

A SEARCH FOR NEUTRINOLESS TAU DECAYS TO THREE LEPTONS

by

JEFFREY A. KOLB

A DISSERTATION

Presented to the Department of Physics  
and the Graduate School of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

June 2008

**University of Oregon Graduate School**

**Confirmation of Approval and Acceptance of Dissertation prepared by:**

Jeffrey Kolb

Title:

"A Search for Neutrinoless Tau Decays to Three Leptons"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Physics by:

David Strom, Chairperson, Physics  
Eric Torrence, Advisor, Physics  
Nilendra Deshpande, Member, Physics  
Michael Raymer, Member, Physics  
Michael Kellman, Outside Member, Chemistry

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2008

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.



## CURRICULUM VITAE

NAME OF AUTHOR: Jeffrey A. Kolb

PLACE OF BIRTH: Williamsport, Pennsylvania

DATE OF BIRTH: December 18, 1979

## GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

Taylor University, Upland, IN

## DEGREES

Doctor of Philosophy in Physics, 2008, University of Oregon

Master of Science in Physics, 2005, University of Oregon

Bachelor of Science in Engineering Physics, 2002, Taylor University

## ACADEMIC INTERESTS

Tau decays

Lepton flavor violation

Distributed computing in high energy physics

## PROFESSIONAL EXPERIENCE

Research Assistant, Department of Physics, University of Oregon, Eugene,  
2005-2008

Teaching Assistant, Department of Physics, University of Oregon, Eugene,  
2002-2004

## PUBLICATIONS

BaBar Collaboration, B. Aubert, *et al.*, *Improved Limits on the Lepton-Flavor Violating Decays  $\tau^- \rightarrow \ell^- \ell^+ \ell^-$* , Phys. Rev. Lett. **99**, 251803 (2007)

K. Kiers, J. Kolb, T. Klein, S. Price, and J.C. Sprott, *Chaos in a Nonlinear Analog Computer*, International Journal of Bifurcation and Chaos **14**, 2867-2873 (2004)

K. Kiers, J. Kolb, J. Lee, A. Soni, and G.-H. Wu, *Ubiquitous CP Violation in a Top-Inspired Left-Right Model*, Phys. Rev. D **66**, 095002 (2002).

## ACKNOWLEDGMENTS

First of all, I would like to acknowledge the support and guidance of my advisor, Eric Torrence, who has been unfailingly happy to share from his impressive knowledge of experimental particle physics. Eric's enthusiasm for doing physics has catalyzed the development of my interest in the field, and his sound advice over the years has led me down a path on which I'm pleasantly surprised to find myself.

I would also like to acknowledge the many friends who have traveled with me for all or part of this physics journey. I am grateful to The Men of the Man Ranch (Adam, Adam, Paul, Chuck, Brian), Al, Iva, Tiffany, the K Center folks, and the rest of my Eugene friends for their support. I also acknowledge all the wonderful people from Solid Start and small group who have been so encouraging.

Finally, I acknowledge my family, who have always cared for and supported me. And Ruthie, who has loved me well.

## DEDICATION

This dissertation is dedicated to Dr. Ken Kiers. Ken's kindness is inspiring and his overwhelming dedication to his students allowed this student to pursue a dream.

## TABLE OF CONTENTS

Chapter		Page
I	INTRODUCTION . . . . .	1
1	Overview . . . . .	1
2	Theoretical Motivation . . . . .	3
	2.1 Lepton Flavor . . . . .	4
	2.2 The Standard Model of Electromagnetic and Weak Interactions . . . . .	4
	2.3 Flavor Structures Beyond the Standard Model . . . . .	8
3	Previous Experimental Work . . . . .	9
	3.1 Search for Neutrinoless Muon Decays . . . . .	9
	3.2 Search for Neutrinoless Tau Decays . . . . .	11
II	THE BABAR EXPERIMENT . . . . .	15
1	Introduction . . . . .	15
2	Particle Acceleration for the <i>BABAR</i> Experiment . . . . .	16
	2.1 Beam Production . . . . .	17
	2.2 Beam Storage . . . . .	17
	2.3 Beam Energy . . . . .	18
	2.4 Interaction Region . . . . .	18
	2.5 Performance . . . . .	19
3	The <i>BABAR</i> Detector . . . . .	19
	3.1 Detector Goals and Constraints . . . . .	22
	3.2 Charged Particle Tracking . . . . .	26
	3.3 Pion and Kaon Identification . . . . .	36
	3.4 Electromagnetic Calorimetry . . . . .	37
	3.5 Muon Detection . . . . .	44
4	Simulations . . . . .	47
5	Data Acquisition and Triggering . . . . .	51

Chapter	Page
5.1	Trigger Requirements and Design . . . . . 51
5.2	Level One Trigger . . . . . 52
5.3	Level Three Trigger . . . . . 57
6	Offline Data Processing . . . . . 57
6.1	Prompt Data Reconstruction . . . . . 58
6.2	Data Skimming . . . . . 59
III	DATA ANALYSIS . . . . . 61
1	Introduction to Analysis . . . . . 61
1.1	Branching Fractions . . . . . 62
1.2	Upper Limits . . . . . 63
1.3	Incorporating Uncertainties into Upper Limits . . . . . 66
1.4	Overview of Analysis Steps . . . . . 69
1.5	Analysis Optimization and Expected Upper Limits . . . . . 70
2	Selection of the Data . . . . . 71
3	Event Preselection . . . . . 72
4	Particle Identification . . . . . 73
5	Mass and Energy Determination . . . . . 80
6	Event Selection . . . . . 82
7	Estimation of Background . . . . . 85
7.1	Backgrounds from $c\bar{c}$ and $uds$ . . . . . 85
7.2	Background from $\tau^+\tau^-$ . . . . . 88
7.3	QED Background . . . . . 88
7.4	Final Background Fit . . . . . 90
8	Systematic Uncertainties . . . . . 97
8.1	Signal Efficiency . . . . . 97
8.2	Background Estimation . . . . . 101
8.3	Other Systematics . . . . . 105
IV	RESULTS AND CONCLUSIONS . . . . . 107
1	Results . . . . . 107
2	Discussion of Results . . . . . 108
2.1	Implications for Theory . . . . . 108
3	Conclusion . . . . . 111
	APPENDICES . . . . . 113
A	TRACK LISTS . . . . . 113

Chapter		Page
1	The CalorClusterNeutral List . . . . .	113
2	The ChargedTracks List . . . . .	113
3	The GoodTracksVeryLoose List . . . . .	113
4	The gammaConversionDefault List . . . . .	113
B	PARTICLE IDENTIFICATION ALGORITHMS . . . . .	115
1	The eMicroTight Selector . . . . .	115
2	The muNNLoose Selector . . . . .	116
3	The KLHTight Selector . . . . .	117
C	AUXILIARY PLOTS . . . . .	118
1	Optimization Plots . . . . .	118
2	N-1 Plots . . . . .	125
3	Plots of Background Fits . . . . .	132
	REFERENCES . . . . .	136

## LIST OF FIGURES

Figure	Page
II.1 The PEP-II interaction region around the <i>BABAR</i> detector. . . . .	20
II.2 Integrated luminosity at <i>BABAR</i> as a function of time. . . . .	21
II.3 Sideview of the <i>BABAR</i> detector. . . . .	24
II.4 Endview of the <i>BABAR</i> detector. . . . .	25
II.5 Longitudinal schematic view of the SVT. . . . .	27
II.6 Transverse schematic view of the SVT. . . . .	28
II.7 Longitudinal view of the DCH, with dimensions in mm. . . . .	29
II.8 Schematic layout of the drift cells for the four innermost superlayers. The lines between the wires have been added to show the cell boundaries. The line below the first layer is the 1-mm-thick beryllium inner wall. . . . .	30
II.9 The magnetic field components $B_z$ and $B_r$ as a function of $z$ for various radial distances $r$ (in meters). . . . .	31
II.10 Relative magnitude of the magnetic field transverse to a high momentum track as a function of track length from the IP for various polar angles (in degrees). The data are normalized to the field at the origin. . . . .	32
II.11 SVT hit reconstruction efficiency, as measured on $\mu^+\mu^-$ events for (a) forward half-modules and (b) backward half modules. Vertical lines delineate the five different layers. . . . .	33
II.12 SVT hit resolution in (left) $z$ and (right) $\phi$ coordinate in microns, plotted as a function of track incident angle in degrees. . . . .	34
II.13 DCH position resolution as a function of the drift distance in layer 18, for tracks on the left and right side of the sense wire. The data are averaged over all cells in the layer. . . . .	35
II.14 Measurements of $dE/dx$ in the DCH as a function of track momentum. The curves show the Bethe-Bloch predictions for each particle type. . . . .	35

Figure	Page
II.15 Transverse momentum resolution, as determined from cosmic ray muons traversing the DCH and SVT. . . . .	39
II.16 Longitudinal view of the DIRC. . . . .	39
II.17 Detail of the DIRC bars and the imaging region. . . . .	40
II.18 Longitudinal view of the EMC. . . . .	41
II.19 A schematic of the wrapped crystal and the readout electronics on the back end. . . . .	42
II.20 Energy resolution of the EMC for photons and electrons, as measured for various processes. The solid line is from the fit (Equation II.4), and the shaded area denotes the fit error. . . . .	43
II.21 Angular resolution of the for photons from $\pi^0$ decays. The solid line is the fit (Equation II.5). . . . .	44
II.22 Overview of the IFR: barrels sectors and forward (FW) and backward (BW) endcaps. . . . .	45
II.23 Cross section of a planar RPC, with schematic of the high voltage connections. . . . .	46
II.24 Track Segment Finder pivot group. . . . .	53
II.25 Single track $z_0$ for all L1 tracks without a cut on $z_0$ . . . . .	55
III.1 The efficiency for muon identification in data and MC by the <b>muNNLoose</b> selector, as a function of muon momentum for (a) positively charged muons, and (b) negatively charged muons. Plot (c) shows the ratio of the data efficiency to the MC efficiency. . . . .	78
III.2 The efficiency for $e^+/e^-$ identification in data and MC by the <b>eMicroTight</b> selector, as a function of particle momentum for (a) positrons, and (b) electrons. Plot (c) shows the ratio of the data efficiency to the MC efficiency. . . . .	78
III.3 The mis-ID rate for pions in data and MC by the <b>eMicroTight</b> selector, as a function of particle momentum for (a) positively charged pions, and (b) negatively charged pions. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate. . . . .	78

Figure	Page
III.4 The mis-ID rate for pions in data and MC by the <b>muNNLoose</b> selector, as a function of particle momentum for (a) positively charged pions, and (b) negatively charged pions. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate. . . . .	79
III.5 The mis-ID rate for kaons in data and MC by the <b>eMicroTight</b> selector, as a function of particle momentum for (a) positively charged kaons, and (b) negatively charged kaons. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate. . . . .	79
III.6 The mis-ID rate for kaons in data and MC by the <b>muNNLoose</b> selector, as a function of particle momentum for (a) positively charged kaons, and (b) negatively charged kaons. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate. . . . .	80
III.7 The $(\Delta M, \Delta E)$ distributions for the signal channels after preselection and particle identification. The box shows the borders of the signal region. The histogram borders correspond to the large box. The z-axis is logarithmically-scaled. . . . .	82
III.8 $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ : $uds$ background. The histograms show the average PID-weight per bin as a function of $\Delta M$ (left) and $\Delta E$ (right). . . . .	86
III.9 $\tau^- \rightarrow e^- e^+ e^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	91
III.10 $\tau^- \rightarrow \mu^- e^+ e^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	92

Figure	Page
III.11 $\tau^- \rightarrow \mu^+ e^- e^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	93
III.12 $\tau^- \rightarrow e^+ \mu^- \mu^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	94
III.13 $\tau^- \rightarrow e^- \mu^+ \mu^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	95
III.14 $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ channel: PDFs with MC-fitted shapes are scaled to data; a) $\Delta E$ projection of the data (points) and the sum of the background PDFs (curve); b) $\Delta M$ projection of the data (points) and the sum of the background PDFs (curve); c) PDF $(\Delta M, \Delta E)$ distribution; d) data $(\Delta M, \Delta E)$ distribution. The filled black boxes and open red box show the signal region (blinded for data). . . . .	96
III.15 $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ a) generated MC Dalitz distribution after preselection; b) efficiency to pass all selection except SB as function of Dalitz distribution; c) selection efficiency as a function of invariant mass squared of the pair of same-sign leptons; d) selection efficiency as a function of invariant mass squared of the pair of opposite-sign leptons; e) efficiency for all selection cuts except PID as a function of invariant mass squared of the pair of same- sign leptons; e) efficiency for all selection cuts except PID as a function of invariant mass squared of the pair of opposite-sign leptons; . . . . .	99

Figure	Page
III.16a) The distribution of $p_1^{cms}$ versus $P_T^{cms}$ for the $\tau^- \rightarrow e^- e^+ e^-$ channel after preselection and PID. The z-axis is logarithmic. The red line shows the cuts applied to select the two-photon control sample. b) The $(\Delta M, \Delta E)$ distribution of these events plotted without the cut on $P_T^{cms}$ . . . . .	104
IV.1 Observed data shown as dots in the $(\Delta M, \Delta E)$ plane and the boundaries of the signal region for each decay mode. The dark and light shading indicates contours containing 50% and 90% of the selected MC signal events, respectively. . . . .	109
IV.2 Feynman diagram of the leading Higgs-induced contribution to $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ in the MSSM. . . . .	111
C.1 Expected upper limit on branching fraction as a function of $p_1^{cms}$ . $\mu^- e^+ e^-$ channel is excluded because the dimuon control sample does not have a requirement on the maximum value of $p_1^{cms}$ (see section 7). Arrows indicate optimized value for the selection cut. . . . .	119
C.2 Expected upper limit on branching fraction as a function of the minimum total transverse momentum $P_T^{cms}$ . Arrows indicate optimized value for the selection cut. . . . .	120
C.3 Expected upper limit on branching fraction as a function of $\Delta M_{min}^{SB}$ . Arrows indicate optimized value for the selection cut. . . . .	121
C.4 Expected upper limit on branching fraction as a function of $\Delta M_{max}^{SB}$ . Arrows indicate optimized value for the selection cut. . . . .	122
C.5 Expected upper limit on branching fraction as a function of $\Delta E_{min}^{SB}$ . Arrows indicate optimized value for the selection cut. . . . .	123
C.6 Expected upper limit on branching fraction as a function of $\Delta E_{max}^{SB}$ . Arrows indicate optimized value for the selection cut. . . . .	124

Figure	Page
<p>C.7 <math>\tau^- \rightarrow e^-e^+e^-</math> channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected <math>\tau^+\tau^-</math> background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected <math>uds</math> background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . .</p>	126
<p>C.8 <math>\tau^- \rightarrow \mu^-e^+e^-</math> channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected <math>\tau^+\tau^-</math> background level, the green histogram shows the expected dimuon background level, and the yellow histogram shows the expected <math>uds</math> background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . .</p>	127
<p>C.9 <math>\tau^- \rightarrow \mu^+e^-e^-</math> channel a) total transverse momentum; b) 1-prong momentum; c) min2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected <math>\tau^+\tau^-</math> background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected <math>uds</math> background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . .</p>	128

- C.10  $\tau^- \rightarrow e^+\mu^-\mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+\tau^-$  background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . . 129
- C.11  $\tau^- \rightarrow e^-\mu^+\mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) minimum 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The green histogram shows the expected Bhabha background level and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . . 130
- C.12  $\tau^- \rightarrow \mu^-\mu^+\mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) minimum 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+\tau^-$  background level and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization. . . . . 131

- C.13 *uds* background: Fit of the *uds* ( $\Delta M, \Delta E$ ) MC distribution with PDF described in the text. column 1)  $\Delta M$  projection of the MC distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the MC distribution (points) and the PDF (curve); column 3) MC ( $\Delta M, \Delta E$ ) distribution; column 4) MC PDF ( $\Delta M, \Delta E$ ) distribution; row 1)  $e^-e^+e^-$ ; row 2)  $\mu^-e^+e^-$ ; row 3)  $e^-\mu^+e^-$ ; row 4)  $\mu^-e^+\mu^-$ ; row 5)  $e^-\mu^+\mu^-$ ; row 6)  $\mu^-\mu^+\mu^-$ .  $\Delta M$  is plotted in ( $\text{GeV}/c^2$ ) and  $\Delta E$  is plotted in ( $\text{GeV}$ ). . . . . 133
- C.14  $\tau^+\tau^-$  background: Fit of the  $\tau^+\tau^-$  ( $\Delta M, \Delta E$ ) MC distribution with PDF described in the text. column 1)  $\Delta M$  projection of the MC distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the MC distribution (points) and the PDF (curve); column 3) MC ( $\Delta M, \Delta E$ ) distribution; column 4) MC PDF ( $\Delta M, \Delta E$ ) distribution; row 1)  $\mu^-e^+e^-$ ; row 2)  $e^-\mu^+e^-$ ; row 3)  $\mu^-e^+\mu^-$ ; row 4)  $e^-\mu^+\mu^-$ ; row 5)  $\mu^-\mu^+\mu^-$ .  $\Delta M$  is plotted in ( $\text{GeV}/c^2$ ) and  $\Delta E$  is plotted in ( $\text{GeV}$ ). Only channels with significant  $\tau^+\tau^-$  contributions in the LB are shown. . . . . 134
- C.15 Bhabha and di-muon backgrounds: Fit of the Bhabha and di-muon ( $\Delta M, \Delta E$ ) MC distributions with PDF described in the text. column 1)  $\Delta M$  projection of the control sample distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the control sample distribution (points) and the PDF (curve); column 3) control sample ( $\Delta M, \Delta E$ ) distribution; column 4) PDF ( $\Delta M, \Delta E$ ) distribution; row 1)  $e^-e^+e^-$ ; row 2)  $\mu^-e^+e^-$ ; row 3)  $e^-\mu^+\mu^-$ .  $\Delta M$  is plotted in ( $\text{GeV}/c^2$ ) and  $\Delta E$  is plotted in ( $\text{GeV}$ ). Only channels with significant QED contributions in the LB are shown. . . . . 135

## LIST OF TABLES

Table	Page
II.1 Approximate production cross sections at <i>BABAR</i> , including experimental acceptance factors. <i>uds</i> refers to the total continuum production to $u\bar{u}, d\bar{d}, s\bar{s}$ . . . . .	17
II.2 Integrated luminosity records for various time periods, in inverse picobarns (pb) and inverse femtobarns (fb), where 1 barn = $10^{-28}$ m <sup>2</sup> . . . . .	19
II.3 Properties of Csl(Tl). . . . .	40
II.4 Primitive trigger objects constructed by the Level 1 trigger. . . . .	56
II.5 IFT trigger patter definitions, where $\mu$ refers to a signal in a sector. . . . .	56
III.1 Background MC samples used in the analysis. . . . .	72
III.2 Preselection efficiencies in percent for signal MC, background MC, and data samples. Cuts are applied sequentially and the marginal efficiencies are quoted. For the signal samples, the loss in efficiency due to the one-prong branching fraction is <i>included</i> in these numbers. ‘Trigger’ means that <b>L3OutDch</b> or <b>L3OutEmc</b> tagbit is set. The $b\bar{b}$ efficiencies include both $B^0\bar{B}^0$ and $B^+B^-$ samples. Uncertainties on the total efficiency numbers are from MC statistics. . . . .	74
III.3 Particle ID selectors used to identify the 3-prong tracks. . . . .	75
III.4 Efficiency for preselected events to pass the PID requirements. . . . .	79
III.5 Signal region boundaries $M_1 < \Delta M < M_2, E_1 < \Delta E < E_2$ for each decay mode. The boundaries of the large box ( <i>LB</i> ) used in the background fits is also shown in the last column. The last row shows the signal efficiencies in percent for these signal regions (for the events passed preselection and PID requirements). . . . .	81

Table	Page
III.6 Signal efficiency for events passing preselection and PID to be in the signal box or in the large box. . . . .	81
III.7 Efficiency for events after PID and LB requirements to pass the selection cuts. As described in Section 7, the Bhabha and dimuon contributions are modeled with data control samples. The corresponding selection efficiencies are not shown. . . . .	84
III.8 Expected number of background events in the grand sideband (GS) and signal box (SB) after the background fits. By construction, the total number of expected background events in the GS is equal to the number of data events in the GS. The luminosity is $376 \text{ fb}^{-1}$ . . . . .	90
III.9 The number of expected (left) and observed (right) events in the boxes neighbor to the signal box. Uncertainties on the sum of the expected background for all four boxes are estimated from the uncertainty on the expected background in the SB. Poisson errors are assigned to the sum of the data for all four boxes. . . . .	103
III.10 Systematic uncertainties expressed in relative percent. . . . .	106
IV.1 The total efficiency $\varepsilon$ , estimated background level in the signal region, expected upper limit, observed number of events in the SB and 90% CL upper limit on $\mathcal{B}(\tau \rightarrow \ell\ell\ell)$ . . . . .	108

# CHAPTER I

## INTRODUCTION

### 1 Overview

Tau lepton decays have always been observed to include at least one neutrino in the final state. Furthermore, neutral leptonic currents such as the photon appear to always generate pairs of leptons of the same type. These sort of observations have led to the postulate that the number of leptons of each type (or flavor) are separately conserved in all reactions. Recent observations of neutrino flavor oscillations provide an unambiguous signature of the non-conservation of lepton flavor, or lepton flavor violation (LFV). But do the interactions of charged leptons still conserve lepton flavor? Many extensions of the standard theory of particle physics naturally predict LFV. In fact, neutrinoless lepton decays could be the first sign of physics beyond the standard theory. In particular, the heaviness of the third generation lepton, the tau, makes it attractive for probing theoretical models with new particles that couple preferentially to more massive particles.

In this work, we present a search for the neutrinoless tau decays  $\tau \rightarrow \ell\ell\ell$ , where  $\ell = e, \mu$ , the two lighter charged leptons. We search for all such decays consistent with the conservation of energy and electric charge:

- $\tau^- \rightarrow e^- e^+ e^-$ ,
- $\tau^- \rightarrow \mu^- e^+ e^-$ ,
- $\tau^- \rightarrow \mu^+ e^- e^-$ ,
- $\tau^- \rightarrow e^+ \mu^- \mu^-$ ,

- $\tau^- \rightarrow e^- \mu^+ \mu^-$ ,
- $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ .

Throughout this work, the charge-conjugate modes are implied.

The construction and operation of the *BABAR* detector and the PEP-II storage rings at the Stanford Linear Accelerator Center presents a unique opportunity to search for rare tau decays such as  $\tau \rightarrow \ell\ell\ell$  and to further test the assumption of lepton flavor conservation. The PEP-II storage rings produce  $e^+e^-$  collisions at a center-of-mass (CM) energy of 10.58 GeV. The *BABAR* detector, constructed around the  $e^+e^-$  collision point, is a multipurpose detector made up of a number of sub-detector systems which are optimized to detect and record the many different decay products of the  $e^+e^-$  collisions. While the primary physics program is based around the decays of B mesons, the accelerator also generates a high rate of  $\tau^+\tau^-$  pairs through the reaction  $e^+e^- \rightarrow \tau^+\tau^-$ . The detector is capable of efficiently identifying and recording the decays of these leptons. Furthermore, the detector's excellent energy and momentum resolution gives the *BABAR* physicists a precise knowledge of the missing energy associated with unobservable neutrinos.

This work is divided into chapters as follows:

- Chapter I begins with a discussion of the role of symmetry in fundamental physics and a presentation of the standard theory of subatomic electromagnetic and weak interactions. Particular attention is paid to the structure of the interaction between the different generations (or families) of leptons. An analogy with the quark interactions is presented and some possibilities for new couplings beyond the standard theory are briefly mentioned. Chapter I concludes with a review of the experimental history of searches for neutrinoless lepton decays.
- Chapter II concerns the *BABAR* experimental apparatus. After an overview of the accelerator and general detector functions, the detector sub-systems are reviewed, with emphasis placed on lepton detection capabilities. Next is a discussion of the computer simulations of the  $e^+e^-$  collisions and of the decay products' subsequent interactions with the detector. The chapter concludes with sections detailing the operation of the trigger and data processing procedure.

- Chapter III deals with the data analysis itself. After an general discussion of branching fractions and limits, the details of the upper limit setting procedure are presented. Next the selection of the data and optimization of the analysis are discussed. Background calculations and estimates of uncertainties complete the chapter.
- Chapter IV presents the final results and a discussion of their merit. The implications are discussed for a variety of models of physics beyond the standard theory.

## 2 Theoretical Motivation

Mathematical symmetries play a fundamental role in theoretical physics. As proven by Emmy Noether in 1917, the existence of a continuous symmetry in a theory implies the existence of a conserved quantity. In other words, if the equations of motion are invariant under some operation, the theory will have a divergence-less current. Therefore, the existence of symmetries in the theories of particle physics is of great interest because they lead to conservation laws, which in turn are powerful tools for predicting and testing. A simple example is the invariance of physical laws under translations in three-dimensional space. Noether's theorem relates this symmetry to the conservation of momentum. Similarly, the invariance of physics under rotations about a point is related to the conservation of angular momentum.

If the operation under which the theory is symmetric is generated by a group, then a conserved current is associated with each generator of the symmetry. This perspective on symmetries is particularly useful for theories with internal symmetries. Such theories contain some number of fields whose interactions do not change under rotations in the space of those fields. An example of this sort of internal symmetry can be seen in the strong interactions of the proton and the neutron. Since these two baryons experience the strong force equally, then there is a symmetry in the abstract isospin space in which the particles are basis vectors. This symmetry leads to the conservation of the isospin quantum number in strong interactions.

## 2.1 Lepton Flavor

Lepton flavor is a quantum number associated with the particular generation or family in which the lepton resides. The charged electron and the neutral electron neutrino (along with their antiparticles) reside in the first lepton generation, while muon-like particles reside in the second generation, and tau-like particles in the third. If lepton flavor is conserved, then the decays of the charged leptons will always involve a neutrino of the same generation. Therefore, the neutrinoless decay of any lepton is a lepton flavor violating process.

In contrast to the symmetry examples mentioned earlier, the conservation of lepton flavor does not appear to be the result of any known symmetry. Due to the almost total absence of right-handed neutrinos, the standard theory described in the next section permits leptonic transitions only within the same generation. But this is the result of the smallness of the neutrino mass and is not due to any fundamental symmetry. Furthermore, the standard theory is known to be an incomplete low-energy effective theory based on a more general model. Without a symmetry to protect lepton flavor, we have no reason to expect its conservation in a more general theory.

## 2.2 The Standard Model of Electromagnetic and Weak Interactions

In the Standard Model (SM) of particle physics, the electromagnetic and weak interactions [1, 2] are described by a quantized field theory which is constructed to be invariant under rotations by elements of the symmetry group  $SU(2) \times U(1)$ .

