the steel plates. Charged particles passing through the RPC ionizes gas, and the resulting discharge is detected by electrodes on the RPC.

Limited Streamer Tubes

Deteriorating efficiencies of a large fraction of the IFR were detected from nearly the beginning of the experiment, as shown in Figure 3.15. The collaboration chose to replace the RPCs in the barrel with a combination of Limited Streamer Tubes (LSTs) and copper z-planes. Of the 18 layers of RPCs, 12 were replaced with this combination; the remaining layers were replaced with brass absorbing layers [31, 32].

The design and operating principles of the LSTs are quite simple. The tubes are made of PVC and divided into 7 or 8 cells. Each cell is 17 mm wide, 15 mm high and 380 mm long; a cross section through a tube is shown in Figure 3.16. In the middle

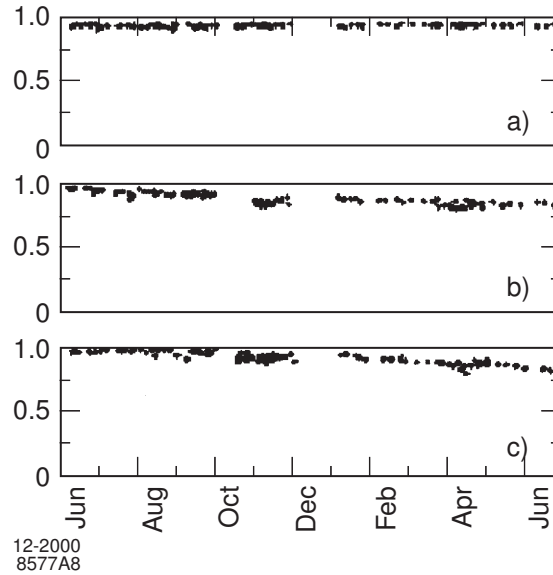


Figure 3.15: Efficiency history for 12 months starting in June 1999 for RPC modules showing different performance: a) highly efficient and stable; b) continuous slow decrease in efficiency; c) more recent, faster decrease in efficiency.

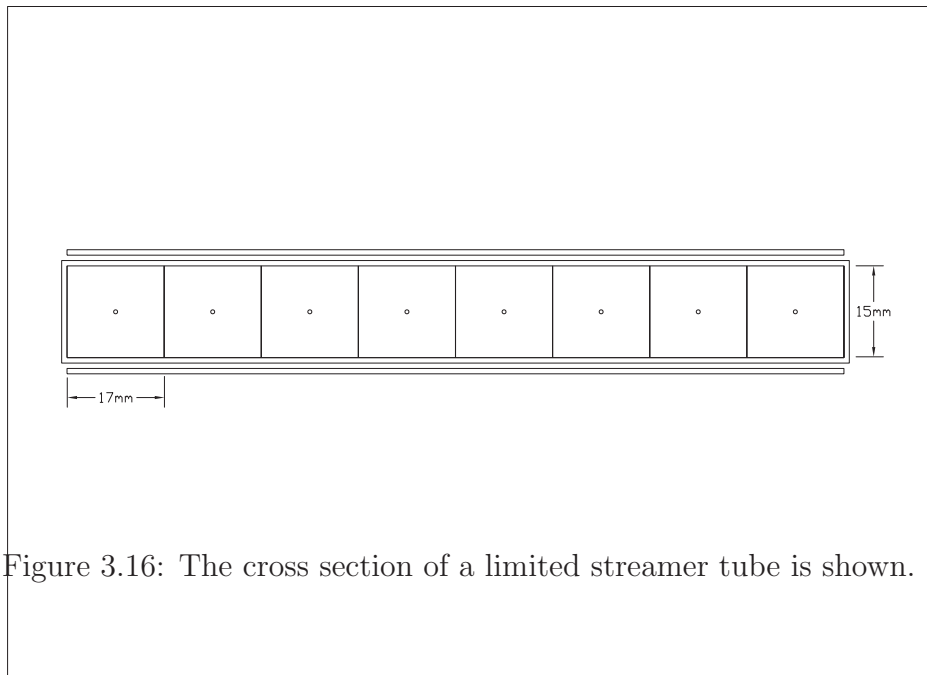


Figure 3.16: The cross section of a limited streamer tube is shown.

of each cell, a gold-plated anode wire is clamped. The interior walls of the tube are painted with a water-based graphite paint and kept at ground potential. The anode, during normal operation, is kept at 5500 V. The z-planes are 4m long and up to 3.8 m wide; each consists of 96 copper strips that are 35 mm wide and separated by 2 mm.

Each LST cell is filled with a gaseous mixture of CO₂, argon, and iso-butane. When a charged particle passes through a cell, it ionizes a few atoms in the gas. The freed electrons are attracted to the anode; as they accelerate towards the wire, they ionize other atoms, causing a cascade called a streamer. The streamer is detected via the anode; it also induces a charge on the z-plane. The wires and planes are oriented at right angles to each other, with the wires providing the ϕ coordinate of the track and the z-planes providing the z coordinate.

The barrel of the detector is a hexagonal prism divided into six sextants. The RPCs were replaced in two of the sextants during the summer of 2004. The remaining four sextants were replaced during August through December of 2006.

We in the Ohio State group played a critical role in the assembly, testing and monitoring of the tubes. Each cell of a tube was scanned with a radioactive source to ensure a suitably low current ($< 1\mu A$). I designed the apparatus that held and scanned the source above the tubes. I also designed the chassis of the high voltage power supplies to keep the tubes at their operating voltage.

The high voltage power supplies for the tubes were designed, tested, and built at Ohio State. Each power supply had a variable numeric label (1-25) and a permanent name. From August of 2005 to December of 2006, I made weekly measurements of the calibration voltage on each supply.

Although muons are identified by the IFR, the information from other detector components are necessary for a complete reconstruction of the particle trajectory. Muon candidates in the IFR are required to be reconstructed by the tracking system and have energy loss consistent with a muon in the EMC.

3.4 Particle Identification

3.4.1 Identification of Individual Particles

The association of charged tracks with actual particles is known as particle identification (PID), and accurate PID is critical to any particle physics analysis. PID is accomplished using information from the SVT, DCH, DIRC, and in the case of muons, the IFR. The DIRC returns the Cerenkov angle (Equation 3.4, and the SVT and DCH measure the energy loss over distance ($\frac{dE}{dx}$) and momentum.

For the condition at *BABAR*, the energy loss can be approximated by the Bethe-Bloch equation:

$$-\frac{dE}{dx} = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left(\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right). \quad (3.5)$$

In this complicated equation, K is a constant, ze is the charge of the incident particle, Z is the atomic number of the absorber, A is the atomic mass of the absorber, T_{\max} is the maximum kinetic energy that can be imparted to a single electron in the absorber during a collision, and I is the mean excitation energy. Equation 3.5 is a good approximation for all particles that are stable enough to reach active detector elements, except for the electron.

In the SVT and DCH, energy loss is modeled by the expressions in Equations 3.6 and 3.7.

$$\left(-\frac{dE}{dx}\right)_{\text{SVT}} = b_1 \beta^{b_2} (\beta\gamma)^{b_3} \quad (3.6)$$

$$\left(-\frac{dE}{dx}\right)_{\text{DCH}} = \frac{a_1}{\beta'} (a_2 - \beta' - \ln(a_3 + (\beta\gamma)^{a_4})), \quad \beta' = \left(\frac{p}{E}\right)^{a_5} \quad (3.7)$$

In these expressions, a_i and b_i are constant parameters determined by the SVT and DCH working groups.

For each charged track, energy loss is calculated in the SVT and DCH as shown above. Its position in Figure 3.7 and a similar plane for the DCH is calculated using momentum information from the DCH.

Multiple particle hypotheses are tested for each charged track. For each hypothesis, a pull is calculated, which is defined as

$$\text{pull} = \frac{\frac{dE}{dx}_{\text{measured}} - \frac{dE}{dx}_{\text{Bethe-Bloch}}}{\sigma} \quad (3.8)$$

The error (σ) is defined separately for the DCH and SVT. The pull distributions for both the SVT and DCH are obtained using pure samples of each type of measured particle. These are fit with a Gaussian. The Gaussian serves as the Probability Density Function (PDF) for the pull distribution, and consequently that a particular particle hypothesis is correct.

When a candidate of unknown type is detected, its pull for each possible particle type is measured. The pull is then related to the PDF obtained using control samples for each type of particle. Using the pull value of the PDF a likelihood value is determined for each particle hypothesis. The DIRC likelihood is quantified using a

List	Requirements
GoodTracksVeryLoose	$ d_0 < 1.5 \text{ cm}, z_0 < 2.5 \text{ cm}$
KLHNotPion	$\mathcal{L}(K)/(\mathcal{L}(K) + \mathcal{L}(\pi)) > 0.20$
KLHLoose	$\mathcal{L}(K)/(\mathcal{L}(K) + \mathcal{L}(\pi)) > 0.8176$ and $\mathcal{L}(K)/(\mathcal{L}(K) + \mathcal{L}(p)) > 0.018$
eBremRecoPidHllMerge	$\mathcal{L}(e)/(\mathcal{L}(e) + 5\mathcal{L}(\pi) + \mathcal{L}(K) + 0.1\mathcal{L}(p)) > 0.95$

Table 3.2: Particle Identification Selectors. The two Tracks lists do not use particle identification. $|d_0|$ is the distance, transverse to the beam axis, from the interaction point to the point of closest approach for the track. $|z_0|$ is the distance, along the beam axis, from the interaction point to the point of closest approach for the track. \mathcal{L} has the same meaning as in Equation 3.9.

lookup table, rather than a Gaussian fit. The probabilities from each of the three sub detectors are formed into a single likelihood value:

$$\mathcal{L} = \mathcal{L}_{DIRC}\mathcal{L}_{DCH}\mathcal{L}_{SVT}. \quad (3.9)$$

Based on the these likelihoods, each track is placed on several predefined lists. The requirements for the lists relevant to this analysis are shown in Table 3.2.

3.4.2 Composite Particle Lists

In addition to individual particles, this analysis makes extensive use of predefined lists of putative particles that have been reconstructed from multiple tracks and neutral clusters. A summary of those lists and their requirements are given in the following list. Note that π^0 decays almost exclusively to $\pi^0 \rightarrow \gamma\gamma$.

- pi0LooseMass: The photons have energy in the lab frame $\geq 30 \text{ MeV}$. The Lateral moment of each must be between 0.0 and 0.8. This list requires $0.10 \text{ GeV} <$

$m_{\gamma\gamma} < 0.16$ GeV and that the energy of the π^0 in the lab frame is greater than 0.2 GeV.

- `MergedPi0Loose`: A merged π^0 is reconstructed from a cluster that has apparently been struck by both photons. This list requires the consistence with that hypothesis to be greater than 1%.
- `pi0AllLoose`: All candidates in either `pi0LooseMass` or `MergedPi0Loose`.
- `rhoCBRecoLoose`: Assumes all tracks from `GoodTracksVeryLoose` are pions and combines them with putative π^0 from `pi0AllLoose`. All ρ^+ candidates with 10 GeV of the known mass are kept.
- `rhoCDefault`: Keeps only those ρ^+ candidates from `rhoCBRecoLoose` that are within 0.32 GeV of the known $m_{\rho^+} = (775.49 \pm 0.34)$ MeV.
- `a1CLoose`: This list consists of all sets of three tracks from `GoodTracksVeryLoose` that have a combined invariant mass greater than zero.

3.5 Blinding and Monte Carlo Simulation of the *BABAR* Detector

The *BABAR* collaboration has adopted the practice of blinding data. This means that a signal region is identified in a variable or combination of variables where data from the targeted decay or process will accumulate. The data points in this region are not examined (“unblinded”) until any analysis procedures that might bias the result have been completed. This is designed to prevent biasing the results of an analysis.

Since blinding forbids us to look at real data when determining and optimizing the analysis procedure, we use simulated data, known as Monte Carlo (MC). MC samples are used to develop analysis procedures, study backgrounds, understand efficiencies, and validate the analysis. The generation of MC samples for *BABAR* is distributed among computing facilities at twenty different sites throughout the collaboration.

MC samples begin with a computer generator that produces events. In each event, particles and their resulting decays occur in accordance with all measured branching fractions and probabilities. Each event is then processed using the GEANT 4 package [33] to simulate detector responses, coverage, and fallibility as accurately as possible. The resulting simulated data is then analyzed as if it were real data.

We use several different MC samples in this analysis. Five are known as generic MC; they simulate all known physics produced by the collider and are named after the physics processes they simulate. The processes $e^+e^- \rightarrow u\bar{u}, d\bar{d}, s\bar{s}$ are simulated by the *uds* MC sample, $e^+e^- \rightarrow c\bar{c}$ by the $c\bar{c}$ MC sample, and $e^+e^- \rightarrow \tau^+\tau^-$ by the $\tau^+\tau^-$ sample. The process $e^+e^- \rightarrow \Upsilon(4S)$ is simulated separately for $\Upsilon(4S) \rightarrow B^+B^-$ and $\Upsilon(4S) \rightarrow B^0\bar{B}^0$.

We also use specialized MC samples that simulate large numbers of our target decays; these are known as signal MC samples. We normalize our signal MC samples to data luminosity assuming certain branching fractions for the signal modes. Based on these predictions, we choose the following branching fractions for the purposes of these normalizations: $\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau) = 1 \times 10^{-4}$, $\mathcal{B}(B^+ \rightarrow \mu^+\nu_\mu) = 5 \times 10^{-7}$, and $\mathcal{B}(B^+ \rightarrow e^+\nu_e) = 1 \times 10^{-11}$.

3.6 Data Quality Monitoring

The beams of electrons and positrons passing through the interaction point inside the *BABAR* detector consist of discrete bunches, which cross at the interaction point. Collisions between e^+ and e^- occur at the rate of ~ 10 MHz. If all information from every collision were to be stored, it would take far more computer memory than we have available.

To reduce the amount of data to a manageable level and preserve as many potentially interesting collision events as possible, *BABAR* employs hardware (Level 1) and software (Level 3) triggers. A Level 2 trigger is present at some other particle physics experiments but was not needed at *BABAR*. The Level 1 trigger output is typically several kHz, depending on luminosity and beam conditions. The level 3 trigger reduces this frequency by approximately an order of magnitude. Events passing the level 3 trigger are stored for review and analysis.

Before being stored, the data are passed through two stages of rudimentary analysis, called Prompt Calibration (PC) and Event Reconstruction (ER). The PC pass is primarily used to monitor the detector and data as they are being collected, and the ER pass is a more rigorous reconstruction of the events.

Before the data are used in physics analysis, the Data Quality Group (DQG) meets weekly to determine the usability of the data collected or processed during the previous week. The data are accepted or rejected on a run-by-run basis. Each sub-detector is represented by an expert who reports which runs are Good, Flawed or Bad for that week.

Every member of the BaBar Collaboration is required to perform service work to ensure that the experiment operates smoothly. I served as the IFR expert on the Data Quality Group (DQG) for sixteen months, from October 2006 to February 2008.

Each sub-detector has specialized standards for determining the condition of data, and the standards I used were largely based on the experience of the previous IFR expert. I used several variables that were most useful and sensitive to potential problems with the IFR. If any run exhibited anomalously large or small values for a given variable, further investigation was performed. Most problems could be understood by looking in the experiment's electronic logbook. I then determined the flag for that event based on the severity of the problem.

3.7 Run Divisions

The data were collected in several periods that are divided by repairs or upgrades to the *BABAR* detector. Each of these periods is called a Run. The dates and amount of data collected for each Run are shown in Table 3.3.

Run	Start Date	End Date	On Resonance Data (fb^{-1})	Off Resonance Data (fb^{-1})
1	Oct. 20, 1999	Oct 29, 2000	20.40	2.62
2	Feb. 6, 2001	June 30, 2002	61.08	6.93
3	Nov. 17, 2002	June 29, 2003	32.28	2.47
4	Sept. 9, 2003	July 31, 2004	100.28	10.12
5	April 4, 2005	Aug. 17, 2006	133.26	14.49
6	Jan. 19, 2007	Apr. 7, 2008	70.31	5.56
Total	Oct. 20, 1999	Apr. 7, 2008	417.6	42.2

Table 3.3: Data samples used in the analysis.

CHAPTER 4

EXPERIMENTAL DETAILS AND DATA SETS

This section describes the analysis strategy employed to study the fully leptonic decays

$$B^+ \rightarrow \ell^+ \nu_\ell \quad (\ell = e, \mu, \tau) \quad (4.1)$$

We discuss the MC samples used to establish and test our analysis procedure, as well as the data used to measure the branching fraction in each mode.

4.1 Event Selection Outline

Since the decays for which we search are rare and difficult to detect, our goal is to whittle the entire data set down to a sample that has as many signal events and as few background events as possible. We must then identify and measure the number of signal events and translate those numbers into branching fractions or upper limits. We accomplish this in several steps, as enumerated below.

1. From the whole data set, a “skim” is performed that separates the events with our desired tag and stores them in a smaller data set. The semileptonic skim uses the tag decay $B^- \rightarrow D^0 \ell^- \bar{\nu} X$.

-
2. The data set is reduced further in size by tightening the tag requirements from the skim and accepting only a finite set of signal B candidate modes. These include $B^+ \rightarrow \tau^+ \nu_\tau$, $B^+ \rightarrow \ell^+ \nu_\ell$, $B^+ \rightarrow K^+ \nu \bar{\nu}$, $B^0 \rightarrow \nu \bar{\nu}$, and $B^0 \rightarrow \nu \bar{\nu} \gamma$. The final states are all mutually exclusive; in other words, each signal B candidate is assigned to *one* final state. Another small data subsample is created that contains all events in which a tag is reconstructed, regardless of whether a signal B is reconstructed. Therefore, it contains no information about the signal side of the events. These data sets are used on several related analyses.

 3. At this step, we narrow the focus of the data set to this analysis. We reduce the data set again via software customized for this analysis. The reduction requires the signal candidate to be in the desired final states. Beyond this selection, no other selection requirements are imposed. New variables are calculated and stored from the available data.

 4. Signal B selection procedures are developed and selectors are optimized. The signal region is defined and blinded.

 5. The signal region is unblinded, and the results are interpreted via the Feldman-Cousins method [34]. Systematic errors are assessed to produce our final result.

4.2 Data

The data samples used in this analysis consist of the Runs 1-6 data sets, which amount to a total integrated on-resonance luminosity of 417.6 fb^{-1} ($(458.9 \pm 5.1) \times 10^6$ $\Upsilon(4S)$ decays) and off-resonance luminosity of 42.2 fb^{-1} , as outlined in Table 3.3.

4.3 Monte Carlo Simulations

To avoid bias, the MC samples are divided into three equal parts. Each event is assigned to one of these three parts by using the remainder when the upperID, which is a unique identifier for each event is divided by 6. This allows us to split the MC into two (remainder = 0-2 or 3-5) or three (remainder = 0-1, 2-3, 4-5) samples of equal size. We use 0-1 as the “Training” sample to generate our PDFs and LHRs, 2-3 as the “Optimization” sample to optimize our final selection requirements, and 4-5 as the “Testing” sample to measure efficiencies and predict background levels. The MC sample with all events included, regardless of the flag value, is called the “Complete” sample.

4.3.1 Generic Background Monte Carlo Samples

Generic MC is used to simulate backgrounds for this analysis. We use B^+B^- , $B^0\bar{B}^0$, $c\bar{c}$, uds , and $\tau^+\tau^-$ generic MC samples. The number and luminosity generated for these samples is shown in Table 4.3.2

4.3.2 Signal Monte Carlo Samples

We use three signal MC samples, one for each of the three leptonic decays (Table 4.2). These are all B^+B^- samples in which one B is forced to decay into our signal mode and the other decays as it would in the generic sample. The skimmed signal MC is used for calculating the tag efficiencies, signal efficiencies, and signal related systematic uncertainties.

4.4 Skim Selection

This section represents Step 1 from §4.1. We have a predefined subset (“skim”) of the data that is highly enriched in events where one B decays into our tagging modes.

This skim chooses only events that contain at least one charged lepton with momentum in the CM frame $1.0 < p^* < 2.8$ GeV. These putative lepton daughters of the Tag B are required to pass particle identification as electrons or muons. A D^0 must also be reconstructed, which is combined with the lepton to form the tag B . In the case where multiple candidates are possible, the candidate with the largest decay vertex probability is selected and an attempt is made to form a signal candidate from the remaining tracks and neutrals in the event.

The D^0 is reconstructed in four modes: $D^0 \rightarrow K^-\pi^+$, $D^0 \rightarrow K^-\pi^+\pi^+\pi^-$, $D^0 \rightarrow K^-\pi^+\pi^0$, and $D^0 \rightarrow K_s^0\pi^+\pi^-$. The branching fractions are listed in 4.3. For the $D^0 \rightarrow K^-\pi^+\pi^0$ mode, the reconstructed mass of the D^0 must be within 70 MeV of the known D^0 mass, which is (1864.84 ± 0.17) MeV. For the remaining three modes, it must be within 40 MeV of the known D^0 mass.

The semileptonic analyses are intended to be used in a complementary fashion with the hadronically tagged analyses. The use of the tags ensures that we have two independent data sets. This allows us to combine the analyses when they are complete and roughly double the statistical power of the data set.

Our tag B is reconstructed in the decay modes $B^\pm \rightarrow D^0\ell^\pm\nu X$, where X is either nothing, a π^0 or a γ from D^{*0} , or other low-momentum transition particles from higher-mass charm resonances, which are not always reconstructed.

4.5 Tag Selection

This section represents Step 2 from §4.1. Any K^\pm daughters of the the D^0 , are required to be KLHNotPion. The K_s is required to be KsTight, and its flight significance is required to be $> 3\sigma$. More refinements of the D^0 occur at tag selection.

4.5.1 Seeding Method

Due to the presence of the neutrino in the products of the tag B , the direction of neither B can be known. Instead, $\text{Cos}\theta_{B-D^0\ell}$ (also known as CosBY) is calculated in the $\Upsilon(4S)$ rest frame using the following equation which assumes that the only missing particle in the tag-side decay is a massless neutrino.

$$\text{Cos}\theta_{B-D^0\ell} = \frac{2E_B E_{D^0\ell} - m_B^2 - m_{D^0\ell}^2}{2|\vec{p}_B||\vec{p}_{D^0\ell}|} \quad (4.2)$$

After the D^0 has been reconstructed, a “seeding” algorithm adds a photon (called the “seed photon”) to the reconstructed D^0 and reevaluates CosBY, which is defined in Equation 4.2. The seeding algorithm performs this procedure with all reconstructed photons that do not overlap with the tag B and have CM energy less than 300 MeV. If a seed photon causes CosBY to become closer to (but not greater than) 1, it is selected. We seek to move CosBY closer to 1 because events containing real D^* mesons usually appear in the low tail of the CosBY distribution. If more than one photon satisfies these conditions, the one which moves $\Delta M \equiv m_{D^{*0}} - m_{D^0}$ closest to the nominal value of 142.12 MeV [1] is used.

The photon is removed from all other relevant calculations, and the event is stored with the modified variables. We studied this procedure using a second photon to

account for possible $D^{*0} \rightarrow D^0\pi^0$ decays, but this did not produce a significant improvement.

In the previous measurement of $\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau)$ with semileptonic tags [25], we concluded that $D^0\ell\nu X$ tags always offer better significance, even at the cost of higher background. Not explicitly reconstructing the D^{*0} (or D^{**} or other higher-mass charm meson) resulted in lower purity but greater statistical power. The increase in statistical power is worth the loss of purity because it produces a net gain in the overall significance of the tag selection.