Left-handed leptons are assigned to  $SU(2)$  doublets in the following way:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L. \quad (\text{I.1})$$

Right-handed charged leptons are observed to not participate in the weak interactions, and are therefore assigned to  $SU(2)$  singlets. Right-handed neutrinos can not exist as massless particles, and are not included in the theory. The quark sector of the theory is similarly constructed, with the six left-handed quarks assigned to 3  $SU(2)$  doublets:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \quad \begin{pmatrix} c \\ s \end{pmatrix}_L \quad \begin{pmatrix} t \\ b \end{pmatrix}_L. \quad (\text{I.2})$$

In contrast to the lepton sector, both up-type and down-type quarks have right-handed and left-handed components. The right-handed components are assigned to SU(2) singlets.

The three generators for the SU(2) symmetry are

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (\text{I.3})$$

while the single generator for the U(1) symmetry is

$$y = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (\text{I.4})$$

A local SU(2)×U(1) rotation on some doublet field  $\psi$  can be parameterized in terms of these generators and the local variables  $\alpha^a$  ( $a = 1, 2, 3$ ) and  $\beta$ ,

$$\psi \rightarrow e^{i\alpha^a \tau^a} e^{i\beta/2} \psi. \quad (\text{I.5})$$

For the theory to be invariant under such rotations, all kinetic terms in the Lagrangian must use the covariant derivative

$$D_\mu \psi = \partial_\mu \psi - ig A_\mu^a \tau^a \psi - i \frac{1}{2} g' B_\mu \psi, \quad (\text{I.6})$$

which couples the  $\psi$  field to the fields  $A_\mu$  and  $B$ . Fields such as  $A_\mu$  and  $B$ , which appear in the covariant derivative and are associated with a symmetry of a theory, are often referred to as gauge fields. In this case, the fields  $A_\mu^a$  and the coupling strength  $g$  are associated with SU(2) symmetry, and the field  $B$  and the coupling strength  $g'$  is associated with the U(1) symmetry.

Because mass terms of the form  $\frac{1}{2} m^2 \bar{\psi} \psi$  are not invariant under SU(2) rotations, all particle masses in the theory must arise through the Higgs mechanism. The complex scalar Higgs field is assigned to an SU(2) doublet. A potential of the form

$$V(\phi) = -\mu^2 \phi^* \phi + \frac{\lambda}{2} (\phi^* \phi)^2 \quad (\text{I.7})$$

results in spontaneous breaking of the SU(2) symmetry, and the Higgs field  $\phi$

acquires a vacuum expectation value (VEV) of the form

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \quad (\text{I.8})$$

After spontaneous symmetry breaking, a rotation with

$$\alpha^1 = \alpha^2 = 0 \quad \alpha^3 = \beta \quad (\text{I.9})$$

still leaves  $\langle \phi \rangle$  invariant. This residual  $U(1)_{EM}$  symmetry is associated with the conservation of electric charge and the corresponding vector boson, the photon ( $A^\mu \propto g' A_\mu^3 + g B_\mu$ ) remains massless. The remaining three gauge boson fields ( $W_\mu^\pm, Z$ ) have mass terms in the covariant derivative which couples the Higgs boson to the gauge bosons.

## SM Flavor Structure for Quarks

Flavor violation, or mixing between different generations of fermions, arises in the  $SU(2) \times U(1)$  theory as a consequence of the so-called Yukawa coupling of the Higgs field  $\phi$  to the left-handed and right-handed fermions. For one generation of quarks ( $u$  and  $d$  only), these couplings in the Lagrangian for the quarks are

$$\Delta \mathcal{L}_q = -\lambda_d \bar{Q}_L \phi d_R - \lambda_u \varepsilon^{ab} \bar{Q}_{La} \phi_b^\dagger u_R + \text{Hermitian conjugate (H.C.)}, \quad (\text{I.10})$$

where  $Q_L$  is the first left-handed quark doublets from Equation I.2,  $d_R$  and  $u_R$  are the associated right-handed quark fields, and  $\lambda$  is the coupling strength. When the vacuum expectation of  $\phi$  (Equation I.8) is inserted, Equation I.10 contains mass terms of the form

$$-\frac{1}{2} m_d \bar{d}_L d_R - \frac{1}{2} m_u \bar{u}_L u_R \quad (\text{I.11})$$

where

$$m_d = \frac{1}{\sqrt{2}} \lambda_d \nu, \quad m_u = \frac{1}{\sqrt{2}} \lambda_u \nu. \quad (\text{I.12})$$

Mass terms of this sort are allowed in the theory because they result from Higgs-quark couplings which are originally invariant under  $SU(2) \times U(1)$  rotations.

Extending the theory to 3 generations of quarks mixes the mass terms. With left-handed quark doublets  $Q_L^i$  and right handed quark singlets  $u_R^j$  and  $d_R^j$ ,

Equation I.10 now becomes

$$\Delta\mathcal{L}_q = -\lambda_d^{ij}\bar{Q}_L^i\phi d_R^j - \lambda_u^{ij}\varepsilon^{ab}\bar{Q}_{La}^i\phi_b^\dagger u_R^j + \text{H.C.}, \quad (\text{I.13})$$

where  $\lambda^{ij}$  is not necessarily diagonal. The matrix  $\lambda$  can always be diagonalized by setting

$$\lambda_u = U_u D_u W_u^\dagger \quad \lambda_d = U_d D_d W_d^\dagger \quad (\text{I.14})$$

and making the rotations

$$u_R^i \rightarrow W_u^{ij} u_R^j, \quad d_R^i \rightarrow W_d^{ij} d_R^j \quad (\text{I.15})$$

and

$$u_L^i \rightarrow U_u^{ij} u_L^j, \quad d_L^i \rightarrow U_d^{ij} d_L^j. \quad (\text{I.16})$$

Inserting the Higgs VEV from Equation I.8 into the rotated form of Equation I.13 leads again to mass terms of the form Equation I.11 for the six quarks.

While the diagonal matrices  $D$  contain the quark masses, and the  $W$  matrices disappear from the theory, the  $U$  matrices show up elsewhere. The rotations of Equation I.15-I.16 must be applied to all left-handed quark fields, and could, in principle, affect in the electromagnetic and weak couplings of the quarks. The unitary rotation matrices cancel in the neutral electroweak interactions. They also cancel from the kinetic terms. However, the effect of the rotations on the charged-current coupling of the  $W^\pm$  to the left-handed quarks takes the form

$$J^{\mu+} = \frac{1}{\sqrt{2}}\bar{u}_L^i\gamma^\mu d_L^i \rightarrow \frac{1}{\sqrt{2}}\bar{u}_L^i\gamma^\mu (U_u^\dagger U_d)^{ij} d_L^j. \quad (\text{I.17})$$

The matrix  $V = U_u^\dagger U_d$  is referred to as the CKM matrix, and describes the flavor structure of the charged-current coupling of left-handed quarks in the basis where the quark mass matrices in Equation I.13 are diagonal. The diagonal elements of  $V$  describe the rate of same-generation quark transitions relative to different-generation transitions, which are described by the off-diagonal elements.

In summary, the presence of mixing between the quark generations in the weak charged-current interaction arises from the fact that the quark mass eigenstates are not the same as the flavor eigenstates which undergo weak interactions. The form of

$V_{CKM}$ , which describes the relationship between these bases, is not predicted by the theory and must be measured experimentally.

## SM Flavor Structure for Leptons

As with the quarks, the left- and right-handed leptons in the  $SU(2) \times U(1)$  theory are coupled by the Higgs field. The most general Higgs coupling to the leptons takes the same form as Equation I.13:

$$\Delta\mathcal{L}_\ell = -\lambda_l^{ij} \bar{L}_L^i \phi l_R^j - \lambda_\nu^{ij} \varepsilon^{ab} \bar{L}_{La}^i \phi_b^\dagger \nu_R^j + \text{H.C.}, \quad (\text{I.18})$$

where the left-handed lepton doublets  $L$  have replaced the left-handed quark doublets  $Q$ . In the case that the neutrinos are absolutely massless, there are no right-handed neutrinos (or left-handed antineutrinos), and the second term in Equation I.18 is zero. The charged lepton mass matrix  $\lambda_l$  can be diagonalized as before,

$$\lambda_\ell = U_\ell D_\ell W_\ell^\dagger \quad (\text{I.19})$$

but the rotations

$$e_L^i \rightarrow U_l^{ij} d_L^j, \quad \nu_L^i \rightarrow U_\ell^{ij} \nu_L^j, \quad e_R^i \rightarrow W_l^{ij} e_R^j \quad (\text{I.20})$$

remove the  $U_\ell$  and  $W_\ell$  matrices from the theory *without* changing the charged current coupling. Thus, with massless neutrinos, the weak interactions only induce transitions within lepton generations and decays such as  $\tau \rightarrow \ell\ell\ell$  are not predicted.

## 2.3 Flavor Structures Beyond the Standard Model

While the existence of neutrino oscillations allows for a complete set of massive left- and right-handed Dirac neutrinos and anti-neutrinos, the sum of the three neutrino masses of this sort are required to be less than 2.0 eV [3]. This situation severely limits the possibility of lepton flavor violation through a CKM-like mechanism.

With this sort of LFV structure, the rate for two-body decays such as  $\tau^- \rightarrow \mu^- \gamma$  is proportional to the ratio  $(\Delta m^2/M_W^2)^2$ , where  $\Delta m$  is the mass difference between the neutrino mass eigenstates and  $M_W$  is the mass of the charged electroweak gauge

boson. This ratio leads to decay rates on the order of  $10^{-40}$ , which would be completely unobservable at any conceivable experiment in the near future.

Three body decays such as  $\tau \rightarrow \ell\ell\ell$  can be generated with light neutrinos at somewhat higher rates via loop diagrams, such as those suggested by Pham [4]. In certain diagrams, the cancellation produced by unitarity of the neutrino mixing matrix is particularly mild, leading to rates of order  $10^{-14}$  for a favorable choice of neutrino mixing angles. Nevertheless, such rates remain well beyond the reach of any foreseeable experiments.

When considering theories beyond the Standard Model with right-handed neutrinos, the possibilities for lepton flavor violation are more numerous. Since these theoretical models must simplify to the SM at the typical collider energy and below, the models usually contain many new particles and lepton couplings. In general, there is no reason to expect all new lepton couplings to be simultaneously diagonal. The new physics model could be as simple as the addition of a new neutral Higgs SU(2) doublet, leading to expanded Yukawa terms such as (tau couplings only)

$$\Delta\mathcal{L}_\ell \simeq \lambda_\tau \bar{\tau}_L \phi_1 \tau_R + \lambda_\tau \bar{\ell}_L^j \phi_2 \Delta_L^{3j} \tau_R + \lambda_\tau \bar{\tau}_L \phi_2 \Delta_R^{3j} \ell_R^j + \text{H.C.}, \quad (\text{I.21})$$

where the  $\Delta_{L,R}^{3j}$  parameters are the source for LFV. In fact, these models must be further constrained just to suppress lepton flavor violating interactions to rates below the current experimental bounds. More complicated theories such as supersymmetry provide numerous new sources of LFV, both directly at tree level or through loop effects where the unitarity suppression is relatively mild [5]. More specific models will be noted in the discussion of the results in Section 2.

### 3 Previous Experimental Work

#### 3.1 Search for Neutrinoless Muon Decays

Searches for non-conservation of lepton number and family have been performed at least as far back as 1948, when E.P. Hincks and B. Pontecorvo performed studies of the decays of what they thought was a meson with a lifetime of  $2.2 \mu\text{s}$  [6]. A search for the decays of this “mu meson” to an electron and a photon, each with energy around 50 MeV, produced a null result. Once the muon was firmly established as a second generation of lepton, other searches [7, 8, 9, 10, 11, 12, 13] for  $\mu^- \rightarrow e^- \gamma$  were made, but none with positive results. In 1959, Lee and Samios [14] used slow

muons in a bubble chamber to search for the decay  $\mu^+ \rightarrow e^+e^-e^+$ . No events were observed, and the experimenters estimated the ratio

$$\frac{\mu^+ \rightarrow e^+e^-e^+}{\mu^+ \rightarrow e^+\nu\bar{\nu}} < 10^{-4}. \quad (\text{I.22})$$

In 1963, a search for the process  $\mu^- \rightarrow e^-e^+e^-$  by the Babaev, Balats, Kaftanov, Landsberg, Lyubimov, and Yu [15] also produced a null result. Two years later, Parker, Anderson, and Rey [16] placed the first upper limit on the branching fraction for the decay  $\mu^- \rightarrow e^-\gamma$ . Based on some technical breakthroughs of Babaev *et al.*, their experiment was designed to observe the simultaneous back-to-back emission of an electron and a gamma ray. The large solid-angle coverage and good angular resolution of their detector allowed significant improvement in background rejection, and led to a 90% C.L. upper limit on  $\mathcal{B}(\mu^- \rightarrow e^-\gamma)$  of  $2.2 \times 10^{-8}$ .

These 1960's era limits stood until the late 1970's. In 1977, a search for  $\mu^- \rightarrow e^-\gamma$  at the Swiss Institute for Nuclear Research (SIN) generated rumors of a positive result. This brought about a resurgence of interest in lepton flavor and number conservation. Careful study of the SIN data led to upper limits on  $\mu^- \rightarrow e^-\gamma$  around  $1 \times 10^{-9}$  [17, 18]. In 1988, the SINDRUM collaboration at SIN placed what is still the best limit on  $\mathcal{B}(\mu^- \rightarrow e^-e^+e^-)$  [19]:

$$\frac{\Gamma(\mu \rightarrow 3e)}{\Gamma(\mu \rightarrow e2\nu)} < 1.0 \times 10^{-12} \text{ (90\% C.L.)}. \quad (\text{I.23})$$

The SINDRUM detector is designed to detect the decays of  $\mu^+$ 's which are stopped at a rate of about  $5 \times 10^6/\text{s}$ . Electrons and positrons are detected in a spectrometer of consisting of five concentric multiwire proportional chambers, and a cylindrical array of 64 scintillation counters in a 0.33 T magnetic field.

Experiments at the Los Alamos Meson Physics Facility (LAMPF) continue to make improvements to the limits on  $\mu^- \rightarrow e^-\gamma$ . In 1988, the physicists used data taken with the Crystal Box detector to set new limits on  $\mu^- \rightarrow e^-\gamma, \mu^- \rightarrow e^-\gamma\gamma$ , and  $\mu^- \rightarrow e^-e^+e^-$  [20]. The MEGA collaboration further lowered the  $\mu^- \rightarrow e^-\gamma$  limit in 1999. Using a similar experimental setup to that of SINDRUM, the detector is designed to observe the characteristic back-to-back positron/photon emission. By making high quality measurements of timing, energy, and position, MEGA set an

upper limit [21] of

$$\frac{\Gamma(\mu^+ \rightarrow e^+\gamma)}{\Gamma(\mu^+ \rightarrow e^+\nu\bar{\nu})} < 1.2 \times 10^{-11} \text{ (90\% C.L.)}. \quad (\text{I.24})$$

### 3.2 Search for Neutrinoless Tau Decays

In 1975, Martin Perl and his collaborators on the MARK-I experiment at SLAC's SPEAR  $e^+e^-$  rings made the somewhat unexpectedly discovery of a third generation of lepton, the tau [22]. One year earlier, in the so-called November Revolution, the charm quark had been observed in the form of the  $J/\psi$   $c\bar{c}$  bound state. Mesons consisting of charm and a lighter quark (D mesons) were soon after observed at slightly higher collision energies. The existence of the charm quark resolved some long-standing theoretical issues involving the observed absence of flavor-changing neutral currents [1]. In contrast, the lack of theoretical motivation for a third generation of leptons<sup>1</sup>, and the fact that the charmed D meson has a mass very close to that of the tau, led to a great deal of initial skepticism over the tau discovery.

A closer analysis of the MARK-I data showed that the increase in cross section around 3.6 GeV CM energy couldn't be easily explained by only D mesons. These charmed mesons were expected to decay primarily to strange particles, yet the experimentalists instead noticed an increase in events with missing momentum and many extra electrons and muons. Further more, there were no peaks in the  $K\pi$  and  $K\pi\pi$  mass spectra. A careful analysis of events with one or more leptons pointed to a new spin-1/2 particle decaying to leptons and invisible neutrinos. The DASP and PLUTO collaborations at the DORIS facility in DESY confirmed the discovery in 1977-78.

Just four years later, collaborators with the MARK-II detector at SPEAR used the world's first large sample of tau decays ( 20000 events) to make searches for radiative, neutrinoless tau decays to an electron or muon [23]. The group also searched for four different three-body, leptonic decays ( $e^-e^+e^-$ ,  $\mu^-e^+e^-$ ,  $e^-\mu^+\mu^-$ ,  $\mu^-\mu^+\mu^-$ ), and for six decays to a lepton and a neutral meson ( $\pi^0$ ,  $K^0$ ,  $\rho^0$ ). The general-purpose, hermetic design of the detector allowed for signal acceptance rates (signal efficiencies) up to 10%. Limits on the branching fractions were set in the

---

<sup>1</sup>Kobayashi and Maskawa noted in 1972 that a third generation of quark would allow for CP violation in the charged weak interactions.

range  $10^{-3} - 10^{-4}$  at the 90% confidence level. All limits cited in this section were calculated at the 90% confidence level.

The next significant search for neutrinoless tau decays was performed with the ARGUS detector and the DORIS accelerator at DESY, and published in 1987 [24]. Like *BABAR*, the ARGUS experiment was conducted at 10.58 GeV CM energy. This energy range is far enough above the tau production threshold to provide good separation of tau and multihadron events, but not so high that the  $1/s$  dependence of the cross section starts to negatively affect the tau production rate. ARGUS achieved exceptionally good signal acceptance rate of around 25% for all search channels. The ARGUS analysts make searches for all six  $\tau \rightarrow \ell\ell\ell$  modes and placed limits of the order  $10^{-5}$ , an order-of-magnitude better than MARK II. They also placed similar limits on the decays  $\tau^- \rightarrow \ell^- h^+ h^-$ ,  $\tau^- \rightarrow \ell^+ h^- h^-$ , and  $\tau^- \rightarrow \ell^- \rho/K^*(892)$ , where  $\ell = e, \mu$  and  $h = \pi^\pm, K^-$ .

New upper limits on the branching fractions for neutrinoless tau decays continued to be placed as collaborations generated tau samples of sufficient size. In 1988, collaborators with the Crystal Ball detector at DORIS placed limits just above  $1 \times 10^{-4}$  on the branching fractions for  $\tau \rightarrow e\gamma$ ,  $\tau \rightarrow e\pi^0$ , and  $\tau \rightarrow e\eta$  [25]. The CLEO collaboration published the first of many limits on neutrinoless tau decays in 1990 [26]. Their search consisted of the  $\tau \rightarrow \ell\ell\ell$  modes and  $\tau \rightarrow \ell hh$ , with zero or one kaons in the final state. While the CLEO limits were still of the order  $10^{-5}$ , they were generally lower than the ARGUS limits. ARGUS responded in 1992 with a large new set of upper limits, including the first searches for non-conservation of baryon number in tau decays [27]. Searches for  $\tau \rightarrow p\gamma$ ,  $\tau \rightarrow p\pi^0$ , and  $\tau \rightarrow p\eta$  yielded upper limits in the range  $(3 - 13) \times 10^{-4}$ . This study also placed new upper limits on  $\tau^- \rightarrow e^- \gamma$  and  $\tau^- \rightarrow \mu^- \gamma$  which were a factor of 2 and 10, respectively, than the previous limits.

CLEO made two other searches for  $\tau \rightarrow \ell\ell\ell$  and  $\tau \rightarrow \ell hh$  in the 1990's. The 1994 results [28] included the first search for  $\tau \rightarrow \ell \bar{K}^*$ , while the 1998 report [29] included the first searches for  $\tau \rightarrow \ell hh$  channels with resonant and non-resonant  $K^+ K^-$  pairs. In the later analysis, CLEO placed limits in the range  $(1 - 8) \times 10^{-6}$ . The DELPHI group at the LEP experiment at CERN published searches for  $\tau^- \rightarrow e^- \gamma$  and  $\tau^- \rightarrow \mu^- \gamma$  in 1995, based on a sample of 81k tau pairs [30]. While their upper limit on  $\tau^- \rightarrow e^- \gamma$  of  $1.1 \times 10^{-4}$  was slightly better than from ARGUS a few years earlier, their limit on  $\tau^- \rightarrow \mu^- \gamma$  was not competitive with CLEO limit of  $4.2 \times 10^{-6}$  from 1993 [31]. Other results from CLEO in the 1990's included new

limits on  $\tau^- \rightarrow e^- \gamma$  and  $\tau^- \rightarrow \mu^- \gamma$  [32], a search for tau decays to a lepton and various combinations of  $\pi^0$ 's and  $\eta$ 's [33], and another search for baryon-number non-conservation  $\tau \rightarrow \bar{p} \gamma / \pi^0 / \eta / 2\pi^0 / \pi^0 \eta$  [34]. Nearly all of these searches resulted in new upper limits.

While the CLEO tau sample led to a steady decrease in the upper limits for many neutrinoless tau decays, two newer experiments would soon acquire tau samples which would dwarf that of CLEO. The B-factories, *BABAR* at SLAC and Belle at the KEK accelerator facility in Japan, began running in 1999. In 2002, CLEO published the results of a search for tau decays to a lepton and one or two  $K_S$  mesons, based on  $13.9 \text{ fb}^{-1}$  of integrated luminosity [35]. For the first time, limits less than  $10^{-6}$  were placed on branching fractions for LFV decays. However, in their first three years of running, both B-factories recorded approximately  $100 \text{ fb}^{-1}$  of integrated luminosity. During this time, both collaborations presented a number of initial results based on a few tens of million tau pairs, though their limits were not yet competitive with those from CLEO.

In early 2004, the B-factories began publishing results based on data from nearly  $100 \text{ fb}^{-1}$  of integrated luminosity. *BABAR* published first, producing limits on  $\tau \rightarrow \ell \ell \ell$  in the range  $(1 - 3) \times 10^{-7}$  [36]. Soon after, Belle published limits in the range  $(2 - 4) \times 10^{-7}$  for the same channels [37]. Later that year, Belle also published new limits for  $\tau \rightarrow \mu \eta$  which were more restrictive than CLEO's previous limits by a factor of nearly 50 [38]. The B-factories limits on canonical LFV channels like  $\tau \rightarrow \ell \ell \ell$  were primarily due to the high luminosity of the machines. Total signal acceptance rates were generally similar to those for previous experiments. If anything, the high backgrounds at *BABAR* and Belle required tighter cuts (particularly in particle identification criteria) and consequently lower signal efficiencies. But in the end, the B-factories' ability to deliver consistently high luminosity allowed the experiment to set increasingly stringent limits on the neutrinoless tau decay branching fractions.

By 2005, *BABAR* and Belle began to publish results based on more than  $100 \text{ fb}^{-1}$ . Belle published a search for  $\tau \rightarrow \ell \pi^0 / \eta / \eta'$  based on  $154 \text{ fb}^{-1}$  of data. With signal detection efficiencies in the range 5 – 9%, Belle placed limits on the branching fractions in the range  $(1.3 - 10) \times 10^{-7}$  [39]. Soon after, *BABAR* published the first of such limits below  $10^{-7}$ . Based on  $221 \text{ fb}^{-1}$  of data, *BABAR* placed a new upper limit on  $\tau^- \rightarrow \mu^- \gamma$  of  $6.8 \times 10^{-8}$  [40]. *BABAR* also placed the first limits from the

B-factories on  $\tau \rightarrow \ell hh$  [41]. Belle responded with new limits on  $\tau^- \rightarrow \Lambda/\bar{\Lambda}\pi^-$  [42] and  $\tau \rightarrow \ell K_S$  [43], the later of which included limits as low as  $4.9 \times 10^{-8}$ .

Recent publications on neutrinoless tau decays include a new limit on  $\tau^- \rightarrow e^- \gamma$  from *BABAR* [44], Belle's first results for  $\tau \rightarrow \ell hh$  and  $\tau \rightarrow \ell \rho^0/K^*/\bar{K}^*/\phi$  [45], and updates on  $\tau \rightarrow \ell \pi^0/\eta/\eta'$  from *BABAR* [46] and Belle [47]. Many of these results place limits on the branching fractions of the order  $10^{-8}$ . The analysis detailed in this paper was published in late 2007 [48]. Using  $376 \text{ fb}^{-1}$  of data, we placed upper limits on  $\tau \rightarrow \ell \ell \ell$  from  $(4 - 8) \times 10^{-8}$  with signal efficiencies in the range  $5.5 - 12\%$ . While the April 2008 shutdown of *BABAR* brings to an end the data taking period for the experiment, final limits based on the full tau data sample are expected in the year following the shutdown. As the size of the data sample at Belle increases, new lower limits will doubtless be set. Plans for a high-luminosity Super-B factory present the possibility of setting limits of the order  $10^{-9} - 10^{-10}$ . This, of course, is under the assumption that no signal is found.

## CHAPTER II

### THE BABAR EXPERIMENT

#### 1 Introduction

Most modern accelerator-based particle physics experiments are conducted by large collaborations of scientists and engineers. The necessary experimental facilities include acceleration and beam guidance devices, which create large numbers of particles in the desired initial state. Detection facilities are also needed to observe and record final states from the reactions. The design and operation of these large facilities requires the dedication of hundreds of highly trained contributors. Most recent large detectors have been constructed as general purpose machines, allowing for the possibility of making many different measurements with the same data. As these experiments typically produce large quantities of data, many scientists are required to extract these measurements from the data.

The B-factory concept was proposed in 1987 to study the decays of B-mesons. In this concept, electrons and positrons are collided at a CM energy of 10.58 GeV, right at the peak of the Upsilon 4(S) resonance in the  $e^+e^-$  total cross section. The Upsilon 4(S) has a mass just slightly greater than twice the mass of the B meson, and it decays almost exclusively to pairs of B mesons. By using asymmetric beam energies to create the 10.58 GeV CM energy system, the B mesons are produced with a boost in the laboratory reference frame. This results in measurable lifetimes and flightlengths of the B mesons. By observing differences in decay properties of B and anti-B ( $\bar{B}$ ) mesons, one can make careful studies of CP violation.

The relatively small branching fraction for B mesons to CP eigenstates necessitates a machine which can produce large numbers of B mesons. The rate of production for a final state  $F$  is  $R_F = \mathcal{L}\sigma_F$ , where  $\mathcal{L}$  is the instantaneous luminosity and  $\sigma_F$  is the cross section for  $e^+e^- \rightarrow F$ . For the collision of two bunches at

frequency  $f$  with  $n_1$  particles in the first bunch and  $n_2$  particles in the second, the instantaneous luminosity is written as

$$\mathcal{L} = f \frac{n_1 n_2}{2\pi A}. \quad (\text{II.1})$$

The cross section  $A$  is the product of the transverse bunch widths in  $x$  and  $y$ , under the assumption that the bunch densities can be described by Gaussian distributions. For a given cross section (see Table II.1), the production rate can be increased by using bunches with more particles, by increasing the bunch-crossing frequency, or by decreasing the bunch cross section. The instantaneous luminosity is a flux, and has units of  $[1/(\text{time} \times \text{area})]$ . The integration of the instantaneous luminosity over the running time gives a measure of the accumulated data. Unless otherwise noted, all further references to the luminosity will refer to the time-integrated luminosity.