We developed the seeding procedure to allow us to combine the statistical power of the $D^0\ell\nu X$ tags with the cleanliness and precision of a fully reconstructed D^{*0} . Unless otherwise stated, for the remainder of this document, E_{extra} refers to the total extra energy in the detector after both B mesons and any seed photon have been reconstructed. Only clusters with more than 30 MeV of energy have been used in the calculation of E_{extra} .

4.5.2 Accepted Signal Modes

Once the best tag B has been chosen, a ranked selection is performed on remaining tracks and neutral clusters. Since the τ decays before it can reach active detector elements, we reconstruct it in five decay modes : $\tau^+ \rightarrow e^+\nu_e\bar{\nu}_\tau$, $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$, $\tau^+ \rightarrow \pi^+\bar{\nu}_\tau$, $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$, and $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$. The relevant branching fractions are shown in Table 4.4.

The list below is abbreviated from the full version used for this and many other analysis. Only those modes relevant to this analysis are shown.

-
1. If signal side has a non-overlapping second tag, it is classified as a Double Tag event.
 2. If the event contains three `GoodTracksVeryLoose` on the `a1CLoose` and that pass loose a_1^+ mass constraints, we classify the event as $B^+ \rightarrow a_1^+ \nu \bar{\nu}$, ($a_1^+ \rightarrow 3\pi$).
 3. If a signal side has only one `GoodTracksVeryLoose`, we place it in one of the following categories in the priority given.
 - (a) If the track passes the selector `muNNLoose` and the event does not contain a photon with momentum greater than 0.5 GeV, we classify it as $B^+ \rightarrow \mu^+ \nu_\mu$.
 - (b) If the signal track passes the selector `eBremRecoPidH11Merge` and the event does not contain a photon with momentum greater than 0.5 GeV, we classify it as $B^+ \rightarrow e^+ \nu_e$.
 - (c) If the single track is found on the list `rhoCDefault`, and the putative ρ^+ satisfies beam flight constraints and tree fitter, we classify the event as $B^+ \rightarrow \rho^+ \nu \bar{\nu}$.
 - (d) If the track is not accepted by any of the above filters, we classify the event as $B^+ \rightarrow \pi^+ \nu \bar{\nu}$ by default.

4.6 Tag Selection for $B^+ \rightarrow \ell^+ \nu_\ell$

This section represents Step 3 from §4.1. Several variables are available to refine our tag further. To determine which ranges of these variables we select, we use a

procedure that seeks to maximize a figure of merit (FOM). A FOM is a number that quantifies how well we are retaining signal events and rejecting background events.

4.6.1 Optimization Procedure

Optimization of tag side cuts was achieved using the `BumpHunterApp` in the `StatPatternRecognition` [35] package. Three statistically independent data text files (known as training, validation, and test) should be used by the application to optimize cut values. Each of these samples has one portion designated signal and one portion designated background. In this analysis, the signal portion was our signal MC and the background portion consisted of our generic MC samples. The output of `BumpHunterApp` is a text file containing cut values that maximize the figure of merit (FOM). `BumpHunterApp` has several FOMs from which the user can choose. We used $\text{FOM} = \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{BG}}}}$, where N_{sig} is the number of signal events and N_{BG} is the number of background events.

In training, several peel parameters are tested to determine which one produces the highest FOM. After `BumpHunterApp` is trained, the actual optimization is performed on the validation sample with the peel parameter determined from training. The cuts are then applied to the test file to ensure that the resulting cut efficiencies are consistent with those derived from the validation file.

We split our training and validation samples into fifths in order to use and understand `BumpHunterApp` with samples that could be run and re-run relatively quickly. Five different training and validation exercises were performed. For training, we scanned `BumpHunterApp` over peel parameters from 0.05 to 0.9 to find the parameter that yielded the maximum FOM in each of the five independent samples.

For validation, `BumpHunterApp` was run again on each of the five samples using the optimum peel parameter determined in training. The cut values produced were checked for consistency among the five samples.

We originally planned to use Run 1 data and MC for training and Run 2 for validation. However, at the time we began performing training and validation for the $D^0\ell\nu X$ tag mode, the MC for charged and neutral B had not been skimmed for Run 2. For these two MC modes, Run 1 was split into tenths; five tenths were used for training and the other five for validation. All normalizations were adjusted accordingly. The number of events used in each sample is shown in Table 4.5.

After the small sample studies described above were complete, our original plan was to split the Run 1 samples in half, use one half for training and the other half for validation. We then would have used Run 2 (when it had all been skimmed) as our test sample. However, for training and validating on half of the Run 1 sample, `BumpHunterApp` exceed the capacity of the *BABAR* computing infrastructure. Therefore, this plan was never completed. The final cut values were determined by averaging the cut values determined from our studies on the five validation exercises on small samples.

The cut values are similar to those used by the Runs 1-4 iteration of the $B^+ \rightarrow \tau^+\nu_\tau$ analysis [24]. Because of this and the memory problems described above, we decided to accept these cut values without using a test sample. This similarity is also a justification for not re-optimizing the cuts for this analysis.

4.6.2 Optimization Results

Several variables are used to refine the list of tags present in the skimmed data and MC.

- **D^0 Decay Mode:** We reconstruct only the D^0 candidates in the decay modes $K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0$, and $K_S^0\pi^+\pi^-$, where $K_S^0 \rightarrow \pi^+\pi^-$.
- **D^0 Mass:** In the procedure used in the analysis on data from Runs 1-5, the D^0 mass distributions in the signal MC sample were fitted with Gaussian curves. The mass acceptance windows were set around the nominal D^0 mass at approximately 3σ from the fits. We have retained these requirements for this analysis. The D^0 mesons reconstructed via the $K^-\pi^+\pi^0$ decay mode are required to have a reconstructed mass between 1.8295 and 1.8995 GeV ($3\sigma \approx 35$ MeV). The other decay modes are required to have a reconstructed mass between 1.8445 and 1.8845 GeV ($3\sigma \approx 20$ MeV).
- **Tag K^+ Particle Identification:** The charged kaon track from D^0 is required to fail both `PidLHElectrons` electron selection and `muNNTight` muon selection.
- **Tag π^+ Particle Identification:** All charged pion tracks from D^0 or its K_S^0 daughter are required to fail the `KLHNotPion`, `PidLHElectrons`, and tight muon selections.
- **γ Lateral Moment:** For the decay mode $D^0 \rightarrow K^-\pi^+\pi^0$, the lateral moment of each of the photon showers is required to be between 0.01 and 0.6. The upper bound is applied to reject hadronic split-offs, beam background etc., which tend

to have lateral moments closer to 1. The lower bound is to exclude the region where data-MC agreement for the lateral shower shape moment is poor.

- B^0 **Veto**: All events in which the a tag-side neutral B is reconstructed, rather than a charged B , are rejected.

4.7 Tag B Efficiency and Yield

Defining the tag and signal efficiencies in a tagged analysis is sometimes confusing for several reasons. The analysts arbitrarily define the cut at which the tag efficiency ends and the signal efficiency begins. In this analysis, we have chosen the tag efficiency to be the efficiency after the B^0 veto is applied. Also, several variables that were used in the tag selection in the previous analysis are now used in signal selection. This means that the tag efficiency of this analysis is not directly comparable to the tag efficiency of any previous analysis.

Tag and Signal efficiencies are measured using the complete signal MC samples for each mode. The $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC has a more populated signal side than the $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$ signal MC samples. For this and other reasons, the three MC samples have different tag efficiencies, as seen in Table 4.6.

4.8 Fitting vs. Cuts

At the final stage of this analysis, we will accept only events that are within certain ranges for several variables. We will then examine the number of events in the signal range and use that to calculate the branching fractions and upper limits. This is called a “cut and count method.”

More sophisticated methods are available and were considered for this analysis. For instance, we considered fitting the data to a functional form and extracting the signal from that fit. This would require an accurate fit function for the background. We could take the shape of this function from our generic MC samples with all other signal selectors applied, but we would also need to verify that this shape agrees with data.

We cannot use the data with our signal selection applied because it might include signal, and we would thus be mistakenly counting signal as background. Ideally, we would have selectors that would produce a sample in data and MC that would reproduce our background shape but be independent of our signal. However, such a sample has proven impossible to find. No known MC sample accurately reproduces the shape of the background as given by the generic MC sample with sufficient statistics to verify agreement between MC and data.

Therefore, any fitting method would produce a very large systematic error that would outweigh any gain obtained by the fit. I do not fit anything and use the simple cut and count method.

MC sample	Number of events $\times 10^6$ (Generated)	Cross section [nb]	Luminosity [fb $^{-1}$]
B^+B^- MC			
SP-1235-BToDlnu-Run1-R22	36.97	0.55	67.21
SP-1235-BToDlnu-Run2-R22	103.1		187.5
SP-1235-BToDlnu-Run3-R22	484.7		88.12
SP-1235-BToDlnu-Run4-R22	168.0		305.4
SP-1235-BToDlnu-Run5-R22	244.2		444.0
SP-1235-BToDlnu-Run6-R22	100.8		183.3
$B^0\bar{B}^0$ MC			
SP-1237-BToDlnu-Run1-R22	37.20	0.55	67.63
SP-1237-BToDlnu-Run2-R22	103.4		187.9
SP-1237-BToDlnu-Run3-R22	48.46		88.12
SP-1237-BToDlnu-Run4-R22	165.3		300.5
SP-1237-BToDlnu-Run5-R22	236.0		429.0
SP-1237-BToDlnu-Run6-R22	198.2		360.4
$c\bar{c}$ MC			
SP-1005-BToDlnu-Run1-R22	58.90	1.30	45.31
SP-1005-BToDlnu-Run2-R22	168.8		129.9
SP-1005-BToDlnu-Run3-R22	83.97		64.60
SP-1005-BToDlnu-Run4-R22	252.8		194.5
SP-1005-BToDlnu-Run5-R22	366.8		282.1
SP-1005-BToDlnu-Run6-R22	156.9		120.7
uds MC			
SP-998-BToDlnu-Run1-R22	47.18	2.09	22.57
SP-998-BToDlnu-Run2-R22	130.9		62.62
SP-998-BToDlnu-Run3-R22	66.72		31.92
SP-998-BToDlnu-Run4-R22	213.4		102.1
SP-998-BToDlnu-Run5-R22	317.8		152.1
SP-998-BToDlnu-Run6-R22	67.96		72.3
$\tau^+\tau^-$ MC			
SP-3429-BToDlnu-Run1-R22	20.38	0.919	22.17
SP-3429-BToDlnu-Run2-R22	55.61		60.51
SP-3429-BToDlnu-Run3-R22	27.99		30.45
SP-3429-BToDlnu-Run4-R22	90.03		97.97
SP-3429-BToDlnu-Run5-R22	132.2		143.9
SP-3429-BToDlnu-Run6-R22	56.43		61.41

Table 4.1: MC samples used in the analysis. Luminosity of each MC sample are obtained by dividing the number of events with the appropriate cross-sections. The cross-sections for all species except $B\bar{B}$ are taken from *The BaBar Physics Book* [2]. The $B\bar{B}$ cross-section is calculated by dividing the number of $\Upsilon(4S)$ in the on-peak data, assuming that the production of B^+B^- and $B^0\bar{B}^0$ are equal.

MC sample	Number of events $\times 10^6$ (Generated)	Cross section [nb]	Luminosity [fb $^{-1}$]
$B \rightarrow e\nu$ VS generic (2803) MC			
SP-2803-BToDlnu-Run1	0.421	1.1×10^{-11}	3.83×10^{10}
SP-2803-BToDlnu-Run2	1.257		1.14×10^{11}
SP-2803-BToDlnu-Run3	0.660		6.00×10^{10}
SP-2803-BToDlnu-Run4	2.024		1.84×10^{11}
SP-2803-BToDlnu-Run5	2.665		2.42×10^{11}
SP-2803-BToDlnu-Run6	0.742		6.75×10^{10}
Total	7.769		
$B \rightarrow \mu\nu$ VS generic (1329) MC			
SP-1329-BToDlnu-Run1	0.421	5.5×10^{-7}	7.65×10^5
SP-1329-BToDlnu-Run2	1.257		2.29×10^6
SP-1329-BToDlnu-Run3	0.660		1.20×10^6
SP-1329-BToDlnu-Run4	2.024		3.68×10^6
SP-1329-BToDlnu-Run5	2.49		4.53×10^6
SP-1329-BToDlnu-Run6	0.742		1.35×10^6
Total	7.594		
$B \rightarrow \tau\nu$ VS generic (1063) MC			
SP-1063-BToDlnu-Run1	1.054	1.1×10^{-4}	9.58×10^3
SP-1063-BToDlnu-Run2	3.144		2.86×10^4
SP-1063-BToDlnu-Run3	1.654		1.50×10^4
SP-1063-BToDlnu-Run4	5.06		4.60×10^4
SP-1063-BToDlnu-Run5	6.662		6.06×10^4
SP-1063-BToDlnu-Run6	1.855		1.69×10^4
Total	19.429		

Table 4.2: Signal MC samples used in the analysis. Luminosity of each MC sample are obtained by dividing the number of events with the appropriate cross-sections.

Decay Mode	Branching Fraction
$B^- \rightarrow D^{*0} \ell^- \bar{\nu}_\ell$	$(6.5 \pm 0.5) \%$
$D^{*0} \rightarrow D^0 \pi^0$	$(61.9 \pm 2.9) \%$
$D^{*0} \rightarrow D^0 \gamma$	$(38.1 \pm 2.9) \%$
$D^0 \rightarrow K^- \pi^+$	$(3.82 \pm 0.07) \%$
$D^0 \rightarrow K^- \pi^+ \pi^+ \pi^-$	$(7.70 \pm 0.25) \%$
$D^0 \rightarrow K^- \pi^+ \pi^0$	$(13.5 \pm 0.6) \%$
$D^0 \rightarrow K_s^0 \pi^+ \pi^-$	$(2.88 \pm 0.19) \%$
$K_s^0 \rightarrow \pi^+ \pi^-$	$(69.20 \pm 0.05) \%$

Table 4.3: Branching fraction of the decays involved in $B^- \rightarrow D^{*0} \ell^- \bar{\nu}_\ell$ tag reconstruction [1].

Decay Mode	Branching Ratio
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	$(17.84 \pm 0.05) \%$
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	$(17.36 \pm 0.05) \%$
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	$(10.90 \pm 0.07) \%$
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	$(25.50 \pm 0.10) \%$
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	$(9.33 \pm 0.08) \%$

Table 4.4: Branching fraction of the τ decay modes used for $B^+ \rightarrow \tau^+ \nu_\tau$ signal search [1].

MC Mode	$D^0 \ell \nu X$	
	Validation	Training
Signal	1306	3772
$B^+ B^-$	2349358	2349358
$B^0 \bar{B}^0$	2124668	2124668
$c\bar{c}$	1275982	2699867
uds	139268	199997
$\tau\tau$	2122	3574

Table 4.5: Number of events in each training and validation MC sample.

Cut	$\epsilon_{\text{tag}}^{B^+ \rightarrow \tau^+ \nu_\tau}$	$\epsilon_{\text{tag}}^{B^+ \rightarrow \mu^+ \nu_\mu}$	$\epsilon_{\text{tag}}^{B^+ \rightarrow e^+ \nu_e}$
Skim Level	-	-	-
D^0 Mass	$(2.933 \pm 0.004)\%$	$(1.419 \pm 0.004)\%$	$(1.548 \pm 0.005)\%$
Tag K^+ PID	$(2.772 \pm 0.004)\%$	$(1.355 \pm 0.004)\%$	$(1.468 \pm 0.004)\%$
Tag π^+ PID	$(1.992 \pm 0.003)\%$	$(1.126 \pm 0.004)\%$	$(1.178 \pm 0.004)\%$
γ Lateral Moment	$(1.990 \pm 0.003)\%$	$(1.125 \pm 0.004)\%$	$(1.176 \pm 0.004)\%$
B^0 Veto	$(1.699 \pm 0.003)\%$	$(1.052 \pm 0.004)\%$	$(1.093 \pm 0.004)\%$
Cut	$B^+ \rightarrow \tau^+ \nu_\tau$	$B^+ \rightarrow \mu^+ \nu_\mu$	$B^+ \rightarrow e^+ \nu_e$
Skim Level	19429000	7594000	7769000
D^0 Mass	569948	107761	120268
Tag K^+ PID	538549	102862	114080
Tag π^+ PID	387032	85475	91481
γ Lateral Moment	386613	85400	91392
B^0 Veto	330082	79922	84939

Table 4.6: Comparison of Tag Efficiencies for different cuts. All of the yields are raw, unweighted yields for the signal MC sample for the mode given, representing Runs 1-6.

CHAPTER 5

EXPERIMENTAL SEARCH FOR $B^+ \rightarrow \ell^+ \nu_\ell$

This chapter represents Step 4 from §4.1. Once our tag B has been refined as much as possible, we turn our attention to the signal B . Our goal is again to select signal events with the highest efficiency possible while selecting background events with the lowest efficiency possible.

We used many variables to separate the signal from background after the tag selection had been completed. Two variables had the most discriminating power. One is the extra energy (E_{extra}). The extra energy in an event is the total energy recorded in the detector that is not assigned to the tag or signal B . In an ideal $B^+ \rightarrow \ell^+ \nu_\ell$ decay, we reconstruct all tracks and clusters associated with the real decay. The only unreconstructed particles would be neutrinos, which leave no energy in the detector. Therefore, we expect our signal to concentrate near zero extra energy.

For the modes $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$, we use another variable with discriminating power on par with E_{extra} : the calculated momentum of the signal lepton in the signal B rest frame ($p'_{\text{sig } \ell}$). The remaining variables, if they have some discriminating power, are used in two likelihood ratios (LHRs). These are products of multiple variables that take advantage of shape differences. They are defined in more detail in Section 5.3.

For the $B^+ \rightarrow \tau^+ \nu_\tau$ analysis, the selection requirements on the LHRs and E_{extra} are simultaneously optimized, and the final event count is the data that remains after the optimized selections are made. Each of the five decay modes is optimized separately. For the $B^+ \rightarrow \ell^+ \nu_\ell$ search, the selection requirements on the LHRs, E_{extra} , and $p'_{\text{sig } \ell}$ are simultaneously optimized, and the final event count is the data that remains after the optimized selections are made.

From §4.5.2, we use the events falling into 2, 3c, 3a, 3b, and 3d for $B^+ \rightarrow \tau^+ \nu_\tau$. Events falling into 3a and 3b are also used for the $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$ decay modes. The two possibilities are distinguished by $p'_{\text{sig } \ell}$. Note that for all variables, a prime (') indicates the variable is in the signal B rest frame, and an asterisk (*) indicates the variable is in the center of mass (CM) frame.

5.1 Separating $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$ from $B^+ \rightarrow \tau^+ \nu_\tau$

In previous and concurrent analyses of $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_\mu$ decays [16, 18], the magnitude of the signal lepton momentum in the signal B rest frame ($p'_{\text{sig } \ell}$) was found to be a powerful discriminating variable. In this analysis, it is used as a separate variable from the LHRs because of its power and because it effectively discriminates between $B^+ \rightarrow \ell^+ \nu_\ell$ and $\tau^+ \rightarrow \ell^+ \nu_\ell$, where $\ell = \mu, e$.

Due to the presence of neutrinos in both B decays, the direction and magnitude of the signal B rest frame cannot be directly measured in this analysis. We studied two methods to estimate the B rest frame and $p'_{\text{sig } \ell}$.

Both methods exploit data from the beams and CM frame to estimate the magnitude and direction of the signal B rest frame relative to the CM frame. The magnitude

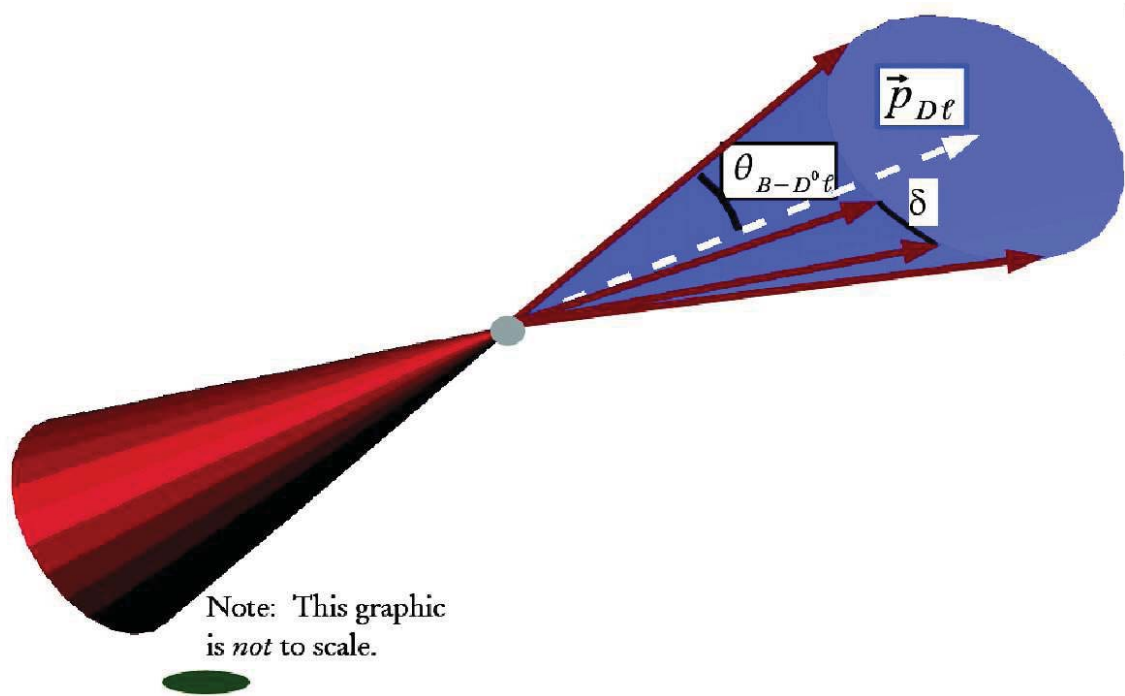


Figure 5.1: In this representation of semileptonic B tagging, the dashed white vector is $\vec{p}_{D^0\ell}^*$. The red vectors, which are constrained to lie on the surface of the blue cone, represent possible directions for the signal B momentum. The angle $\theta_{B-D^0\ell}$ defines the cone with respect to $\vec{p}_{D^0\ell}^*$. The gray spot is the B vertex, and the red cone represents the possible paths of the signal B projected back through the vertex. The green ellipse represents the beam spot. This graphic is *not* to scale.

of the momentum of the B relative to the CM frame ($|\vec{p}_B^*|$) can be calculated from the beam information

$$|\vec{p}_B^*| = \sqrt{E_B^2 - m_B^2} = \sqrt{\left(\frac{E_{\text{CM}}}{2}\right)^2 - m_B^2} = 341.0 \text{ MeV}, \quad (5.1)$$

where E_{CM} is the center of mass energy of the collider and the remaining variables have the same meanings as in Equation 4.2.

The flight direction of the tag B in the CM frame lies on a cone with angle $\theta_{B-D^0\ell}$ around the reconstructed $\vec{p}_{D^0\ell}^*$, as shown in Figure 5.1. The difference between the two methods is the way in which this conical rotational ambiguity is broken. Once the ambiguity is broken, the reference frame of the tag B is completely determined. The signal B rest frame can be easily determined relative to the CM frame using conservation of momentum and beam information.

In the cases where $\cos(\theta_{B-D^0\ell}) < -1$, $\theta_{B-D^0\ell}$ is set to $-\pi$ for the purposes of the calculation. In both methods, the vectors on the cone are identified by an azimuthal angle δ defined with respect to the central axis of the cone.