As seen in Equation II.1, the instantaneous luminosity is a general property of the colliding beam system and is not dependent on the final state  $F$ . Thus, the high instantaneous luminosity required for B meson studies provides a high rate for other final states as well. In fact, the cross sections at 10.58 GeV for  $e^+e^- \rightarrow u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}, \tau^+\tau^-$  are all similar to that for  $B\bar{B}$ , effectively making the B-factory a tau and charm factory as well (see Table II.1). In fact, Babar has recorded significantly more tau decays than any previous experiment. This large sample of tau decays leads to better precision on SM measurements and opportunities to place more stringent limits on unobserved processes, including lepton-flavor violating decays.

In Section 2, we examine the production and collision of  $e^+e^-$  pairs at 10.58 GeV CM energy, and in Section 3 we consider the detection of the decay products of those collisions. In Section 6, we will focus on the computing and data processing components of the experiment. Much of the discussion in this chapter is based on material from reference [49]. All figures are taken from this reference.

## 2 Particle Acceleration for the *BABAR* Experiment

The *BABAR* experiment makes use of SLAC's 2 mile linear accelerator (linac) facility to produce beams of 9 GeV electrons and 3.1 GeV positrons. The beams of particles are then injected into the 800 m diameter PEP-II storage rings. At the IR-2 interaction region, the beams are brought into collision. The *BABAR* detector,

$e^+e^- \rightarrow$	cross-section/nb
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$uds$	2.09
$\tau^+\tau^-$	0.89
$\mu^+\mu^-$	1.16
$e^+e^-$	$\simeq 40$

Table II.1: Approximate production cross sections at *BABAR*, including experimental acceptance factors.  $uds$  refers to the total continuum production to  $u\bar{u}, d\bar{d}, s\bar{s}$ .

constructed around this interaction point (IP) detects the long-lived particles coming from the final state of the  $e^+e^-$  interaction.

## 2.1 Beam Production

The Stanford Linear Accelerator has been used to accelerate particles for collisions since its construction was completed in 1966. In its current role, linac is used to generate 9.0 GeV electrons and 3.1 GeV positrons to be collided inside the *BABAR* detector. Electrons are produced with a polarized electron gun at the far end of the linac. The electrons are collected into bunches of about 500 billion particles apiece, and magnetically steered through damping rings to optimize the shape of the bunches. Oscillating electric and magnetic fields then accelerate the bunches down the 2-mile-long linac. Before being injected into the PEP-II rings, some electrons are diverted for positron production. A fixed tungsten target is bombarded with these electrons, producing  $e^+e^-$  pairs. The resulting positrons are returned to the far end of the linac, collected into bunches of similar size and shape to those of the electrons, and accelerated back down the linac, out of phase with the electrons.

## 2.2 Beam Storage

At the near end of the linac, bunches of electrons at 9 GeV and bunches of positrons at 3.1 GeV are injected into the PEP-II storage rings. A tunnel contains two beampipes for the counter-rotating beams, along with steering magnets and acceleration stations. The tunnel circles the SLAC research yard at a radius of 400 meters. The high energy ring (HER) contains the electron beam rotating clockwise,

while in the low energy ring (LER) the positrons flow counterclockwise. Particles in both rings are kept in orbit by a combination of magnets and radio frequency (RF) acceleration.

Bunches in the rings have a longitudinal length (along the direction of travel) of about 1 cm. For a given bunch spacing in the ring, only a certain number of bunches can be circulating at any time. Furthermore, the quality of the beams in the rings deteriorates over time due to a number of factors including random  $e^+e^-$  collisions and beam-gas interactions. With no further injection of bunches, this situation leads to an effective beam lifetime of 2-4 hours. The luminosity also decreases as the beam quality deteriorates. For the first few years of the *BABAR* experiment, the solution was to dump the beam and refill the rings with fresh bunches from the linac. Unfortunately, data could not be taken during the refill, which often took 40 minutes. The current solution, known as *trickle injection*, is to continuously inject small numbers of bunches into the ring. Under trickle injection, the detector records data almost continuously, with only a brief inhibit window where the detector ignores data around a recently refilled bunch. Trickle injection was implemented for the LER  $e^+$  beam in November 2003 and for the HER  $e^-$  beam in March 2004.

### 2.3 Beam Energy

About 10% of the data recorded at *BABAR* is taken with the  $e^+e^-$  CM energy lowered by about 40 MeV to 10.54 GeV. At this *off-peak* energy, the  $e^+e^-$  cross section is sufficiently far below the  $\Upsilon(4S)$  resonance that the production is effectively free of B-mesons. The cross section for  $c\bar{c}$ ,  $uds$ , and  $\tau^+\tau^-$  production is nearly flat through the entire energy region from the  $\Upsilon(4S)$  peak though the off-peak energy range. The data recorded at the off-peak energy allow physicists studying B meson decays to understand the contribution of the continuum production to the total cross section at the  $\Upsilon(4S)$  resonance. For physicists studying  $\tau$  decays, data recorded on- and off-resonance are equally useful, since the  $\tau^+\tau^-$  cross section is essentially the same at both energies.

### 2.4 Interaction Region

The HER and LER beams are brought into collision at the interaction region (IR) inside the *BABAR* detector. The incoming beams are focused and brought into

collision by a combination a dipole and quadrupole magnets. After colliding head-on, the bunches are quickly separated so as not to disrupt the next incoming bunch from the opposite beam. Figure II.1 shows the layout of the beams and the PEP-II magnets around the interaction region. The beampipe around the IP is 27.8 mm in radius, and constructed of double-walled beryllium. Beryllium is one of the lightest elements, giving it a small radiation length, but it is also very stiff. Water circulated in the 1.5 mm gap between the walls of the beampipe provides cooling. The inner surface of the beampipe is coated with a 4  $\mu\text{m}$  layer of gold, which reduces synchrotron radiation at the IP. A support tube encloses the beampipe, the innermost detector component, and the innermost magnets. The total material corresponds 0.019 radiation lengths, with the beryllium, the gold, and the support tube contributing approximately equal amounts.

## 2.5 Performance

The PEP-II B-factory has capably delivered luminosity to the *BABAR* detector for the duration of the experiment. The record-high instantaneous luminosity of  $1.21 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-2}$  was reached on August 16, 2007. Integrated luminosity records are shown in Table II.2 for individual 8-hour shifts, days, and months. The total integrated luminosity is shown in Figure II.2.

## 3 The *BABAR* Detector

The *BABAR* detector is a general purpose detector which must provide tracking capabilities for electrons, muons, protons, and charged kaons and pions. The detector must also provide good angular and energy resolution for electrons and

Table II.2: Integrated luminosity records for various time periods, in inverse picobarns (pb) and inverse femtobarns (fb), where 1 barn =  $10^{-28} \text{ m}^2$ .

Time Period	Integrated Luminosity
8 hours	329.7 pb <sup>-1</sup>
24 hours	891.2 pb <sup>-1</sup>
7 days	5.25 fb <sup>-1</sup>
30 days	18.84 fb <sup>-1</sup>

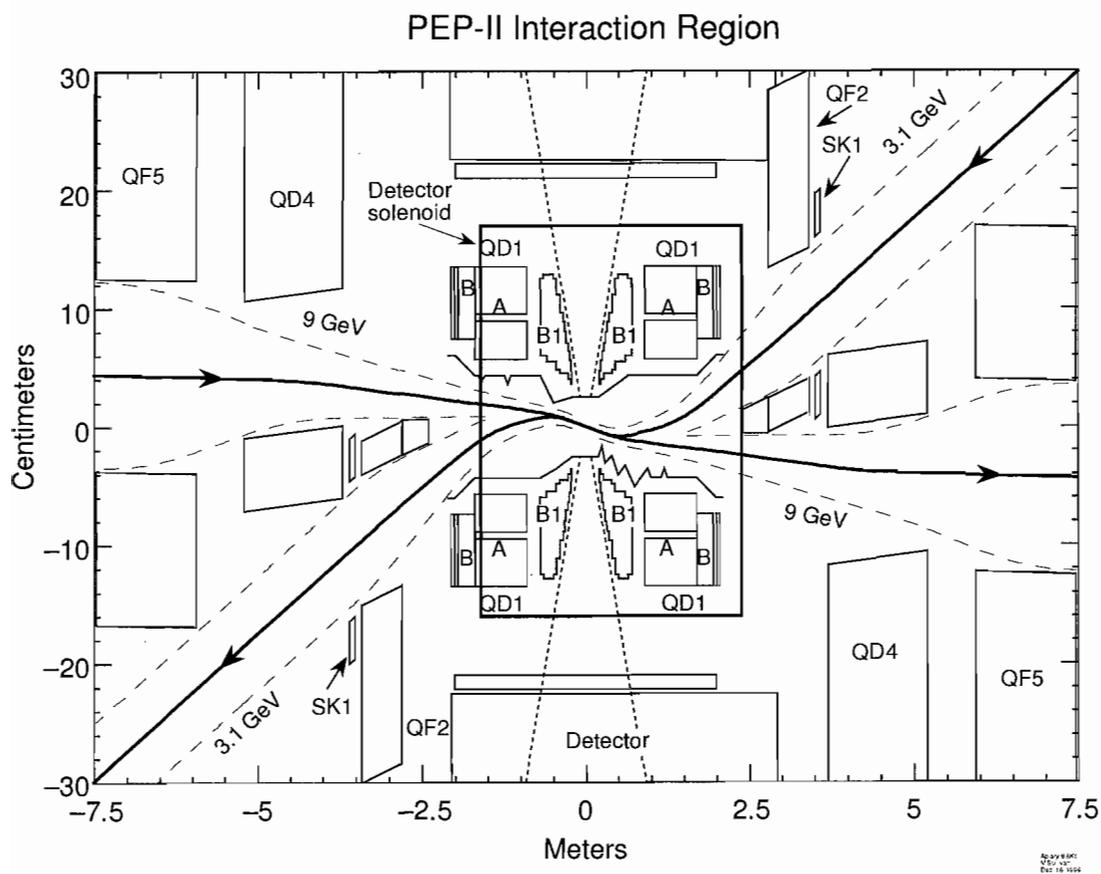


Figure II.1: The PEP-II interaction region around the *BABAR* detector.

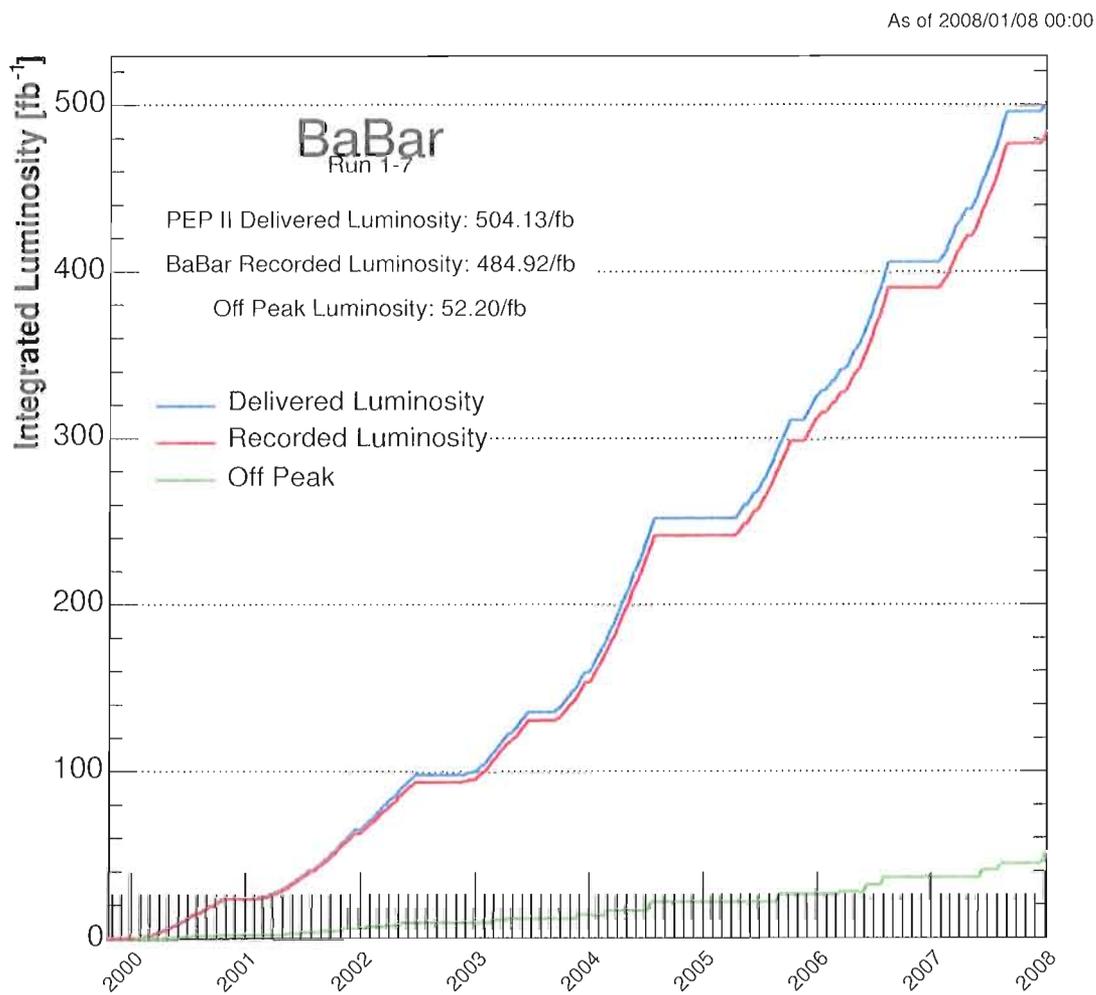


Figure II.2: Integrated luminosity at *BABAR* as a function of time.

photons. Identification of long-lived particles is important, particularly differentiation between charged pions and kaons. Finally, the detector must be able to reconstruct decay vertices, especially those of short-lived B mesons.

### 3.1 Detector Goals and Constraints

The design of the *BABAR* detector is driven by physical constraints from the interaction region (IR) layout and by performance goals for specific physics processes. The PEP-II focusing magnets nearest to the IP limit the total length along the beam axis. The detector is offset by 37 cm in the direction of boost. This offset increases the acceptance of the detector components in the CM frame. In order to reduce perturbation of the beams by the tracking system solenoid, the detector axis is offset by 20 mrad with respect to the beam axis in the horizontal plane.

The high luminosity of the  $e^+e^-$  interactions creates an environment with high levels of machine background signals unrelated to primary  $e^+e^-$  collisions. The dipole and quadrupole magnets which steer the beams into collision produce large amounts of synchrotron radiation. Though these high energy photons are generally diverted away from the detector, they still provide the primary source of machine background. Other sources of background include beam-gas interaction due to imperfect vacuum conditions in the beampipe, and the interaction with the machine of low energy particles from radiative Bhabha scattering. The detector components have been designed to withstand background rates at nearly 10 times the design luminosity for the duration of the expected 10-year lifetime.

The general physics program of the *BABAR* Collaboration sets some basic performance goals for the detector:

- large and uniform acceptance, down to small polar angles,
- high efficiency for charged track reconstruction,
- good track momentum resolution,
- good energy and angular resolution in calorimeter, and
- efficient particle identification and low mis-ID rates.

The B-physics program brings about further constraints. Branching fractions to CP eigenstates are of the order  $10^{-4}$ , and the final states typically include two or more

charged particles and several  $\pi^0$ 's. Therefore, further goals from precision B-physics measurements include:

- good resolution on the displaced B decay vertex, and
- significant  $\pi$ /kaon separation.

While the  $\tau \rightarrow \ell\ell\ell$  search does not require the full set of *BABAR* detector capabilities, the significance of the resulting measurement is still constrained by a number of performance issues. The most important factors are:

- tracking efficiency and resolution, and
- electron and muon identification and hadron rejection.

The innermost component of the *BABAR* detector is the Silicon Vertex Tracker (SVT). Moving radially outward, the cylindrical Drift Chamber (DCH) surrounds the SVT and completes the inner tracking system. The Detector of Internally Reflected Cherenkov light (DIRC) is located outside the tracker and provides particle identification information. The Electromagnetic Calorimeter (EMC) lies outside the DIRC and just inside the superconducting magnet. The outermost system is the Instrumented Flux Return (IFR), which completes the magnetic circuit and provides muon detection. Figure II.3 shows a  $y - z$  cross-section of the *BABAR* detector, and Figure II.4 shows a  $x - y$  view. The *BABAR* coordinate system follows the diagram in the upper left-hand corner of II.3, and is defined as follows:

- the origin is located at the center of the detector (not the IP).
- the  $y$  axis points radially upward.
- the  $x$  axis points radially outward in a plane parallel to the ground.
- the  $z$  axis point in the direction of the CM boost (in the direction of the electron beam).
- the angle  $\theta$  is the polar angle, measured from the positive  $z$  axis toward the positive  $y$  axis.
- the angle  $\phi$  is the azimuthal angle, measured from the positive  $y$  axis toward the positive  $x$  axis.

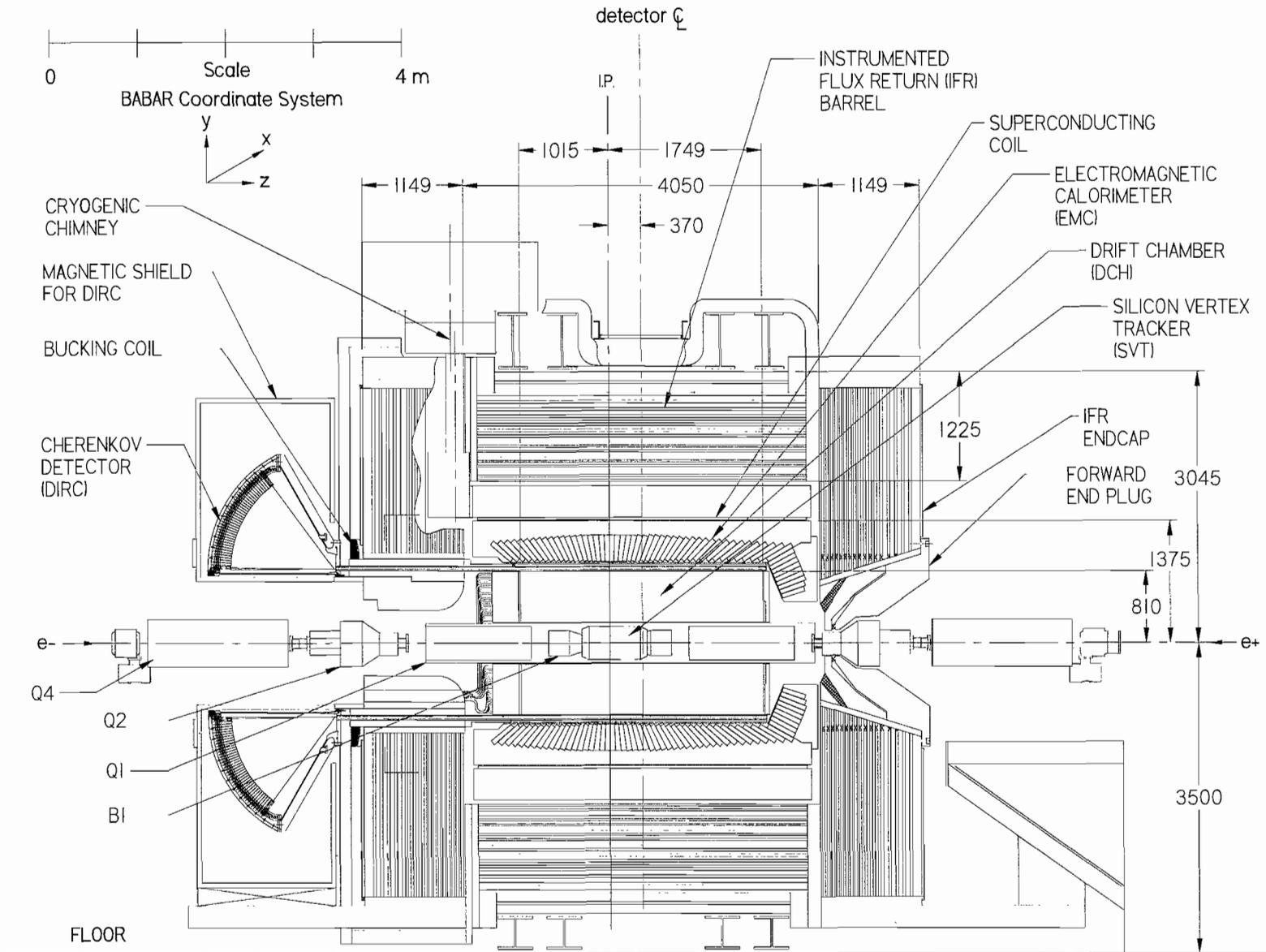


Figure II.3: Sideview of the *BABAR* detector.

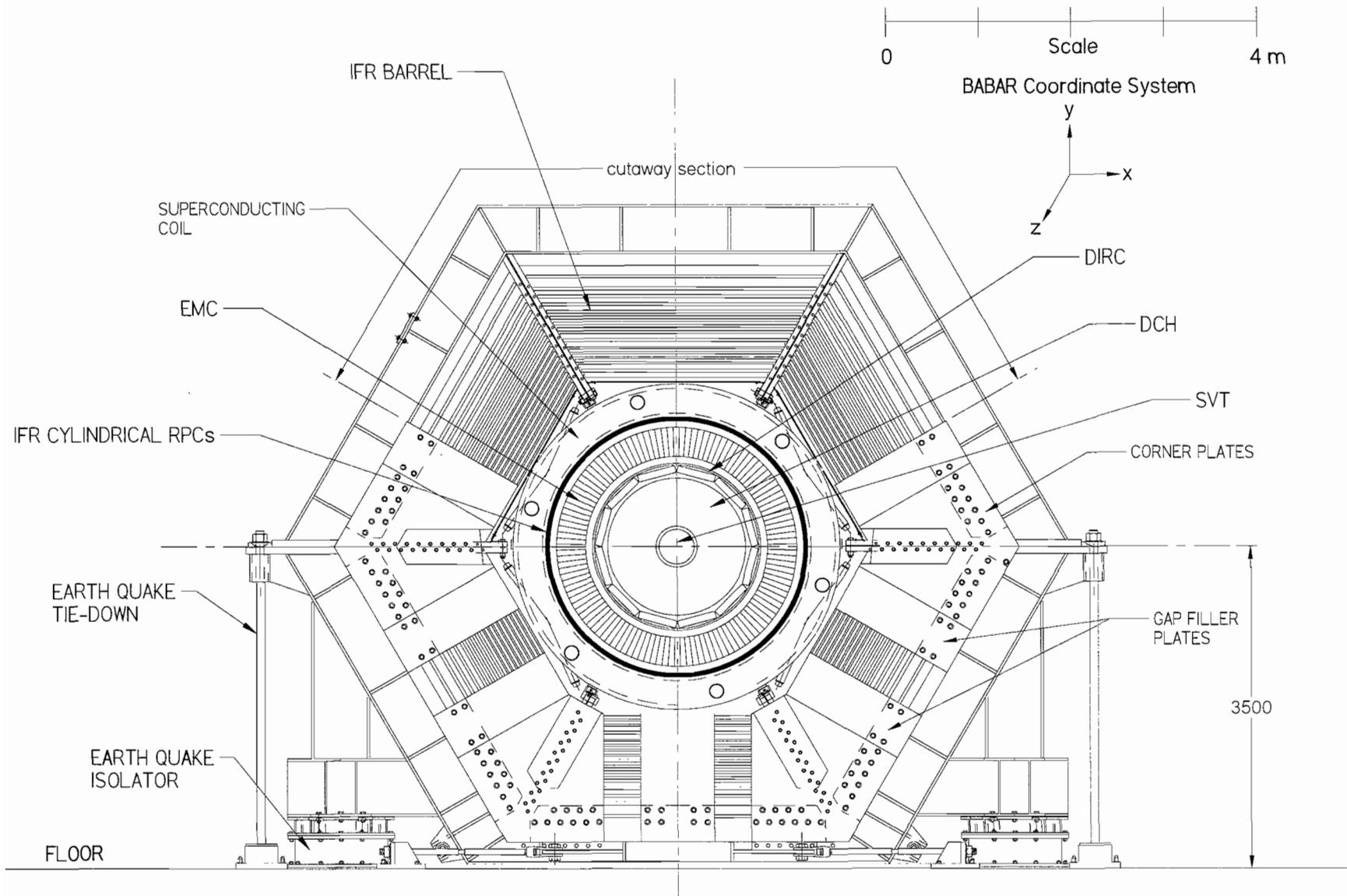


Figure II.4: Endview of the *BABAR* detector.

## 3.2 Charged Particle Tracking

The inner components of the *BABAR* detector are surrounded by a super-conducting solenoid that produces a 1.5 Tesla magnetic field. Charged particles follow helical trajectories in this field, and the component of the particle momenta transverse to the field lines can be calculated from the curvature of the trajectories. These trajectories are reconstructed from the interactions of the particles with the SVT and the DCH.

The SVT records the trajectories of particles within approximately 10 cm of the interaction point. As the name implies, the SVT plays an important role in measuring the decay vertices of short lived particles. In fact, the design of the SVT was primarily driven by the need to accurately measure the lifetime and flightlength of B mesons, along with constraints from the PEP-II magnets and beampipe. The SVT also provides an initial measurement of the energy loss due to ionization ( $dE/dx$ ). Because of the magnetic field, charged particles with transverse momentum less than 100 MeV/c do not reach the DCH, and the SVT provides the only trajectory measurements for such particles. The vertexing capability of *BABAR* is relatively unimportant for tau studies such as the search for  $\tau \rightarrow \ell\ell\ell$ . However, initial trajectories from the SVT still play an important role in the tracking of charged particles.

The DCH, with its capability of measuring charged particle trajectories throughout most of its 800 mm radius tracking volume, plays a primary role in the tracking of charged particles. Measurements of track curvature provide momentum and  $dE/dx$  information and the DCH is capable of making these measurement for particles with momentum greater than 100 MeV/c and with  $0.1 \text{ GeV}/c < p_t < 5.0 \text{ GeV}/c$ . Because different particles such as electrons, pions, and muons have different energy loss characteristics,  $dE/dx$  information acts as a form of particle identification. For low momentum particles, this information is generally complimentary to other particle identification information derived from the DIRC. At extreme forward and backward angles, the DCH  $dE/dx$  measurement is the only source of particle identification. Finally, the data acquisition system uses signals from the DCH to create primitive trigger signals, as described in Section 5.

## Silicon Vertex Tracker

The SVT consists of five layers of double-sided silicon strips. The inner three layers are built from six modules in  $\phi$  and are flat along the direction of the beam. The inner layers are built as close as possible to the beam pipe to minimize the effect of multiple scattering on vertex measurements. The outer two layers are constructed as arches and contain 16 (layer 4) or 18 (layer 5) modules. These layers are close to the inner radius of the DCH, and allow for better linking of hits in the SVT to tracks in the DCH. A longitudinal schematic view of the SVT can be seen in Figure II.5. The strips on opposite sides of each layer are oriented orthogonally to each other, with the  $\phi$  strips running parallel to the beam and the  $z$  strips oriented transversely to the beam axis. To provide full azimuthal coverage and to aid in alignment, the inner layers are tilted by a small amount in  $\phi$ , and the outer layers are divided into two sub-layers. Figure II.6 shows a transverse schematic view of the SVT. Each layer is divided into half-modules, which are read out at each end of the detector by radiation-hard circuits. The total number of readout channels is approximately 150,000. Radiation is a major factor for any component so close to the beam pipe. The SVT is required to withstand radiation doses of 1 Rad/day for layers 1-3 and 0.1 Rad/day for the outer two layers.