Y-Average Method

This method³ uses averaging to resolve the rotational ambiguity. It has been used in previous analyses [36, 37]. First, an arbitrary δ is chosen and designated δ_0 . $p'_{\text{sig } \ell}$ is calculated for each of δ_0 through δ_{49} , where $\delta_n - \delta_{n-1} = \frac{2\pi}{50}$. The final value of $p'_{\text{sig } \ell}$ is taken as the average of these 50 values.

A previous analysis [37] found that using more than four δ values in this method frame does not significantly improve its effectiveness. However, 50 values were used to insure a consistent and fair comparison with the other method studied.

³Thanks to Robert Kowalewski for suggesting this method.

Beam Spot Method

Independent of the Y-Average Method, we developed the Beam Spot Method to resolve the rotational ambiguity using beam spot information. The beam spot is the physical region in which the beams of electrons and positrons overlap. All collisions that produce $\Upsilon(4S)$ occur in this spot. A single point is recorded in data for the spot, although in reality it is an extended volume. Its dimensions are roughly $120 \mu\text{m}$ by $5.6 \mu\text{m}$ by 0.9cm .

All real tag B mesons should originate in the beam spot and decay at the B -vertex. Rather than averaging over the vectors on the cone, this method extrapolates the path of the signal B through the B -vertex for each δ_n . The vector corresponding to the path that brings the tag B closest to the beam spot is used to calculate $p'_{\text{sig } \ell}$.

For each path, the distance to the beam spot is given by

$$\sqrt{\frac{\begin{vmatrix} y_0 - y_1 & z_0 - z_1 \\ b & c \end{vmatrix}^2 + \begin{vmatrix} z_0 - z_1 & x_0 - x_1 \\ c & a \end{vmatrix}^2 + \begin{vmatrix} x_0 - x_1 & y_0 - y_1 \\ a & b \end{vmatrix}^2}{a^2 + b^2 + c^2}}, \quad (5.2)$$

which is a general equation for the distance between the point (x_0, y_0, z_0) and the line through (x_1, y_1, z_1) in direction (a, b, c) [38]. In this case, (x_0, y_0, z_0) is the location of the beam spot, (x_1, y_1, z_1) is the B -vertex, and (a, b, c) is the direction corresponding to the given δ . The vertical lines in the numerator represent the determinants of the variables they enclose.

Method Comparison and Choice

For true signal events, $p'_{\text{sig } \ell}$ should exhibit a peak at

$$p'_{\text{sig } \ell} = \frac{m_B^2 - m_\ell^2}{2m_B} \approx \frac{m_B}{2} = 2.64 \text{ GeV}. \quad (5.3)$$

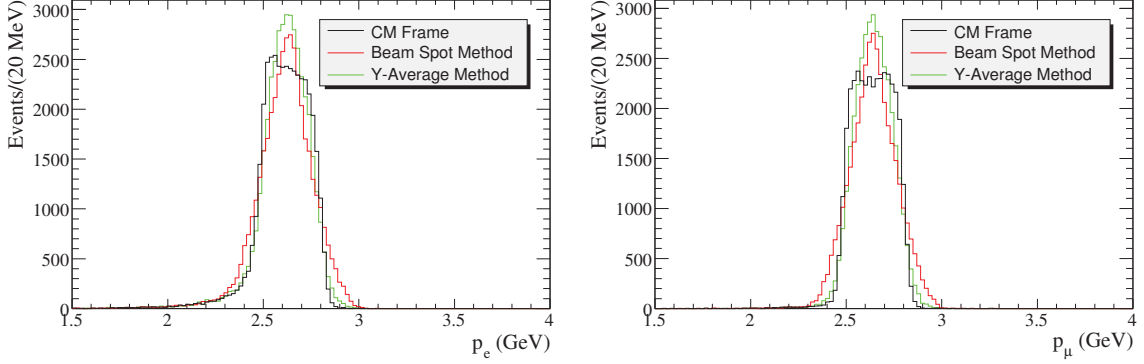


Figure 5.2: This plot is of three different calculations of the signal lepton momentum in the $B^+ \rightarrow e^+ \nu_e$ (left) and $B^+ \rightarrow \mu^+ \nu_\mu$ (right) modes. The black solid line shows p_{sig}^* ; the red solid line shows $p'_{\text{sig } \ell}$ calculated using the Beam Spot method; the green dashed line shows $p'_{\text{sig } \ell}$ calculated using the Y-Average Method. These distributions were made using the signal MC samples representing Runs 1-6. Tag selection is applied, but no signal selection is applied beyond requiring that the selected signal mode is reconstructed.

Figure 5.2 shows that the two methods produce a peak at the proper place in the distribution, but the distribution from the Y-Average method is clearly narrower and has a sharper peak than the Beam Spot method. Figure 5.3 shows that the Y-Average method produces no significant difference in the background shape compared to the Beam Spot method. Therefore, we choose to use the Y-Average method as the preferred estimate of the signal lepton momentum in the signal B rest frame.

Cross-feed and Mode Separation

Studying all three fully leptonic decays of the charged B requires us to choose a value of $p'_{\text{sig } \ell}$ to divide signal events with a single reconstructed lepton into $B^+ \rightarrow \ell^+ \nu_\ell$ candidates and leptonic τ decay candidates.

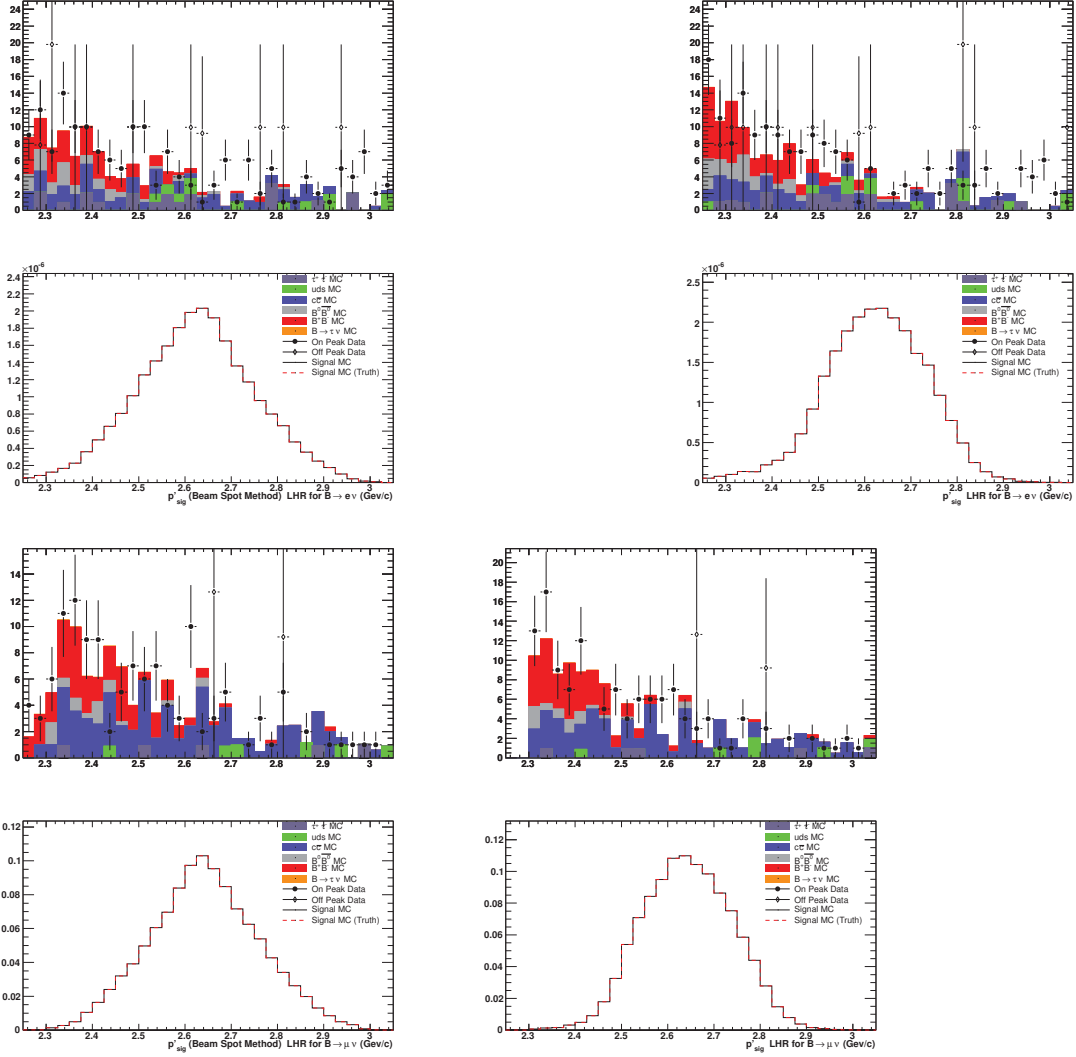


Figure 5.3: Background and signal distributions for $p'_{\text{sig } \ell}$ from both the Beam Spot and Y-Average methods. No significant difference is seen in background distributions for the two methods. All MC and data samples normalized to Run 1-6 data luminosity.

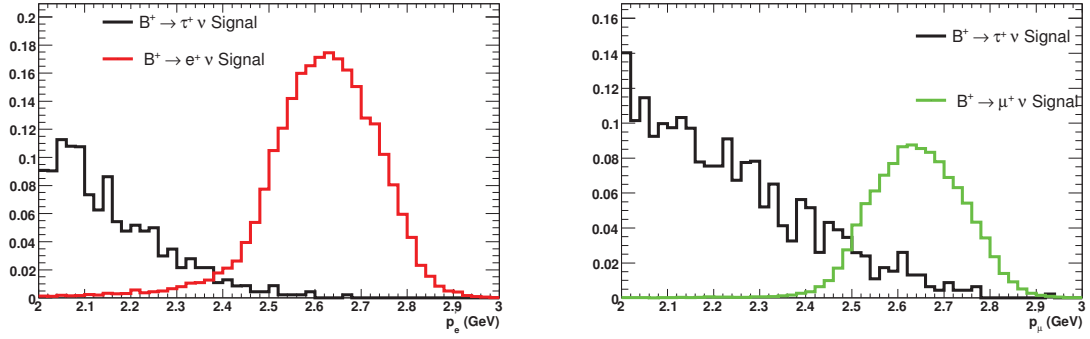


Figure 5.4: Shown are the $p'_{\text{sig } \ell}$ distributions for the regions of overlap between the $B^+ \rightarrow \tau^+ \nu_\tau$, $B^+ \rightarrow e^+ \nu_e$ (left), and $B^+ \rightarrow \mu^+ \nu_\mu$ (right) signal MC samples. All distributions are scaled to Runs 1-6 data luminosity, except $B^+ \rightarrow e^+ \nu_e$ which has its branching fraction scaled up to 1×10^{-7} .

If an event has a reconstructed signal muon candidate and $p'_{\text{sig } \ell} > 2.3 \text{ GeV}$, it is classified as a $B^+ \rightarrow \mu^+ \nu_\mu$ candidate; otherwise it is classified as a $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ candidate. If an event has a reconstructed signal electron candidate and $p'_{\text{sig } \ell} > 2.25 \text{ GeV}$, it is classified as a $B^+ \rightarrow e^+ \nu_e$ candidate; otherwise it is classified as a $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ candidate. These values are derived from Figure 5.4.

We also see that the leptonic τ decays can be a background for the other two leptonic B decays, and vice-versa. We expect to find 0.156 $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ decays misreconstructed as $B^+ \rightarrow \mu^+ \nu_\mu$ and 0.0194 $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ decays misreconstructed as $B^+ \rightarrow e^+ \nu_e$. We also expect to see 5.38×10^{-8} $B^+ \rightarrow e^+ \nu_e$ decays misreconstructed as $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and 8.04×10^{-4} $B^+ \rightarrow \mu^+ \nu_\mu$ decays misreconstructed as $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$.

5.2 Extra Energy

For this analysis, we require that all neutral clusters in the EMC have a minimum energy of 30 MeV. This requirement removes clusters from beam photons and

Mode	Figure of Merit	
	$E_{\text{extra}}^{\text{seed } 30}$ All	$E_{\text{extra}}^{\text{seed } 20}$ All
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	0.69	0.69
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	0.54	0.54
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	0.23	0.23
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	0.20	0.20
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	0.044	0.043
$B^+ \rightarrow \mu^+ \nu_\mu$	0.080	0.079
$B^+ \rightarrow e^+ \nu_e$	1.4×10^{-6}	1.4×10^{-6}

Table 5.1: Comparison of two different minimum neutral cluster energy requirements for E_{extra} . After our data sample has passed the tag selection requirements and the signal mode in question has been reconstructed, we optimize a requirement on *only* E_{extra} . Comparing the resulting Figures of Merit (same as used in §5.4) indicates which definition has the most discriminating power. The number in the superscript indicates the minimum neutral cluster energy requirement in MeV.

noise in the calorimeter crystals. In addition to the physical reasons for making the 30 MeV minimum cluster energy requirement, we perform a study to determine which definition of E_{extra} results in the best performance. We add the photon seeding method described in §4.5.1 to E_{extra} for both requirements. For each decay mode, we then make our tag selection requirements. We then optimize by scanning over the requirements on *only* E_{extra} ; this prevents bias because E_{extra} is also used in CosBY. The resulting Figures of Merit, produced using the optimization algorithm on the optimization MC sample, are shown in Table 5.1. We see a slight preference for the 30 MeV requirement in some modes and no significant difference in the others.

5.3 Likelihood Ratio

To take advantage of shape differences between variables that are not strong enough to be useful in a cut-and-count analysis, we use two Likelihood Ratios (LHRs)

that consist of several probability density functions (PDFs). Separate LHRs are generated for $B\bar{B}$ (B^+B^- , $B^0\bar{B}^0$) and continuum ($c\bar{c}$, $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $\tau^+\tau^-$) generic MC.

LHRs consist of several PDFs. A PDF is a histogram of a variable distribution that is normalized to unit area. Each bin of the histogram is treated as a probability. Two PDFs are generated for each variable in a LHR. One uses truth-matched signal MC and is treated as a signal probability distribution $P_s(x)$. The other uses generic MC and is treated as a background probability distribution $P_b(x)$. These two PDFs are combined to form a probability distribution $P_i(x)$:

$$P_i(x) = \frac{P_s(x)}{P_s(x) + P_b(x)}, \quad (5.4)$$

where each i represents a different variable. Bins in a distribution that are more likely to be background have $P_i(x)$ closer to 0; bins that are more likely to be signal tend to have $P_i(x)$ closer to 1. Each LHR is formed by multiplying all $P_i(x)$ together:

$$\text{LHR}(x) \equiv \prod_i P_i(x). \quad (5.5)$$

Ideally, a LHR is a doubly peaked distribution with background events forming a peak near zero and signal events forming a peak near 1. The PDFs are created using MC after all tag cuts have been applied. Separate PDFs are generated for each Run of each decay mode. Any given event will have one LHR for $B\bar{B}$ MC and one for Continuum MC, where the PDFs are selected based on the run and reconstructed decay mode of that event.

5.3.1 Variables Included in LHR

Two variables are reserved from use in the LHR. They are E_{extra} (§5.2) and $p'_{\text{sig } \ell}$ (§5.1). Every other variable that we thought might have some discriminating power, including obvious (eg. R_{All}^2) and more obscure variables (e.g. M_3^{min}), was considered as a PDF. Some of these are usually used as discriminating variables with selection requirements placed on them. However, we wish to take advantage of shape differences more subtle than can be accounted for with selection requirements on individual variables.

In all, 27 variables were considered for inclusion in the LHRs. Some of these are only relevant for certain decay modes, such as the mass of the putative ρ^0 resonance in the decay $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$. For each LHR in each of the reconstructed decay modes, the PDFs were generated and visually inspected for any noticeable shape difference. Four of these variables were combined to form 2-D PDFs. The remainder were used as 1-D PDFs.

The full technical descriptions of the variables follow below. Note that these descriptions, which continue until the beginning of §5.3.2, contain details and technicalities intended to be understood only by a specialist in particle physics.

One Dimensional PDFs

- **Separation Between the Signal and Tag B Vertices $\Delta z / \sigma_{\Delta z}$:**

Due to the neutrinos on both sides of the event, the vertices of the reconstructed signal tracks and neutral clusters do not correspond exactly to the true B decay points. However, the reconstructed vertices are still displaced in space while tracks from continuum processes tend to point back to the interaction point.

We calculate the displacement between the putative B vertices divided by the uncertainty on that displacement. Continuum events are distributed more strongly towards zero than true $B\bar{B}$ events.

- **Net Event Charge:**

For properly reconstructed events, the sum of all electric charges of all tracks should be zero. We noted a drop in tag efficiency from Runs 1-3 through Runs 4 and 5. We found that approximately half of this drop was due to the requirement that net charge be zero, which was enforced at the tag level. In order to avoid this drop in efficiency, the net charge of the event was moved from tag selection to signal LHR.

- **Ratio of the Second to Zeroth Fox-Wolfram Moment R_{All}^2 :**

The Fox-Wolfram moments are rotationally invariant kinematic quantities designed to quantify the shapes of events resulting from e^+e^- events. They represented as H_l , where l is the number of the moment. They are defined as

$$H_l \equiv \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} \left| \sum_i Y_l^m(\theta_i) \frac{\vec{p}_i}{\sqrt{s}} \right|^2, \quad (5.6)$$

where i runs over all hadrons in the event, Y_l^m are the spherical harmonics, theta is the angle of the momentum with respect to the z axis, and \sqrt{s} center-of-mass energy of the collision [39]. This variable is the ratio

$$R_{\text{All}}^2 \equiv \frac{H_2}{H_0}. \quad (5.7)$$

- **$\text{Cos}\theta_{B-D^0\ell}$:**

Since this variable presumes the existence of a true signal B decaying with unreconstructed neutrinos, signal events are more common in the physical region, but background events have a larger tail into the non-physical region.

- **D^0 Decay Mode:**

Signal and background events have different distributions of D^0 decay mode. For instance, true signal events are found recoiling against a $D^0 \rightarrow K^- \pi^+$ more often than $B\bar{B}$ background events.

- **Center of Mass Momentum for the Tag K^+ ($p_{\text{Tag}K^+}^*$)**

- **Center of Mass Momentum for the Tag ℓ^+ ($p_{\text{Tag}\ell^+}^*$)**

- **Tag K^\pm Selector:**

If the tag D^0 decay produces a charged K , we do not simply impose a specific particle identification selector. Instead, we create a PDF of the strictest selector that allows the track to pass. This PDF consists of integers in the range 10-16. If the putative tag K passes only the loosest selection, it is assigned the value 10. If it passes stricter selectors, it is assigned the values 11 through 14 depending on the strictness of the selector passed. The value 16 is assigned to K_s^0 candidates. A higher fraction of true signal than background events pass `KLHVeryTight`; conversely, a higher fraction of background than signal events pass `KLHNotPion`.

- **Minimum Invariant Mass of Any 2 Reconstructed Tracks (M_2^{min}):**

Since the minimum invariant mass of any three tracks M_3^{min} was a useful variable in our analysis based on Runs 1-5 data, we decided to try using M_2^{min} . As the

name suggests, it is the smallest invariant mass produced by any combination of two tracks used to reconstruct the signal B .

- **$m_{\ell\ell}$:**

If an event contains two putative lepton tracks on the signal side, their invariant mass is calculated and stored as $m_{\ell\ell}$. This variable was originally developed to remove pair-produced leptons. It is obviously highly correlated with M_2^{\min} , so both variables are never used in the same LHR.

- **Tight μ Selector:**

If an event is reconstructed as option 3a from §4.5.2 and does not have a $p'_{\text{sig } \ell}$ high enough to classify it as $B^+ \rightarrow \mu^+ \nu_\mu$, it is classified as $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$. We considered placing a tighter cut on μ PID, but found that a cut produced an unacceptable loss in efficiency. Instead, we create a binary PDF. If the track passes `muNNTight`, this PDF is set to 1, otherwise it is set to zero. More Continuum background than signal accumulates at zero.

- **Signal K^\pm Selector:**

If charged kaons that pass the `KLHTight` particle identification requirements are reconstructed in the signal B of an event, that event is assigned to another decay mode. This PDF is set to zero if a signal track passes `KLHVeryLoose`. Note that this PDF was only tested for the four single-track τ decay modes. The $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ uses the $a_1 \ell^+ / K^+$ Selector.

- **a_1 Resonance ℓ^+ / K^+ Selector:**

In order to reject any leptons or kaons that are misreconstructed as pions in the decay $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$, a K and lepton selector is included in the LHR. If any of the three putative π^+ tracks passes `PidLHElectrons`, `muNNTight`, or `KLHTight`, this PDF is set to zero.

- **Vertex Status:**

For those τ decay channels that involve multiple tracks or neutral clusters, a vertex is created. The quality of that vertex is reported as an integer from zero through four, which is included as PDF. Zero indicates that the vertex fit was successful. One indicates that the fit did not converge. Two indicates an error due to bad input. Three indicates that the fit failed for other reasons. Four indicates that the vertex fit was not performed.

- **Reconstructed Mass of the τ Daughter:**

The decays $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$ and $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ often proceed through the ρ^+ and a_1^+ resonances, respectively. For true signal events, a peak at the resonance mass appears in the invariant mass distribution of the signal tracks and neutrals. Background events yield a flat or linear distribution. This PDF is only useful for τ reconstructed from multiple tracks and neutrals.

- **m_{ρ^0} :**

In the case of $\tau \rightarrow a_1^+\bar{\nu}_\tau$, we begin with events from the list `a1CLoose` and refine them using `TreeFitter`. This list assumes that the decay proceeds through the resonance $a_1^+ \rightarrow \rho^0\pi^+$, where $\rho^0 \rightarrow \pi^+\pi^-$. In the mass of this resonance (m_{ρ^0}), a peak appears for signal events that does not appear for background events.

-
- $p_{\rho^0}^*$:

We can also calculate the center of mass momentum for the putative resonant ρ^0 . Signal events tend to have larger values for $p_{\rho^0}^*$ than background events.

- **Center of Mass Momentum for the π^+ in $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$ ($p_{\pi^+}^*$)**
- **Center of Mass Momentum for the π^0 in $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$ ($p_{\pi^0}^*$)**
- **Center of Mass Momentum for the first π in $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ ($p_{\pi_1}^*$)**
- **Center of Mass Momentum for the second π in $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ ($p_{\pi_2}^*$)**
- **Center of Mass Momentum for the third π in $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ ($p_{\pi_3}^*$)**

Two Dimensional PDFs

As the name suggests, 2-D PDFs are two-dimensional histograms that contain two variable distributions. The histogram is normalized so that all of its content integrates to unity. Each bin is treated as a probability, and those probabilities are multiplied into the LHRs for each event.

We use 2-D PDFs in cases where we want to exploit two variables that are highly correlated or that have a stronger separation when combined than when separate.

- $R_{\text{cont.}}$:

In the earlier versions of the semileptonic $B^+ \rightarrow \tau^+\nu_\tau$ analysis, we found that Continuum background accumulated in a small region of the $\left|\cos(\theta_{\vec{T}_{D\ell},\text{signal}})\right|$ - M_3^{min} plane [25]. $\cos(\theta_{\vec{T}_{D\ell},\text{signal}})$ is the cosine of the angle between the momentum of the signal candidate and the thrust vector of the tag B in the CM frame.

M_3^{min} is the minimum invariant mass of any three tracks in the event, regardless of whether they are part of the tag or signal B .