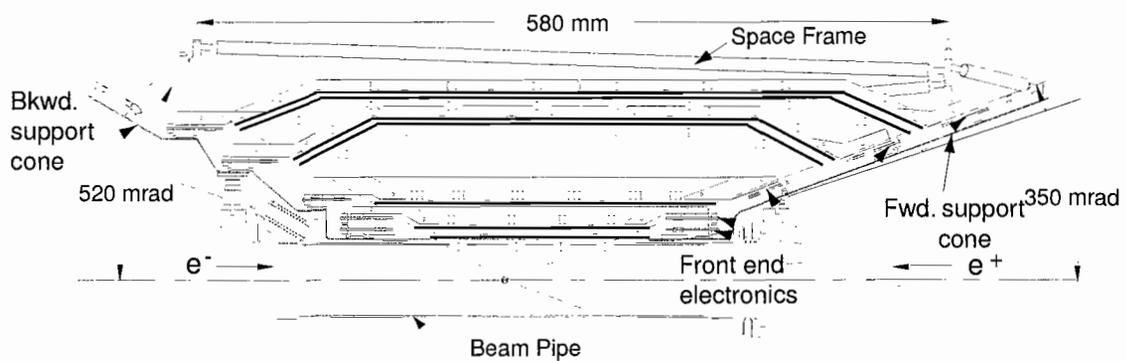


Figure II.5: Longitudinal schematic view of the SVT.

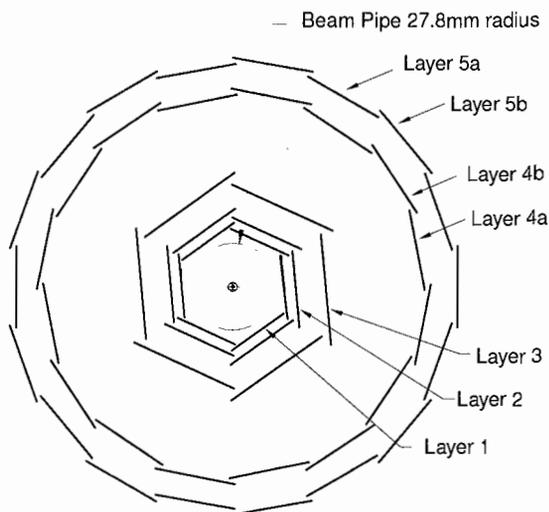


Figure II.6: Transverse schematic view of the SVT.

## Drift Chamber

The DCH measures 276 cm in length, with an inner radius of 23.6 cm and an outer radius of 80.9 cm. Figure II.7 shows a longitudinal view of the drift chamber. The chamber is filled with a 80:20 mixture of helium:isobutane gas at 4 mbar above atmospheric pressure, and consists of 7104 hexagonal drift cells arranged in 40 cylindrical layers. The layers are grouped by four into superlayers (see Figure II.8), with alternating superlayers offset by  $\pm 45$  to  $\pm 76$  mrad azimuthally to provide longitudinal position information. Each cell consists of one grounded tungsten-rhenium sense wire surrounded by six aluminum field wires held at approximately  $+1900$  V<sup>1</sup>. Charged particles passing through the chamber ionize the gas, and the ionization shower, guided by the field created by the field wires, drifts to the sense wire to be read out at the backward end-plate. Each signal on the sense wire gives a measurement of drift time, which is used to calculate track trajectories, as well as a measurement of integrated charge, from which energy loss can be calculated. The choice of low-mass wires and helium-based gas mixture leads to minimal electromagnetic scattering in the DCH, about 0.2% of the radiation length for the material.

<sup>1</sup>DCH voltage has been set to slightly different values over the course of the experiment. Early data were recorded at 1900 V and 1960 V. The majority of the data has been recorded at 1930 V.

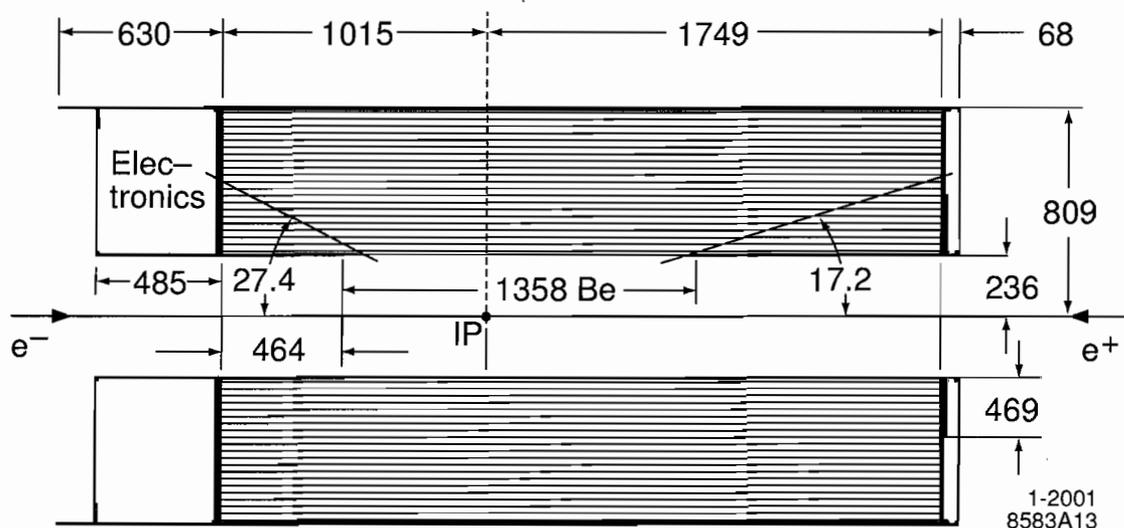
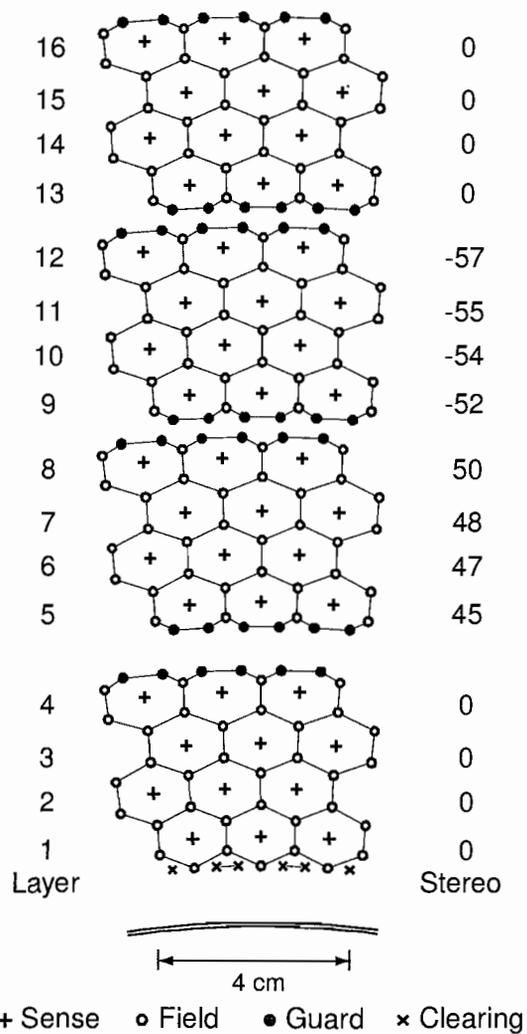


Figure II.7: Longitudinal view of the DCH, with dimensions in mm.

## The Magnet

The *BABAR* magnet system consists of a super-conducting solenoid, along with a flux return which is instrumented from muon detection, and a field-compensating bucking coil. The magnet also provides structural support for many of the detector components. Figures II.3 and II.4 show many of the key components of the magnet system, as well as some of the nearby PEP-II magnets. To optimize the detector acceptance for the asymmetric collision energy, the detector is offset by 370 mm in the direction of the electron beam. The  $z$ -component of the magnetic field lies along the  $z$ -axis of the detector coordinate system; this is also the approximate direction of the electron beam.

The solenoid contains 10.3 km of cable made of filaments of super-conducting niobium-titanium wound into wires and co-extruded with aluminum. The 4.5K operating temperature is maintained by the circulation of liquid helium through channels welded to the solenoid support cylinder. While the flux return provides important support for the detector components, its design was driven by the much larger magnetic forces and by earthquake considerations. Asymmetries and imperfections in the flux return steel result in large axial forces, and a quench of the magnet could generate sizable forces via eddy currents. Four earthquake isolators limit horizontal acceleration to 0.4 g, and the detector components have been



1-2001  
8583A14

Figure II.8: Schematic layout of the drift cells for the four innermost superlayers. The lines between the wires have been added to show the cell boundaries. The line below the first layer is the 1-mm-thick beryllium inner wall.

designed to tolerate vertical acceleration up to 0.6 g. The bucking coil is a water-cooled copper coil placed around the beampipe at the backward end of the detector. This coil reduces field leakage into the PEP-II components and shields the DIRC photomultipliers. The magnetic field has been carefully studied (see Figure II.9), and shown to be of uniformly high quality. Within the tracking volume, the azimuthal component  $B_\phi$  does not exceed 1 mT. The variation of the field transverse to the trajectory of a high momentum track is at most 2.5% from maximum to minimum within the tracking volume, as seen in Figure II.10

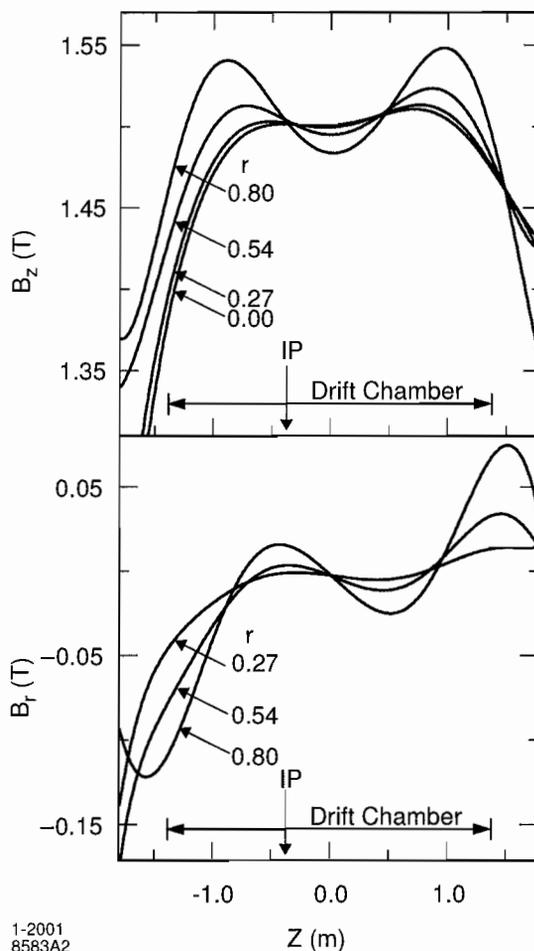


Figure II.9: The magnetic field components  $B_Z$  and  $B_r$  as a function of  $z$  for various radial distances  $r$  (in meters).

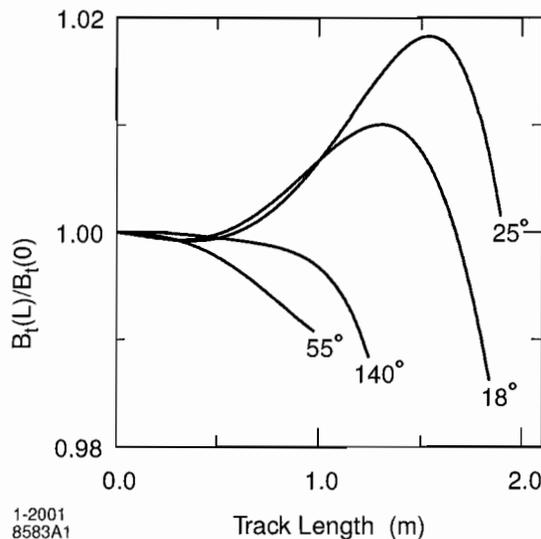


Figure II.10: Relative magnitude of the magnetic field transverse to a high momentum track as a function of track length from the IP for various polar angles (in degrees). The data are normalized to the field at the origin.

## Tracking Performance

The performance of the SVT under normal running conditions can be studied in terms of efficiency and spatial resolution. The efficiency can be calculated for each half-module by comparing the number of tracks crossing the active area of the detector to the number of associated hits read out. The efficiency for a sample of events recorded in July 2000 is plotted in Figure II.11. The combined hardware and software efficiency for these events is 97%, excluding the effect of defective readout sections.

The spatial resolution is calculated from hit information for events with two high-momentum tracks. The track momentum is compared to the hit location, and the difference is projected onto the wafer plane along either the  $z$  or  $\phi$  direction. The width of these distributions gives the resolution, which are shown in Figure II.12. Averaged over the whole SVT, the spatial resolution for normal tracks ranges from 10 – 15  $\mu\text{m}$  for the inner layers to 30 – 40  $\mu\text{m}$  for the outer layers.

The position resolution of the DCH can be measured by studying the precise relationship between measured drift time and drift distance in  $e^+e^-$  and  $\mu^+\mu^-$  events. This drift distance is computed from an estimate of the distance of closest

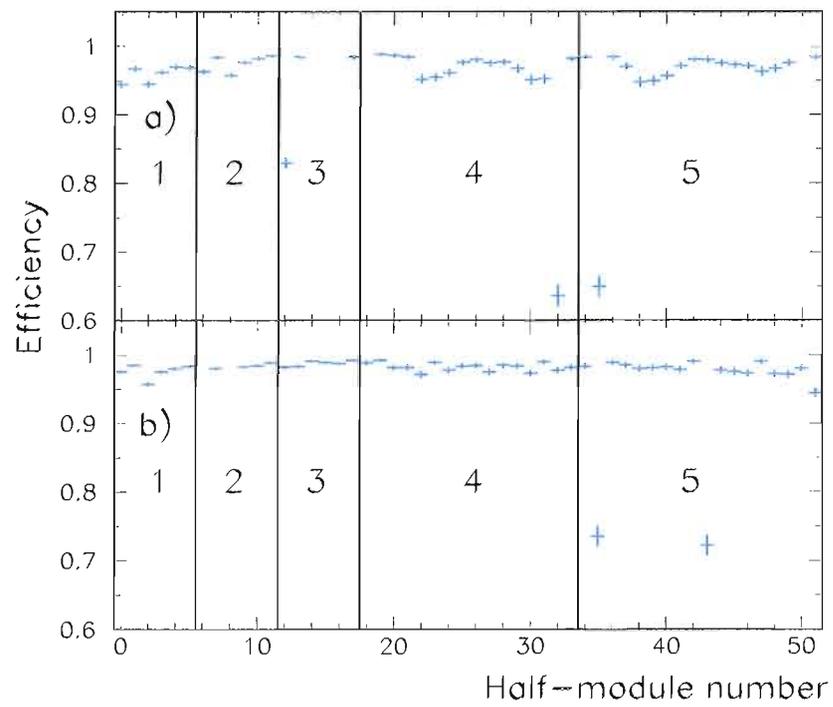


Figure II.11: SVT hit reconstruction efficiency, as measured on  $\mu^+\mu^-$  events for (a) forward half-modules and (b) backward half modules. Vertical lines delineate the five different layers.

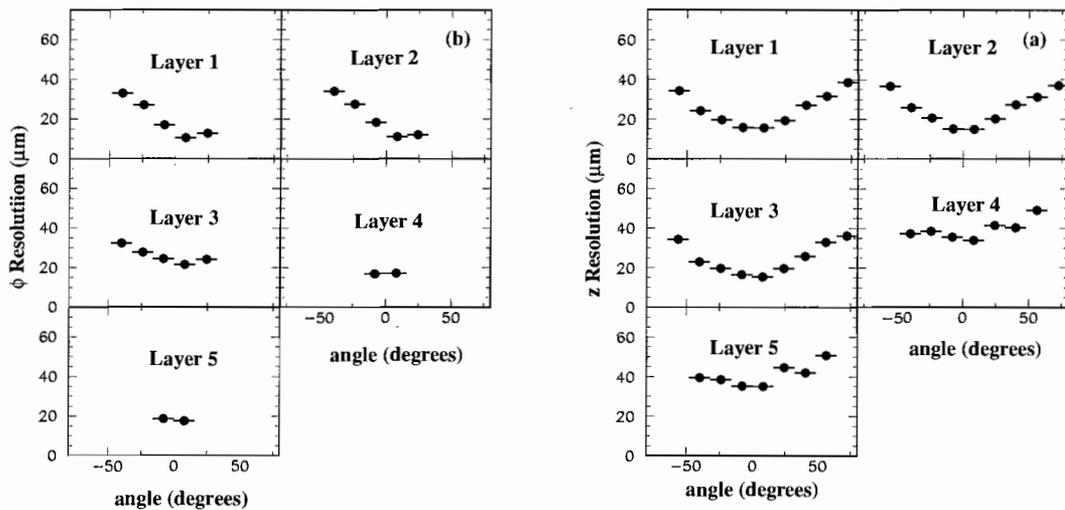


Figure II.12: SVT hit resolution in (left)  $z$  and (right)  $\phi$  coordinate in microns, plotted as a function of track incident angle in degrees.

approach between the track and the sense wire. The drift distances and drift times are averaged all the wires in the layer, but are separated by into two sets: those tracks passing to the left of the sense wire, and those tracks passing to the right. Figure II.13 shows the position resolution as a function of drift distance, for tracks on the left and right side of the sense wire.

The specific energy loss for charged particles traversing the DCH is computed by measuring the total charge deposited. The charge from each traversed cell is corrected for gain variations, pedestal-subtracted, and integrated over a time range of about  $1.8 \mu\text{s}$ . Further corrections are made on account of variations in gas pressure and temperature, cell geometry, signal saturation, and entrance angle to the cell. Measurements of  $dE/dx$  in the DCH are plotted as a function of momentum in Figure II.14. Resolution of just over 7% is achieved.

The total tracking efficiency is based on the combined performance of the SVT, the DCH, and the algorithms used in the software reconstruction of the tracks. While relatively simple track finding algorithms are used to quickly generate input signals for the trigger, the offline reconstruction of charged particle tracks (see Section 6.1) makes use of a variety of sophisticated search methods which refit each track multiple times, searching for stub tracks and missed hits to better the

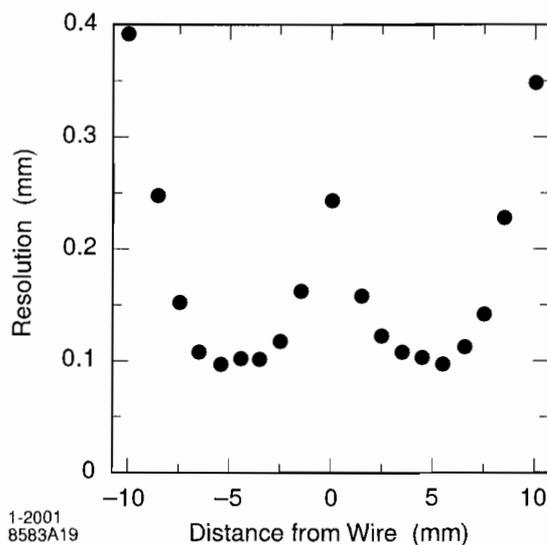


Figure II.13: DCH position resolution as a function of the drift distance in layer 18, for tracks on the left and right side of the sense wire. The data are averaged over all cells in the layer.

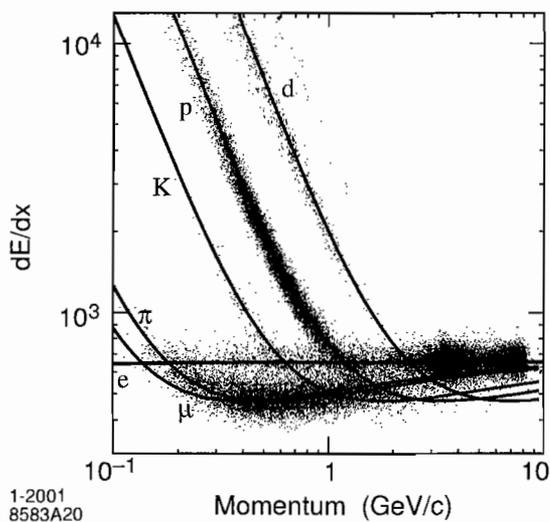


Figure II.14: Measurements of  $dE/dx$  in the DCH as a function of track momentum. The curves show the Bethe-Bloch predictions for each particle type.

resolution. The absolute tracking efficiency for the DCH can be measured by simply comparing the number of tracks detected in the SVT to the number of reconstructed tracks in the DCH. This efficiency varies based on the voltage of the sense wire, with a maximum efficiency of 89% for the initial voltage of 1960 V, and a slightly lower efficiency for the final voltage choice, 1930 V.

Fully reconstructed tracks are parameterized by five values (and the associated error matrix) which are measured at the point of closest approach to the  $z$ -axis. The distances from the origin of the coordinate system are  $d_o$  and  $z_0$ , in the  $x - y$  plane and along the  $z$ -axis, respectively. The angle  $\phi_0$  is the azimuthal angle, while  $\lambda$  is the dip angle relative to the transverse plane and  $\omega = 1/p_t$  is the curvature.

As measured for Bhabha ( $e^+e^-$ ) and di-muon events, the resolutions on the first four of these parameters are

$$\begin{aligned}\sigma_{d_o} &= 23 \mu\text{m} & \sigma_{\phi_0} &= 0.43 \text{ mrad} \\ \sigma_{z_0} &= 29 \mu\text{m} & \sigma_{\tan\lambda} &= 0.53 \cdot 10^{-3}.\end{aligned}$$

The most important resolution for the purpose of the  $\tau \rightarrow \ell\ell\ell$  analysis is that of the transverse momentum  $p_t$ . This resolution can be parameterized by the linear function

$$\sigma_{p_t}/p_t = (0.13 \pm 0.01)\% + (0.45 \pm 0.03)\%, \quad (\text{II.2})$$

where  $p_t$  is measured in GeV/c. This curve and the resolution data are shown in Figure II.15.

### 3.3 Pion and Kaon Identification

The DIRC is an innovative detector which provides particle identification via a measurement of the particle's velocity. Velocity measurements, when coupled with the momentum measured in the DCH, provide discrimination between particles of different mass, particularly charged hadrons such as pions and kaons. While  $\pi/K$  separation is extremely important for the flavor-tagging of B meson decays and the identification of rare two-body B decays, the  $\tau \rightarrow \ell\ell\ell$  searches are naturally less sensitive to the quality of charged hadron identification. Nevertheless, electron identification makes some use of the DIRC <sup>2</sup> and kaon rejection is highly dependent

---

<sup>2</sup>The algorithm used for electron identification in this analysis does not actually use information from the DIRC. However, an updated electron selection procedure based on likelihood ratios does

on this detector. The momentum range of the DIRC is set in part by the need for  $\pi/K$  separation in time-dependent asymmetry measurements, for which the typical hadron momentum is below 1 GeV. For rare two-body B meson decays, the hadron momenta lie between 1.7 and 4.2 GeV. The DIRC is designed to provide  $4\sigma$   $\pi/K$  separation over the full momentum range.

Figure II.16 shows a sideview of the major DIRC components. The detector contains a layer of rectangular silica (quartz) bars oriented parallel to the beams with an inner radius of 810 mm. The 144 bars are arranged in a 12-sided polygonal barrel. Each bar is 4.9 m long and constructed from four 1.225 m pieces glued end-to-end. The bars are 17.25 mm thick and 35 mm wide.

As charged particles with velocity exceeding the Cherenkov threshold pass through the bars, Cherenkov photons are emitted in a cone about the track momentum vector with an opening angle  $\theta_C$  given by

$$\cos(\theta_C) = \frac{1}{\beta n}, \quad (\text{II.3})$$

where  $\beta$  is the velocity divided by the speed of light, and  $n$  is the index of refraction for the bars. The photons are transmitted down the bars and the angle  $\theta_C$  is maintained via total internal reflection (TIR). At the forward end of the detector, mirrors reflect the light toward the opposite end. At the back end of the detector, the bars terminate at the conical, water-filled standoff box (SOB). Photomultiplier tubes (PMTs) line the rear of the SOB and detect the photons coming from the bars. A trapezoidal wedge of silica is fixed to the end of each bar. By reflecting the photons at large angle with respect to the bar axis, the silica wedge reduces losses due to TIR at the silica/water interface, and reduces the density of PMTs needed for a given resolution. Figure II.17 shows the details of the bar end region, including the wedge and the SOB.

### 3.4 Electromagnetic Calorimetry

While the inner detectors (the vertex tracker, drift chamber, and Cherenkov counter) are specifically designed to have a minimal and predictable impact on a particle's momentum, the electromagnetic calorimeter does just the opposite. The

---

make use of the DIRC. This updated selector is used for essentially all electron identification in recent *BABAR* analyses.

EMC is constructed of a material which induces electromagnetic showers, the products of which are read out and used to make measurements of energy and angular position. For the general program of B meson physics at *BABAR*, the design and performance of the EMC is driven by the need to detect photons from the decays of neutral pions and  $\eta$  mesons. In many analyses, including the  $\tau \rightarrow \ell\ell\ell$  analysis, the energy deposition in the EMC is used to identify electrons.

## Calorimetry Requirements and Design

The EMC is designed to measure electromagnetic showers over the range of energy from 20 MeV to 9 GeV. The lower bound comes from the need for efficient reconstruction of B meson decays containing neutral pions and  $\eta$  mesons decaying to photons. The upper bound on the energy range is set by the need to measure high energy electrons from the  $e^+e^- \rightarrow e^+e^-e^+e^-$  and  $e^+e^- \rightarrow \gamma\gamma$  processes which are used for calibration. Energy resolution of 1 – 2% is required for rare processes involving neutral mesons decaying to high energy photons. Measurement of these rare processes also requires angular resolution of a few mrad at energies above 2 GeV. The EMC must also fulfill a number of physical and mechanical requirements, including the ability to operate inside the 1.5 T magnetic field. Temperature and radiation exposure must be carefully monitored and controlled, energy calibrations must be easily performed over the full energy range, and the whole detector must operate reliably over the expected ten-year lifetime of the machine.

To meet the stated physics requirements, the EMC was constructed from thallium-doped cesium-iodide (CsI(Tl)) crystals in a finely-segmented array. The crystals have a high light yield and a short radiation length relative the crystal depth. The transverse size of the crystals is approximately the Moliere radius of the material, which optimizes the angular resolution while appropriately minimizing the number of readout channels for each shower. The relevant properties of CsI(Tl) are shown in Table II.3.

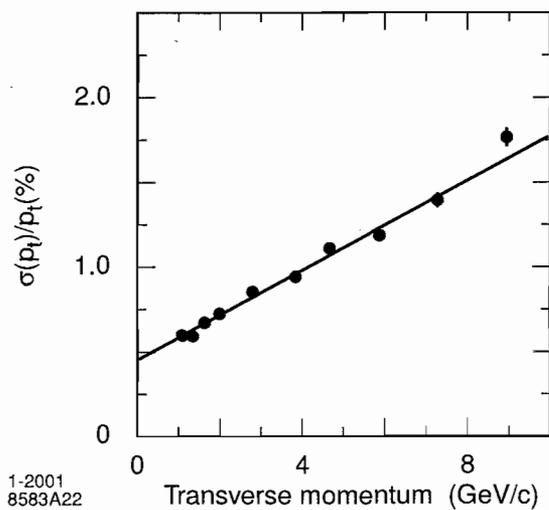


Figure II.15: Transverse momentum resolution, as determined from cosmic ray muons traversing the DCH and SVT.

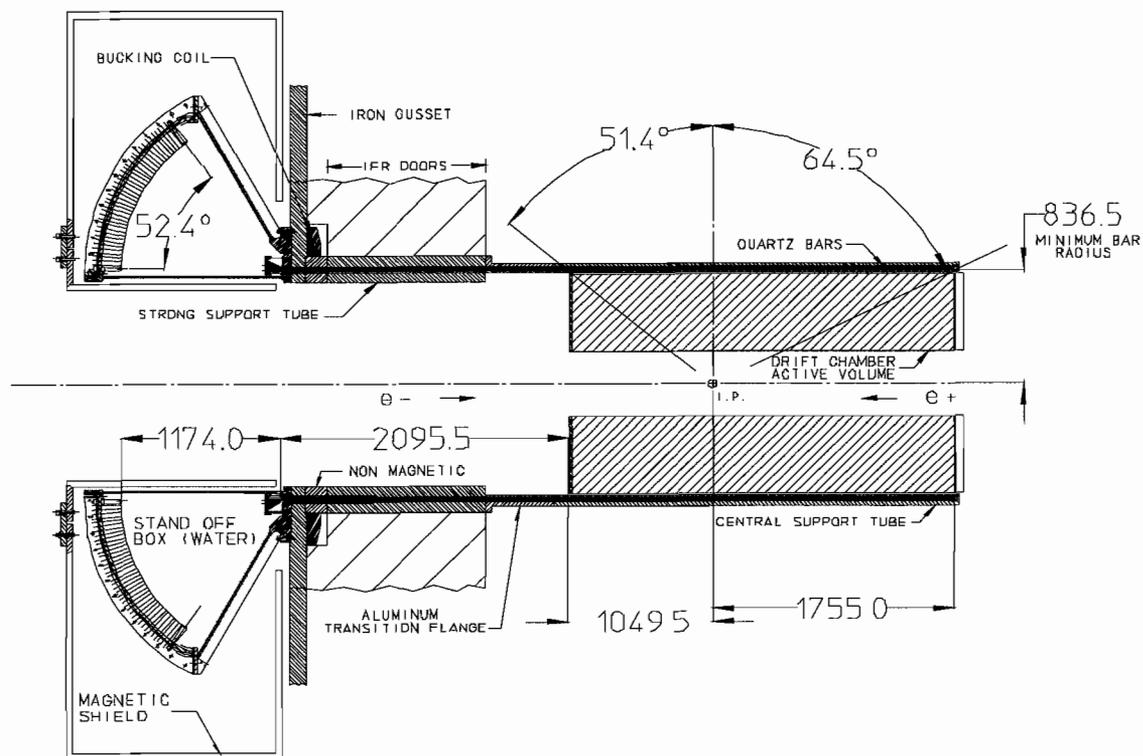


Figure II.16: Longitudinal view of the DIRC.

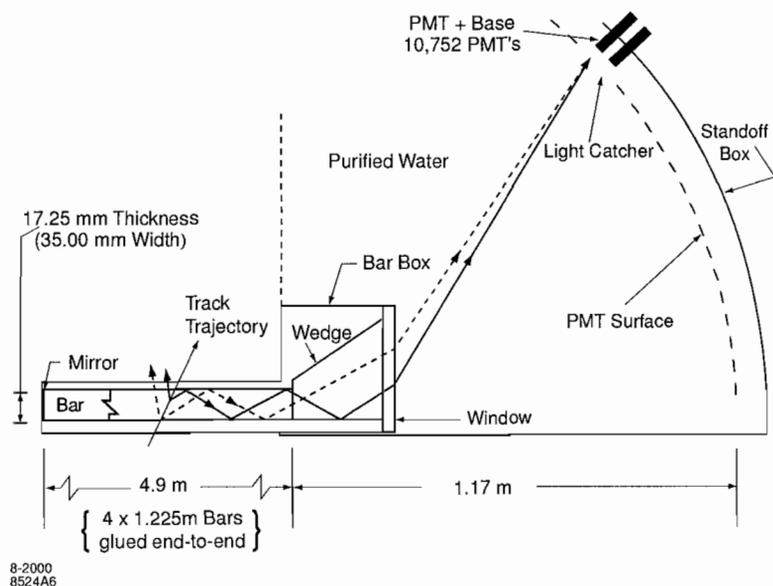


Figure II.17: Detail of the DIRC bars and the imaging region.

Table II.3: Properties of CsI(Tl).

Parameter	Values
Radiation Length	1.85 cm
Moliere Radius	3.8 cm
Density	4.53 g/cm <sup>3</sup>
Light Yield	50000 $\gamma$ /MeV
Light Yield Temp. Coeff.	0.28%/°C
Peak Emission $\lambda_{max}$	565 nm
Refractive Index ( $\lambda_{max}$ )	1.80
Signal Decay Time	680 ns (64%) 3.34 $\mu$ s (36%)

The EMC consists of a cylindrical barrel and a conical forward endcap. The detector has full  $360^\circ$  azimuthal coverage and polar coverage from  $15.8^\circ$  to  $141.8^\circ$ , corresponding to 90% coverage in the CM system. The crystals have a tapered trapezoidal cross section and lengths which vary according to the polar position of the crystal. A longitudinal cross section is shown in Figure II.18. The barrel contains 5760 crystals arranged in 48 rings in  $\theta$ , each containing identical 120 crystals evenly spaced in  $\phi$ . The endcap contains 820 crystals arranged in 8 rings in  $\theta$ . The innermost two rings in the endcap are primarily for shower containment, and electrons at the corresponding polar angles are difficult to identify. To minimize the amount of pre-showering, the crystals are supported from the outside and only a thin gas seal separates the EMC from the DIRC.

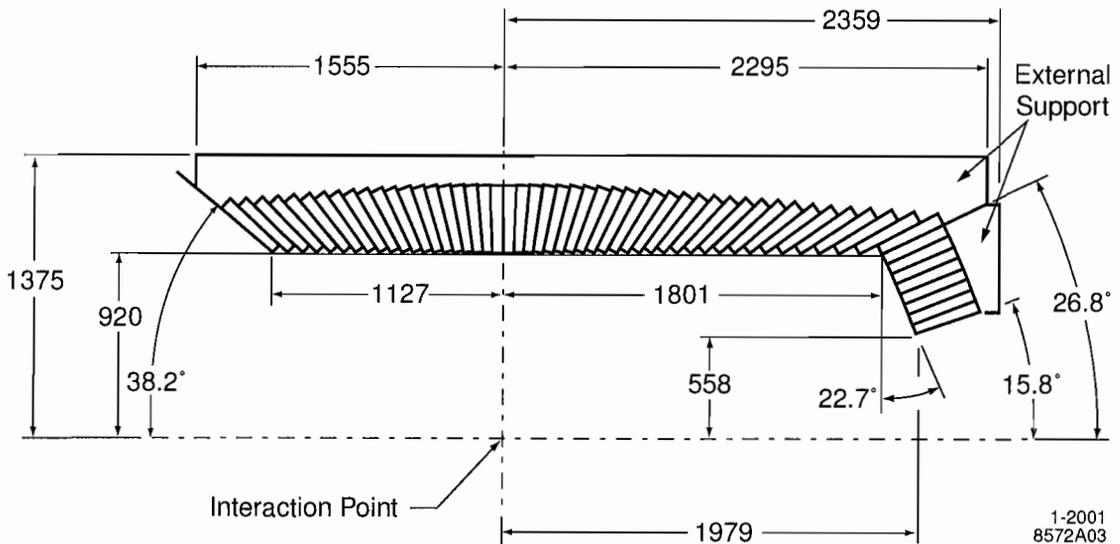


Figure II.18: Longitudinal view of the EMC.

Each crystal is read out by a pair of photodiodes at the back of the crystal. While most light is internally reflected by the crystal surfaces, each crystal is wrapped in two layers of reflective material to enhance the number of photons which reach the back of the crystal. Further layers of foil and epoxy provide shielding and electrical isolation. Each photodiode is connected to a low-noise preamplifier. The amplified signal is passed on to a custom auto-range encoding circuit, which provides different gains for different ranges of energy. Upon the reception of an L1 accept signal, features extraction is performed on a  $\pm 2\mu\text{s}$  window around the

waveform peak. A schematic of the wrapped crystal and some of the readout electronics is shown in Figure II.19.

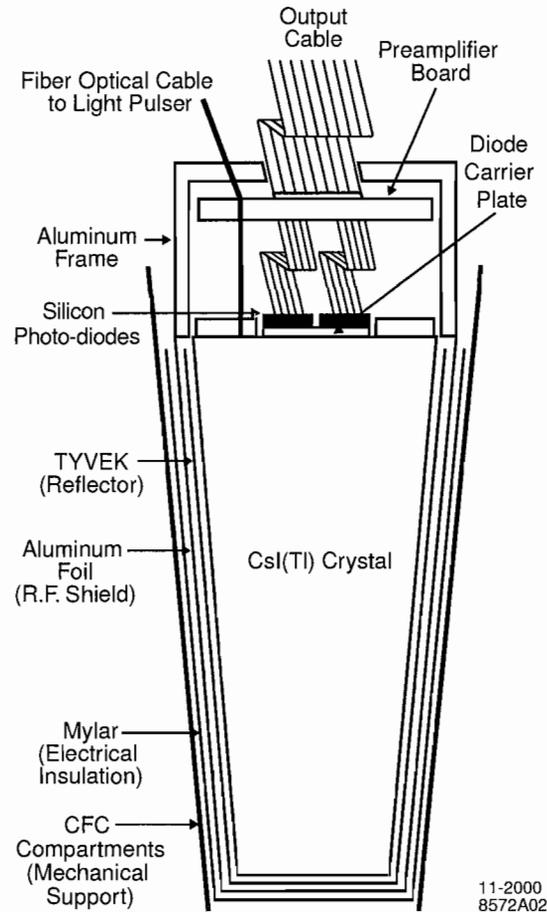


Figure II.19: A schematic of the wrapped crystal and the readout electronics on the back end.

## Calorimetry Performance

The energy resolution of the EMC is measured with a number of different sources over a wide range of energy, and can be parameterized as

$$\frac{\sigma_E}{E} = \frac{(2.3 \pm 0.3)\%}{\sqrt[4]{E(\text{GeV})}} \oplus (1.85 \pm 0.12)\%. \quad (\text{II.4})$$

Figure II.20 shows the measured energy resolution as well as the fitted function (Equation II.4). The first term in the fit comes from statistical fluctuations in the

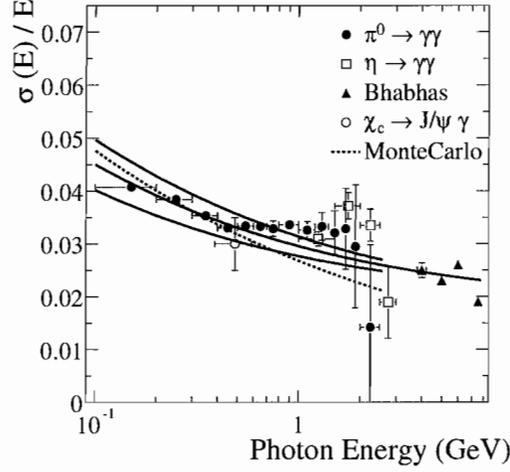


Figure II.20: Energy resolution of the EMC for photons and electrons, as measured for various processes. The solid line is from the fit (Equation II.4), and the shaded area denotes the fit error.

number of photons and other electronic noise. The constant term  $b$ , which dominates at high energies, is associated with light collection, leakage, and absorption between and in front of the crystals. The angular resolution, which due to the crystal cross section is the same in  $\theta$  and  $\phi$ , is parameterized similarly:

$$\sigma_{\theta} = \frac{(2.3 \pm 0.3)\%}{\sqrt{E(\text{GeV})}} \oplus (1.85 \pm 0.12)\%. \quad (\text{II.5})$$

Figure II.21 shows the angular resolution for photons from  $\pi^0$  decays.

Electron identification make significant use of the EMC. While the  $dE/dx$  loss in the DCH and the Cherenkov angle in the DIRC can be used to separate charged hadrons from electrons, the most important variable for positive electron identification is the ratio of the energy deposited in the calorimeter to the momentum of the charged track, as measured in the DCH and SVT. This ratio should be very near unity for electrons. The details of the electron identification algorithm used in this analysis can be found in Appendix B. Plots of the electron selection efficiency and hadron mis-ID rates for the electron identification algorithm used in the analysis can be found in Section 4.

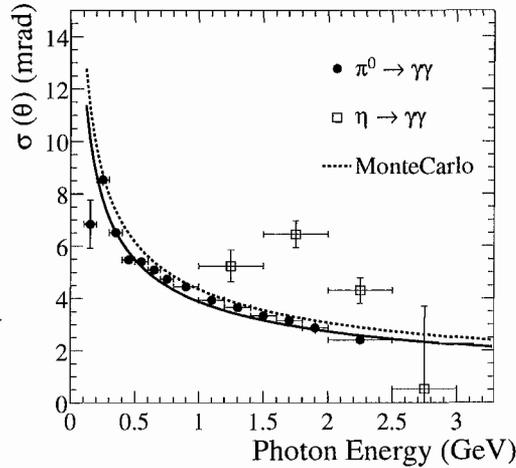


Figure II.21: Angular resolution of the for photons from  $\pi^0$  decays. The solid line is the fit (Equation II.5).

### 3.5 Muon Detection

The Instrumented Flux Return must efficiently identify muons over a wide range of momenta and angles. The IFR was also designed to detect neutral hadrons, such  $K_L^0$ 's and neutrons, although in practice the IFR is rarely used to detect anything but muon identification. In terms of the general *BABAR* physics program, muon detection plays an important role in the measurement of leptonic and semileptonic decays of B and D mesons. The IFR aids in the measurement of missing momentum and can be used to veto charm decays. Muon identification plays a particularly important role in the  $\tau \rightarrow \ell\ell\ell$  searches. Because some of the searches (particularly  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ ) are expected to produce muons with momentum lower than 500 MeV, the performance of the IFR for identification of muons over a large momentum range is of great importance.

The IFR is an integration of the magnetic flux return and the muon identification instrumentation. It consists of a central hexagonal barrel, which covers 50% of the solid angle in the CM frame, and two endcaps (Figure II.22). The IFR is 3.75 m long, and has an outer radius of 3.01 m and an inner radius of 1.78 m.

The finely-segmented steel of the IFR provides a path for the solenoid's magnetic field. The steel is segmented into layers ranging in thickness from 2 cm to 10 cm. The thickness was chosen based simulations of muon penetration and hadron

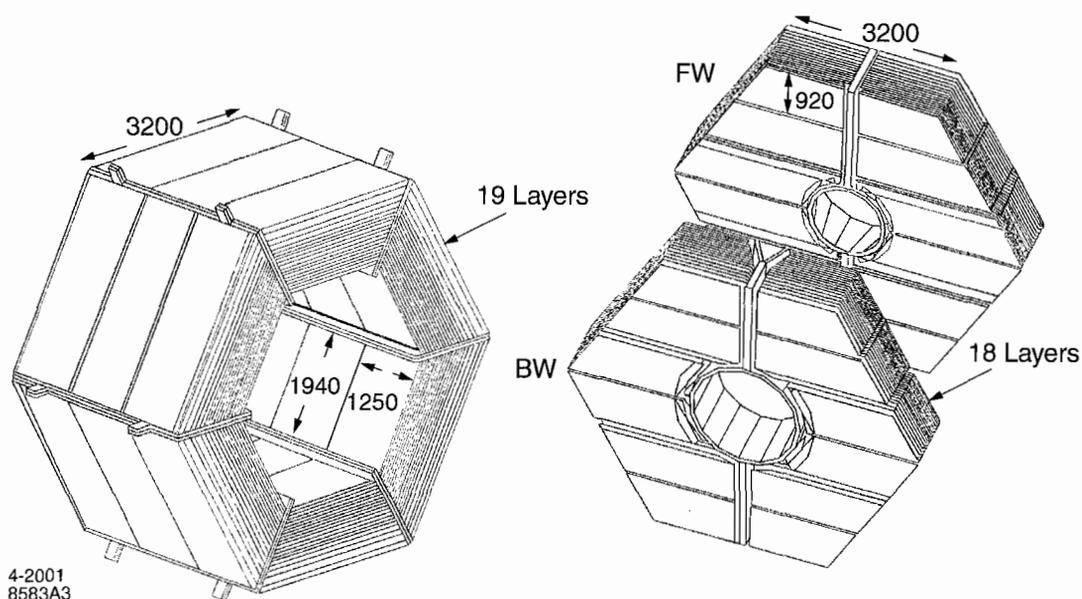


Figure II.22: Overview of the IFR: barrels sectors and forward (FW) and backward (BW) endcaps.

shower shapes in the steel. The IFR steel is interspersed with layers of detector material: either resistive plate chambers (RPCs) or limited streamer tubes (LSTs). Two-dimensional position measurements are made by using orthogonal readout strips oriented in  $\phi$  and  $z$ . In the original detector, the IFR instrumentation consisted solely of RPCs. Currently, the hexagonal barrel is instrumented with LSTs, while the forward and backward endcaps still contain RPCs.

The RPCs identify muons by detecting streamers between a high-voltage gap. Each RPC strip consists of two Bakelite sheets, 2 mm thick and separated by a gas-filled 2 mm gap. The inside surfaces of the Bakelite are treated with linseed oil. The outside surfaces are painted with graphite and kept at high voltage (about 8 kV and ground). An insulating mylar film separates the high voltage graphite from the aluminum readout layer. The cross-section of an RPC is shown in Figure II.23. A charged particle crossing the RPC gap initiates an electric discharge which is read out via capacitive coupling to the aluminum strips. The readout strips on opposite sides of the RPC are oriented orthogonally to give a two-dimensional position measurement.

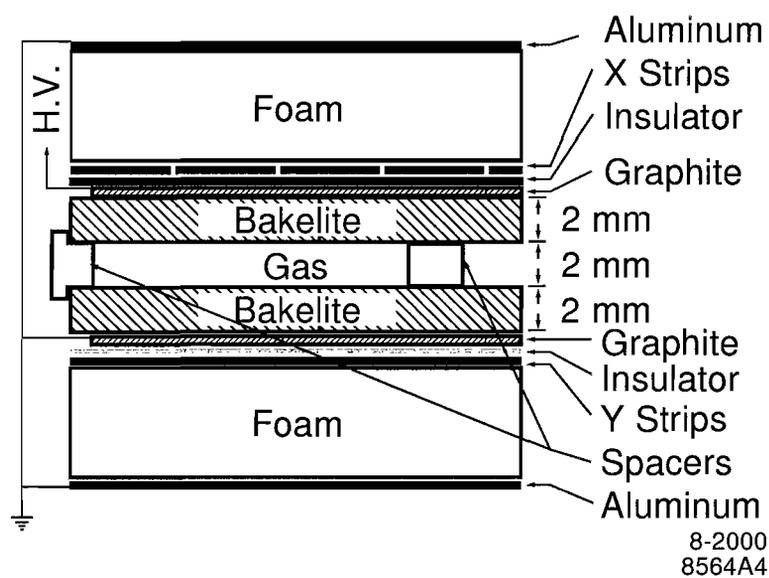


Figure II.23: Cross section of a planar RPC, with schematic of the high voltage connections.

The LSTs identify muons by detecting streamer ionization on a high-voltage wire. LSTs are constructed of a single  $100\ \mu\text{m}$  diameter sense wire running down the center of a  $9\text{mm} \times 9\text{mm}$  plastic section. Plastic structures, or profiles, contain 8 such sections side-by-side, with one side open. These profiles are coated with graphite and inserted into plastic tubes of matching dimensions for gas containment. Signals on the wires themselves provide a  $\phi$  measurement, and strips on the outside of the tubes running perpendicularly to the wire provide a  $z$  measurement.

The original IFR contained 19 layers of RPCs in the barrel, 18 layers in the endcaps, and 2 cylindrical layers between the EMC and the magnet. The RPCs in 12 of the barrel layers were replaced with LSTs over the period 2004-2006. Six of the remaining layers were filled with brass to compensate for the loss of absorbing material.

Muon identification relies almost entirely on the IFR, although other systems can provide limited information. Muons are detected as tracks in the SVT and the DCH, and must behave like a minimum-ionizing particle in the EMC. The tracks from the inner detector are extrapolated to the IFR, taking into account the non-uniform magnetic field, multiple scattering, and the average energy loss. Extrapolated tracks for real muons must appropriately intersect the observed clusters of hits in the IFR. The depth of penetration into the IFR must also be consistent with a muon of the given momentum and angle.

When developing selection criteria for the identification of muons, there is always a trade-off between efficiency for muon and mis-ID rates for pions and other hadrons. These numbers are parameterized in terms of particle momentum in the lab frame, polar angle, and azimuthal angle. The IFR efficiency for identifying low-momentum muons is one of the limiting factors for the  $\tau \rightarrow \ell\ell\ell$  searches. In particular, muons with momentum less than 1 GeV rarely reach the IFR. The actual performance of the IFR for detecting muons from LFV tau decays is discussed in Section 4.

## 4 Simulations

In the Babar experiment, simulations of  $e^+e^-$  collisions and the subsequent detector and trigger response play an important role. These simulations can be used to compensate for detector inefficiencies, as well as providing theoretical predictions for distributions. It is useful to simulate both the signal events for which one is

searching as well as the background events which mimic the signal. Simulations of signal processes allow one to carefully study the effect of one's analysis on the signal efficiency. Simulations of the background processes allow precise comparisons between the distributions of the simulated events and distributions of the data themselves. Once the validity of the background simulations are confirmed for a general situation, they can then be used to make predictions of specific background contributions. This method of background prediction is in contrast with predictions made directly from data, in which biases can be introduced by extrapolating from a potentially small number of data events. For the  $\tau \rightarrow \ell\ell\ell$  analysis, simulations of the expected background events are compared to real data events in a kinematical region *near* where the signal is expected. Once the background simulation is verified, these events can be used to predict the expected number of background events *in* the signal region - all without actually counting data events in the signal region. This procedure reduces sensitivity to potentially large statistical fluctuations in the number of background events seen in the small signal region of the data sample.

The simulation of events at Babar starts with piece of software called an event generator. The goal of such a generator is to reproduce the behavior of the colliding  $e^+e^-$  pair. A minimal set of desirable behaviors includes the accurate simulation of differential and total cross sections as well as initial and final state radiation, and the proper treatment of spin, particularly for short-lived particles. Initial and final state radiation refers to the emission of one or more photons from the initial (incident) electron and positron or from the final (outgoing) particles. For the  $\tau \rightarrow \ell\ell\ell$  search at Babar, the  $\tau$  particles are produced in pairs via the reaction  $e^+e^- \rightarrow \tau^+\tau^-$ . Consequently, the final-state particles in the simulation of the reaction are  $\tau$  particles, which have a lifetime of about 0.29 picoseconds and which travel on average less than a tenth of a millimeter before decaying. A secondary piece of software simulates the decay of the  $\tau$  particles and the radiation from the  $\tau$  decay products. Other software simulates the detector response and the trigger.

The individual momentum four vectors for particles in simulated events are generated using Monte Carlo (MC) techniques. In fact, the data sets containing the simulated events are often referred to as "Monte Carlo". There are a number of different elementary MC algorithms but the basic goal is the same: to use randomly generated numbers to create data which follow a specified distribution. This process can also be thought of as numerical integration of the distribution. A trivially simple example is that of a flat, bounded distribution in one variable. In this case,  $n$

random numbers are drawn from a uniform distribution,

$$f(x; a, b) = \begin{cases} 1/(b - a) & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \quad (\text{II.6})$$

As  $n$  becomes large, a histogram of the generated values reproduces the original distribution.

Because it is computationally cheap to generate random numbers which follow a uniform distribution, such numbers are often used as a seed for MC events which are to follow a more complicated distribution. In the Acceptance-Rejection method of Von Neumann, the desired probability density function (PDF)  $f(x)$  is enclosed by a function  $C h(x)$ , where  $C$  is a constant greater than 1 and  $h(x)$  is typically a uniform distribution or a normalized sum of uniform distributions. To generate data distributed according to  $f(x)$ , a candidate  $x$  is first generated according to  $h(x)$ . A second candidate  $u$  is then drawn from a uniform distribution  $(0, 1)$ . The candidate  $x$  is accepted into the data set if  $u C h(x) \leq f(x)$ ; otherwise  $x$  is rejected and the process starts over.

The event generator used for the simulation of  $\tau$  pair production at Babar is **KK2f** [50]. Conceptually, the algorithm is simple: the differential cross section is given by the squared, spin-summed matrix element times the phase space. Random numbers are used to draw a specific value from the PDF for each independent quantity, such as  $|\vec{p}_\tau|$ ,  $\phi$ ,  $\theta$ , etc. In practice, the allowance for arbitrary numbers of initial and final state photons which can interfere with each other, plus the inclusion of higher order QED and EW corrections, makes for a very complicated calculation. The **KK2f** generator achieves significantly better precision than the previous generators of its kind (e.g. **KORALZ**[51], **KORALB**[52]). For the simulation of other background processes such as  $e^+e^- \rightarrow q\bar{q}$  ( $q = u, d, s, c, b$ ), Babar uses the **EvtGen** [53] and **Jetset** [54] packages.

The decay of the  $\tau$  particles is simulated by the **TAUOLA** software package [55]. For the  $\tau \rightarrow \ell\ell\ell$  analysis, **TAUOLA** must generate two different classes of  $\tau$  decays: generic decays in which the  $\tau$  decays according to SM branching fractions and differential decay widths, and specific LFV decays for which the distributions are not known.

Generic  $\tau$  decay rates are defined in **TAUOLA** by a **DECAY.DEC** file. The file lists the most recent values of the  $\tau$  branching fractions from the Particle Data

Group (PDG) [56]. For a given  $\tau$ , a specific decay mode is selected randomly with weights given by the measured branching fractions. Then, an algorithm to specify the outgoing particle momenta and angle must be chosen. For leptonic decays of the  $\tau$ , the SM matrix element is known. From the square of the matrix element one can calculate the differential decay width, which leads directly to PDFs for parameters of the outgoing leptons and neutrinos. Complete QED corrections of  $\mathcal{O}(\alpha)$  are included in TAUOLA. For two-body semileptonic decays ( $\tau \rightarrow K\nu_\tau$ ,  $\tau \rightarrow \pi\nu_\tau$ ), SM calculations give the differential decay widths to zeroth order, with the pion and kaon decay constants taken from experiment. Radiative corrections are included in the leading logarithmic approximation.