- **m_{miss} vs. $\cos(p_\theta^{miss})$ (m_{miss} -dir):**

The total invariant mass and initial momentum of each event are well known from beam information. Since neutrinos escape undetected from each event, we expect the total reconstructed invariant mass to be less than what the beam provides. The difference is called, rather uncreatively, the missing mass (m_{miss}). The missing momentum (p^{miss}) is similarly defined. In the spherical coordinate system used by *BABAR*, θ is the angle with respect to the beam line, and (p_θ^{miss}) is the angle between the missing momentum vector and the beam line.

The PEP-II beam pipe corresponds to values of $\cos(p_\theta^{miss})$ near ± 1 . One source of background are events in which real particles are lost down the beam pipe, which is outside of the detector coverage. Since they are not reconstructed, they can be misinterpreted as neutrinos. This PDF allows the LHR to account for background events that have high missing mass but are likely to have lost particles down the beam pipe.

- **$\cos \theta'_{\tau-Y}$ vs. $|\vec{p}'_Y|$ (Cos τ Y-pY):**

$\cos \theta'_{\tau-Y}$ is the equivalent of $\cos \theta_{B-D^0\ell}$ for the signal B . Y represents all of the reconstructed daughters of the signal B , and $\theta'_{\tau-Y}$ is the calculated angle between Y and τ in the signal B rest frame.

$|\vec{p}'_Y|$ is the calculated momentum of Y in the signal B rest frame; it is the same as $p'_{sig\ell}$ defined to include the hadronic τ decay modes. Since $p'_{sig\ell}$ is such a powerful variable for selected $B^+ \rightarrow e^+\nu_e$ and $B^+ \rightarrow \mu^+\nu_\mu$, we tested it for

$B^+ \rightarrow \tau^+ \nu_\tau$. We found that it was not very useful unless combined in this 2D PDF. This PDF is not used in $B^+ \rightarrow e^+ \nu_e$ or $B^+ \rightarrow \mu^+ \nu_\mu$ reconstruction; $p'_{\text{sig } \ell}$ is kept as a separate variable.

5.3.2 Final PDF Selection

All of the relevant PDFs were visually inspected to see if any shape differences were apparent. A few were rejected based on this inspection. Those that remained were tested quantitatively. For each decay mode, we tested each PDF in both LHRs using Runs 1-6 MC samples.

For each of the 14 LHRs (7 decay modes \times 2 background types), a fixed signal yield was chosen. Each LHR was tested using signal MC samples and the background MC samples it was designed to reject. The tests were performed on MC after the appropriate decay mode was selected and with E_{extra} required to be less than 1.5 GeV.

For a given LHR, a baseline performance was calculated by using all variables remaining after the visual inspection. A cut was placed on the LHR in question to produce the chosen signal yield, and a figure of merit (FOM) calculated. No cut was placed on the other LHR. Each PDF was tested by removing it (and *only* it) from the LHR. The LHR was scanned again until the chosen signal yield was reached, and the FOM was recalculated. If removing the PDF increased the PDF by a statistically significant amount, it was not included in the final LHR.

If any two PDFs were correlated with each other (e.g. $m_{\ell\ell}$ and M_2^{min}), we selected only one to be included in the LHR. We also rejected any PDFs that, for physical reasons, should have no discriminating power. For instance, all tag-side variables were removed from all $LHR_{B\bar{B}}$.

As a cross-check, we performed the same test where the goal of the scan was maximizing significance rather than matching a signal yield. This produced no significant discrepancies. However, it required us to change our FOM. Maximizing standard significance, which is defined as

$$\text{FOM}_{\text{Significance}} = N_{\text{sig}} / \sqrt{N_{\text{sig}} + N_{\text{BG}}}, \quad (5.8)$$

can lead to unreasonably tight cuts that reduce the background (N_{BG}) to zero and leave very little signal.

Instead, we use the Punzi Figure of Merit [40]

$$\text{FOM}_{\text{Punzi}} = \frac{N_{\text{sig}}}{N_{\sigma}/2 + \sqrt{b}}. \quad (5.9)$$

N_{sig} is the signal yield; N_{BG} is the background yield. N_{σ} is the desired number of standard deviations desired from the result.

For the LHRs used for $B^+ \rightarrow \tau^+ \nu_{\tau}$, $N_{\sigma} = 6$. We do not use the Punzi FOM to optimize the final cuts; we use Significance. For the LHRs used for $B^+ \rightarrow e^+ \nu_e$ and $B^+ \rightarrow \mu^+ \nu_{\mu}$, $N_{\sigma} = 3$.

The results of this selection process are shown in Table 5.2.

$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	$B^+ \rightarrow \mu^+ \nu_\mu$	$B^+ \rightarrow e^+ \nu_e$
Net Charge R_{all}^2 $p_{\text{Tag}\ell}^*$ Tag K^+ Sel. Level D^0 Dec. Mode						
$p_{\text{Tag}K^+}^*$ CosBY						
$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$	$\Delta z / \sigma_{\Delta z}$
$m_{\ell\ell}$	m_2^{\min}	-	m_2^{\min}	-	$m_{\ell\ell}$	$m_{\ell\ell}$
$m_{\text{miss}}\text{-dir}$	$m_{\text{miss}}\text{-dir}$	-	-	m_2^{\min}	$m_{\text{miss}}\text{-dir}$	$m_{\text{miss}}\text{-dir}$
-	Signal K^+ Sel.	-	Signal K^+ Sel.	$a_1 \ell^+ / K^+$ Sel.	-	-
-	Tight μ Sel.	-	-	-	-	-
Cos τ Y-pY	Cos τ Y-pY	Cos τ Y-pY	-	Cos τ Y-pY	-	-
-	-	-	$p_{\pi^+}^*$	$p_{\pi_1^+}^*$	-	-
-	-	-	$p_{\pi^0}^*$	-	-	-
-	-	-	-	$p_{\pi_3^+}^*$	-	-
-	-	-	m_{ρ^+}	$m_{a_1^+}$	-	-
-	-	-	Vertex Status	Vertex Status	$R_{\text{cont.}}$	-
-	-	-	-	-	-	-

Table 5.2: This is a list of all variables used as signal PDFs for the various decay modes. Variables in **bold** are used for both Continuum and $B\bar{B}$ background. Variables in plain roman font are used for Continuum only. Variables in *italics* are used for $B\bar{B}$ only.

5.4 Optimization Procedure

We use three (E_{extra} , $LHR_{B\bar{B}}$, and $LHR_{\text{cont.}}$) variables in our final selection of the five τ decay modes. For $B^+ \rightarrow e^+\nu_e$ and $B^+ \rightarrow \mu^+\nu_\mu$, we add a fourth variable $p'_{\text{sig } \ell}$. The final requirements for these variables are obtained by optimizing on a figure of merit. For the τ decay modes, we choose the FOM to be significance ($N_{\text{sig}}/\sqrt{(N_{\text{sig}} + N_{\text{BG}})}$). For the other two leptonic B decay modes, we use the Punzi FOM (Equation 5.9). The Punzi FOM is better suited to searches for small signals on small backgrounds. It is designed to prevent optimization algorithms from reducing the background to zero and creating a undesirably low signal yield. This produces a superficially high FOM with standard significance but unacceptably reduces the signal yield. In the $B^+ \rightarrow \tau^+\nu_\tau$ mode, both signal and background are large enough that optimizations perform well with standard significance. In the Punzi FOM, we set $N_\sigma = 3$.

N_{sig} is defined as the number of events in the MC signal samples that were generated by the channel in question. Our optimization procedure scans over each variable. By definition, the LHR s are constrained to values in the interval $[0.0,1.0]$. We scan E_{extra} from 0.0 to 1.0 GeV and $p'_{\text{sig } \ell}$ from 2.2 to 3.2 GeV. The actual maximum values scanned are increased by 0.01 to account for any binning effects, and the full ranges are divided into 101 bins.

The chosen FOM is calculated based on the number of events that pass when the full range of all variables is allowed. One variable is then modified by sequentially removing bins beginning at the upper (E_{extra}) or lower ($p'_{\text{sig } \ell}$, $LHR_{B\bar{B}}$, $LHR_{\text{cont.}}$) end of the range, with the other variables fixed. The point that maximizes the FOM for this variable is then saved. The next variable is scanned in the same manner with

the first variable constrained to be within the saved range. This process is repeated for all variables.

The procedure is taken through two more iterations. The second iteration begins with all the variables in the ranges constrained from the first iteration. One variable is scanned again with all remaining variables constrained to the ranges from the first iteration. This process is repeated for all remaining variables. Each one is scanned with the remaining variables constrained by the previous optimization, whether it is from the first iteration or this one. A third iteration is performed in the same manner as the second. The results are shown in Table 5.3.

The values shown in Table 5.3 were produced by keeping the variables in the same order for each iteration. This could cause us to find a local, rather than global, maximum because the phase space is always limited by the fixed optimum value of the first variable. In order to remove this flaw, the optimization was re-run with the order of variables randomized in each iteration. For the $\tau^+ \rightarrow \pi^+\bar{\nu}_\tau$, $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$, $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$, and $B^+ \rightarrow e^+\nu_e$ modes, this produced no change in cuts or optimized FOM. For the $\tau^+ \rightarrow e^+\nu_e\bar{\nu}_\tau$ and $\tau^+ \rightarrow \mu^+\nu_\mu\bar{\nu}_\tau$ modes, some optimization values changed. However, the changes were by no more than one bin, and the variation in the optimized FOM was well within uncertainty. For the $B^+ \rightarrow \mu^+\nu_\mu$ mode, we see significant changes in the optimized values on E_{extra} and $p'_{\text{sig } \ell}$. No upper bound is placed on $p'_{\text{sig } \ell}$ by the randomized optimization. The FOM from the randomized optimization (0.105) is lower than the FOM produced by the original optimization (0.112). Therefore, we shall keep the optimized values from the original optimization.

Mode	E_{extra}	$LHR_{B\bar{B}}$	$LHR_{\text{cont.}}$	$p'_{\text{sig } \ell}$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	[0,0.24] GeV	[0.74,1]	[0.16,1]	[0.00,2.25] GeV
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	[0,0.24] GeV	[0.14,1]	[0.72,1]	[0.00,2.30] GeV
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	[0,0.35] GeV	[0.57,1]	[0.8,1]	-
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	[0,0.24] GeV	[0.97,1]	[0.95,1]	-
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	[0,0.31] GeV	[0.97,1]	[0.93,1]	-
$B^+ \rightarrow \mu^+ \nu_\mu$	[0,0.72] GeV	[0.33,1]	[0.75,1]	[2.45,2.92] GeV
$B^+ \rightarrow e^+ \nu_e$	[0,0.57] GeV	[0.00,1]	[0.01,1]	[2.52,3.02] GeV

Table 5.3: Optimized ranges from which we accept signal candidates, which are mostly given by our optimization procedure. The exceptions are the $p'_{\text{sig } \ell}$ ranges for the two leptonic τ decay modes, which were chosen as described in §5.1, and the upper limits of the $p'_{\text{sig } \ell}$ ranges for the other two leptonic B decays, which were chosen to incorporate almost all of the signal MC samples.

5.4.1 Photon Pair Conversion

After we examined the data in the signal region, we discovered an excess of data above our MC simulations at low values of $m_{\ell\ell}$, which is the minimum invariant mass of any two leptons. To remove these events, which are probably due to unmodeled photon pair conversion, we decided remove all events below a certain value of $m_{\ell\ell}$ after all other analysis cuts had been applied. This value was chosen using only signal and background MC simulations with the optimization technique described in this section. The result excludes events in the $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ channel with $m_{\ell\ell} < 0.29 \text{ GeV}/c^2$. All efficiencies and yields in this note have been updated to include the effects of this requirement.

Mode	Signal Prediction	BG Prediction from E_{extra} SB	FOM
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	15.14 ± 0.33	91.4 ± 12.7	1.47 ± 0.09
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	12.09 ± 0.29	137.2 ± 13.3	0.99 ± 0.05
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	18.96 ± 0.37	233.0 ± 18.9	1.19 ± 0.05
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	6.56 ± 0.22	59.2 ± 8.8	0.81 ± 0.06
$B^+ \rightarrow \tau^+ \nu_\tau$	53.03 ± 0.63	533.9 ± 31.3	2.19 ± 0.06
$B^+ \rightarrow \mu^+ \nu_\mu$	0.74 ± 0.01	15.2 ± 9.9	0.14 ± 0.09
$B^+ \rightarrow e^+ \nu_e$	$(1.84 \pm 0.02) \times 10^{-5}$	24.0 ± 11.2	$(0.29 \pm 0.15) \times 10^{-5}$

Table 5.4: Signal predictions using optimized cuts. BG predictions from the E_{extra} sideband, as described in 6.2.2. The figure of merit (FOM) is calculated from the second and third columns.

5.4.2 Signal and Overall Efficiency

In an analysis with a tag and signal B , one can define efficiencies for the tag (ε_{tag}) and signal (ε_{sig}) sides of the event. The separation between these two efficiencies is somewhat arbitrary. Both efficiencies are measured using the signal MC sample.

In general ε_{tag} is defined as the quotient of the unnormalized yield after the final tag selection and the total number of events generated in the MC sample. Similarly, ε_{sig} is defined as the quotient of the yield after the optimized signal selection and the yield after final tag selection. Thus, we can define an overall efficiency $\varepsilon \equiv \varepsilon_{\text{sig}} \times \varepsilon_{\text{tag}}$ that is the quotient of the yield after final signal selection and the total number of events generated in the MC sample.

The arbitrariness results from freedom of the analyst to choose the division between the tag and signal selections. In this analysis, we choose to define the tag selection to include only the variables defined in §4.6.2. This constraint is looser than

what we placed on the tag side in the Runs 1-5 analysis, so the efficiency is much higher now, but the two are not comparable.

Mode	Yield	Number Generated ($\times \frac{1}{3}$)	ε_{sig}	$\varepsilon(\times 10^{-4})$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	2186	6476333	$(1.987 \pm 0.043)\%$	3.38 ± 0.07
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	1771	6476333	$(1.610 \pm 0.038)\%$	2.73 ± 0.06
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	2729	6476333	$(2.48 \pm 0.05)\%$	4.21 ± 0.08
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	945	6476333	$(0.859 \pm 0.028)\%$	1.46 ± 0.05
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	332	6476333	$(0.302 \pm 0.017)\%$	0.51 ± 0.03
$B^+ \rightarrow \tau^+ \nu_\tau$	8005	6476333	$(7.28 \pm 0.08)\%$	12.36 ± 0.14
$B^+ \rightarrow \mu^+ \nu_\mu$	8236	2531333	$(30.92 \pm 0.36)\%$	32.54 ± 0.36
$B^+ \rightarrow e^+ \nu_e$	10470	2589667	$(36.98 \pm 0.38)\%$	40.43 ± 0.40

Table 5.5: Overall efficiency ($\varepsilon \equiv \varepsilon_{sig} \times \varepsilon_{tag}$) of optimized signal selection for all modes in Runs 1-6. These are unnormalized numbers of events for the relevant signal MC Test samples. The Yield is the number of events passing the optimized selection. The third column is the total number of events generated for the signal MC sample divided by 3; the division counterbalances the reduction in yield caused by using only the Test MC samples. The fourth column is the quotient of the Yield over the Number Generated, which is, by definition the same as the product of the tag and signal efficiencies.

Table 5.5 shows the overall efficiency for all modes with all runs combined for each signal MC sample. The lower efficiency for $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ relative to $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ results from muon reconstruction being less efficient than electron reconstruction from low momenta, as seen in Figure 5.5. Also shown in Figure 5.5 are the various true contributions to the $\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$. Since this is the default category for single charged tracks, we see that several other particles, including misidentified muons, contribute to the low-momentum part of the spectrum.

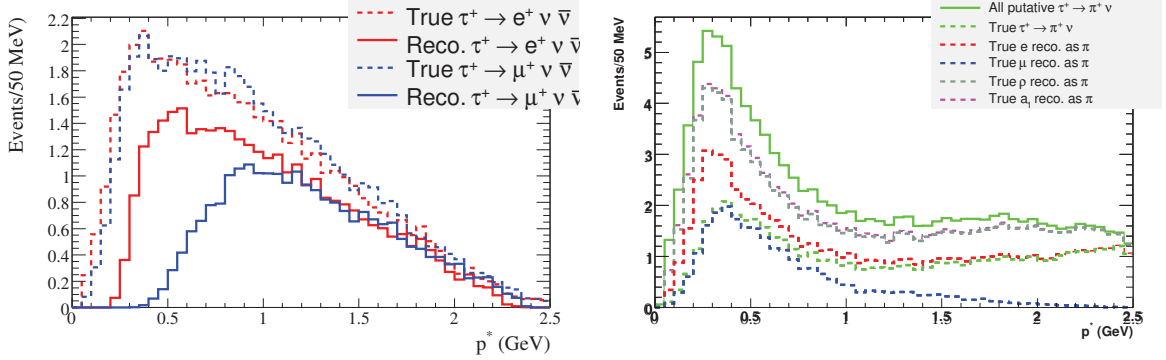


Figure 5.5: At left, we have the momentum spectra, in the center of mass frame, for the signal leptons in the $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ and $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ modes. We can see that muon identification becomes less efficient than electron identification below about 1.2 GeV. At right, we have the same momentum spectrum with distributions of the different modes reconstructed as pions. This shows that, at low momenta many of the true muons and electrons are reconstructed as pions. Also, we see a sizable contribution from real $\rho^+ \rightarrow \pi^+ \pi^0$ where we have missed the π^0 . All distributions are taken from the signal MC sample and scaled to Run 1-6 data luminosity. The dashed lines are the true distributions for the indicated particle that pass the tag selection. The solid lines show the distributions that pass the tag cuts and are reconstructed as the given mode.

CHAPTER 6

CONTROL SAMPLES AND SYSTEMATIC CORRECTIONS

In order to determine if we see real signal events when we examine the signal region of data, we need a reliable background prediction. We would like to use data as much as possible to make this prediction. This is accomplished using data with variable values adjacent to our signal region. These are called sidebands. The data in the sidebands are used as control samples.

6.1 Control Samples

Control Samples are subsets of the generic background MC samples (§4.3.1) and on-peak data that do not pass some or all of our final optimized signal requirements. They are used to make background predictions, assess the level of agreement between simulated and real data, and assess other systematic uncertainties. All of our control samples and background predictions are defined in this section. Systematic uncertainties are derived and described in §8.

6.1.1 D^0 Mass Sideband Sample

The combinatoric background, after all selection criteria have been applied, can be estimated by extrapolating the events in the D^0 mass sideband into the D^0 mass signal

region. The lower sideband is defined as the region of D^0 mass $[1.7835, 1.8185]$ GeV ($[1.8175, 1.8375]$ GeV) window around the nominal D^0 mass in $D^0 \rightarrow K^-\pi^+\pi^0$ (all other) decays. The upper sideband is defined as the region of D^0 mass $[1.9015, 1.9455]$ GeV ($[1.8915, 1.9115]$ GeV) window around the nominal D^0 mass in $D^0 \rightarrow K^-\pi^+\pi^0$ (all other) decays.

This sample is also an excellent means to check the agreement of the extra energy simulation in a mass sideband region, while examining the full range of extra energy.

6.1.2 E_{extra} Sideband Sample

We define this “sideband” region as $E_{\text{extra}} \geq 0.6$ GeV, except for $B^+ \rightarrow \mu^+\nu_\mu$ where it is defined as $E_{\text{extra}} \geq 0.72$ GeV. The “signal region” is defined separately for each signal mode.

6.1.3 LHR Sideband Samples

Both likelihood ratios have the range $[0,1]$ by definition. The signal regions, as optimized, are shown in Table 5.3. The sideband regions are defined as the range within $[0,1]$ and outside of the signal region. Due to loose optimized criteria, not all modes have useful sidebands.

6.1.4 $p'_{\text{sig } \ell}$ Sideband Sample

The $p'_{\text{sig } \ell}$ sideband only exists for the $B^+ \rightarrow e^+\nu_e$ and $B^+ \rightarrow \mu^+\nu_\mu$ modes. For these two modes, the sidebands are defined by removing the signal region and the region assigned to the leptonic τ decay mode. This gives a lower and upper sideband, and we use all events in both.

6.1.5 Double Tag Sample

Events where both of the B-mesons are reconstructed in tagging modes, $B^- \rightarrow D^{(*)0} \ell^- \bar{\nu}_\ell$ vs $B^+ \rightarrow \bar{D}^{(*)0} \ell^+ \nu_\ell$, are referred to as “double tag” events. Due to both the large branching fraction of $D\ell\nu X$ decays and the high tagging efficiency for reconstructing these events, a sizable sample of such events are available in the on-resonance dataset.

For double tag events, we first applied the tag selection requirements described in section 4.6 to both of the tag candidates. We require that the two tag-B candidates do not share any tracks or neutrals. If there are more than two such non-overlapping tag B candidates in the event, then the best two are selected using vertex probability selection. This procedure resulted in noticeable shape and yield discrepancies between data and MC, as seen in Figure 6.1.

In order to improve agreement, several additional selection criteria were imposed. These are based on the selection criteria from the previous analysis (Runs 1-5). We required that $-2.0 < \cos \theta_{B-D^0\ell} < 1.1$ for both B decays and the event has zero net charge. The resulting distribution is shown in Figure 6.1. The yield disagreement has improved, and the agreement in shape is better. To test how much of the disagreement is due to yield and how much is due to shape, we normalize data and MC to unit area with the same cuts as in Figure 6.1. The resulting distribution is shown in Figure 6.1; the excellent agreement shows that disagreement is only due to yield. Since this sample will be used to calculate Data/MC ratios rather than absolute yields, we will use this modified double tag sample as a control sample for tag efficiency and E_{extra} shape.

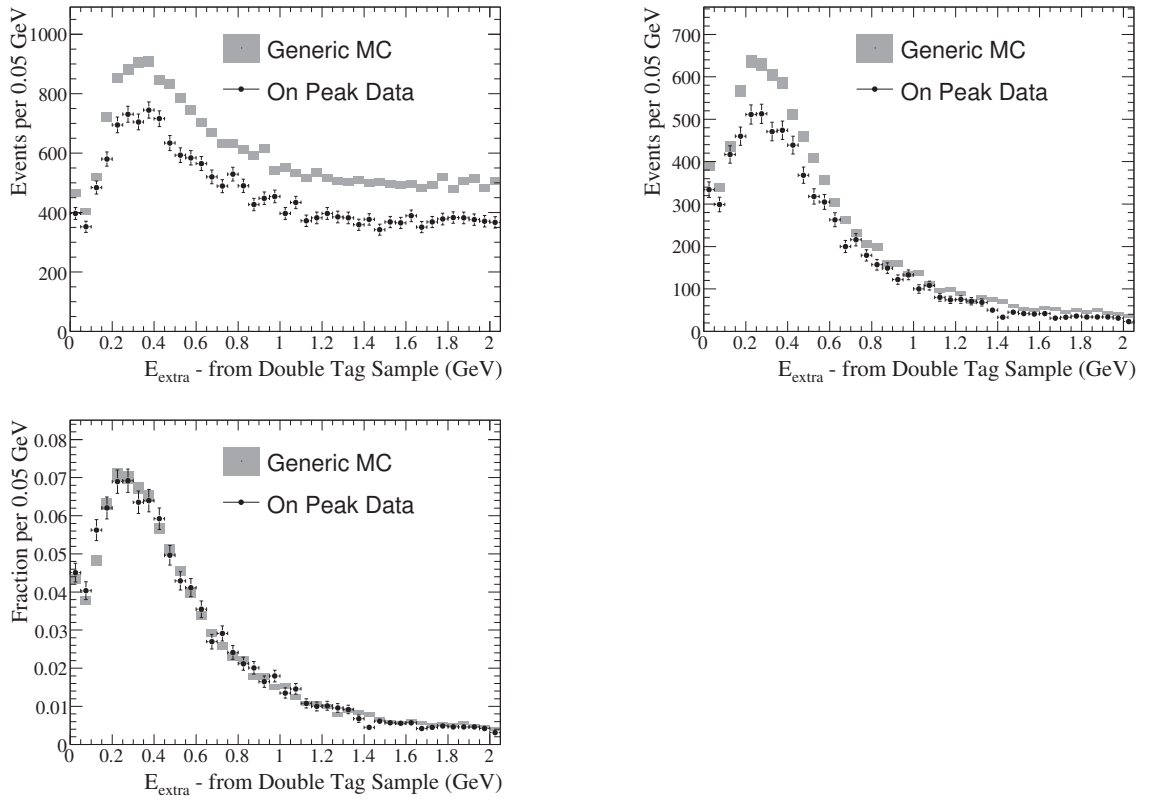


Figure 6.1: Distribution of E_{extra} in double tags with different sets of selection requirements. Data and MC are normalized to Run 1-6 luminosity in the top left and top right graphs. In the bottom graph graph the same cuts are applied as in the top right graph, but Data and MC are normalized to unity.