For  $\tau$  decays with two or more hadrons in the final-state, one must choose a specific parameterization for the hadronic portion of the matrix element. The choice of this form factor is influenced by the observation that hadronic  $\tau$  decays are dominated by intermediate resonances decaying to pions, kaons, and other pseudoscalars. In TAUOLA, these form factors are thus parameterized as Breit-Wigner functions corresponding to the intermediate vector and axial-vector resonances<sup>3</sup>. The masses and widths of these resonances must be taken from experiment. For high-multiplicity decays, chains of these resonances are used, with heavier intermediate particles decaying to lighter resonances along with final-state pseudoscalars. For decays where the same final state can occur via different decay chains, the relative contribution of each path is fixed to the experimental value.

In the search for neutrinoless  $\tau$  decays to three leptons, the TAUOLA program is also used to simulate the LFV decays. Since these decays have never been observed, and few (if any) models exist which predict the dynamics of the final-state leptons, the choice is made to model these decays in the simplest way possible. The matrix element is set to unity and the differential decay widths are proportional to only the Lorentz-invariant phase space for three particles. This choice explicitly removes any resonant behavior and does not allow for relative angular momentum between any two outgoing leptons.

Radiation from the leptonic decay products of the  $\tau$  particles must be simulated as well. For the *BABAR* experiment, this is done by the PHOTOS software package [57].

---

<sup>3</sup>In the case of some higher-multiplicity  $\tau$  decays, these resonances could also be pseudoscalars.

The output of the event simulation is a set of four-vectors which describe the kinematics of the long-lived particles in the event<sup>4</sup>. The four-vectors are used as inputs to GEANT4[58], a software package which simulates the passage of particles through the Babar detector. This simulation models multiple scattering, leptonic and hadronic ionization of the traversed material, leptonic bremsstrahlung and pair production, positron annihilation, the photoelectric effect, and Compton scattering. The simulation also incorporates the effect of background noise in the detector by mixing in signals taken from real snapshots of the detector subsystem electronics. Finally, the simulated detector output is passed to the L1 trigger simulation (see Section 5). If the trigger simulation generates an Accept signal, the detector simulation output is passed onto the L3 trigger and the reconstruction software, just as if it were data corresponding to a real event.

## 5 Data Acquisition and Triggering

The high luminosity of PEP-II is achieved in part by shortening the space between bunches, which corresponds to a higher bunch-crossing frequency at the IP. This high event rate amounts to an essentially continuous stream of collisions, preventing the synchronization of the detector readout with the bunch-crossings. The actual physics rate, by which we mean  $q\bar{q}$ ,  $\mu^+\mu^-$ , and  $\tau^+\tau^-$  events, is only about 65 Hz at an instantaneous luminosity of  $10^{34}$  cm<sup>-2</sup> s<sup>-1</sup>. Bhabha scattering, which is generally uninteresting for physics purposes, contributes around 500 Hz, and random interactions of the beam produce detectable tracks and clusters at nearly 20 kHz. Since the data storage rate is limited to 100-200 Hz, the triggering mechanism must provide an event rate reduction of around two orders of magnitude.

### 5.1 Trigger Requirements and Design

The *BABAR* trigger is designed as a two-level system: a hardware-based Level One (L1) trigger, and a software-based Level Three (L3) trigger. *BABAR* has no Level Two trigger. The trigger is required to operate with very high efficiency for physics processes of interest, and with good stability and easily measured and reproducible

---

<sup>4</sup>In defining which particles are long-lived, some care must be taken with particles of intermediate lifetimes, such as  $K_S^0$  mesons, which decay a measurable distance from the IP. For the  $\tau \rightarrow \ell\ell\ell$  such particles are relatively unimportant, as they only occur in the background and are not part of the signal.

behavior. Specifically, the efficiency for triggering on  $B\bar{B}$  pairs must exceed 99%, and deadtime must not exceed 1%.

To achieve the necessary event-rate reduction, event data for the entire detector is read into storage buffers every 67 ns. This time interval corresponds to 16 bunch-crossings, most of which are empty events with no interesting physics. The storage buffers can hold data for up to 193 events. In parallel, a small subset of the event data is sent to the trigger for processing. The size of the event buffer sets the limit on the total time for the L1 trigger to make the choice to store an event for further processing. This latency is about 13  $\mu$ s.

The trigger algorithms must be sufficiently simple to allow for relatively easy and accurate simulation. In order to meet these requirements, the trigger was designed to recognize general topologies rather than specific physics processes. Orthogonal selection criteria allow for independent calibrations of different components and robustness against missing and fake signals. The data objects calculated as part of the trigger algorithm are stored and made available for efficiency studies. Finally, a small number of events are passed and stored regardless of the trigger decision. These events provide further data for performance studies. The trigger is made to as flexible as possible, with a maximum amount of configurable parameters.

## 5.2 Level One Trigger

The Level One trigger samples a small set of the DCH and EMC signals every 269 ns. The IFR is sampled every 134 ns. A decision whether to store the event for further processing must be made within the 13  $\mu$ s latency window. The L1 trigger consists of three sub-triggers working in parallel: the Drift Chamber Trigger (DCT), the EMC trigger (EMT), and the IFR trigger (IFT). A global L1 trigger (GLT) collects outputs from the 3 sub-triggers and forms a number of configurable trigger lines. The values of these lines are passed to the Fast Control and Timing System (FCTS), which makes the final decision to read out the event buffers and send the event for further processing. In order to limit the load on L3, the L1 output rate is configured to be no more than 1-2 kHz.

## Level One Drift Chamber Trigger

The input to the DCT consists of a single bit for each DCH sense wire. The output is a set of 16-bit  $\phi$ -maps which represent candidate tracks. These maps are generated through use of three different modules. First, DCH signals are combined to form track segments by set of 24 Track Segment Finder (TSF) modules. Information about these segments is then passed to the Binary Link Tracker (BLT) module, where the segments are linked to form complete tracks. In parallel with the BLT, TSF outputs are also sent to a set of eight  $z_0/p_T$  Discriminator (ZPD) modules, which select tracks based on a fit to their transverse momentum ( $p_T$ ) and distance of closest approach to the z-axis ( $z_0$ ). Prior to 2004, Transverse Momentum Discriminator (PTD) modules were used to select tracks with high  $p_T$ . PDT modules did not fit for  $z_0$ . With the projected increase in background in mind, the ZPD modules were designed to better reject backgrounds by discarding events with  $z_0 > 20$  cm.

The Track Segment Finder modules are responsible for finding track segments in the 1776 overlapping groups of eight DCH cells called *pivot groups* (see Figure II.24). Each group contains one *pivot cell* and each cell contains one sense wire. The signals on every DCH sense wire are sampled every 269 ns. Each signal found increments a two-bit counter for the cell and the counters for all eight cells in the group form a 16-bit value that is used to address a lookup table. In the case that the group value corresponds to a valid segment, the lookup tables provide position and time information which form the basis of the output data. The TSF algorithm is capable of refining the event time and its uncertainty such that the output data can be forwarded to the BLT and the ZPD every 134 ns.

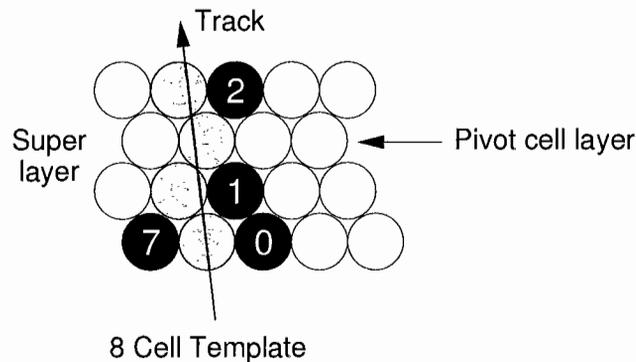


Figure II.24: Track Segment Finder pivot group.

The Binary Link Tracker receives hit information from the TSF and maps it onto the DCH geometry in terms of a map of supercells: 32 sectors in  $\phi$  and 10 radial superlayers (SL). The segments are combined in such a way that dead or inefficient supercells do not degrade the track-finding efficiency. The linking algorithm is based on the CLEO-II trigger [59], and starts from the innermost superlayer and works its way outward. Linked track segments are classified by the outermost superlayer reached. Short tracks are defined by reaching the middle superlayer, and long tracks must reach the outermost superlayer. See Table II.4 for the definition of these and other DCT output objects. These tracks are sent to the GLT in the form of a 16-bit  $\phi$ -map.

The  $z_0/p_T$  Discriminator modules provide further background rejection by evaluating candidate tracks according to their  $z_0$  value. Figure II.25 shows the distribution in  $z_0$  of tracks reconstructed by L3 *without* a cut on  $z_0$ . The ZPD algorithm first searches seed track segments from the TSF and fits them for an initial measurement of  $p_T$  and the dip angle ( $\lambda$ ). Other segments are added to the candidate track and, by using information from the DCH stereo superlayers, subsequent fits give a value for  $z_0$  and refined values for  $p_T$  and  $\lambda$ . Tracks reaching SL 7 and with  $p_T$  and  $z_0$  values within an adjustable range are sent on to the GLT. Table II.4 shows these and other DCT output objects.

## Level One Calorimeter Trigger

The Electromagnetic Calorimeter Trigger searches for calorimeter showers above specified energy levels, and sends corresponding location information to the GLT. The EMT operates in terms of *towers*,  $240 \times 8 \times 3$  ( $\theta \times \phi$ ) arrays of crystals in the barrel and 40 19-22 crystal wedges in the endcap. Every 269 ns, all crystal energies above 20 MeV are summed over each tower and sent to the EMT. The conversion of the tower energy to  $\phi$ -maps for the GLT is done by 10 Trigger Processor Boards. These boards determine the total energy in the 40 sectors in  $\phi$ , while summing over different  $\theta$  ranges. These energy sums are compared against the trigger objects shown in Table II.4. After an estimation of the time of the energy deposit and a correction for timing jitter, the results are sent to the GLT.

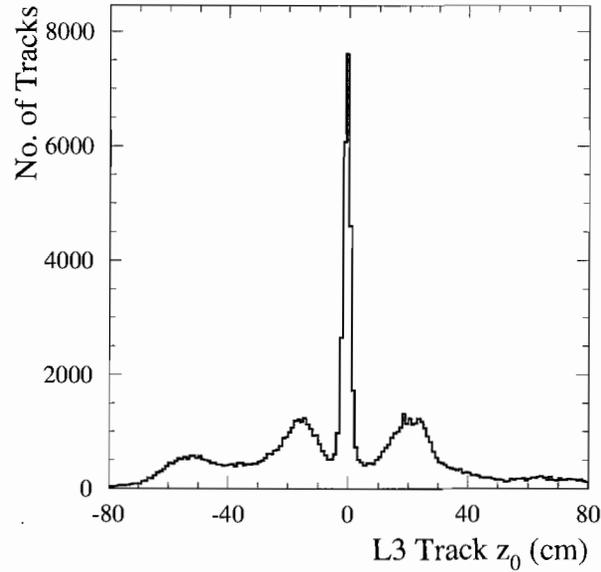


Figure II.25: Single track  $z_0$  for all L1 tracks without a cut on  $z_0$ .

## Level One Muon Trigger

The Level One Muon Trigger (IFT) is used to trigger on muon pairs from the IP and cosmic rays. The output of the IFT is used primarily for calibration and diagnostic purposes. For the purpose of the trigger, the IFT is split into ten sectors: one for each of the six barrel sextants, and one for each half end-door. The input to the IFT is an OR of all  $\phi$ -strips in eight selected layers in each sector. The IFT module samples these sectors every 134 ns and generates a three-bit trigger word (U) in which is encoded the values for the seven IFT trigger conditions (see Table II.5).

## Global Level One Trigger

The inputs to the Global Level One Trigger are the 11 trigger objects (stored in the form of  $\phi$ -maps) listed in Table II.4, plus the IFT summary word U. First, the GLT synchronizes the incoming signals by accounting for different latencies in the L1 components. Next, an additional set of  $\phi$ -maps are formed from back-to-back tracks and clusters, and pairs of tracks and clusters with similar  $\phi$  values. The  $\phi$ -maps are used to address lookup tables for a count of the number of each trigger object in the

Table II.4: Primitive trigger objects constructed by the Level 1 trigger.

	DCT Object	$p_T$ cut	$z_0$ cut
A	Short tracks reaching SL5	120 MeV/c	-
B	Long tracks reaching SL5	150 MeV/c	-
Z	Standard Z tracks reaching SL7	200 MeV/c	12 cm
Zt	Tight Z tracks reaching SL7	200 MeV/c	10 cm
Z'	High $p_T$ tracks reaching SL7	800 MeV/c	15 cm
Zk	Tracks reaching SL7 (asymmetric cut)	200 MeV/c ( $e^-$ ) 800 MeV/c ( $e^+$ )	12 cm
	EMT Objects	energy cut	
M	Minimum ionizing clusters	130 MeV	-
G	Intermediate energy clusters	350 MeV	-
E	High energy electron/photon	900 MeV	-
X	M object in forward endcap	130 MeV	-
Y	Backward barrel electron	1000 MeV	-
U	IFT hit pattern from IFR	-	-

Table II.5: IFT trigger patter definitions, where  $\mu$  refers to a signal in a sector.

U	Trigger condition
1	$\geq 2\mu$ topologies other than U = 5,6,7
2	1 $\mu$ in backward endcap
3	1 $\mu$ in forward endcap
4	1 $\mu$ in barrel
5	2 back-to-back $\mu$ s in barrel + 1 forward $\mu$
6	1 $\mu$ in barrel + 1 forward $\mu$
7	2 back-to-back $\mu$ s in barrel

event, and trigger lines are formed by logical combinations of these counts. The time of the trigger is derived from the timing distribution of the highest priority trigger. The GLT sends the value of each trigger line to the FCTS, which can optionally mask or scale any of those triggers. If a valid trigger remains, the FCTS issues an L1 Accept signal, and the FEE buffers are read out for the appropriate time.

### 5.3 Level Three Trigger

The L3 trigger performs fast track finding and fitting by using segments from the TSF and then taking actual DCH information to better the resolution on the track parameters. First, the time of the track is determined from the TSF segments. Then, the segments compatible with this time are used to address a lookup table. This table is populated with data corresponding to simulated tracks above a cutoff transverse momentum and which originate within a certain adjustable distance from the IP. Tracks from the table are then refit for all five track parameters. This full fit allows for tracks which do not originate from the IP.

The L3 trigger processes the EMC data in two steps: first, crystals with energy above a threshold are identified, and second, clusters are formed from these crystals. Individual crystal energy measurements are rejected if they are below 20 MeV or if their timestamp lies outside the 1.3  $\mu$ s event window. The remaining crystal energy measurements are added, along with their times, to a list. Clusters are formed by using a lookup table addressed by the crystal positions and energies in the list. Clusters must have energy greater than 100 MeV.

Based on the fitted tracks and the reconstructed clusters, the L3 trigger performs a variety of filtering processes that classify events and reduce backgrounds. A Bhabha filter identifies and vetos one- and two-prong events with E/p measurements consistent with the expectation for Bhabha events. Other filters flag radiative Bhabha,  $\gamma\gamma$ , and cosmic events for calibration and luminosity monitoring.

## 6 Offline Data Processing

Events which are selected by the Level Three trigger are stored for further processing. These events are grouped into *runs*, with each run representing approximately an hour's worth of data-taking. The full set of detector signals for a

run of events is written to a single data file, usually referred to as an extended tagged container (XTC) file. The raw size of each event in the XTC is about 30 KB, and XTC files are typically a few tens of GB in size.

## 6.1 Prompt Data Reconstruction

In recent years, the full processing of the event data has been conducted offline, meaning that the data are not fully processed in real-time (i.e. as the signals are recorded by the detector). Instead, all subsequent processing operates groups of events corresponding to one run (and one XTC file). These data are processed in a two-pass system. First, calibration conditions are calculated from a subset of the events in the run and written to the conditions database. This step is referred to as Prompt Calibration (PC). Secondly, all the events are reconstructed based on the conditions in the database, and are written out to event collections. This step is referred to as Event Reconstruction (ER).

The PC step of the data processing makes use of only a subset of the events in a run. For technical reasons, these events are also stored in a secondary data file called a *calib-XTC*. The *calib-XTC* file for each run is filled with events passing a particular set of L3 trigger output lines, all of which are designed to provide a constant output rate of 1 or 2 Hz, depending on the trigger line. These output lines select Bhabha events, di-muon events, cosmic muons, and low-multiplicity hadronic events. The PC processing software runs on these events and writes out a set of calibrations which give a picture of the detector conditions at the moment the events were recorded.

To better account for changing detector conditions, the PC processing software makes use of *rolling calibrations*. In this method, the calibration constants from previous runs are stored and used as additional input information for the calculation of the current run's calibrations. This method also effectively provides for larger statistics without actually increasing the sampling rate for the calibration events. When all the calibration events in a run have been processed, the results are collected and passed to a final processing module. This module calculates the final calibration constants for that run and writes them to a temporary database where they are made available for the next run. The calibration constants are also written to the main conditions database, and assigned a validity interval corresponding to the time interval over which the run was recorded. During a period of high

luminosity typical of the later years of *BABAR* running, the PC processing step was performed on computers at SLAC, utilizing around 30 CPUs.

The Event Reconstruction step processes the full set of events in the XTC file. Because the detector has already been calibrated for the run period, the event reconstruction can process the events in any order. This task is accomplished by a farm of a few hundred multi-CPU computers at Padova, Italy, along with (more recently) a similar farm at SLAC.

The actual reconstruction of an event (both PC and ER) is done by a software application called Elf. In contrast to the trigger algorithms, this software uses the full event data to reconstruct tracks in the DCH and SVT and clusters in the EMC and IFR. Elf also creates lists of different particles by running particle identification (PID) algorithms on the reconstructed tracks (see Section 4 for a full description of the algorithms used in the analysis). Finally, Elf fills a set of Boolean variables called *tags* which provide a way of quickly classifying events based on very general characteristics. Background filter and trigger information are also stored as tags. The charged-track lists, neutral particle lists, PID lists, and tag variables are written out to files called event collections, which are made available for further processing and analysts' use. Simulated data are also reconstructed with Elf, but are not run through the prompt reconstruction system. The event generator software is bundled with the detector simulation software and with Elf to form one integrated production package which directly outputs event collections.

## 6.2 Data Skimming

Most physics measurements made with the *BABAR* data involve only a specific type of event. Often these events constitute only a small fraction of the total data set. To facilitate the many *BABAR* analysts, one final step of centralized data processing takes place before the typical user sees the data. Once a run is processed by the prompt reconstruction system, the output collections are *skimmed*. A *skim* refers to a subset of reconstructed events which fulfill some basic criteria. Groups of physicists working with similar analyses define a skim by choosing a simple set of criteria that selects an acceptably large fraction of the events of interest. *BABAR* analysts have defined hundreds of skims over the years, and some number of these skims are chosen to be calculated for the data and MC events. Using a large farm of computers at SLAC, each event in an event collection is processed and assigned a

true or false value for each skim being run. A *deep-copy skim* is a physical copy of the reconstructed data for each event that passes a particular skim. A *pointer skim* is a collection of pointers to the data for event that passes the skim. Pointer skims are much smaller, but the redundant data of deep-copy skims provides better computing performance with large numbers of users.

The  $\tau \rightarrow \ell\ell\ell$  analysis only uses data and MC events which pass the **Tau1N** skim. This skim selects events for which the following criteria are true (see Appendix A for track and neutral list definitions):

- Event passes either DCH L3 trigger or EMC L3 trigger (always true for data, not necessarily true for MC).
- Event passes one or more of the following background filters: BGFMultihadron, BGFNeutralHadron, BGFTau, BGFMuMU, BGFTwoProng. Analysis will later require BGFMultiHadron (see Section 3).
- The number of entries in the ChargedTracks list is less than eleven.
- The thrust is defined as the vector which minimizes the transverse momentum for all entries in the ChargedTracks and CalorClusterNeutral lists. The thrust axis is used to divide the event into hemispheres in the CM frame. The number of EMC clusters with energy greater than 50 MeV in each hemisphere must be less than or equal to six.
- Using tracks from the GoodTracksVeryLoose list, one hemisphere must contain one track, while the other must contain at least three.

## CHAPTER III

### DATA ANALYSIS

#### 1 Introduction to Analysis

Measurements made with the *BABAR* data can be classified by the expected statistical significance of the signal. *BABAR* is generally considered a precision experiment, meaning that one expects to find many events corresponding to the measurement that one wishes to make. In some cases, such as the measurement of a branching fraction, it is simply the number of these signal events which one wishes to measure. In other cases, the measurement is derived from information contained in the signal events. In either case, the large number of these events allows the physicist to make a measurement with relatively little statistical uncertainty. Many *BABAR* analyses are of this type, where systematic uncertainties dominate the total uncertainty.

The  $\tau \rightarrow \ell\ell\ell$  search is one of a complimentary sort of analyses in which the statistical significance of the expected number of signal events is small. In many cases the observed number of signal events is small enough that the experiment can be said to have a null result. In this case, rather than making a measurement, a limit is placed on the quantity in question. For the  $\tau \rightarrow \ell\ell\ell$  searches, the goal is to place an upper limit on the branching fractions.

A few *BABAR* analyses are of a hybrid sort. These measurements yield statistically insignificant numbers of signal events based on data from the early years of running, but are expected to result in statistically significant measurements with most or all of the final *BABAR* data set. The search for the leptonic B meson decay  $B \rightarrow \tau\nu_\tau$  is one of these experiments. The  $\tau \rightarrow \ell\ell\ell$  processes are not predicted by the SM, but they are predicted at varying levels by untested extensions of the SM. Therefore, the search must be conducted in such a way as to naturally

incorporate the observation of a statistically significant number of signal events. In other words, the analysis should not be biased toward a null result, and the choice to quote a measurement of the branching fraction or an upper limit on the branching fraction must be well-justified.

## 1.1 Branching Fractions

Even though we expect not to measure a branching fraction for  $\tau \rightarrow \ell\ell\ell$ , the essential elements of a branching fraction must be measured before an upper limit can be calculated. Theoretically, the branching fraction for  $\tau \rightarrow X$  is defined by the ratio of the  $\tau$  partial width for final state  $X$  to the total width,

$$\mathcal{B}(\tau \rightarrow X) = \frac{\Gamma_X}{\Gamma_{tot}}. \quad (\text{III.1})$$

A measurement of the branching fraction is given by

$$\mathcal{B}(\tau \rightarrow X) = \frac{N_X}{N_\tau}, \quad (\text{III.2})$$

where  $N_X$  is the measured number of  $\tau \rightarrow X$  decays and  $N_\tau$  is the total number of  $\tau$  decays.

From Equation III.2, we can now see the quantities necessary for a measurement of the branching fraction. The total number of  $\tau$  decays can be calculated from the cross section  $\sigma(e^+e^- \rightarrow \tau^+\tau^-)$  and the time-integrated luminosity  $\mathcal{L}$  of the data set. The number of decays  $\tau \rightarrow X$  is estimated by the ratio

$$N_X = \frac{N_{meas}}{\varepsilon} = \frac{N_{obs} - N_{bkg}}{\varepsilon}, \quad (\text{III.3})$$

where the number of decays measured  $N_{meas}$  is given by the number of decays observed  $N_{obs}$  minus the number of background events  $N_{bkg}$ . A background event is an  $e^+e^-$  collision with decay products that look like  $\tau \rightarrow X$ , but in fact are a different process. The quantity  $\varepsilon$  is an estimate of the signal efficiency, the probability that a real  $\tau \rightarrow X$  decay is identified as such. This quantity incorporates detector, trigger, and software reconstruction inefficiencies, as well as the effects of event selection criteria in the analysis itself. In summary, five quantities must be measured or estimated to calculate the branching fraction: the number of events observed ( $N_{obs}$ ), the number of expected background events  $N_{bkg}$ , the signal

efficiency ( $\varepsilon$ ), the luminosity  $\mathcal{L}$ , and the cross section  $\sigma_{\tau^+\tau^-}$ . Uncertainties must be estimated as well for all quantities (except  $N_{obs}$ ).

With the five quantities previously mentioned (and the associated errors), one can calculate the branching fraction for  $\tau \rightarrow X$ . Because the decays  $\tau \rightarrow \ell\ell\ell$  have never been observed, we approach the analysis with the expectation of a null result, where  $N_{obs}$  is statistically consistent with a Poisson fluctuation around  $N_{bgd}$ . Therefore, we must define a procedure to calculate an upper limit on the branching fraction.

## 1.2 Upper Limits

Before we discuss upper limits, we must first introduce some basic tools for describing the uncertainty associated with a measurement. These methods could be applied to statistically significant measurements, as well as upper or lower limits on unmeasured quantities. Much of the material in this section can be found in greater detail in Cowan [60].

### Confidence Intervals for a Continuous Variable

In the classical (or frequentist) interpretation of statistics, the uncertainty of a measurement can be expressed through the construction of Neyman confidence intervals [61]. In this method, one defines the estimator  $\hat{\theta}$ , which is the outcome (or a function of the outcome) of the experiment, and an estimate of the true value  $\theta$ . Furthermore, one knows the probability density function  $G(\hat{\theta}; \theta)$  as a function of  $\hat{\theta}$ , with the true (but unknown) value  $\theta$  as a parameter. In many practical cases, this PDF is a Gaussian function or a Poisson function. From this PDF, one can define a value  $u_\alpha$  such that there is a fixed probability  $\alpha$  to observe  $\hat{\theta} \geq u_\alpha$ . Similarly, one can define a value  $u_\beta$  such that there is a fixed probability  $\beta$  to observe  $\hat{\theta} \leq u_\beta$ . Thus,  $u_\alpha$  and  $u_\beta$  can be determined from

$$\alpha = P(\hat{\theta} \geq u_\alpha(\theta)) = \int_{u_\alpha(\theta)}^{\infty} G(\hat{\theta}; \theta) d\hat{\theta}, \quad (\text{III.4})$$

$$\beta = P(\hat{\theta} \leq u_\beta(\theta)) = \int_{-\infty}^{u_\beta(\theta)} G(\hat{\theta}; \theta) d\hat{\theta}. \quad (\text{III.5})$$

The functions  $u_\alpha(\theta)$  and  $u_\beta(\theta)$  can be plotted in the  $\theta/\hat{\theta}$  plane and the bounded region is called the *confidence belt*. For a true value  $\theta$  and PDF  $G(\hat{\theta}; \theta)$ , the probability for the experiment to yield an estimator  $\hat{\theta}$  that is in the interval  $[u_\alpha(\theta), u_\beta(\theta)]$  is  $1 - \beta - \alpha$ .

A line corresponding to a given measurement  $\theta_{obs}$  intersects the curves  $u_\alpha$  and  $u_\beta$ . Call the corresponding values of  $\theta$   $a$  and  $b$ , respectively. The interval  $[a, b]$  is called the *confidence interval* at a *confidence level*  $1 - \beta - \alpha$ . If the experiment were repeated  $n$  times, yielding  $n$  values of  $\hat{\theta}$  and  $n$  confidence intervals, the resulting intervals would include the true value  $\theta$  in a fraction  $1 - \beta - \alpha$  of the experiments.