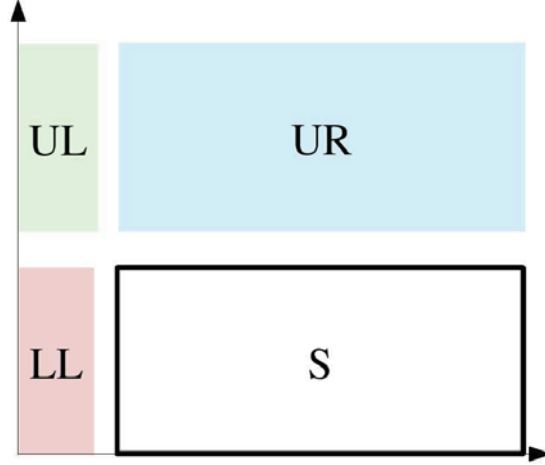


Figure 6.2: Proposed construction for 2D background prediction

6.2 Background Prediction

6.2.1 Background Prediction using 2D Side Bands

We construct a plane, as seen in figure 6.2. The axes of the plane could be any two of the four final selection variables (E_{extra} , $LHR_{\text{cont.}}$, $LHR_{B\bar{B}}$, and $p'_{\text{sig } \ell}$). The plane would be divided into a signal region and three sidebands (UL, UR, and LL). The prediction would be given by

$$S = \frac{UR}{UL} LL \quad (6.1)$$

This method has the advantage of taking the background prediction entirely from data. Most of these results are systematically lower than the other sidebands; the cause of this is unknown, but we will not be using this method any further.

6.2.2 Background Prediction using 1D Side Bands

D^0 Mass

The sideband regions must be relatively flat if this method is to work. If the sideband region below the D^0 mass peak (the lower sideband) and the sideband region above the D^0 mass peak (the upper sideband) have different yields, we average the two when performing the extrapolation.

Since we average the yields in the upper and lower sidebands, this provides us with a sideband region that is half as wide as the D^0 mass signal region. If the average of the upper and lower sidebands is denoted $\langle N^{SB} \rangle$, then we can compute the estimated combinatoric background in the signal region as follows:

$$2 \times \langle N^{SB} \rangle = N_{bkg}^{comb.} \quad (6.2)$$

From the preceding discussion we have the purely combinatoric background (meaning the D^0 tag meson is fake) from the D^0 mass sideband projections. In both data and MC, the total background in the extra energy signal box from the D^0 mass distribution is given by:

$$N_{bkg}^{total} = N_{bkg}^{peaking} + N_{bkg}^{comb.}, \quad (6.3)$$

where $N_{bkg}^{peaking}$ ($N_{bkg}^{comb.}$) is the part of the total background which arises from events with real (fake) D^0 mesons in the tag B. We can write this equation for both data and MC. From MC, we can subtract the extrapolated combination background in the D^0 signal region from the total background in that region to obtain the peaking component in the MC. We can then add that peaking MC component to the data

Mode	MC Full	MC From SB	Data From SB	BG Prediction
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	98 ± 11	52 ± 8	56 ± 7	102 ± 15
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	136 ± 12	42 ± 8	52 ± 7	146 ± 16
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	212 ± 17	40 ± 7	67 ± 8	239 ± 20
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	62 ± 9	23 ± 5	19 ± 4	58 ± 11
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	123 ± 12	68 ± 9	65 ± 8	120 ± 17
$B^+ \rightarrow \mu^+ \nu_\mu$	12 ± 5	2.7 ± 1.9	5.0 ± 2.2	14 ± 6
$B^+ \rightarrow e^+ \nu_e$	15 ± 5	14 ± 6	13.0 ± 3.6	14 ± 8

Table 6.1: BG predictions using the D^0 mass sideband and peaking generic MC.

combinatoric component, to get an effective total background estimate,

$$N_{bkg}^{peaking}(MC) = N_{bkg}^{total} - N_{bkg}^{comb.}, \quad (6.4)$$

$$N_{bkg}^{total}(predicted) = N_{bkg}^{peaking}(MC) + N_{bkg}^{comb.}(DATA). \quad (6.5)$$

We show the results of this prediction in Table 6.1. This estimate can then be compared to that obtained from the extra energy sideband extrapolation. Consistency from these two methods would suggest a robust prediction of the background. The assumption in this method is that the MC models the D^0 mass peaking background accurately, and that this D^0 peaking background corresponds to the “peaking” background in the extra energy signal regions.

We present the results of this comparison in Table 6.6. The results suggest that in the four modes we intend to use for this result, the background prediction from the extra energy sideband is reliable.

E_{extra}

For each mode, after applying the optimized final selections (except E_{extra} in the case of the sideband), the number of MC events in the signal region ($N_{\text{MC,Sig}}$) and side band ($N_{\text{MC,SideB}}$) are counted and their ratio (R^{MC}) is obtained:

$$R^{\text{MC}} = \frac{N_{\text{MC,Sig}}}{N_{\text{MC,SideB}}}$$

Using the number of data events in the side band ($N_{\text{data,SideB}}$) and the ratio R^{MC} , the number of expected background events in the signal region in data ($N_{\text{exp,Sig}}$) is estimated:

$$N_{\text{exp,Sig}} = N_{\text{data,SideB}} \cdot R^{\text{MC}}$$

Table 6.2 shows the background predictions from the E_{extra} sideband. The results of all sideband background predictions are compared in Table 6.6.

$LHR_{\text{cont.}}$ Sideband Sample

With the definitions of sideband and signal regions from §6.1.3, the background predictions are obtained in the same way as with the E_{extra} sideband (§6.2.2). The results are shown in Table 6.3 and compared with the other sideband predictions in Table 6.6.

$LHR_{B\bar{B}}$ Sideband Sample

The background predictions for the $LHR_{B\bar{B}}$ sidebands are obtained in the same way as the $LHR_{\text{cont.}}$ sidebands. The results are shown in Table 6.4 and compared with the other sideband predictions in Table 6.6.

Mode	R_{MC}	$N_{\text{data,SideB}}$	$N_{\text{exp,Sig}}$	$N_{MC,Sig}$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	0.322 ± 0.040	284 ± 17	91 ± 13	98 ± 11
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	0.128 ± 0.012	1070 ± 33	137 ± 13	136 ± 12
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	0.033 ± 0.003	6994 ± 84	233 ± 19	212 ± 17
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	0.035 ± 0.005	1684 ± 41	59 ± 9	62 ± 9
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	0.052 ± 0.005	2339 ± 48	122 ± 13	123 ± 12
$B^+ \rightarrow \mu^+ \nu_\mu$	1.1 ± 0.6	14.0 ± 3.7	15 ± 10	12 ± 5
$B^+ \rightarrow e^+ \nu_e$	0.57 ± 0.25	42 ± 6	24 ± 11	15 ± 5

Table 6.2: BG Predictions from E_{extra} sideband. R_{MC} is the ratio of events in the sideband to events in the signal region of E_{extra} in the background MC. $N_{\text{data,SideB}}$ is the number of events in the E_{extra} sideband in data. $N_{MC,Sig}$ is the number of normalized events in the E_{extra} signal region of the background MC samples. This is the background prediction taken solely from the MC samples. $N_{\text{exp,Sig}}$ is the product of R_{MC} and $N_{\text{data,SideB}}$; it is the background prediction extrapolated from the data sideband using the MC samples.

Mode	R_{MC}	$N_{\text{data,SideB}}$	$N_{\text{exp,Sig}}$	$N_{MC,Sig}$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	42 ± 31	3.0 ± 1.7	127 ± 119	98 ± 11
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	3.9 ± 0.8	49 ± 7	192 ± 49	136 ± 12
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	0.374 ± 0.038	611 ± 25	229 ± 25	212 ± 17
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	0.54 ± 0.11	97 ± 10	53 ± 12	62 ± 9
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	0.44 ± 0.05	268 ± 16	117 ± 16	123 ± 12
$B^+ \rightarrow \mu^+ \nu_\mu$	0.9 ± 0.5	13.0 ± 3.6	12 ± 7	12 ± 5

Table 6.3: BG Predictions from LHR_{cont} sideband. All variables are as described in the caption of Table 6.2.

Mode	R_{MC}	$N_{\text{data,SideB}}$	$N_{\text{exp,Sig}}$	$N_{MC,\text{Sig}}$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	0.874 ± 0.126	114.000 ± 10.677	99.681 ± 17.148	98 ± 11
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	19.822 ± 7.701	4.000 ± 2.000	79.286 ± 50.203	136 ± 12
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	8.742 ± 1.969	32.000 ± 5.657	279.755 ± 80.106	212 ± 17
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	0.819 ± 0.152	79.000 ± 8.888	64.662 ± 14.023	62 ± 9
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	0.147 ± 0.015	864.000 ± 29.394	126.966 ± 14.032	123 ± 12
$B^+ \rightarrow \mu^+ \nu_\mu$	7.382 ± 7.988	2.000 ± 1.414	14.763 ± 19.084	12 ± 5

Table 6.4: BG Predictions from $LHR_{B\bar{B}}$ sideband. All variables are as described in the caption of Table 6.2.

Mode	R_{MC}	$N_{\text{data,SideB}}$	$N_{\text{exp,Sig}}$	$N_{MC,\text{Sig}}$
$B^+ \rightarrow \mu^+ \nu_\mu$	0.796 ± 0.432	16.000 ± 4.000	12.731 ± 7.608	11.515 ± 4.986
$B^+ \rightarrow e^+ \nu_e$	0.875 ± 0.424	40.000 ± 6.325	35.010 ± 17.826	14.562 ± 5.316

Table 6.5: BG Predictions from $p'_{\text{sig } \ell}$ sideband. All variables are as described in the caption of Table 6.2.

$p'_{\text{sig } \ell}$ Sideband Sample

With the definition of sideband and signal regions from §6.1.4, the background predictions are obtained in the same way as with the E_{extra} sideband (§6.2.2). The results are shown in Table 6.5 and compared with the other sideband predictions in Table 6.6.

6.2.3 Choice of Background Prediction Method

The 1D sideband predictions are highly consistent for each mode and thus can be used to validate and support one another. We must choose one BG prediction to use in interpreting our results. We choose the E_{extra} sideband because it relies more

on data than the D^0 sideband or direct counting from the signal MC samples and because it can be used consistently across all seven modes.

Mode	MC Counting	D^0 Mass	E_{extra}	$LHR_{\text{cont.}}$	$LHR_{B\bar{B}}$	$p_{\text{sig } \ell}$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	98.4 ± 10.8	102.6 ± 15.3	91.4 ± 12.8	127.4 ± 118.8	99.7 ± 17.1	-
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	136.1 ± 11.8	146.1 ± 16.0	137.2 ± 13.3	192.4 ± 48.8	79.3 ± 50.2	-
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	212.1 ± 16.8	239.2 ± 20.0	233.0 ± 18.9	228.6 ± 24.8	279.8 ± 80.1	-
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	62.4 ± 9.0	57.7 ± 11.4	59.2 ± 8.8	52.8 ± 11.7	64.7 ± 14.0	-
$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	122.9 ± 12.2	119.8 ± 17.3	121.7 ± 12.6	116.6 ± 16.2	127.0 ± 14.0	-
$B^+ \rightarrow \mu^+ \nu_\mu$	11.5 ± 5.0	13.9 ± 5.8	15.1 ± 9.9	11.5 ± 7.3	14.8 ± 19.1	12.7 ± 7.6
$B^+ \rightarrow e^+ \nu_e$	14.6 ± 5.3	13.6 ± 8.5	24.0 ± 11.2	-	-	35.0 ± 17.8

Table 6.6: Comparison of BG predictions from various sidebands.

CHAPTER 7

THE FELDMAN-COUSINS METHOD

When we examine the signal region in data, we must translate the number of events seen into a branching fraction or upper limit. This is not as simple as it may seem. We must use a method that accounts for potential statistical fluctuations and can set a proper upper limit or errors on a branching fraction.

7.1 The Feldman Cousins Method

We use the method developed by Feldman and Cousins [34], which is designed to produce an upper limit for null results and a two-sided confidence interval for non-null results. The Feldman Cousins method begins with the construction of a confidence belt, which is a two-dimensional histogram N_{true} vs. N_{sig} . N_{sig} represents the extracted signal yield for an ensemble of experiments for each value of N_{true} . We generate this distributions using a random number generator.

For each value of N_{true} , we generate two sets of 100,000 random numbers. One set is based on the predicted signal yield from Table 5.4; a Poisson random number generator is used with the central value set to the value from the Table. The second set is based on the background predictions from Table 6.2. A Gaussian random number generator is used with the center set to the central value from the table (N_{BG}) and

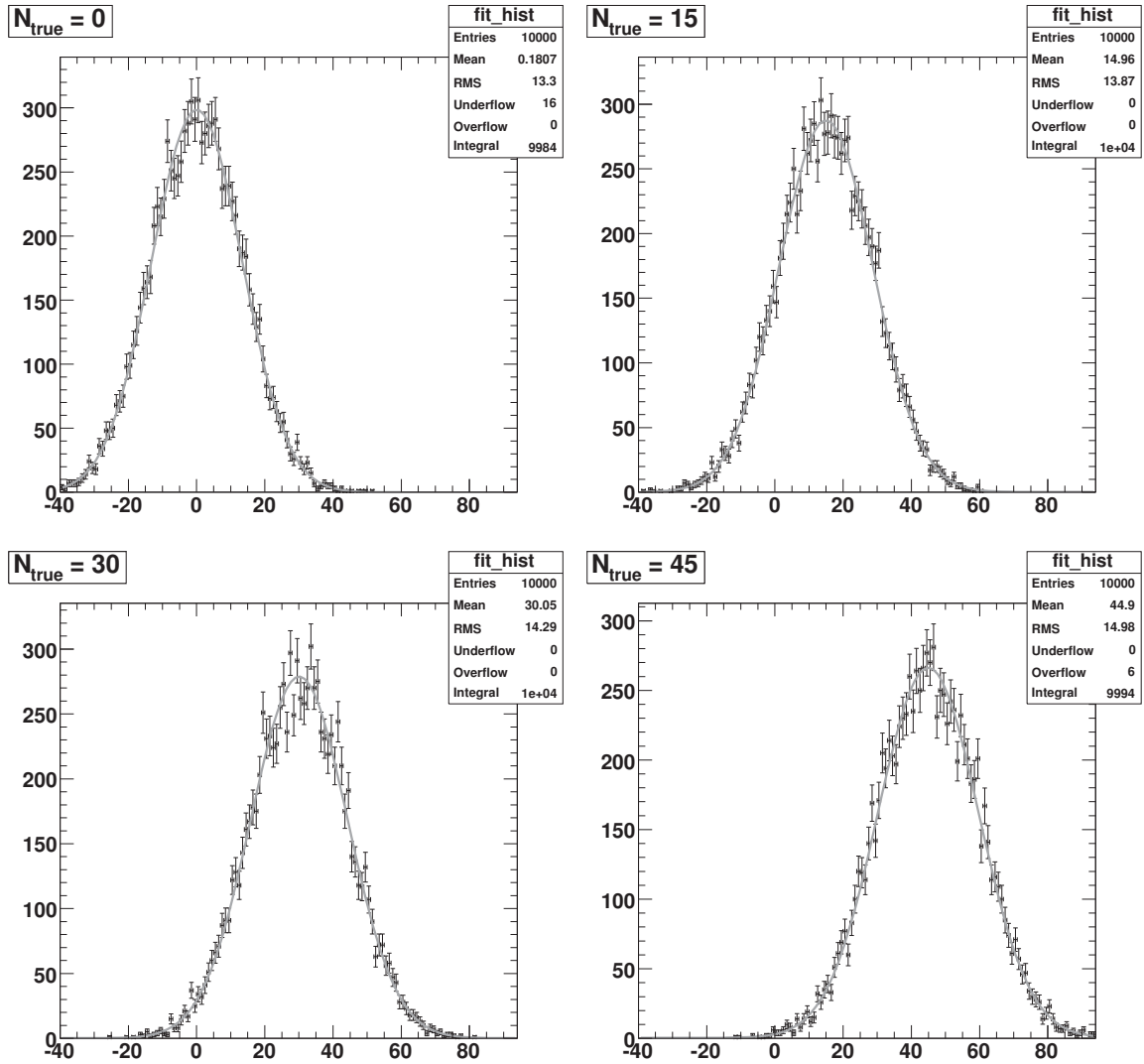


Figure 7.1: The fits to N_{sig} for several values of N_{true} .

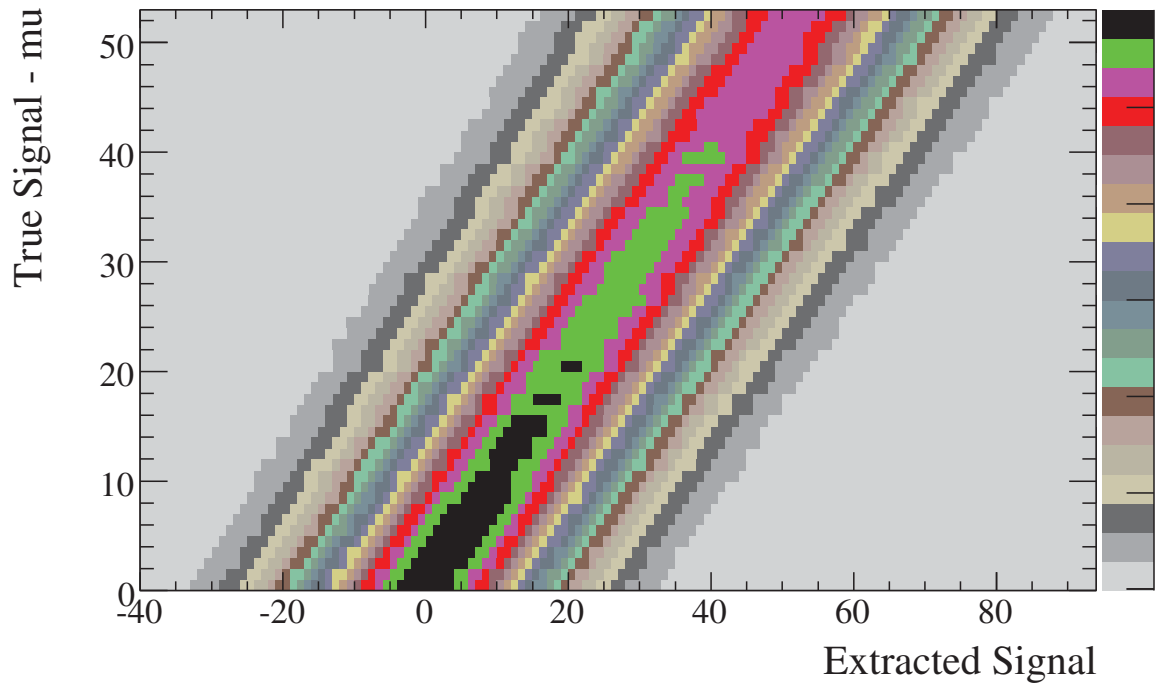


Figure 7.2: The distribution of N_{true} vs. N_{sig} . For each integer value of N_{true} , we generated 100,000 random numbers and fit the resulting N_{sig} distribution to the sum of the Gaussian functions.

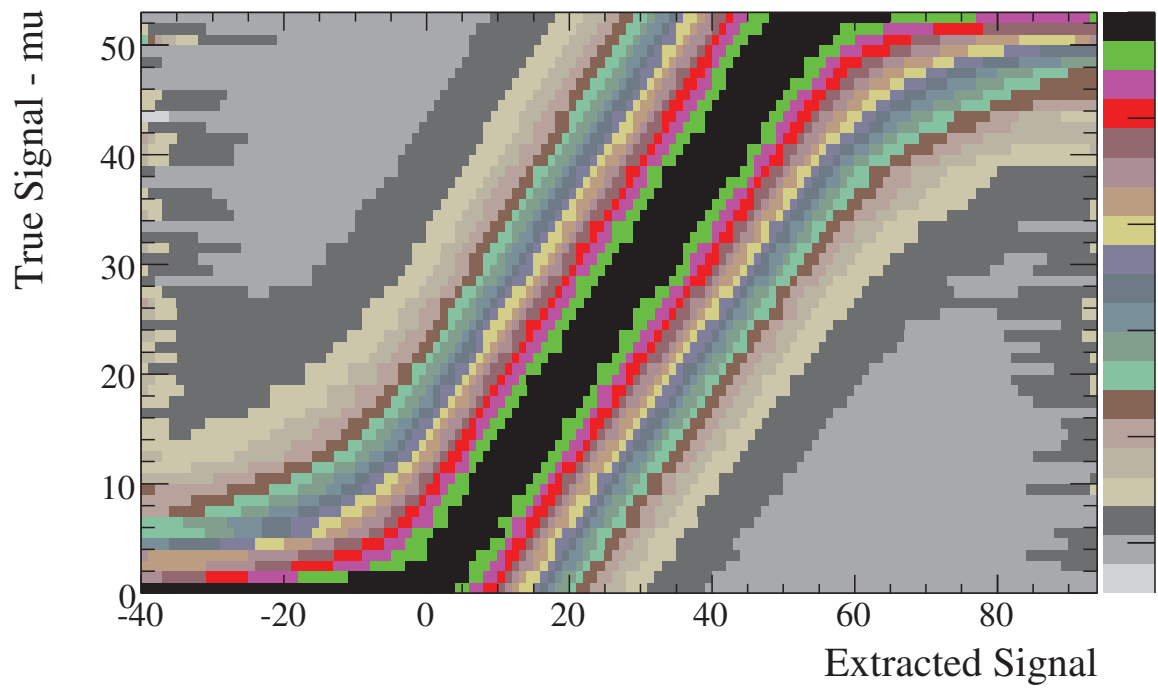


Figure 7.3: An example of the distribution of R

the width equal to $\sqrt{N_{\text{BG}}}$. This error accounts for the statistical error on the total number of observed events.

The resulting random number is then used as the center of another random number distribution (N_{back}), with a width equal to the error on (N_{BG}) from Table 6.2. This error accounts for the uncertainty on the background prediction. These two numbers are summed for each experiment to form N_{obs} .

$$N_{\text{sig}} = N_{\text{obs}} - N_{\text{back}}. \quad (7.1)$$

To smooth the statistical fluctuations resulting from the random number generation, we fit the distribution for each value of N_{true} to the sum of two Gaussians; examples are shown in Figure 7.1. N_{sig} is then reassigned its value from the fitting functions. The resulting smoothed distribution is shown in Figure 7.2 for $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$.

For each N_{true} , we must define an acceptance region in N_{sig} that will determine our upper limit or central value with uncertainty. The Feldman Cousins method defines this acceptance region without referring to data or any bias regarding whether we seek an upper limit or branching fraction. For each bin in Figure 7.2, we calculate the ratio

$$R \equiv \frac{P(N_{\text{sig}}|N_{\text{true}})}{P(N_{\text{sig}}|N_{\text{best}})}, \quad (7.2)$$

where N_{best} is the value of N_{true} that maximizes the probability of observing N_{sig} . Thus, R ranges between 0 and 1. The distribution of R for $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ is shown in Figure 7.3.

For each value of N_{true} , we sort the bins of N_{sig} in order of descending R . The probabilities $P(N_{\text{sig}}|N_{\text{true}})$ are summed in this order until the desired confidence level is reached. The confidence bands for all three B modes are shown in Figure 9.10.

When the data is unblinded, we calculate N_{sig} and draw a vertical line through the confidence bands at N_{sig} . The upper and lower limits will be determined by the intersection of that line with the appropriate confidence bands.

7.2 Application of Systematic Errors

For calculating upper limits and the significance of the the $B^+ \rightarrow \tau^+ \nu_\tau$ result, we add the total systematic uncertainty in quadrature with the other uncertainties on the background. Thus, we widen the background distribution by the systematic uncertainties and proceed with the remaining steps.

For calculating branching fractions, we do not use the systematic errors in the Feldman-Cousins calculation. We multiply the fractional systematic uncertainty by the branching fraction central value and report that as a separate uncertainty.

CHAPTER 8

SYSTEMATICS

The branching fraction for any of the decay modes in this analysis is given by

$$\mathcal{B}(B^+ \rightarrow \ell^+ \nu_\ell) = \frac{N_{\text{sig}}}{N_{BB} \varepsilon_{\text{tag}} \varepsilon_{\text{sig}}}, \quad (8.1)$$

$$N_{\text{sig}} \equiv N_{\text{obs}} - N_{\text{BG}}, \quad (8.2)$$

where N_{obs} is the total number of events observed in the signal region, N_{BG} is the predicted number of events from background in the signal region, N_{BB} is the total number of $\Upsilon(4S)$ decays in the data set, and the efficiencies can have different values for each mode. Each of these variables, except N_{obs} , brings a systematic error into the branching fraction.

8.1 Systematic Error from Background Prediction (N_{BG})

We use the ratio of data to MC samples in our background prediction, and the statistical error on that prediction is already large. The shape of our data and MC samples agree well in the sideband region, so we have no need to apply an additional systematic error. Therefore, we accept this as the total error and apply no further systematic correction or uncertainty.

8.2 Systematic Error from B-Counting (N_{BB})

The method for determining this systematic error is found in Ref. [29]. We find a correction of 1.1% for all Runs.

8.3 Systematic Error from Tagging Efficiency (ε_{tag})

We pursue a procedure that attempts to combine what we know about both the single- and double-tagged samples.

The single-tagged sample consists of all events in which a tag B is reconstructed, regardless of what (if anything) is found in the signal B . In this sample, we have large D^0 mass sidebands which we can use to subtract combinatoric D^0 events. That leaves events with (a) correctly reconstructed $B^+ \rightarrow D^0 + X$, (b) neutral B meson events with a correctly reconstructed D^0 meson, and (c) charm non-resonant events with real D^0 mesons. This is very different from the double-tags, which are about 90% charged B, about 10% neutral B, and a handful of charm events. We only use sideband subtraction for the single tags. We do not subtract the sideband from the first or second tag when obtaining the number of double tags N_2 .

The efficiency of “perfect” single-tagged events is given by

$$\varepsilon_1 = \frac{N_1}{N_{B^+B^-}}. \quad (8.3)$$

The efficiency of reconstructing double-tagged events can be written as

$$\varepsilon_1 \times \varepsilon_2 = \frac{N_2}{N_{B^+B^-}}. \quad (8.4)$$

Taking the ratio of Equations 8.3 and 8.4, we obtain:

$$\varepsilon_2 = N_2/N_1.$$

We can then compare this efficiency between data and MC. This eliminates the assumption that the single and double-tag efficiencies are the same, while at the same time admitting effects from a sample with inherently higher (and potentially poorly modeled) backgrounds.

8.3.1 $D^0 \rightarrow K^- \pi^+$ tags

The $D^0 \rightarrow K^- \pi^+$ are the “cleanest” tags, given the low multiplicity and the fully charged nature of the final state. We perform the computation of the single tags (via sideband subtraction) using these tags first. The results of the single-tag computation, and the double-tags, are in Table 8.1. In both single and double tags, only the D^0 decay mode of the first tag is restricted; the second tag is allowed to decay into any of the four modes used in this analysis.

We perform a separate test of the sideband subtraction method on just uds and $\tau^+ \tau^-$ MC in Run 4, and find a “peaking” yield of -91.0 ± 171.2 events, which is consistent with zero. This suggests the sideband subtraction, when applied to purely combinatoric events, does not leave a “fake” peaking component in the D^0 mass signal box.

We determine from the comparison of the double-to-single-tag yields in both data and MC that, in $K\pi$ tags, a downward correction is required. The scaling of MC to data is 89.1%. This scale factor has a 2.4% uncertainty, which we take as a systematic uncertainty on the tag efficiency.

	Single Tags	Double Tags	N_2/N_1
Data	812182.0 ± 1132.4	2278.0 ± 47.7	$(0.280 \pm 0.006) \times 10^{-2}$
MC	2611475.0 ± 1945.4	8225.0 ± 90.7	$(0.315 \pm 0.003) \times 10^{-2}$
Data/MC	-	-	0.891 ± 0.021

Table 8.1: Single and double tag yields in the full data and MC samples of $D^0 \rightarrow K^- \pi^+$ events.

	Single Tags	Double Tags	N_2/N_1
Data	602762.0 ± 2034.0	2138.0 ± 46.2	$(0.355 \pm 0.008) \times 10^{-2}$
MC	2051028.0 ± 3468.0	8549.0 ± 92.5	$(0.417 \pm 0.005) \times 10^{-2}$
Data/MC	-	-	0.851 ± 0.021

Table 8.2: Single and double tag yields in full data and MC samples of $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ events.

8.3.2 $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ tags

We repeat the exercise for a less experimentally clean sample of tags, one which contains more combinatoric D mesons than the $K\pi$ tags: those with three charged pion in the final state. The results are shown in Table 8.2. The correction found from this mode is consistent with the correction from $D^0 \rightarrow K^- \pi^+$ to within 1.3σ .

8.4 Systematic Error from Signal Efficiency (ε_{sig})

8.4.1 Modeling of E_{extra} shape

The systematic correction and uncertainty is taken from the double tag E_{extra} distribution described in Section 6.1.5. The double-tagged events provide us with a means of comparing data and simulation, using an independent control sample, to

extract this uncertainty. A cut is imposed on the E_{extra} distributions shown in Figure 6.1 to extract the yield of candidates satisfying $E_{\text{extra}} \leq 0.6 \text{ GeV}$. This yield is then compared to the number of candidates in the full sample. Comparing the the ratio extracted from MC to that extracted from data yields a correction factor, the error on which is taken as the systematic uncertainty for E_{extra} . We extract a correction and uncertainty of 1.015 ± 0.021 .

8.4.2 Tracking Efficiency for Signal Side

Since the tagging efficiency systematic correction and uncertainty encompass the systematic effects for tracks on the tag side, the correction and systematic uncertainties associated with the signal track are evaluated. The Tracking Efficiency Task Force on *BABAR* provides recipes for assigning systematic uncertainties to charged tracks. They have a correction of 1.0 and place all systematic effects in the uncertainty. They are based on a data sample that consists of events in which $e^+e^- \rightarrow \tau^+\tau^-$; one τ decays with three signal tracks, and the other decays with only one signal track.

For analyses where the multiplicity of the final state is fixed, the Task Force recommends using an overall correction factor added in quadrature with an uncertainty per track. We will need to apply these corrections separately to the single prong modes and the $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ mode. Thus, the multiplicity of each final state will remain fixed.

The average for all runs for `GoodTracksVeryLoose` is 0.27% for the overall factor and 0.23% uncertainty per track. Adding in quadrature yields 0.36%. The total correction for the six modes with a single charged track is 0.36%; for $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$

we take three times the uncertainty per track to yield a 0.69% uncertainty. When combined with the overall correction, the tracking systematic uncertainty is 0.74%.

8.4.3 π^0 Selection

We only select neutral pions in the $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$ mode, so this correction only applies to that mode. We apply the correction for this list as recommended by the Neutral Reconstruction Analysis Working Group of *BABAR*. They are the section of the collaboration in charge of the development, optimization and evaluation of the reconstruction and selection of neutral particles. This correction is (0.984 ± 0.030) , and is only applied to the efficiency of the $\tau^+ \rightarrow \pi^+\pi^0\bar{\nu}_\tau$ mode.

The same group recommends using a different correction for merged π^0 candidates. Merged π^0 candidates are reconstructed from photons that both are detected by the same EMC crystal. However, we do not store information on which π^0 candidates are merged in our data set. Also, since this analysis will be dominated by statistical errors, we assume that the number of merged π^0 candidates in this mode is insignificant and apply only the correction described above.

8.4.4 Particle Identification

Since the particle identification algorithms have not changed since our previous search for $B^+ \rightarrow \tau^+\nu_\tau$ [25], we use the same values as conservative estimates of our current systematic errors. Since the previous analysis only used the four single-prong τ decay modes, we must extrapolate. For the $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$ mode, we apply the π correction three times and take triple the uncertainty as our a_1 uncertainty. For the $B^+ \rightarrow \mu^+\nu_\mu$ and $B^+ \rightarrow e^+\nu_e$ modes, we use the μ and e systematics, which

is a conservative choice because particle identification is more effective at the high momenta that characterize two-body B decays.

8.5 Summary of Systematic Corrections and Uncertainties

All contributions to the systematic uncertainty are summarized in Table 8.3. The corrected efficiencies are shown in Table 8.4. We extract a total systematic uncertainty of 3.6% for $B^+ \rightarrow \tau^+ \nu_\tau$, 4.4% for $B^+ \rightarrow \mu^+ \nu_\mu$, and 4.0% for $B^+ \rightarrow e^+ \nu_e$.

Source	Applicable Mode(s)	Correction	Fractional Uncertainty (%)
B Counting	All	1.0	1.1
Tag efficiency	All	0.891 ± 0.021	2.4
E_{extra}	All	1.015 ± 0.021	2.1
π^0 Reconstruction	$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	0.984 ± 0.030	3.0
Tracking Efficiency	$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	1.0	0.36
	$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	1.0	0.36
	$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	1.0	0.36
	$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	1.0	0.36
	$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	1.0	0.74
	$B^+ \rightarrow \mu^+ \nu_\mu$	1.0	0.36
	$B^+ \rightarrow e^+ \nu_e$	1.0	0.36
Particle Identification	$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	1.01	2.5
	$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	0.92	3.1
	$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	1.02	0.8
	$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	1.00	1.5
	$\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$	1.06	2.4
	$B^+ \rightarrow \mu^+ \nu_\mu$	0.92	3.1
	$B^+ \rightarrow e^+ \nu_e$	1.01	2.5

Table 8.3: Summary of systematic corrections, uncertainties, and fractional uncertainties

Efficiency	Corrected	Fractional Systematic Error (%)
Tag($B^+ \rightarrow \tau^+ \nu_\tau$)	$(1.514 \pm 0.003 \pm 0.036)\%$	2.4
Tag($B^+ \rightarrow \mu^+ \nu_\mu$)	$(0.937 \pm 0.003 \pm 0.022)\%$	2.4
Tag($B^+ \rightarrow e^+ \nu_e$)	$(0.974 \pm 0.003 \pm 0.023)\%$	2.4
$\varepsilon_{\text{sig}}^{(\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau)}$	$(2.08 \pm 0.04 \pm 0.07)\%$	3.3
$\varepsilon_{\text{sig}}^{(\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau)}$	$(1.50 \pm 0.04 \pm 0.06)\%$	3.7
$\varepsilon_{\text{sig}}^{(\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau)}$	$(2.57 \pm 0.05 \pm 0.06)\%$	2.2
$\varepsilon_{\text{sig}}^{(\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau)}$	$(0.86 \pm 0.03 \pm 0.03)\%$	4.0
$\varepsilon_{\text{sig}}^{(\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau)}$	$(0.33 \pm 0.02 \pm 0.01)\%$	3.3
$\varepsilon_{\text{sig}}^{(B^+ \rightarrow \tau^+ \nu_\tau)}$	$(7.33 \pm 0.08 \pm 0.21)\%$	2.8
$\varepsilon_{\text{sig}}^{(B^+ \rightarrow \mu^+ \nu_\mu)}$	$(28.87 \pm 0.33 \pm 1.08)\%$	3.7
$\varepsilon_{\text{sig}}^{(B^+ \rightarrow e^+ \nu_e)}$	$(37.91 \pm 0.39 \pm 1.24)\%$	3.3

Table 8.4: The corrected tag and signal efficiencies. Two errors are quoted: the first is the MC statistical uncertainty, and the second is the systematic error computed from the sources in this section.

CHAPTER 9

RESULTS

This chapter represents Step 5 from §4.1. We examine the signal region and use the Feldman-Cousins method to determine upper limits and branching fractions. The total number of events observed is compared with background predictions in Table 9.4.

9.1 $B^+ \rightarrow e^+ \nu_e$

We observe 17 events with an expected background of 24.0 ± 11.2 events. This observation is consistent with a null result. Using the Feldman-Cousins method, we determine the following upper limit (at 90% confidence level):

$$\mathcal{B}(B^+ \rightarrow e^+ \nu_e) < 7.7 \times 10^{-6}.$$

This is nearing competitiveness with the upper limit achieved by the hadronically tagged Runs 1-5 analysis [18].

9.2 $B^+ \rightarrow \mu^+ \nu_\mu$

We observe 11 events with an expected background of 15.2 ± 9.9 events. This observation is consistent with a null result. We determine the following upper limit (at 90% confidence level):

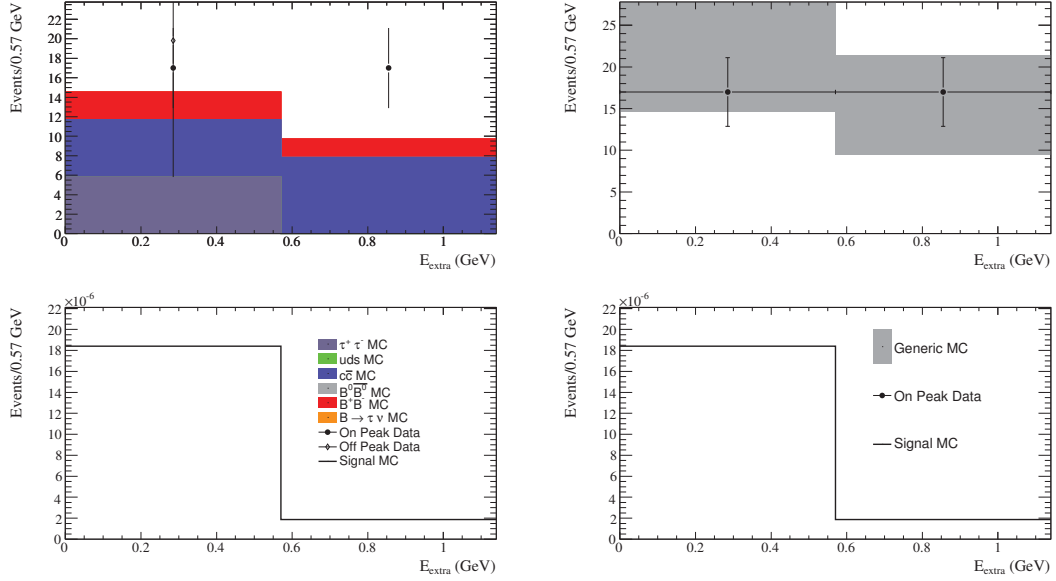


Figure 9.1: Total extra energy is plotted after all cuts have been applied to the $B^+ \rightarrow e^+ \nu_e$ mode. Off-resonance data and MC have been normalized to the on-resonance luminosity. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC samples have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2 and summed together. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow e^+ \nu_e$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

$$\mathcal{B}(B^+ \rightarrow \mu^+ \nu_\mu) < 11 \times 10^{-6}.$$

This is within a factor of 2 of the hadronically tagged upper limit but an order of magnitude above the inclusive measurement. However, the inclusive measurement observed approximately 600 background events [16]; this analysis had 11 total events. When enough data is available, this low background will be much more conducive to discovery and measurement.

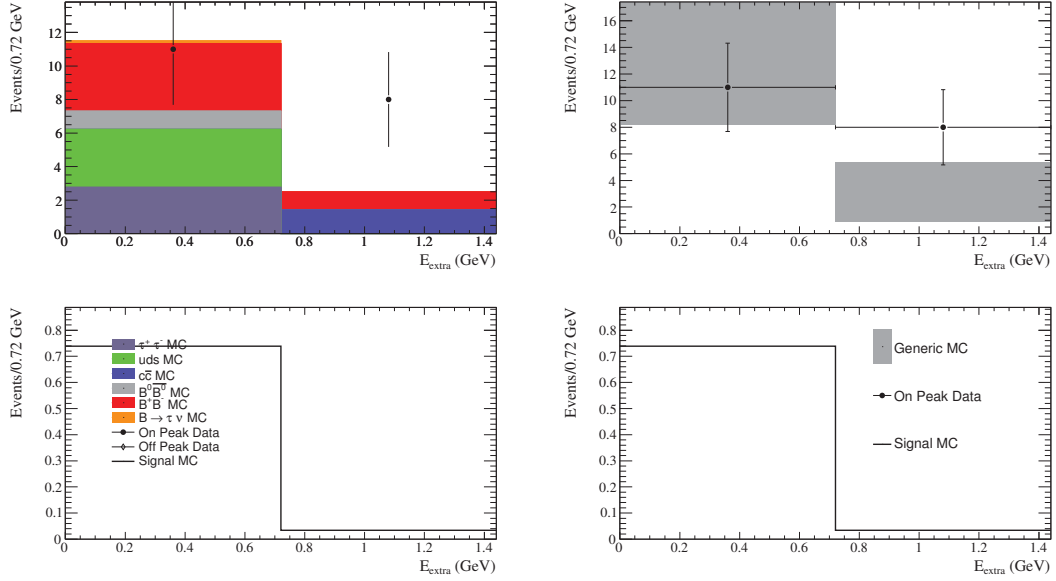


Figure 9.2: Total extra energy is plotted after all cuts have been applied to the $B^+ \rightarrow \mu^+ \nu_\mu$ mode. Off-resonance data and MC have been normalized to the on-resonance luminosity. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC samples have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2 and summed together. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \mu^+ \nu_\mu$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

9.3 $B^+ \rightarrow \tau^+ \nu_\tau$

Recall that the τ is reconstructed in several decay modes. The analysis method for each is optimized separately, and the systematic errors are also calculated separately for each.

9.3.1 Exclusion of the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ mode

Tables 9.5, 9.6, 9.7, and 9.8 give the expected branching fractions or upper limits with all systematic corrections and uncertainties included. Tables 9.5 and 9.6 show that including the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ channel in the τ branching fraction decreases the significance of the result. From Equations 5.8 and 8.2, we see that significance is $N_{\text{sig}}/\sqrt{N_{\text{obs}}}$. With the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ mode, we expect $55.4/\sqrt{698} = 2.10$. Without the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ mode, we expect $53.2/\sqrt{574} = 2.22$. Therefore, we use only the four single-prong channels to reconstruct the τ . This also makes the determination of systematic errors easier because all remaining modes have exactly one charged track in the signal.

9.3.2 Excess seen in $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ mode

We notice that the electron mode has a much larger excess than the other τ decay modes. After we examined the data in the signal region, we discovered an excess of data above our MC simulations at low values of $m_{\ell\ell}$ (Figure 9.3), which is the minimum invariant mass of any two leptons. To remove these events, which are probably due to unmodeled photon pair conversion, we decided to remove all events below a certain value of $m_{\ell\ell}$ after all other analysis cuts had been applied. This value was chosen using only signal and background MC simulations with the optimization technique described in this section. The result excludes events in the $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ channel with $m_{\ell\ell} < 0.29$ GeV. All efficiencies and yields in this note include the effects of this requirement.

To test the probability of a statistical fluctuation producing the excess seen in the $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ channel, we fill a histogram with the branching fraction calculated

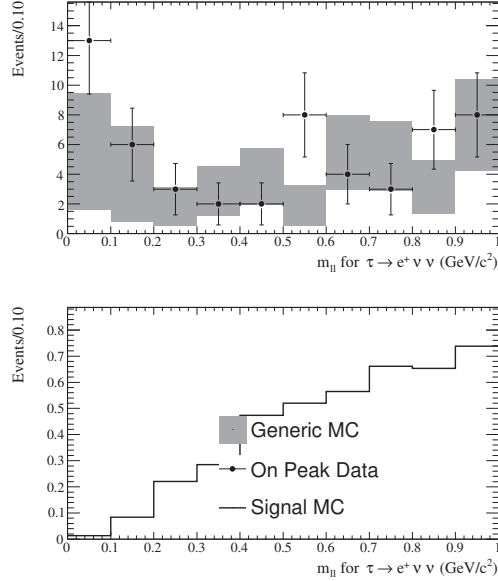


Figure 9.3: $m_{\ell\ell}$ is plotted after all cuts have been applied for the $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ mode. Events in this distribution are required to pass all selection criteria, except the $m_{\ell\ell}$ requirement. The background MC sample has been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

from each τ decay channel separately. The values and uncertainties are taken from the Feldman-Cousins method; systematic uncertainties are not included. We fit a constant to the branching fractions; the result is $(1.8 \pm 0.6) \times 10^{-4}$. The fit has a reduced χ^2 of 1.64, which corresponds to a probability of 18%. The results of this study are shown graphically in Figure 9.4.

We have tested for the following other potential background contributions: two-photon fusion QED events, “events” that contain two overlapping e^+e^- collisions, overzealous Bremsstrahlung recovery, photon pair production where the e^+ and e^-

Mode	$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	$(4.0 \pm 1.2) \times 10^{-4}$
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	$(1.0^{+1.2}_{-0.9}) \times 10^{-4}$
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	$(0.6^{+1.1}_{-0.5}) \times 10^{-4}$
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	$(2.0^{+1.4}_{-1.3}) \times 10^{-4}$
$B^+ \rightarrow \tau^+ \nu_\tau$	$(1.8 \pm 0.8 \pm 0.1) \times 10^{-4}$

Table 9.1: Each branching fraction is calculated for each of the four τ decay modes separately.

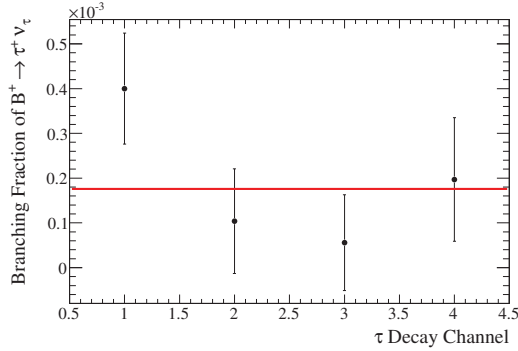


Figure 9.4: The branching fractions calculated from each τ decay channel are plotted. 1 = $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$, 2 = $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$, 3 = $\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$, 4 = $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$. The horizontal line is a fit to a constant.

are reconstructed as the tag and signal lepton, and photon pair production events where one lepton is lost and the other is reconstructed as the signal electron.