For a specific value of  $\hat{\theta}$  that leads to an interval  $[a, b]$ ,

$$\hat{\theta}_{obs} = u_\alpha(a) = u_\beta(b). \quad (\text{III.6})$$

Equations III.4 and III.5 then become

$$\alpha = \int_{\hat{\theta}_{obs}}^{\infty} G(\hat{\theta}; a) d\hat{\theta}, \quad (\text{III.7})$$

$$\beta = \int_{-\infty}^{\hat{\theta}_{obs}} G(\hat{\theta}; b) d\hat{\theta}. \quad (\text{III.8})$$

In addition to the two-sided interval  $[a, b]$ , the values  $a$  and  $b$  alone correspond to one-sided confidence intervals, or upper and lower limits. The value  $a$  is the hypothetical value of the unknown parameter  $\theta$  for which a fraction  $\alpha$  of repeated estimates for  $\hat{\theta}$  would be *higher* than the current value  $\hat{\theta}_{obs}$ . Similarly, the value  $b$  is the hypothetical value of  $\theta$  for which a fraction  $\beta$  of repeated estimates for  $\hat{\theta}$  would be *lower* than the current value  $\hat{\theta}_{obs}$ . Therefore,  $b$  represents an upper limit on  $\theta$  given the observation  $\hat{\theta}_{obs}$  at the  $1 - \beta$  confidence level. In many high-energy physics experiments,  $\beta$  is chosen to be 0.1, so that upper limits are reported at 90% confidence levels.

A physicist must choose to report a one-sided or two-sided confidence interval, but this choice can introduce some complications. As noted by Feldman and Cousins [62], it is undesirable for this choice to be made on the basis of the outcome of the experiment. For example, a physicist may decide to quote a two-sided interval if the number of signal events observed is greater than some number, and otherwise quote a limit (or one-sided interval). Under closer scrutiny, this procedure can lead to intervals which do not satisfy Equations III.4, III.5, III.7, or III.8. The

procedure also fails by producing empty confidence intervals for certain values of  $\hat{\theta}$ . This failure often occurs in counting experiments for which there is an expected background such that  $\hat{\theta}_{obs} = N_{obs} - N_{bgd}$ . The Feldman and Cousins prescription for the construction of confidence intervals involves an ordering principle by which individual values of  $\hat{\theta}$  are added to the interval until the confidence level reaches or exceeds the desired value. Such intervals naturally shift from one- to two-sided intervals as the observed number of events increases. For the case where the observed number of events is significantly lower than the background, these intervals also have the desirable behavior of remaining non-zero. Feldman and Cousins also suggest that the experimenters should quote the *sensitivity*, or the average upper limit that would be obtained for an ensemble of experiments with the quoted background and no true signal. This quantity is discussed in more detail in Section 1.5. In the case of the  $\tau \rightarrow \ell\ell\ell$  analysis, the values for  $N_{obs}$  are consistent with the background expectations. Therefore, we make use of the sensitivity to give a sense of the most probable outcome of the experiment, but the upper limits can be accurately calculated without constructing the full Feldman and Cousins intervals.

## Confidence Intervals for Discrete Variables

When searching for rare or unobserved processes such as  $\tau \rightarrow \ell\ell\ell$ , the outcome of the analysis is an observation of a small number of events which pass all selection criteria. For any rare reaction which leads to a small number of expected events  $N_{exp}$ , the number of observed events  $N_{obs}$  is sampled from a Poisson distribution in the variable  $n$  with mean  $\nu = N_{exp}$ ,

$$P(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}. \quad (\text{III.9})$$

Next, we rewrite Equations III.7, III.8 for the Poisson variable  $n = \hat{\theta}$ , with  $N_{obs} = \hat{\theta}_{obs}$  and  $\nu = \theta$ ,

$$\alpha = P(n \geq N_{obs}; a) = \sum_{n=N_{obs}}^{n=\infty} \frac{a^n}{n!} e^{-a}, \quad (\text{III.10})$$

$$\beta = P(n \leq N_{obs}; b) = \sum_{n=0}^{n=N_{obs}} \frac{b^n}{n!} e^{-b}. \quad (\text{III.11})$$

The Poisson distribution is a function of the discrete variable  $n$ , so the integrals in Equations III.7 and III.8 have been replaced by sums in Equations III.10 and III.11. With Equation III.11 and the outcome of the experiment  $N_{obs}$ , one can calculate an upper limit  $b$  at the  $1 - \beta$  confidence level.

Now that we have defined the upper limit on the number of signal events observed in terms of a one-sided classical confidence interval, it is important to clarify what this upper limit actually means. The upper limit at the  $1 - \beta$  confidence level on the number of events observed  $s^{CL} = b$  is the hypothetical value of the number of signal events for which the probability of observing  $N_{obs}$  events or less is  $\beta$ . This is an important point: the calculated upper limit is nothing more than the hypothetical value of the true signal strength for which the experimental observation is unlikely to a particular degree.

### 1.3 Incorporating Uncertainties into Upper Limits

All measured quantities come with associated uncertainties and it is important to take these into account in a consistent manner. As listed in Section 1.1, there are five quantities needed for the calculation of a branching fraction or upper limit, and four of these quantities can have significant uncertainties. It is helpful to define a quantity known as the *sensitivity*  $S = 2\epsilon\mathcal{L}\sigma_{\tau\tau}$ . The uncertainties on the cross section, the luminosity, and the signal efficiency can be combined into the uncertainty on the sensitivity, effectively leaving two uncertainties in the problem. We'll start with the simplest case of a 90% confidence level upper limit on the signal observed, with no background and no uncertainty. We will then generalize to the case of an upper limit on the branching fraction, with uncertainties on all measured quantities.

In the absence of background events,

$$0.1 = \sum_{n=0}^{N_{obs}} P(n; s^{90}), \quad (\text{III.12})$$

defines the upper limit on the number of signal events observed ( $s^{90}$ ) at the 90% confidence level, where  $N^{obs}$  is the number of events observed in the experiment, and  $P(n; s)$  is the Poisson function of  $n$  with expectation value  $s$ . Equation (III.12) is typically solved numerically.

To add an expected background, we introduce the constant background parameter  $b$ , which must be estimated from the experiment. Now the outcome of the experiment is assumed to be drawn from a Poisson distribution with mean  $s + b$ . Then, the upper limit on the number of signal events observed ( $s^{90}$ ) is given by

$$0.1 = \sum_{n=0}^{N_{obs}} P(n; s^{90}, b) = \sum_{n=0}^{N_{obs}} \frac{(s^{90} + b)^n}{n!} \exp[-(s^{90} + b)]. \quad (\text{III.13})$$

To include the effect of an uncertainty on the number of expected background events, we follow the Cousins and Highland [63] procedure and “smear” the expected background over a range about the central value. The width of the smearing is given by the uncertainty on the background. Under the assumption of a Gaussian distribution for the background, the upper limit on the number of signal events is given by

$$0.1 = \sum_{n=0}^{N_{obs}} \int_{\alpha}^{\beta} G(b'; b, \sigma_b) P(n; s^{90}, b') db', \quad (\text{III.14})$$

where  $G(b'; b, \sigma_b)$  is a Gaussian function of  $b'$  with mean  $b$  and width  $\sigma_b$ . The limits of integration  $\beta$  and  $\alpha$  can be set to  $+\infty$  and  $-\infty$ , respectively, in the case of an analytical solution, or to the desired precision in the case of a numerical solution.

When the upper limit is to be set on a branching fraction, uncertainties in the signal efficiency  $\varepsilon$  can be accounted for in a similar way [64]. Recall that the branching fraction  $\mathcal{B}$  is the ratio of the number of signal events to the total number of events,

$$\mathcal{B} = \frac{s}{N_{tot}} = \frac{\mu - b}{S}, \quad (\text{III.15})$$

where  $S = \varepsilon N_{tot} = 2\varepsilon \mathcal{L}\sigma$ , and  $\mu = S\mathcal{B} + b$  is the mean number of events expected to be observed for a signal with branching fraction  $\mathcal{B}$ , sensitivity  $S$ , and background  $b$ . Thus, in the case of significant uncertainty on both the expected background and on the sensitivity to the signal, the upper limit on the branching fraction ( $\mathcal{B}^{90}$ ) is defined by the following:

$$0.1 = \sum_{n=0}^{N_{obs}} \int \int G(S'; S, \sigma_S) G(b'; b, \sigma_b) P(n; b', S, \mathcal{B}^{90}) db' dS', \quad (\text{III.16})$$

where

$$P(n; b, S, \mathcal{B}^{90}) = (S\mathcal{B}^{90} + b)^n \frac{\exp[-(S\mathcal{B}^{90} + b)]}{n!} \quad (\text{III.17})$$

As mentioned at the beginning of this section, the total uncertainty on  $S$  may need to take into account the uncertainty on the total number of events  $N_{tot} = 2\mathcal{L}\sigma$ , as well as the uncertainty on the signal efficiency.

A numerical solution to equation (III.16) is implemented in the TCousinsHighland C++ class. The basic strategy is to find the zero of the function

$$f(\mathcal{B}) = \langle \text{Prob}(N_{obs}, \mathcal{B}) \rangle - 0.1, \quad (\text{III.18})$$

where  $\langle \text{Prob}(N_{obs}, \mathcal{B}) \rangle$  is the *average* Poisson probability as defined below. The average is taken over a distribution of values for  $\text{Prob}(N_{obs}, \mathcal{B})$ , where each value is calculated with a sample background count  $b$  and a sample sensitivity  $S$ , both drawn from Gaussian distributions. Here,  $S$  is defined as before and is proportional to the signal efficiency.

The root of equation (III.18) is estimated with the Bisection method, in which two values of  $f$  ( $f_0, f_1$ ) are varied until they differ from zero by no more than the desired precision, and the conditions  $f_0 < 0$  and  $f_1 > 0$  remain true. Values for  $f$  are calculated from values for  $\mathcal{B}$  via the following method:

1. Chose a value for  $\mathcal{B}$ .
2. Draw a value for the background  $b_i$  from a Gaussian distribution with mean  $b$  and width  $\sigma_b$ .
3. Draw a value for the sensitivity  $S_i$  from a Gaussian distribution with mean  $S$  and width  $\sigma_S$ .
4. Calculate the expected number of events  $\mu_i$  for this point:  $\mu_i = \mathcal{B}S_i + b_i$ .
5. Calculate the  $i$ -th Poisson probability  $\text{Prob}_i(N_{obs}, \mu_i)$  for  $n \leq N_{obs}$ :

$$\text{Prob}_i(N_{obs}, \mu_i) = \sum_{n=0}^{N_{obs}} P(n, \mu_i) \quad (\text{III.19})$$

6. Repeat steps 2-5  $j$  times.

7. The Poisson probability for  $\mathcal{B}$  is the average over the values  $\text{Prob}_i(n, \mu_i)$ , and  $f$  is given by

$$f(\mathcal{B}) = \frac{1}{j} \sum_{i=0}^{j-1} \text{Prob}_i(N_{obs}, \mu_i) - 0.1 \quad (\text{III.20})$$

The value for the upper limit is given by

$$\mathcal{B}^{90} = \frac{\mathcal{B}_0 - \mathcal{B}_1}{2}, \quad (\text{III.21})$$

where  $f_0(\mathcal{B}_0)$  and  $f_1(\mathcal{B}_1)$  fulfill the conditions for the Bisection method listed above.

## 1.4 Overview of Analysis Steps

In the previous section, we discussed the ingredients necessary for placing upper limits on  $\mathcal{B}(\tau \rightarrow \ell\ell\ell)$ . Now, we outline the major steps in the analysis which lead to these final ingredients. These steps are nothing more than a very carefully chosen set of selection criteria by which the set of all the events in the *BABAR* dataset is filtered down to a few final events. Each of the six  $\tau \rightarrow \ell\ell\ell$  searches employs a different set of selection criteria, although the variables used are generally the same. Data events, as well as signal and background MC events, are all run through the same selection procedure. Unqualified references to “events” should be assumed to refer to both data and MC events.

In the first step, we select events which pass a very broad selection called a *skim* (see Section 6.2). The selected events are then required to pass a set of *preselection cuts*, which reject poorly reconstructed events and other events which look very little like the  $\tau \rightarrow \ell\ell\ell$  signal. We next ensure that the preselected events contain the three leptons appropriate to the particular search channel. We define two important variables, a mass variable and an energy variable, which provide some of the most precise separation of signal and background events. Rather than immediately using cuts in these variables to reject background events, the distributions of the MC and data events in the plane of the mass and energy variables is used to estimate the final background contribution in the region where the signal is expected. Lastly, a final set of selection criteria is applied to further reduce backgrounds. These criteria are chosen separately for each search channel, and tuned to address channel-specific backgrounds.

A signal region which contains most of the signal MC events is defined in terms of the mass and energy variables. This analysis is conducted in a *blinded* fashion, meaning that the number of data events in this region is left unknown until all selection criteria are fixed and all systematic uncertainties are studied. This technique avoids bias by ensuring that the selection criteria are not tuned to a statistical fluctuation in the number of data events in the signal region. The distribution of the remaining background MC events in the mass and energy plane provides the final estimate of the background contribution in the signal region. Systematic uncertainties are studied and errors are assigned to the background estimate and the signal efficiency. Finally, the methods described in Section 1.3 are used to calculate the upper limits on the six  $\tau \rightarrow \ell\ell\ell$  branching fractions.

## 1.5 Analysis Optimization and Expected Upper Limits

The choice of selection criteria in this analysis should be based on an optimization of the result. One would typically choose the upper limit for this figure of merit, and optimize for the lowest limit. Because we are blind to the number of data events in the signal region  $N_{obs}$ , we need to optimize some other quantity which does not depend on  $N_{obs}$ . We choose to optimize the analysis to produce the lowest *expected* upper limit, as suggested by Feldman and Cousins [62]. This expected upper limit is defined as the mean upper limit expected in the background-only hypothesis for a given sensitivity  $S = 2\varepsilon\mathcal{L}\sigma_{\tau\tau}$  and background contribution  $N_{bgd}$ . This expected upper limit on the branching fraction is calculated as

$$\mathcal{B}_{exp}^{90} = \sum_{n=0}^{\infty} P(n; N_{bgd}) \mathcal{B}^{90}(n, N_{bgd}, S), \quad (\text{III.22})$$

where  $\mathcal{B}^{90}(n, N_{bgd})$  is the upper limit on the branching fraction based on the observation of  $n$  events with background contribution  $N_{bgd}$  and sensitivity  $S$ , and  $P(n; N_{bgd})$  is the probability of observing  $n$  events from a Poisson distribution with mean  $N_{bgd}$ . The upper limit  $\mathcal{B}$  is calculated by the method described in Section 1.3, which incorporates all uncertainties.

## 2 Selection of the Data

This analysis is performed using data recorded from June 1999 through August 2006. The *BABAR* Collaboration divides the data-taking period of the experiment into *Run Cycles*, or simply *Runs*. The data used in the analysis comprise the full dataset for Runs 1-5, with a total luminosity of  $376 \text{ fb}^{-1}$ . These data include  $339 \text{ fb}^{-1}$  recorded at the  $\Upsilon(4S)$  resonance with a CM system energy of 10.58 GeV. The rest of the data,  $36.6 \text{ fb}^{-1}$ , were recorded off-resonance at a CM energy of 10.54 GeV. To speed up the data processing, only data included in the **Tau1N** skim (described in Section 6.2) were used for the initial data and Monte Carlo samples.

The signal MC explicitly includes one (*signal*) tau lepton decaying to three lighter leptons, while the second tau decays according to the standard (*generic*) tau decay tables. These decay tables include the latest values of the tau branching fractions from the Particle Data Group [65]. The signal sample is divided into 6 subsamples according the LFV mode (inclusion of charge conjugates is implied):

- $\tau^- \rightarrow e^- e^+ e^-$ ,
- $\tau^- \rightarrow e^- \mu^+ \mu^-$ ,
- $\tau^- \rightarrow e^+ \mu^- \mu^-$ ,
- $\tau^- \rightarrow \mu^+ e^- e^-$ ,
- $\tau^- \rightarrow \mu^- e^+ e^-$ ,
- $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ .

Each subsample contains a total of 286k events with equal numbers of  $\tau^+$  and  $\tau^-$  LFV decays. The signal events are generated with the **KK2f** generator [50] which simulates the initial state radiation and final state photon production. The LFV decays in the signal modes are produced using a flat phase-space distribution in the decay products, while the generic tau decays are simulated with **TAUOLA**[55]. Radiation from the final-state leptons has been simulated with **PHOTOS**[57].

Background estimations are made using large MC samples which simulate the types of background events expected to be seen in the analysis. These backgrounds can be grouped into three broad classes:  $b\bar{b}$ ,  $c\bar{c}$ ,  $u\bar{u}/d\bar{d}/s\bar{s}$  ( $q\bar{q}$  background); Bhabha,  $\mu^+\mu^-$  (QED background); and generic  $\tau^+\tau^-$  events with no LFV decays ( $\tau^+\tau^-$  background).

Due to an insufficient quantity of Bhabha and  $\mu^+\mu^-$  MC events, these MC samples have been ignored and the QED contribution is estimated with data samples. Exact process names, MC statistics used and cross sections assumed for the processes are given in Table III.1. The cross sections used are taken from [66] except for  $\tau^+\tau^-$ , which is calculated with **KK2f** [67] As described in Section 7, the overall background normalization for each background type is determined from the data, so the absolute cross sections are not actually used in this analysis. Generic  $\tau^+\tau^-$  events have been generated with **KK2f**.

Table III.1: Background MC samples used in the analysis.

Sample	MC Process Name	$\sigma$ (nb)	$N_{events}^{MC}, 10^6$	$\mathcal{L}_{MC}/\mathcal{L}_{data}$
$b\bar{b}$	half <b>B+B-</b> , half <b>B0B0bar</b>	1.05	1025	2.60
$c\bar{c}$	<b>e+e- <math>\rightarrow</math> ccbar</b>	1.30	275.2	0.56
$uds$	<b>e+e- <math>\rightarrow</math> uubar/ddbar/ssbar</b>	2.09	398.8	0.51
$\tau^+\tau^-$ -bkgr	<b>e+e- <math>\rightarrow</math> tau+ tau- (KK2f)</b>	0.89	184.4	0.55

### 3 Event Preselection

The LFV tau decay to three leptons produces three charged track. To reduce the background contribution of high multiplicity  $q\bar{q}$  events, we require the other tau in the event to decay to one charged track. Therefore, the first step of the analysis is to select events with a 1-3 topology that is characteristic of the signal tau events.

The Tau1N skim is used for all data and MC samples. The criteria for the Tau1N selection are described in Chapter II, Section 6.2. The further preselection requirements are listed below. Efficiencies for each cut are shown in Table III.2.

- Event has **BGFMultiHadron** filter bit set.

This bit is set true for events with more than two tracks in the **ChargedTracks** (see Appendix 2) list and with  $R2 < 0.98$ .  $R2$  is the ratio of the 2nd to the 0th Fox-Wolfram moment.

- Exactly 4 ‘good tracks’ are required in the event.

For this analysis, we acquire our *good tracks* from the **GoodTracksVeryLoose** list (see Appendix 3). These good tracks are additionally required to point to the collision region ( $docaXY < 1$  cm,

do  $caZ < 5$  cm) and have a momentum in the range  $p_T > 0.1$  GeV/c,  $p < 10$  GeV/c in the Lab frame. Good tracks must have value of the polar angle  $\theta$  which allows for good particle identification ( $0.41 < \theta < 2.46$ , driven by the range of the PidTables for lepton selectors, and by the EMC acceptance). The tracks identified as a part of a converted photon candidate (found in **gammaConversionDefault** list, described in Appendix 4) are not counted as good tracks. No attempt has been made to reconstruct  $K_S$  decays.

- The total charge of the good tracks in the event is equal to 0.
- The event has a ‘reconstructed 1-3 topology’.

The event is divided on two hemispheres using the plane perpendicular to the thrust<sup>1</sup> axis of the event. The sign of scalar product of the given track momentum with the thrust direction determines the hemisphere to which this track belongs. The thrust of the event is calculated using charged and neutral (with energy greater than 50 MeV) particle candidates in the CM frame. One hemisphere must have exactly one good track, while other 3 must belong to the second hemisphere. This defines a *reconstructed 1-3 topology*.

## 4 Particle Identification

After events with a 1-3 topology have been selected, particle identification (PID) criteria are applied to the tracks in the 3-prong hemisphere. Except for a few cases to be addressed in Section 6, tracks and neutral clusters in the 1-prong hemisphere are not subject to particle identification. Like the pre-selection criteria described in Section 3 and the more specific selection cuts described in Section 6, the particle identification step is designed to reject background events, while maintaining a high efficiency for signal events.

To identify tracks and neutrals as specific types of long-lived particles, analysts at *BABAR* have developed a set of algorithms called PID *selectors*. These selectors take input information from many components of the *BABAR* detector. Typical inputs are  $dE/dx$  energy loss in the drift chamber, energy loss and shower shape in the calorimeter, and hits in the IFR. The output of a selector which is run on a particular track or neutral cluster is always a *true* or *false* signal. Appendix B lists

---

<sup>1</sup>The thrust axis is defined as the axis which minimizes the transverse momentum in the event.

Table III.2: Preselection efficiencies in percent for signal MC, background MC, and data samples. Cuts are applied sequentially and the marginal efficiencies are quoted. For the signal samples, the loss in efficiency due to the one-prong branching fraction is *included* in these numbers. ‘Trigger’ means that **L3OutDch** or **L3OutEmc** tagbit is set. The  $b\bar{b}$  efficiencies include both  $B^0\bar{B}^0$  and  $B^+B^-$  samples. Uncertainties on the total efficiency numbers are from MC statistics.

Sample	Tau1N	BGFMH	4 tracks	Zero Charge	1-3 topology	Total
Signal MC						
EEE	43.0	99.5	89.6	98.5	99.5	$37.6 \pm 0.1$
EEMr	42.0	99.5	89.9	98.8	99.5	$38.6 \pm 0.1$
EEMw	44.1	99.5	89.7	98.7	99.5	$38.7 \pm 0.1$
EMMr	45.6	99.5	91.1	98.9	99.5	$40.7 \pm 0.1$
EMMw	45.6	99.6	90.9	99.0	99.5	$40.6 \pm 0.1$
MMM	47.2	99.7	92.3	99.0	99.6	$42.7 \pm 0.1$
Background MC						
$b\bar{b}$	0.41	98.5	39.2	65.1	83.7	0.18
$c\bar{c}$	4.49	97.8	46.0	72.6	89.1	1.28
$uds$	6.06	97.9	52.2	80.2	90.7	2.11
$\tau^+\tau^-$	16.0	95.1	77.1	98.4	99.5	11.5
Run 1-5 Data						
DATA On-peak	3.63	93.5	51.9	85.5	93.3	1.51
DATA Off-peak	4.13	93.8	52.4	86.1	93.4	1.72

the specific criteria for the selectors used in this analysis. No modifications have been made to these standard *BABAR* particle identification algorithms.

Information about particle identification is stored in lists called PID *lists*. The PID list is a list of all the tracks or neutrals in the event which meet the criteria for a particular PID selector. In this analysis, we do not directly identify any neutral particles, so all particles mentioned in this section will be charged tracks. A given track is said to be identified as a particular particle type when the track is included in the list for the selector of that particle type. The PID lists are not exclusive, and a track which meets the criteria for more than one particle type will appear in both lists.

Analysts often need flexibility and control over the certainty of the identification of a given track or neutral cluster. Therefore, multiple lists are generated for each particle type, with each list corresponding to a different level of certainty. “Tighter” selectors have lower efficiencies to identify a particle of the correct type. They also have lower probabilities to incorrectly select a particle of a different type. The selector names generally reflect three properties of the selector: the particle being selected, the type of algorithm used, and the tightness or looseness of the selector. For instance, this analysis make use of the **muNNLoose** selector, which selects muons based on a neural network (NN) algorithm, using a loose selection which has a relatively high efficiency to identify real muons. Because different  $\tau \rightarrow \ell\ell\ell$  search channels are populated by different background types, it is helpful to apply looser particle identification criteria to some search channels, and tighter criteria to others. The choice of PID selector is that which, when applied along with a set of nominal selection cuts (Section 6), provides the best expected upper limit on the branching fraction (see Section 1). The PID selectors used for the analysis are listed in Table III.3.

Table III.3: Particle ID selectors used to identify the 3-prong tracks.

Search channel	Electron selector	Muon selector
$e^-e^+e^-$	<b>eMicroTight</b>	N/A
$\mu^-e^+e^-$	<b>eMicroTight</b>	<b>muNNLoose</b>
$e^-\mu^+e^-$	<b>eMicroTight</b>	<b>muNNLoose</b>
$e^-\mu^+\mu^-$	<b>eMicroTight</b>	<b>muNNLoose</b>
$\mu^-e^+\mu^-$	<b>eMicroLoose</b>	<b>muNNLoose</b>
$\mu^-\mu^+\mu^-$	N/A	<b>muNNLoose</b>

The *BABAR* PID group generates PID *tables* which reflect the performance of the PID selectors for tracks with a wide range of kinematic properties. The group starts with high *purity* samples by selecting data events with a very high probability of containing a particular set of particles. For instance, the muon sample comes from the easily-identified process  $e^+e^- \rightarrow \mu^+\mu^-\gamma$ . The group then runs all selectors on these samples and calculates the selection efficiency for each particle type as a function of  $\theta$ ,  $\phi$ , and  $p = |\vec{p}|$ . The PID tables allow one to calculate the efficiency to pass any selector for any particle type with any value for  $\theta$ ,  $\phi$ , and  $p$ . Because selectors for all particles are run over all the samples, these tables include not only efficiencies but also mis-identification rates.

To ensure that the MC samples accurately reproduce the particle identification performance observed in data, most *BABAR* analyses apply a correction factor to compensate for the observed difference between the MC and data PID efficiencies and mis-identification rates. In this analysis, we avoid this correction by ignoring MC PID information all together. Instead, we make use of the fact each MC track was generated as a specific particle type. Each MC track is re-weighted by a PID probability for a particle of its type and values of  $\theta$ ,  $\phi$ , and  $p$ . This PID probability is given by the efficiency or mis-identification probability for *data* as obtained from the PID tables. Take, for example, a MC particle generated as a muon with  $(\theta, \phi, p)^{MC}$ , which is being identified as a loose muon with the **muNNLoose** selector. This MC particle is re-weighted by the efficiency obtained from the entry for  $(\theta, \phi, p)^{MC}$  in the **muNNLoose** PID table, which was created from real muons in data. The original information regarding which MC tracks are in which PID lists is completely ignored. This procedure makes much more efficient use of the available MC statistics by not explicitly rejecting any MC tracks or events. It also avoids the need to correct for the differences between data and MC PID selector efficiencies.

The final MC event weight is given by the product of the MC track probabilities in the 3-prong hemisphere of the event. The 1-prong track PID information does not contribute to the event weight. Data events are accepted or rejected in the traditional manner by requiring that all tracks in the 3-prong hemisphere are found in the appropriate PID lists.