In two-photon fusion events, the e^+ and e^- each emit a photon; the two photons interact to produce multiple hadrons, see Figure 1.4. Two-photon fusion events rarely produce real D^0 mesons, so they should populate the sidebands and peak of the D^0 mass distribution. Therefore, we plot the E_{extra} distribution for $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ taken

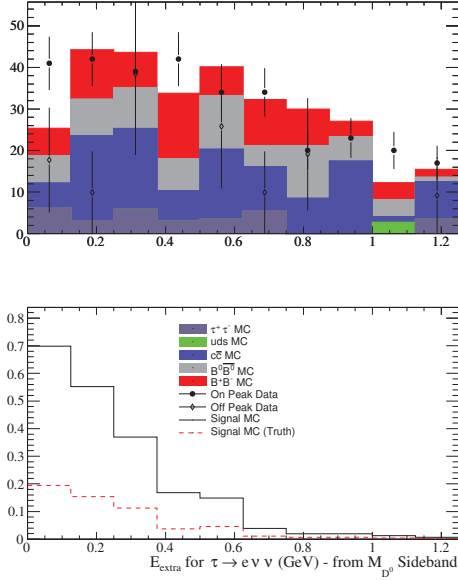


Figure 9.5: Total extra energy is plotted for the mode $B^+ \rightarrow \tau^+ \nu_\tau (\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau)$. This distribution is taken from the sidebands of the D^0 mass distribution (§6.1.1). All other optimized signal selection criteria, have been applied. Comparing this to Figure 9.6 shows that the excess of signal events in this mode occurs in the peak of the D^0 mass distribution, which suggests that it is not due to two-photon fusion. The generic MC samples are scaled to data luminosity. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

from the sidebands of the D^0 mass distribution in Figure 9.5. We expect an excess of two-photon fusion events to cause our background prediction from the D^0 mass sidebands to be different from the other sidebands, but this is not the case, as seen in Table 6.6. Comparing this to Figure 9.6 shows that the excess of signal events in this mode occurs mostly in the peak of the D^0 mass distribution, which suggests that it is not due to two-photon fusion.

Since the signal lepton is an e^\pm , these events will only populate the electron τ decay mode. Two-photon fusion events will also only be reconstructed as events with

an e^\pm on the tag side. To check for them, we separate events by the identity of the lepton from the tag B . The numerical results of this check are shown in Table 9.2.

We examined the distribution of Δz , which is the separation between the putative B vertices. Overlapping events should have wider separation than real events, and thus we should see an excess of data above the MC at high values for Δz . No such excess was found.

Bremsstrahlung radiation is emitted by all charged particles as interactions with detector materials and electromagnetic fields decelerate them. In *BABAR*, only electrons and positrons emit significant amounts of Bremsstrahlung. To improve reconstruction accuracy and efficiency, we employ an algorithm to identify photons probably produced by Bremsstrahlung radiation and combine them with the charged tracks that generated them.

Another suggested source of background in this channel is Bremsstrahlung recovery that assigns more photons to an electron than it actually generated. This could move events into the signal range of E_{extra} undeservedly. To test for this, we compare the E_{extra} distributions for electrons with and without Bremsstrahlung recovery. When we turn off Bremsstrahlung recovery, we do so for electrons in both the tag and signal B . Comparing Figures 9.6 (with Bremsstrahlung recovery) and 9.9 (without Bremsstrahlung recovery) shows that data and MC simulation have very similar shapes regardless of whether Bremsstrahlung recovery is used. No suspicious excess appears when Bremsstrahlung recovery is activated, so this is not a likely source of the excess.

Tag Lepton	Background Prediction (N_{BG})	Observed Events (N_{obs})	Excess (σ)
e^+	66 ± 13	106	3.2σ
μ^+	35 ± 8	64	3.6σ

Table 9.2: Comparison of background predictions (from E_{extra} sideband) and observed yields in on-peak data for the two tag leptons. The fourth column is given by $(N_{\text{obs}} - N_{\text{BG}})/\sigma_{N_{\text{BG}}}$. Note that these studies were performed *before* the $m_{\ell\ell}$ requirement was applied.

E_{extra} Range	Background Prediction (N_{BG})	Observed Events (N_{obs})	Excess (σ)
[0.0, 0.12] GeV	52 ± 10	85	3.5σ
[0.12, 0.24] GeV	53 ± 9	85	3.7σ

Table 9.3: Comparison of background predictions (from E_{extra} sideband) and observed yields in on-peak data for the two subsets of the signal region of E_{extra} . The fourth column is given by $(N_{\text{obs}} - N_{\text{BG}})/\sigma_{N_{\text{BG}}}$. Note that these studies were performed *before* the $m_{\ell\ell}$ requirement was applied.

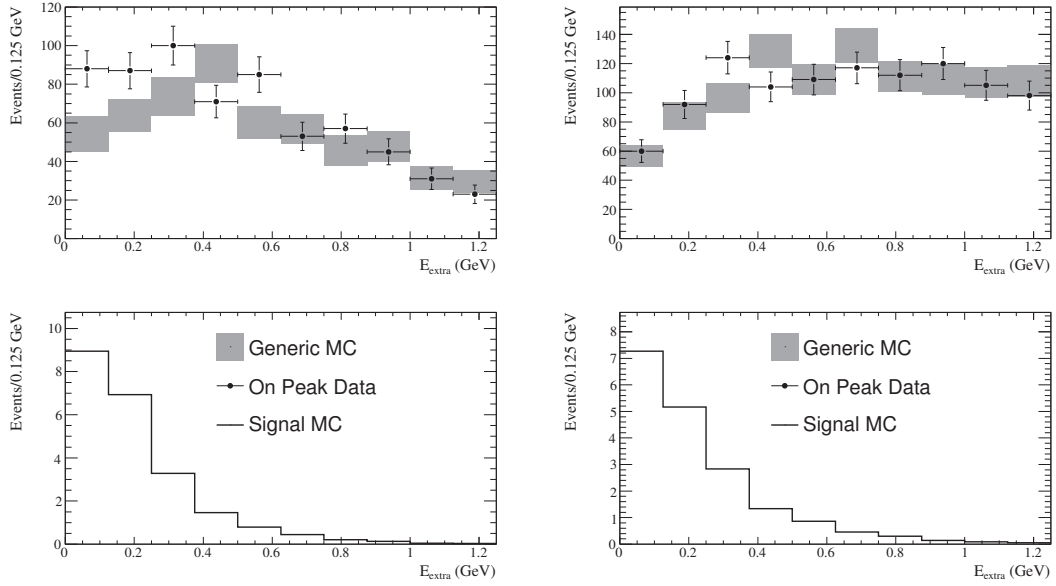


Figure 9.6: Total extra energy is plotted after all cuts have been applied in the modes (left) $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and (right) $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$. The background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

9.3.3 Results for $B^+ \rightarrow \tau^+ \nu_\tau$

For all four τ decay channels, we observe 610 events with a background prediction of 520.8 ± 30.6 events. This is inconsistent with the null result. Therefore, we determine the following branching fraction:

$$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) = (1.8 \pm 0.8 \pm 0.1) \times 10^{-4}.$$

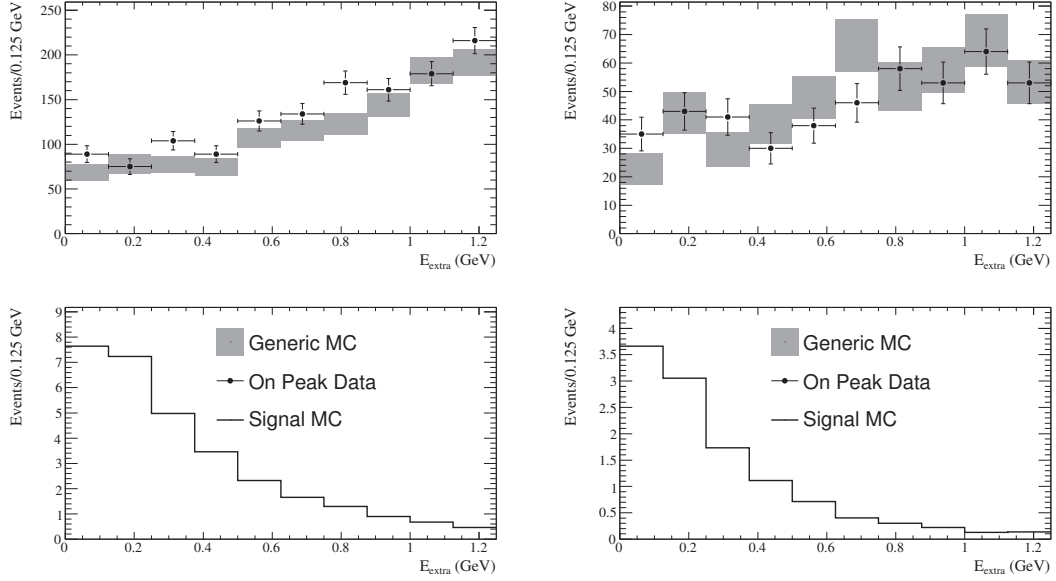


Figure 9.7: Total extra energy is plotted after all cuts have been applied in the mode (left) $\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$ and (right) $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$. The background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

Using the Feldman Cousins method, we draw successively wider confidence bands until one is wide enough for our result to be consistent with zero. Via this approach, we find that we exclude the null result at the level of 2.4σ .

This is within 1.3σ of our measurement with Run 1-5 [25]. consistent with the *BABAR*[19] and *Belle*[20] hadronically tagged measurements of $(1.8_{-0.9}^{+1.0}) \times 10^{-4}$ and $(1.8 \pm 0.7) \times 10^{-4}$, respectively. It is also consistent with the *Belle* semileptonically tagged measurement of $(1.65_{-0.37}^{+0.38+0.35}) \times 10^{-4}$ [21].

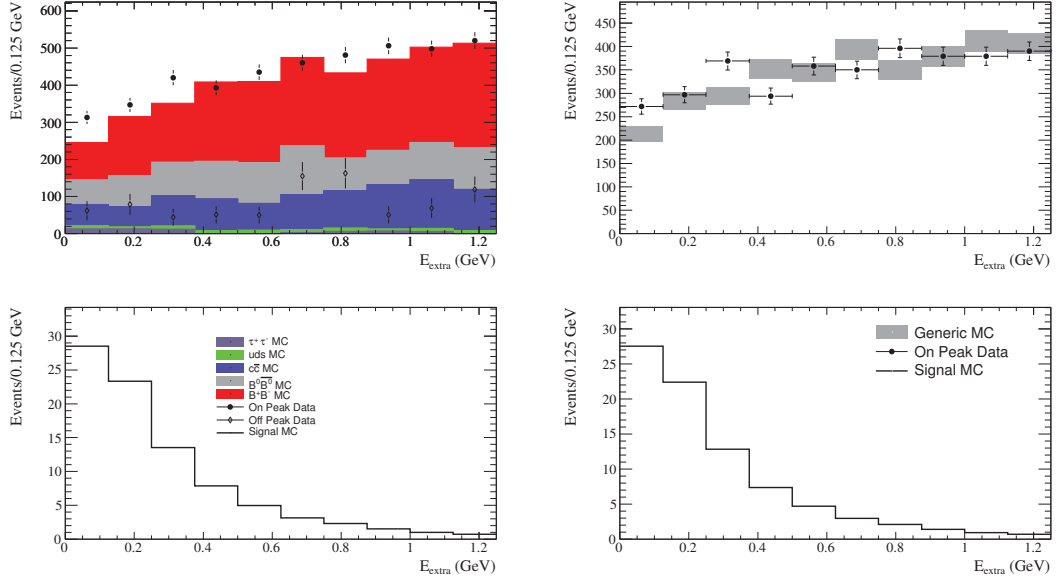


Figure 9.8: Total extra energy is plotted after all cuts have been applied with all $B^+ \rightarrow \tau^+ \nu_\tau$ modes combined. Events in this distribution are required to pass all selection criteria. At left, the background MC samples are scaled to the total data luminosity. At right, the background MC have been scaled according to the ratio of predicted backgrounds from data and MC as presented in section 6.2.2. The gray rectangles represent the extent of the error bars on the MC histogram. Simulated $B^+ \rightarrow \tau^+ \nu_\tau$ signal MC is plotted (lower) for comparison. The signal MC yield is normalized to the branching fractions chosen at the end of §3.5.

9.3.4 $f_B |V_{ub}|$

Using the value of the branching fraction and the same value of $|V_{ub}| = (3.96 \pm 0.15_{-0.23}^{+0.20}) \times 10^{-3}$ used to predict the branching fraction in §2.2, we extract a preliminary value for the B meson decay constant:

$$f_B = 255 \pm 58 \text{ MeV}.$$

Statistical and systematic errors have been combined. This is consistent with the lattice QCD prediction and shows no evidence of physics beyond the standard model.

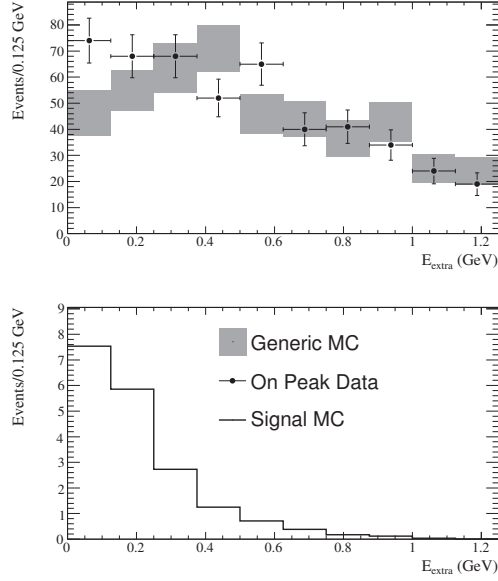


Figure 9.9: Total extra energy is plotted after all cuts have been applied for $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$ and with no Bremsstrahlung recovery. Events in this distribution are required to pass all selection criteria. The MC is not scaled.

9.3.5 Combined Result

We can combine this branching fraction with the measurement obtained using the hadronically tagged analysis [19]. The combined branching fraction is given by

$$\mathcal{B} = \frac{\sum_j \mathcal{B}_j / \sigma_j^2}{1 / \sigma^2}, \quad (9.1)$$

where j represents the two tagging methods, σ_j represents the combined statistical and systematic errors for each tag, and the overall error for the combined branching fraction σ is given by

$$\frac{1}{\sigma^2} = \sum_j \frac{1}{\sigma_j^2} \quad (9.2)$$

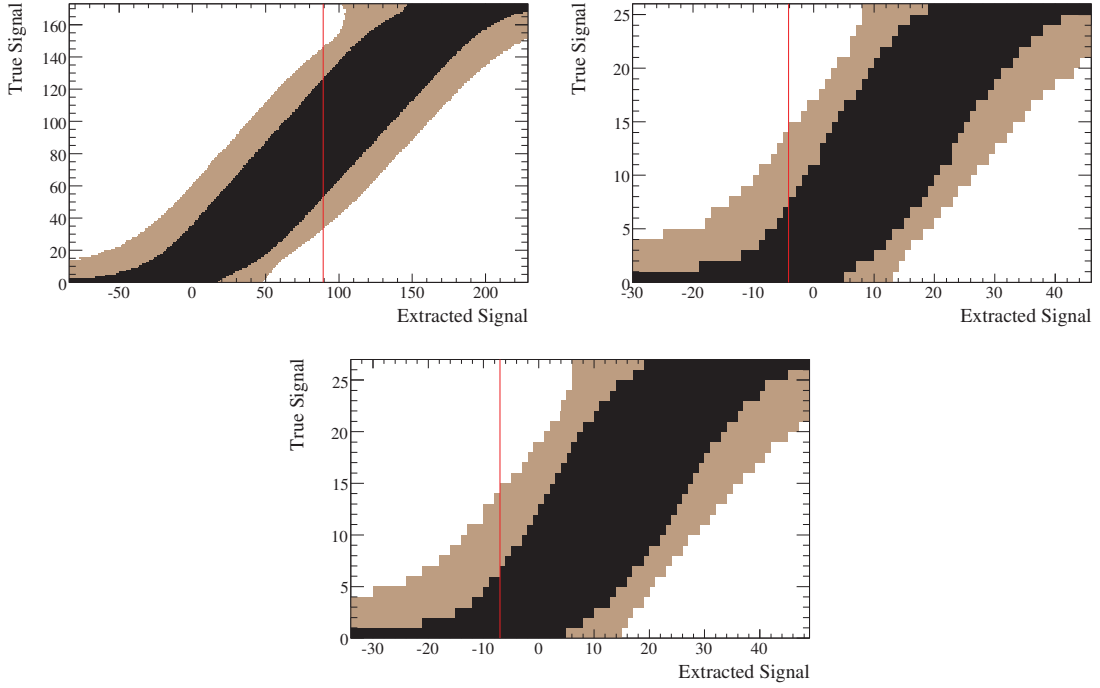


Figure 9.10: The confidence band produced by the Feldman Cousins method. The central band represents the 1σ confidence level; the next band out represents the 90% confidence level. The vertical red line represents the value of N_{sig} based on the number of observed events. The top left plot is for $B^+ \rightarrow \tau^+ \nu_\tau$; top right is for $B^+ \rightarrow \mu^+ \nu_\mu$, and the bottom plot is for $B^+ \rightarrow e^+ \nu_e$.

The combined result is

$$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau) = (1.8 \pm 0.6) \times 10^{-4}. \quad (9.3)$$

This excludes the null hypotheses at the level of 3.2σ .

9.3.6 Constraints on the Type II Two Higgs Doublet Model

As mentioned in §1.2.4, the possible existence of charged Higgs bosons (H^\pm) can be constrained by measuring $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$. The type II Two Higgs Doublet Model

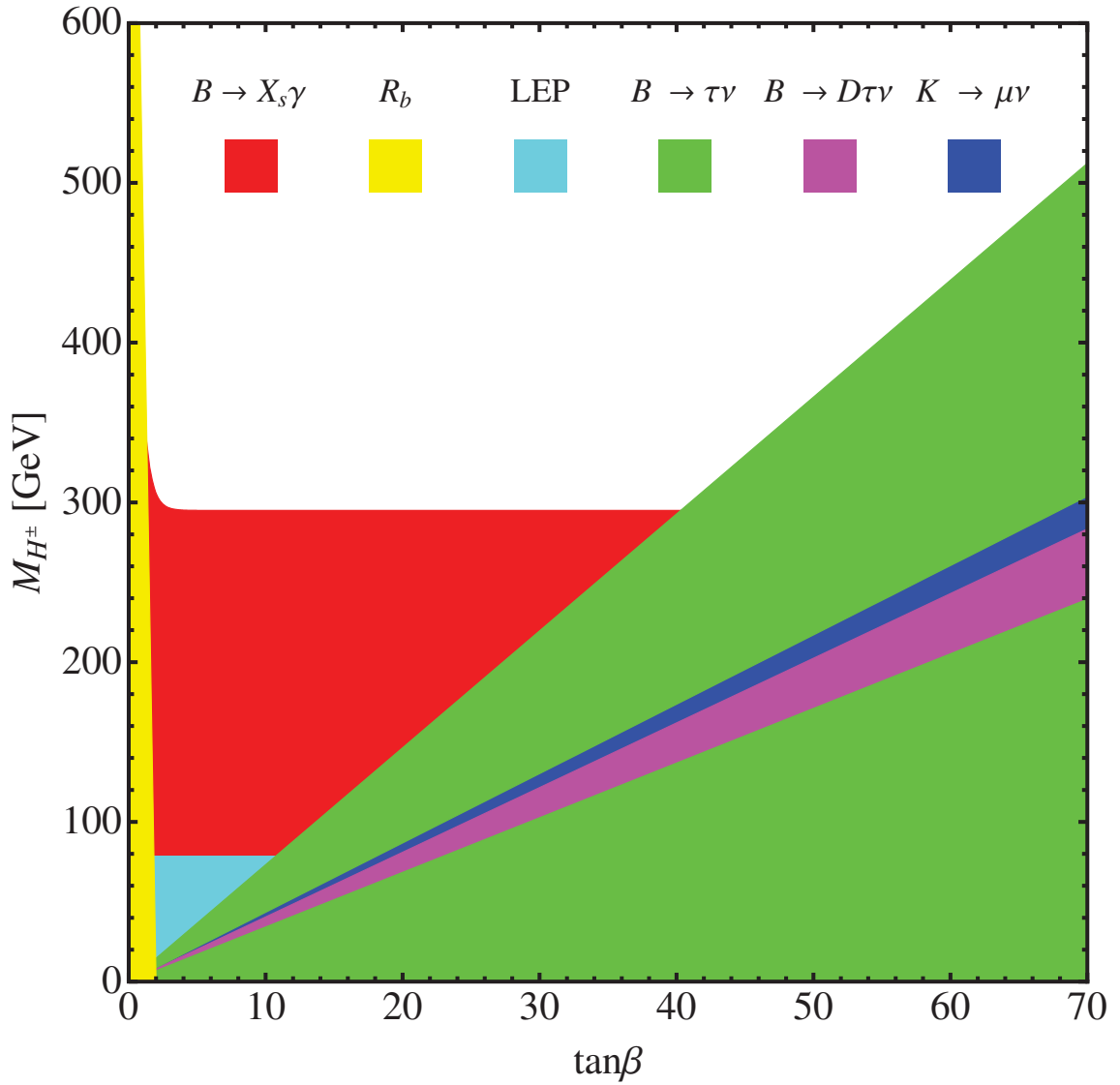


Figure 9.11: Shown are the regions in phase space ($\tan\beta$ vs. m_{H^\pm}) of the Two Higgs Doublet Model that are excluded by all relevant measurements and searches, including this one. The colored regions are excluded at the 95% confidence level [3]. The two green wedges are the region excluded by the world average of $\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$ before my measurement was made public. Since my measurement is consistent with the average, these wedges will not change greatly when it is included.

Mode	Expected Background	Observed Events	FOM
$\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$	91.4 ± 12.7	148	4.65
$\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$	137.2 ± 13.3	148	0.89
$\tau^+ \rightarrow \pi^+ \bar{\nu}_\tau$	233.0 ± 18.9	243	0.64
$\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}_\tau$	59.2 ± 8.8	71	1.4
$B^+ \rightarrow \tau^+ \nu_\tau$	520.8 ± 30.6	610	3.61
$B^+ \rightarrow \mu^+ \nu_\mu$	15.2 ± 9.9	11	-2.77
$B^+ \rightarrow e^+ \nu_e$	24.0 ± 11.2	17	-4.67

Table 9.4: The observed number of on-resonance data events in the signal region are shown, together with the number of expected background events and the number of expected signal events, which is taken from our signal MC samples. The FOM for the τ modes is $N_{\text{sig}}/\sqrt{N_{\text{obs}}}$ and $N_{\text{sig}}/(1.5 + \sqrt{N_{\text{BG}}})$ for the other two modes. Note that N_{sig} and hence FOM is negative for some modes.

(2HDM) postulates the existence of H^+ in addition to the neutral Higgs boson predicted by the SM (H^0). A charged Higgs could enhance or suppress the branching fraction of $B^+ \rightarrow \ell^+ \nu_\ell$ by playing the same role as the W^+ in Figure 2.1.

In this model u -type (charge $\frac{2}{3}$) quarks receive mass from one of the Higgs doublets, and d -type (charge $-\frac{1}{3}$) quarks receive mass from the other. The ratio of the vacuum expectation values of these two doubles is $\tan \beta$.

If the 2HDM is correct, the $\mathcal{B}(B^+ \rightarrow \ell^+ \nu_\ell)$ becomes

$$\mathcal{B}(B^+ \rightarrow \ell^+ \nu_\ell)_{\text{total}} = \mathcal{B}(B^+ \rightarrow \ell^+ \nu_\ell)_{\text{SM}} \left(1 - \tan \beta \left(\frac{m_B^2}{m_{H^+}^2} \right) \right)^2, \quad (9.4)$$

where $\mathcal{B}(B^+ \rightarrow \ell^+ \nu_\ell)_{\text{SM}}$ is the prediction from the standard model, given in Equation 2.1 and m_{H^+} is the mass of the charged Higgs boson. Note that the 2HDM could still be true with no noticeable effect if

$$\tan \beta \left(\frac{m_B^2}{m_{H^+}^2} \right) \approx 2. \quad (9.5)$$

Since my measurement is consistent with previous measurements and the SM prediction, it constrains the possible values of $\tan\beta$ and m_{H^+} . Other searches and measurements have also constrained these values. All of these constraints are summarized in Figure 9.11. The two green wedges are the region excluded by the world average of $\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau)$ before my measurement was made public. The space between the wedges represents the case given in Equation 9.5. Since my measurement is consistent with the average, these wedges will not change greatly when it is included.

9.4 Summary

I have measured the branching fraction of $B^+ \rightarrow \tau^+\nu_\tau$ with semileptonic tags in the full dataset of *BABAR*. This measurement is consistent with all other contemporary measurements and the prediction of the standard model of particle physics. It also rules out a large area of phase space in the type II Two Higgs Doublet Model. I have also set the first upper limits on $\mathcal{B}(B^+ \rightarrow \mu^+\nu_\mu)$ and $\mathcal{B}(B^+ \rightarrow e^+\nu_e)$ using semileptonic tags.

N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$	N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$
662	19.43	$< 17.57 \times 10^{-5}$ (0.90 CL)	696	53.4	$(11.0^{+8.4}_{-7.3} \pm 0.4) \times 10^{-5}$
663	20.43	$< 17.99 \times 10^{-5}$ (0.90 CL)	697	54.4	$(11.3^{+8.2}_{-7.3} \pm 0.4) \times 10^{-5}$
664	21.43	$< 18.19 \times 10^{-5}$ (0.90 CL)	698	55.4	$(11.5^{+8.2}_{-7.5} \pm 0.4) \times 10^{-5}$
665	22.43	$< 18.40 \times 10^{-5}$ (0.90 CL)	699	56.4	$(11.7^{+8.2}_{-7.5} \pm 0.4) \times 10^{-5}$
666	23.43	$< 18.61 \times 10^{-5}$ (0.90 CL)	700	57.4	$(11.9^{+8.2}_{-7.5} \pm 0.4) \times 10^{-5}$
667	24.43	$< 18.81 \times 10^{-5}$ (0.90 CL)	701	58.4	$(12.1^{+8.4}_{-7.5} \pm 0.4) \times 10^{-5}$
668	25.43	$< 19.02 \times 10^{-5}$ (0.90 CL)	702	59.4	$(12.3^{+8.4}_{-7.5} \pm 0.4) \times 10^{-5}$
669	26.43	$< 19.02 \times 10^{-5}$ (0.90 CL)	703	60.4	$(12.5^{+8.4}_{-7.7} \pm 0.4) \times 10^{-5}$
670	27.43	$< 19.23 \times 10^{-5}$ (0.90 CL)	704	61.4	$(12.7^{+8.4}_{-7.7} \pm 0.5) \times 10^{-5}$
671	28.43	$< 19.85 \times 10^{-5}$ (0.90 CL)	705	62.4	$(12.9^{+8.2}_{-7.7} \pm 0.5) \times 10^{-5}$
672	29.43	$< 19.85 \times 10^{-5}$ (0.90 CL)	706	63.4	$(13.1^{+8.2}_{-7.7} \pm 0.5) \times 10^{-5}$
673	30.43	$< 20.05 \times 10^{-5}$ (0.90 CL)	707	64.4	$(13.3^{+8.2}_{-7.7} \pm 0.5) \times 10^{-5}$
674	31.43	$< 20.26 \times 10^{-5}$ (0.90 CL)	708	65.4	$(13.5^{+8.2}_{-7.9} \pm 0.5) \times 10^{-5}$
675	32.43	$< 20.47 \times 10^{-5}$ (0.90 CL)	709	66.4	$(13.7^{+8.4}_{-7.9} \pm 0.5) \times 10^{-5}$
676	33.43	$< 20.67 \times 10^{-5}$ (0.90 CL)	710	67.4	$(13.9^{+8.4}_{-7.9} \pm 0.5) \times 10^{-5}$
677	34.43	$< 20.88 \times 10^{-5}$ (0.90 CL)	711	68.4	$(14.1^{+8.2}_{-7.9} \pm 0.5) \times 10^{-5}$
678	35.43	$< 21.09 \times 10^{-5}$ (0.90 CL)	712	69.4	$(14.4^{+8.2}_{-7.9} \pm 0.5) \times 10^{-5}$
679	36.43	$< 21.29 \times 10^{-5}$ (0.90 CL)	713	70.4	$(14.6^{+8.4}_{-8.2} \pm 0.5) \times 10^{-5}$
680	37.43	$< 21.71 \times 10^{-5}$ (0.90 CL)	714	71.4	$(14.8^{+8.2}_{-8.2} \pm 0.5) \times 10^{-5}$
681	38.43	$< 21.71 \times 10^{-5}$ (0.90 CL)	715	72.4	$(15.0^{+8.4}_{-8.2} \pm 0.5) \times 10^{-5}$
682	39.43	$< 21.91 \times 10^{-5}$ (0.90 CL)	716	73.4	$(15.2^{+8.4}_{-8.2} \pm 0.5) \times 10^{-5}$
683	40.43	$< 22.12 \times 10^{-5}$ (0.90 CL)	717	74.4	$(15.4^{+8.4}_{-8.2} \pm 0.6) \times 10^{-5}$
684	41.43	$< 22.53 \times 10^{-5}$ (0.90 CL)	718	75.4	$(15.6^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
685	42.43	$< 22.53 \times 10^{-5}$ (0.90 CL)	719	76.4	$(15.8^{+8.4}_{-8.2} \pm 0.6) \times 10^{-5}$
686	43.43	$< 22.74 \times 10^{-5}$ (0.90 CL)	720	77.4	$(16.0^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
687	44.43	$< 23.15 \times 10^{-5}$ (0.90 CL)	721	78.4	$(16.2^{+8.4}_{-8.4} \pm 0.6) \times 10^{-5}$
688	45.43	$< 23.15 \times 10^{-5}$ (0.90 CL)	722	79.4	$(16.4^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
689	46.43	$< 23.57 \times 10^{-5}$ (0.90 CL)	723	80.4	$(16.6^{+8.2}_{-8.4} \pm 0.6) \times 10^{-5}$
690	47.43	$< 23.57 \times 10^{-5}$ (0.90 CL)	724	81.4	$(16.8^{+8.2}_{-8.4} \pm 0.6) \times 10^{-5}$
691	48.43	$< 23.77 \times 10^{-5}$ (0.90 CL)	725	82.4	$(17.0^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
692	49.43	$< 23.98 \times 10^{-5}$ (0.90 CL)	726	83.4	$(17.2^{+8.4}_{-8.2} \pm 0.6) \times 10^{-5}$
693	50.43	$< 24.19 \times 10^{-5}$ (0.90 CL)	727	84.4	$(17.5^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
694	51.43	$< 24.39 \times 10^{-5}$ (0.90 CL)	728	85.4	$(17.7^{+8.4}_{-8.2} \pm 0.6) \times 10^{-5}$
695	52.43	$< 24.60 \times 10^{-5}$ (0.90 CL)	729	86.4	$(17.9^{+8.2}_{-8.2} \pm 0.6) \times 10^{-5}$
			730	87.4	$(18.1^{+8.4}_{-8.2} \pm 0.7) \times 10^{-5}$

Table 9.5: $B^+ \rightarrow \tau^+ \nu_\tau$ with the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ channel.

N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$	N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow \tau^+ \nu_\tau)$
541	20.18	$< 16.75 \times 10^{-5}$ (0.90 CL)	569	48.2	$(10.0^{+7.6}_{-6.7} \pm 0.4) \times 10^{-5}$
542	21.18	$< 16.75 \times 10^{-5}$ (0.90 CL)	570	49.2	$(10.2^{+7.6}_{-6.7} \pm 0.4) \times 10^{-5}$
543	22.18	$< 17.16 \times 10^{-5}$ (0.90 CL)	571	50.2	$(10.4^{+7.4}_{-6.7} \pm 0.4) \times 10^{-5}$
544	23.18	$< 17.37 \times 10^{-5}$ (0.90 CL)	572	51.2	$(10.6^{+7.4}_{-6.7} \pm 0.4) \times 10^{-5}$
545	24.18	$< 17.37 \times 10^{-5}$ (0.90 CL)	573	52.2	$(10.8^{+7.4}_{-6.9} \pm 0.4) \times 10^{-5}$
546	25.18	$< 17.78 \times 10^{-5}$ (0.90 CL)	574	53.2	$(11.0^{+7.4}_{-6.9} \pm 0.4) \times 10^{-5}$
547	26.18	$< 17.99 \times 10^{-5}$ (0.90 CL)	575	54.2	$(11.2^{+7.4}_{-6.9} \pm 0.4) \times 10^{-5}$
548	27.18	$< 18.19 \times 10^{-5}$ (0.90 CL)	576	55.2	$(11.4^{+7.4}_{-7.1} \pm 0.4) \times 10^{-5}$
549	28.18	$< 18.19 \times 10^{-5}$ (0.90 CL)	577	56.2	$(11.6^{+7.4}_{-7.1} \pm 0.4) \times 10^{-5}$
550	29.18	$< 18.61 \times 10^{-5}$ (0.90 CL)	578	57.2	$(11.8^{+7.4}_{-7.1} \pm 0.4) \times 10^{-5}$
551	30.18	$< 18.81 \times 10^{-5}$ (0.90 CL)	579	58.2	$(12.0^{+7.4}_{-7.1} \pm 0.4) \times 10^{-5}$
552	31.18	$< 19.02 \times 10^{-5}$ (0.90 CL)	580	59.2	$(12.2^{+7.6}_{-7.1} \pm 0.5) \times 10^{-5}$
553	32.18	$< 19.23 \times 10^{-5}$ (0.90 CL)	581	60.2	$(12.4^{+7.4}_{-7.1} \pm 0.5) \times 10^{-5}$
554	33.18	$< 19.43 \times 10^{-5}$ (0.90 CL)	582	61.2	$(12.6^{+7.6}_{-7.1} \pm 0.5) \times 10^{-5}$
555	34.18	$< 19.64 \times 10^{-5}$ (0.90 CL)	583	62.2	$(12.9^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
556	35.18	$< 20.05 \times 10^{-5}$ (0.90 CL)	584	63.2	$(13.1^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
557	36.18	$< 20.05 \times 10^{-5}$ (0.90 CL)	585	64.2	$(13.3^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
558	37.18	$< 20.26 \times 10^{-5}$ (0.90 CL)	586	65.2	$(13.5^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
559	38.18	$< 20.47 \times 10^{-5}$ (0.90 CL)	587	66.2	$(13.7^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
560	39.18	$< 20.67 \times 10^{-5}$ (0.90 CL)	588	67.2	$(13.9^{+7.4}_{-7.3} \pm 0.5) \times 10^{-5}$
561	40.18	$< 20.88 \times 10^{-5}$ (0.90 CL)	589	68.2	$(14.1^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
562	41.18	$< 21.09 \times 10^{-5}$ (0.90 CL)	590	69.2	$(14.3^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
563	42.18	$< 21.29 \times 10^{-5}$ (0.90 CL)	591	70.2	$(14.5^{+7.6}_{-7.3} \pm 0.5) \times 10^{-5}$
564	43.18	$< 21.50 \times 10^{-5}$ (0.90 CL)	592	71.2	$(14.7^{+7.4}_{-7.5} \pm 0.5) \times 10^{-5}$
565	44.18	$< 21.71 \times 10^{-5}$ (0.90 CL)	593	72.2	$(14.9^{+7.6}_{-7.3} \pm 0.6) \times 10^{-5}$
566	45.18	$< 22.12 \times 10^{-5}$ (0.90 CL)	594	73.2	$(15.1^{+7.6}_{-7.5} \pm 0.6) \times 10^{-5}$
567	46.18	$< 22.12 \times 10^{-5}$ (0.90 CL)	595	74.2	$(15.3^{+7.6}_{-7.5} \pm 0.6) \times 10^{-5}$
568	47.18	$< 22.33 \times 10^{-5}$ (0.90 CL)	596	75.2	$(15.5^{+7.6}_{-7.3} \pm 0.6) \times 10^{-5}$
			597	76.2	$(15.7^{+7.4}_{-7.5} \pm 0.6) \times 10^{-5}$
			598	77.2	$(16.0^{+7.6}_{-7.3} \pm 0.6) \times 10^{-5}$
			599	78.2	$(16.2^{+7.6}_{-7.5} \pm 0.6) \times 10^{-5}$
			600	79.2	$(16.4^{+7.6}_{-7.5} \pm 0.6) \times 10^{-5}$

Table 9.6: $B^+ \rightarrow \tau^+ \nu_\tau$ without the $\tau^+ \rightarrow \pi^+ \pi^- \pi^+ \bar{\nu}_\tau$ channel.

N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow \mu^+ \nu_\mu)$
6	-9.15	$< 0.72 \times 10^{-5}$ (0.90 CL)
7	-8.15	$< 0.80 \times 10^{-5}$ (0.90 CL)
8	-7.15	$< 0.80 \times 10^{-5}$ (0.90 CL)
9	-6.15	$< 0.89 \times 10^{-5}$ (0.90 CL)
10	-5.15	$< 0.97 \times 10^{-5}$ (0.90 CL)
11	-4.15	$< 1.05 \times 10^{-5}$ (0.90 CL)
12	-3.15	$< 1.13 \times 10^{-5}$ (0.90 CL)
13	-2.15	$< 1.13 \times 10^{-5}$ (0.90 CL)
14	-1.15	$< 1.21 \times 10^{-5}$ (0.90 CL)
15	-0.15	$< 1.29 \times 10^{-5}$ (0.90 CL)
16	0.85	$< 1.37 \times 10^{-5}$ (0.90 CL)

Table 9.7: $B^+ \rightarrow \mu^+ \nu_\mu$

N_{obs}	N_{sig}	$\mathcal{B}(B^+ \rightarrow e^+ \nu_e)$
13	-11.01	$< 0.59 \times 10^{-5}$ (0.90 CL)
14	-10.01	$< 0.59 \times 10^{-5}$ (0.90 CL)
15	-9.01	$< 0.71 \times 10^{-5}$ (0.90 CL)
16	-8.01	$< 0.71 \times 10^{-5}$ (0.90 CL)
17	-7.01	$< 0.77 \times 10^{-5}$ (0.90 CL)
18	-6.01	$< 0.77 \times 10^{-5}$ (0.90 CL)
19	-5.01	$< 0.83 \times 10^{-5}$ (0.90 CL)
20	-4.01	$< 0.89 \times 10^{-5}$ (0.90 CL)
21	-3.01	$< 0.89 \times 10^{-5}$ (0.90 CL)
22	-2.01	$< 0.94 \times 10^{-5}$ (0.90 CL)

Table 9.8: $B^+ \rightarrow e^+ \nu_e$

BIBLIOGRAPHY

- [1] **Particle Data Group** Collaboration, W.-M. Yao *et al.*, “Review of particle physics,” *J. Phys. G: Nucl. Part. Phys.* **33** (2006) 1.
- [2] **BABAR** Collaboration, P. F. Harrison and H. R. Quinn, eds., *The BaBar physics book: Physics at an asymmetric B factory*. 1997. 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.
- [3] U. Haisch, “anti-B \rightarrow X(s) gamma: Standard Model and Beyond,”
0805.2141.
- [4] H. Geiger and E. Marsden, “On a diffuse reflection of the α -particles,”
Proceedings of the Royal Society of London. Series A **82** (July, 1909) 495–500.
- [5] R. P. Feynman, M. Gell-Mann, and G. Zweig, “Group U(6) x U(6) generated by current components,” *Phys. Rev. Lett.* **13** (1964) 678–680.
- [6] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*. John Wiley and Sons, Inc., New York, NY, fifth ed., 1997.
- [7] *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984.

- [8] J. Ellis, G. Giudice, M. L. Mangano, I. Tkachev, and U. Wiedemann, “Review of the Safety of LHC Collisions,” *J. Phys.* **G35** (2008) 115004, 0806.3414.
- [9] W.-S. Hou, “Enhanced charged Higgs boson effects in $B^- \rightarrow \tau\bar{\nu}, \mu\bar{\nu}$ and $b \rightarrow \tau\bar{\nu} + X$,” *Phys. Rev.* **D48** (1993), no. 5, 2342.
- [10] **Heavy Flavor Averaging Group (HFAG)** Collaboration, E. Barberio *et al.*, “Averages of b-hadron properties at the end of 2006,” 0704.3575.
- [11] S. S. Gershtein and M. Y. Khlopov, “SU(4) Symmetry Breaking and Lepton Decays of Heavy Pseudoscalar Mesons,” *JETP Lett.* **23** (1976) 338.
- [12] M. Y. Khlopov, “Effects of symmetry breaking in semileptonic decays of mesons,” *Sov. J. Nucl. Phys.* **28** (1978), no. 4, 583.
- [13] **HPQCD** Collaboration, A. Gray *et al.*, “The b meson decay constant from unquenched lattice qcd,” *Phys. Rev. Lett.* **95** (2005) 212001, hep-lat/0507015.
- [14] **CKMfitter Group** Collaboration, J. Charles *et al.*, “CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories,” *Eur. Phys. J.* **C41** (2005) 1–131, hep-ph/0406184.
- [15] **UTfit** Collaboration, M. Bona *et al.*, “The unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: A reappraisal after the measurements of Δ_{m_s} and $\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau)$,” *JHEP* **0610** (2006) 081.
- [16] **BABAR** Collaboration, B. Aubert *et al.*, “Search for $B^+ \rightarrow \ell^+\nu_\ell$ with inclusive reconstruction at *BABAR*,” 0807.4187.

- [17] **Belle** Collaboration, N. Satoyama *et al.*, “A search for the rare leptonic decays $B^+ \rightarrow \mu^+\nu$ and $B^+ \rightarrow e^+\nu$,” *Phys. Lett.* **B647** (2007) 67–73, hep-ex/0611045.
- [18] **BABAR** Collaboration, B. Aubert *et al.*, “Searches for the decays $B^0 \rightarrow \ell^+\tau^-$ and $B^+ \rightarrow \ell^+\nu$ ($\ell = e, \mu$) using hadronic tag reconstruction,” 0801.0697.
- [19] **BABAR** Collaboration, B. Aubert *et al.*, “A search for $B^+ \rightarrow \tau^+\nu$ with Hadronic B tags,” *Phys. Rev.* **D77** (2008) 011107, 0708.2260.
- [20] K. Ikado *et al.*, “Evidence of the purely leptonic decay $B^- \rightarrow \tau^-\bar{\nu}_\tau$,” *Phys. Rev. Lett.* **97** (2006) 251802, hep-ex/0604018.
- [21] **The Belle** Collaboration, I. Adachi, “Measurement of $B^- \rightarrow \tau^-\bar{\nu}_\tau$ Decay With a Semileptonic Tagging Method,” 0809.3834.
- [22] **ALEPH** Collaboration, D. Buskulic *et al.*, “Measurement of the $b \rightarrow \tau^-\bar{\nu}_\tau X$ branching ratio and an upper limit on $B^- \rightarrow \tau^-\bar{\nu}_\tau$,” *Phys. Lett.* **B343** (1995) 444–452.
- [23] **Cleo** Collaboration, M. Artuso *et al.*, “A Search for $B \rightarrow \ell\bar{\nu}_\ell$,” *Phys. Rev. Lett.* **75** (1995) 785–789.
- [24] **BABAR** Collaboration, B. Aubert *et al.*, “A search for the decay $B^+ \rightarrow \tau^+\nu_\tau$,” *Phys. Rev.* **D73** (2006) 057101.
- [25] **BABAR** Collaboration, B. Aubert *et al.*, “A search for $B^+ \rightarrow \tau^+\nu$,” *Phys. Rev.* **D76** (2007) 0520020, arXiv:0705.1820 [hep-ex].

- [26] “PEP-II: An Asymmetric B Factory. Conceptual Design Report. June 1993,” SLAC-418.
- [27] **BABAR** Collaboration, B. Aubert *et al.*, “The BABAR detector,” *Nuclear Instruments and Methods in Physics Research* **A479** (2002) 1–116, hep-ex/0105044.
- [28] **Belle** Collaboration, A. Abashian *et al.*, “The Belle detector,” *Nuclear Instruments and Methods in Physics Research* **A479** (2002) 117–232.
- [29] **BABAR** Collaboration, B. Aubert *et al.*, “Study of inclusive production of charmonium mesons in B decay,” *Phys. Rev.* **D67** (2003) 032002, hep-ex/0207097.
- [30] S. Banerjee, B. Pietrzyk, J. M. Roney, and Z. Was, “Tau and muon pair production cross-sections in electron- positron annihilations at $\sqrt{s} = 10.58$ GeV,” *Phys. Rev.* **D77** (2008) 054012, 0706.3235.
- [31] G. Benelli, K. Honscheid, E. A. Lewis, J. J. Regensburger, and D. S. Smith, “The BaBar LST detector high voltage system: Design and implementation,” *IEEE Nucl. Sci. Symp. Conf. Rec.* **2** (2006) 1145–1148.
- [32] W. Menges, “The BABAR muon system upgrade,” *IEEE Nucl. Sci. Symp. Conf. Rec.* **5** (2006) 1470–1474, physics/0609039.
- [33] **GEANT4** Collaboration, S. Agostinelli *et al.*, “Geant4: A simulation toolkit,” *Nuclear Instruments and Methods in Physics Research* **A506** (2003) 250–303.

- [34] G. J. Feldman and R. D. Cousins, “A Unified approach to the classical statistical analysis of small signals,” *Phys. Rev.* **D57** (1998) 3873–3889, physics/9711021.
- [35] I. Narsky, “StatPatternRecognition: A C++ Package for Statistical Analysis of High Energy Physics Data,” *Phys. Rev. Lett.* **97** (2005) 251802, physics/0507143.
- [36] **BABAR** Collaboration, B. Aubert *et al.*, “Measurement of $B \rightarrow D^*$ form factors in the semileptonic decay $\overline{B^0} \rightarrow D^{*+}\ell^-\bar{\nu}$,” hep-ex/0409047.
- [37] **BABAR** Collaboration, B. Aubert *et al.*, “Measurement of the $B^0 \rightarrow \pi^-\ell^+\nu$ form-factor shape and branching fraction, and determination of $|V_{ub}|$ with a loose neutrino reconstruction technique,” *Phys. Rev. Lett.* **98** (2007) 091801, hep-ex/0612020.
- [38] D. Zwillinger, ed., *CRC Standard Mathematical Tables and Formulae*. Chapman and Hall/CRC, 31 ed., 2003.
- [39] G. C. Fox and S. Wolfram, “Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes,” *Phys. Rev. Lett.* **41** (1978) 1581.
- [40] G. Punzi, “Sensitivity of searches for new signals and its optimization,” physics/0308063.