One potential deficiency in such a PID weight scheme for MC is that only tracks with a MC truth match are assigned a PID efficiency. All MC particles are generated with a definite particle type, but the MC truth match, which associates a generated MC track with a reconstructed MC track in the detector simulation, does

not always exist. The remaining tracks without definite particles types (usually pions or ghosts<sup>2</sup>, but could also be leptons) are not assigned a PID efficiency at all. Secondly, PID efficiencies can only be assigned to tracks with parameters that fall within the range of the PID table bins. Thirdly, some low momentum bins in the PID tables have no entries, in which case the low momentum MC track would be assigned zero weight by default. The truth-matching problem affects about 0.3% of the pre-selected *uds* MC tracks, primarily low momentum tracks in the range 100 – 300 MeV/c. The difference between assigning these tracks zero weight and assigning them the average track weight has been studied, and the impact is negligible. The requirements on the polar angle  $\Theta$  in the pre-selection (see Section 3) ensure that all tracks lie within the range of the `PidTables`. The effect of low momentum tracks (for which the corresponding PID table bin is empty) is more significant, as 35.3% of the  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  events have at least one slow muon below 500 MeV/c. These tracks account for the low PID efficiency for channels with muons. The average PID efficiency for muons is 65%, which includes the effect of zero-weight slow muons. Figure III.1 shows the muon efficiency as a function of momentum, over a wide range of polar angles. The average PID efficiency for electrons is 91%, including the small effect of electron tracks which have no truth match. Figure III.2 show the electron identification efficiency. The corresponding electron (muon) mis-identification rate for pions in 3-prong SM  $\tau^+ \tau^-$  decays is 2.7(2.9)%, Figures III.3 and III.4 show the pion fake rates as a function of momentum. The mis-identification rates for kaons in 3-prong *uds* events are 4.6% and 2.3% for electron selection and muon selection, respectively. The kaon fake rates are shown in Figures III.5 and III.6.

As described in Section 7, the Bhabha and di-muon backgrounds are modeled with data control samples. For channels  $\tau^- \rightarrow e^- e^+ e^-$  and  $\tau^- \rightarrow e^- \mu^+ \mu^-$ , the PID efficiency for these samples is the same as that for data. For all other channels, QED control samples of sufficient statistics are obtained through a procedure that does not involve particle identification. The rejection factors for different sources of the background are given in Table III.4.

---

<sup>2</sup>A ghost is a second track reconstructed from the same physical track.

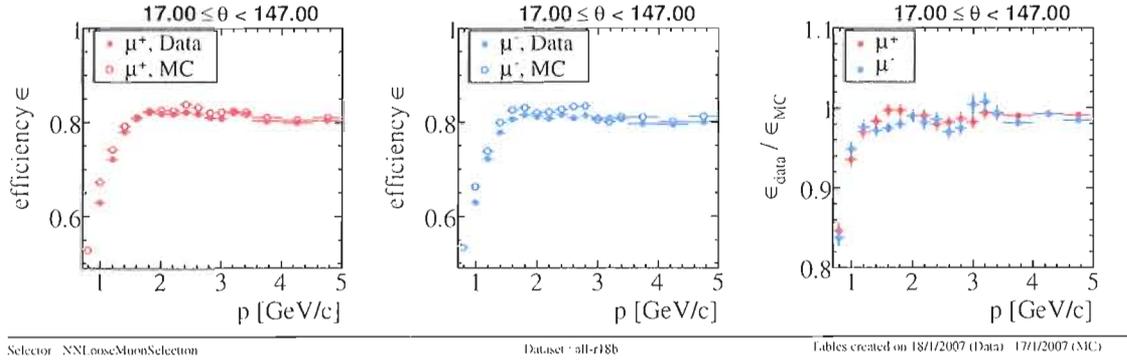


Figure III.1: The efficiency for muon identification in data and MC by the **muNNLoose** selector, as a function of muon momentum for (a) positively charged muons, and (b) negatively charged muons. Plot (c) shows the ratio of the data efficiency to the MC efficiency.

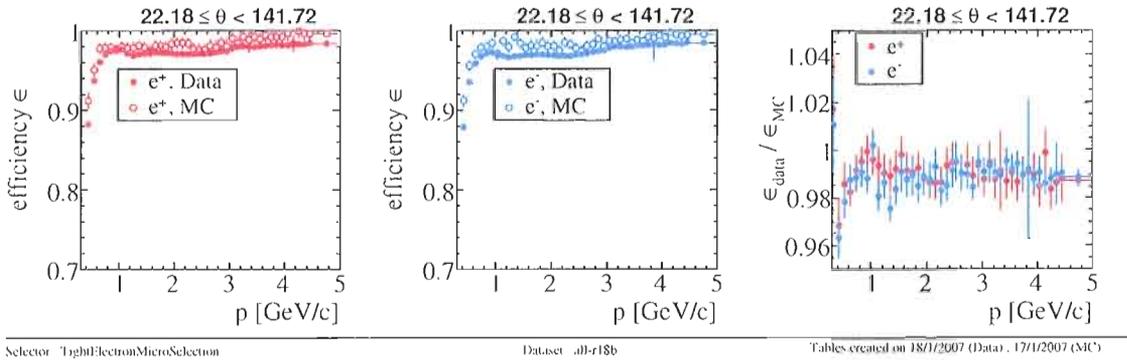


Figure III.2: The efficiency for  $e^+/e^-$  identification in data and MC by the **eMicroTight** selector, as a function of particle momentum for (a) positrons, and (b) electrons. Plot (c) shows the ratio of the data efficiency to the MC efficiency.

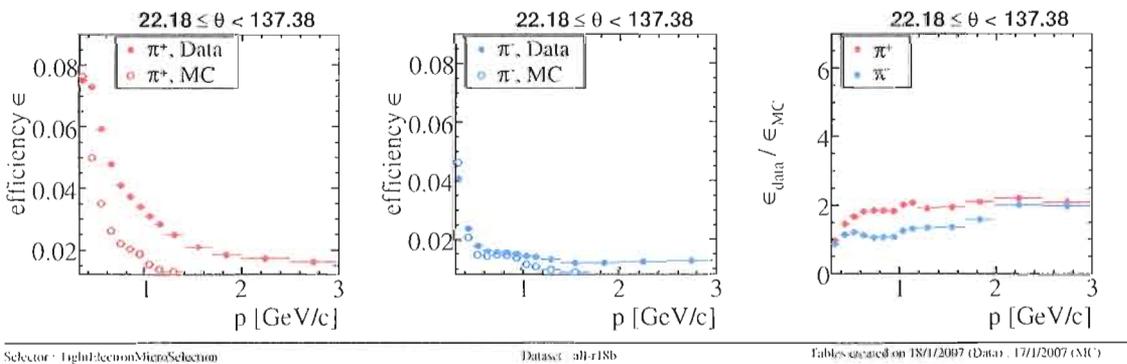


Figure III.3: The mis-ID rate for pions in data and MC by the **eMicroTight** selector, as a function of particle momentum for (a) positively charged pions, and (b) negatively charged pions. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate.

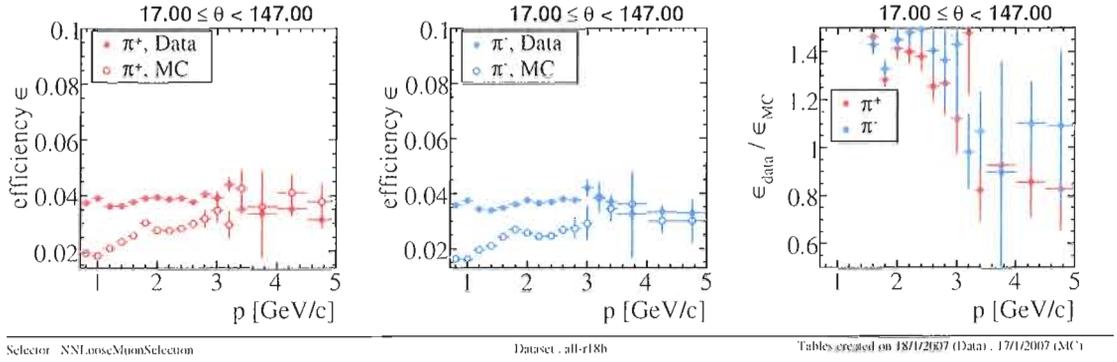


Figure III.4: The mis-ID rate for pions in data and MC by the **muNNLoose** selector, as a function of particle momentum for (a) positively charged pions, and (b) negatively charged pions. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate.

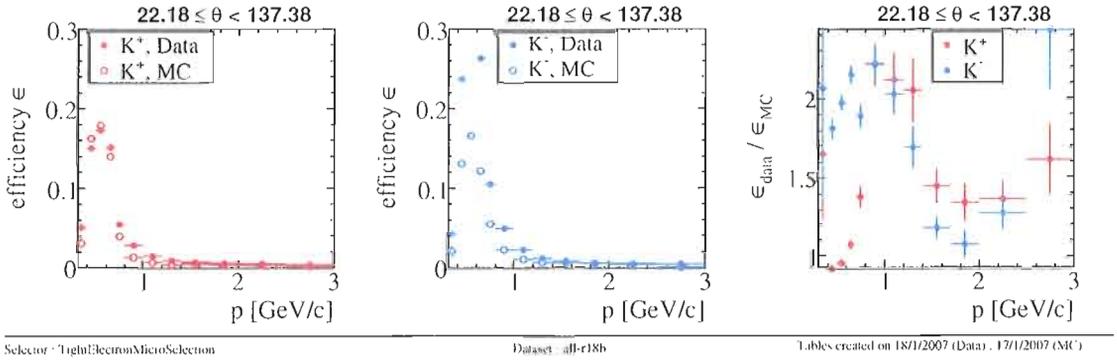


Figure III.5: The mis-ID rate for kaons in data and MC by the **eMicroTight** selector, as a function of particle momentum for (a) positively charged kaons, and (b) negatively charged kaons. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate.

Table III.4: Efficiency for preselected events to pass the PID requirements.

	Signal	$bb$	$c\bar{c}$	$uds$	$\tau^+\tau^-$	DATA
$e^-e^+e^-$	0.775	$9.2 \cdot 10^{-7}$	$6.9 \cdot 10^{-9}$	$1.9 \cdot 10^{-8}$	$6.4 \cdot 10^{-7}$	$9.9 \cdot 10^{-4}$
$\mu^-e^+e^-$	0.531	$1.2 \cdot 10^{-6}$	$9.0 \cdot 10^{-8}$	$3.2 \cdot 10^{-7}$	$5.8 \cdot 10^{-7}$	$5.0 \cdot 10^{-5}$
$e^-\mu^+e^-$	0.533	$2.2 \cdot 10^{-6}$	$2.0 \cdot 10^{-6}$	$1.6 \cdot 10^{-6}$	$1.1 \cdot 10^{-6}$	$2.1 \cdot 10^{-6}$
$\mu^-e^+\mu^-$	0.368	$8.7 \cdot 10^{-7}$	$2.0 \cdot 10^{-6}$	$5.9 \cdot 10^{-6}$	$7.2 \cdot 10^{-6}$	$8.0 \cdot 10^{-6}$
$e^-\mu^+\mu^-$	0.359	$1.7 \cdot 10^{-6}$	$9.6 \cdot 10^{-7}$	$2.5 \cdot 10^{-7}$	$1.4 \cdot 10^{-7}$	$1.4 \cdot 10^{-4}$
$\mu^-\mu^+\mu^-$	0.235	$2.4 \cdot 10^{-6}$	$1.2 \cdot 10^{-6}$	$2.1 \cdot 10^{-6}$	$3.4 \cdot 10^{-6}$	$1.0 \cdot 10^{-5}$

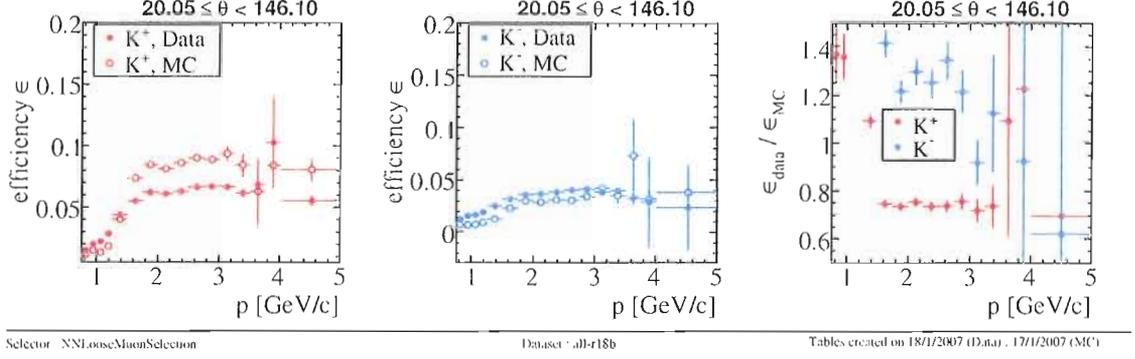


Figure III.6: The mis-ID rate for kaons in data and MC by the **muNNLoose** selector, as a function of particle momentum for (a) positively charged kaons, and (b) negatively charged kaons. Plot (c) shows the ratio of the data mis-ID rate to the MC mis-ID rate.

## 5 Mass and Energy Determination

Since no neutrino is present in the LFV decay mode, the signal events are expected to have the same total energy and invariant mass as the parent tau lepton. The total energy difference

$$\Delta E = E_{rec}^* - E_{beam}^* \quad (\text{III.23})$$

and the invariant mass difference

$$\Delta M = m_{rec} - m_{\tau} \quad (\text{III.24})$$

are calculated from the momentum of the three observed tracks in the 3-prong hemisphere, with the track mass hypotheses corresponding to the search channel. In the study of B meson decays, the energy substituted mass ( $m_{ES} = \sqrt{m_B^2 - |p|^2}$ , where  $m_B$  is the B meson mass) provides better resolution by taking into account the low CM momentum of the B mesons. For  $\tau$  events, the decaying particles are not nearly at rest, and the reconstructed mass  $m_{rec}$  provides better resolution. The energy-constrained mass  $m_{ec}$ , while having slightly better resolution, is not used due to technical difficulties. It is also expected that the use of  $m_{ec}$  would not decrease the expected upper limit by more than 10%, and that only for the channels with expected backgrounds above one event.

The search for LFV decay modes proceeds by considering the two-dimensional distribution in the  $(\Delta M, \Delta E)$  plane where the signal events should peak around the

origin. The quantities  $\Delta E$  and  $\Delta M$  tend to be smeared out somewhat due to tracking resolution and radiative effects from initial-state, final-state, or bremsstrahlung photon emission. As the initial tau energy is unobservable and must be inferred from the beam energy, energy lost to these radiative effects tends to preferentially push events toward lower values of both  $\Delta E$  and  $\Delta M$ .

Since electrons have larger radiative losses than muons, the radiative tail in the  $(\Delta M, \Delta E)$  distribution depends upon the decay channel considered (see Figure III.7). For this reason, the optimal signal region is defined separately for each signal channel. The selected signal region, as well as the borders of the Large Box (LB) used for background studies, are shown in Table III.5. The Grand Sideband region (*GS*) is defined as the large box minus the signal region. The choice of a box for the signal region over something more complicated (like an ellipse) is primarily for technical convenience, as it is easier to perform a 2D integration over a rectangular region. The signal efficiencies to pass SB and LB cuts are given in Table III.6.

Table III.5: Signal region boundaries  $M_1 < \Delta M < M_2$ ,  $E_1 < \Delta E < E_2$  for each decay mode. The boundaries of the large box (*LB*) used in the background fits is also shown in the last column. The last row shows the signal efficiencies in percent for these signal regions (for the events passed preselection and PID requirements).

Sample	$e^-e^+e^-$	$e^-\mu^+e^-$	$\mu^-e^+e^-$	$\mu^-e^+\mu^-$	$e^-\mu^+\mu^-$	$\mu^-\mu^+\mu^-$	<i>LB</i>
$M_1, \text{ GeV}/c^2$	-0.07	-0.10	-0.05	-0.05	-0.05	-0.02	-0.6
$M_2, \text{ GeV}/c^2$	0.02	0.02	0.02	0.02	0.02	0.02	0.4
$E_1, \text{ GeV}/c^2$	-0.20	-0.35	-0.20	-0.20	-0.20	-0.20	-0.7
$E_2, \text{ GeV}/c^2$	0.05	0.05	0.05	0.05	0.05	0.05	0.4

Table III.6: Signal efficiency for events passing preselection and PID to be in the signal box or in the large box.

Sample	SB efficiency [%]	LB efficiency [%]
$e^-e^+e^-$	52.0 %	81.2 %
$\mu^-e^+e^-$	59.5 %	86.6 %
$e^-\mu^+e^-$	69.8 %	86.6 %
$e^-\mu^+\mu^-$	67.3 %	90.4 %
$\mu^-e^+\mu^-$	70.6 %	91.0 %
$\mu^-\mu^+\mu^-$	82.3 %	94.5 %

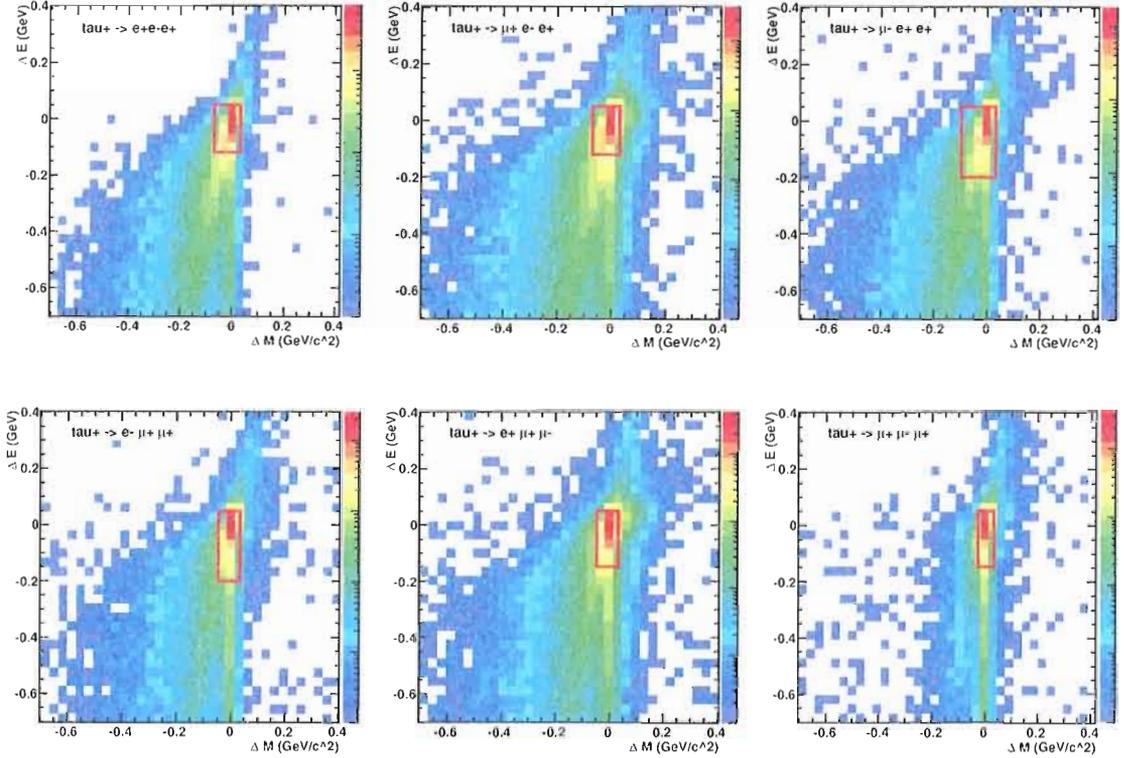


Figure III.7: The  $(\Delta M, \Delta E)$  distributions for the signal channels after preselection and particle identification. The box shows the borders of the signal region. The histogram borders correspond to the large box. The z-axis is logarithmically-scaled.

## 6 Event Selection

After the preselection and particle ID requirements, there is still a significant contribution of background events expected in the signal region. A final set of selection cuts are applied separately for each signal hypothesis to further reduce the remaining background and improve the sensitivity of the analysis.

Since  $\tau \rightarrow \ell\ell\ell$  events have never been observed and the best limits [36, 37] previous to this analysis are of the order of  $10^{-7}$ , it is expected that this analysis will find a null result. For this reason, the cuts have been optimized to minimize the expected upper limit on the branching fraction. This expected upper limit on the number of signal events is defined as the mean upper limit expected for the background-only hypothesis for a given background contribution  $N_{bgd}$  and signal efficiency  $\varepsilon$  (see Section 1.5).

As described more fully in Section 7, the background estimates are extracted from the data itself in the sideband region. There is a danger that statistical fluctuations in the data will favor a particular cut value and the background estimates will be biased toward a value which is too low. To avoid being sensitive to this kind of bias, the background normalizations for the optimization procedure are estimated and fixed with a nominal set of cuts applied. The data are not refit as the cuts are moved, but rather the MC and control samples are used to predict the relative background change as a function of a particular cut value. Cut optimization is considered for each channel separately. In some cases the optimal cut value does not change significantly for different channels, or there is a wide range of optimal cut values. In these cases, a single cut value is chosen for all channels.

The selection cuts applied to all channels are the following:

- Mass of the one-prong hemisphere ( $m_{1pr}$ ) is calculated as the invariant mass of the charged candidates and neutrals in the the 1-prong hemisphere and the total missing momentum in the event. The charged track is assigned the most-likely mass hypothesis. This one-prong mass is required to be  $m_{1pr} \in (0.3, 3.0)$  GeV/ $c^2$  for all channels except  $e^-e^+e^-$  and  $\mu^-e^+e^-$ , for which the requirement is  $m_{1pr} \in (0.5, 2.5)$  GeV/ $c^2$ .
- Momentum of one-prong track ( $p_1^{cms}$ ) is less than 4.8 GeV/ $c$ .
- No tracks on the 3-prong side may pass tight kaon criteria (see Appendix B for definition of the **KLHTight** selector).

The following selections cuts are not effective in all search channels, and are applied to individual channels as noted:

- Total transverse momentum in the CM frame ( $p_T^{cms}$ ) is greater than 0.4 GeV/ $c$  for channels  $e^-e^+e^-$  and  $e^-\mu^+\mu^-$  and greater than 0.2 GeV/ $C$  for  $\mu^-e^+e^-$ .
- The invariant mass is calculated for the two possible pairs of opposite-sign tracks on the 3-prong side. The smallest of these values ( $m_{2tr}^{min}$ ) must be greater that 0.25 GeV/ $c^2$ . Applied to channels  $e^-e^+e^-$  and  $\mu^-e^+e^-$  as a cut against conversions in Bhabha and di-muon events. This cut is tighter than the conversion cut in the preselection criteria.
- One-prong track must not be identified as a loose electron (**eMicroLoose**) (see Appendix B for definition of the **eMicroLoose** selector). To ensure that

the veto works, the track is additionally required to have non-zero EMC information <sup>3</sup>. Applied as a cut against Bhabha events in channels  $e^-e^+e^-$  and  $e^-\mu^+\mu^-$ .

- One-prong track must not be identified as a loose muon (**muNNLoose**). (see Appendix B for definition of the **muNNLoose** selector). Applied as a cut against di-muon events in channels  $\mu^-e^+e^-$  and  $\mu^-\mu^+\mu^-$ .

The efficiency of the selection is given in Table III.7. Optimization plots, showing the expected upper limit on the branching fraction for different cut values, are shown in Figures C.1- C.6 in Appendix 1. Other cuts which have been considered include:

- polar angle of missing momentum in LAB frame.
- # of photon candidates on 1-prong and 3-prong hemispheres.
- minimum track momentum in the 3-prong hemisphere.
- the acollinearity angle between the 1-prong and 3-prong momentum vectors in the CM frame.

The distributions of the MC and data events in the selection variables are shown in Figures C.7-C.12 in Appendix 2. The events are plotted with all selection criteria applied *except* the cut in the plotted variable.

Table III.7: Efficiency for events after PID and LB requirements to pass the selection cuts. As described in Section 7, the Bhabha and dimuon contributions are modeled with data control samples. The corresponding selection efficiencies are not shown.

	Signal[%]	$b\bar{b}$ [%]	$c\bar{c}$ [%]	$uds$ [%]	$\tau^+\tau^-$ [%]	DATA [%]
$e^-e^+e^-$	68.6	48.6	13.8	9.21	0.130	0.215
$\mu^-e^+e^-$	72.3	60.3	38.1	35.7	8.12	1.7
$e^-\mu^+e^-$	94.3	79.2	23.7	50.9	90.1	40.0
$\mu^-e^+\mu^-$	93.7	56.8	27.7	51.1	87.1	63.2
$e^-\mu^+\mu^-$	71.7	57.8	21.6	45.4	65.5	0.700
$\mu^-\mu^+\mu^-$	77.2	50.0	21.8	50.7	71.7	18.3

<sup>3</sup>We mean that the software object **BtaCalQual** exists. This object will not exist if the prompt reconstruction found no EMC energy deposit associated with the track. This requirement ensures that events with electrons which hit cracks in the EMC do not pass the 1-prong electron veto.

## 7 Estimation of Background

To estimate the expected background contribution in the signal region, a background fitting procedure has been developed which uses the data directly to estimate the background levels in the two-dimensional  $(\Delta M, \Delta E)$  plane. For  $c\bar{c}$ ,  $uds$ , and generic  $\tau^+\tau^-$  backgrounds, Monte Carlo samples are used to construct an analytic two-dimensional PDF as a function of  $\Delta M$  and  $\Delta E$ . Due to lack of suitable MC events, the QED (Bhabha and di-muon) background is estimated directly from the data using the procedure described in Section 7.3.

The final background rates are estimated by performing an unbinned likelihood fit over the large box region excluding the signal region (also known as the *grand sideband*). Each of the background classes (QED,  $c\bar{c}$ ,  $uds$ , and  $\tau^+\tau^-$ ) has a single analytic PDF which describes the shape of that background in the  $(\Delta M, \Delta E)$  plane for each signal hypotheses. The normalization of each PDF is determined from the fit to the grand sideband data, and the final background estimate is then made by integrating the normalized PDFs over the signal box region.

Systematic uncertainties due to the background estimation, including dependence upon the exact PDF functions used and variations of the shape parameters, are discussed in Section 8.

### 7.1 Backgrounds from $c\bar{c}$ and $uds$

The shapes of the  $c\bar{c}$  and  $uds$  backgrounds in the signal region are estimated using MC samples. These two backgrounds have very similar distributions in  $\Delta M$  and  $\Delta E$ . Since the overall rate is determined in a fit to the data sidebands, the  $uds$  MC sample is used to simulate both  $uds$  and  $c\bar{c}$  backgrounds. Background estimates which include fits to the  $c\bar{c}$  sample differ negligibly from background estimates which use only the  $uds$  MC sample.

An unbinned likelihood fit with weights is used to constrain the parameters of an analytic two-dimensional PDF to the observed MC distributions of  $(\Delta M, \Delta E)$ . The weight of the events corresponds to the probability of the Particle Identification (taken from the PID tables). As one can see from the Figure III.8, the average PID-weight is not constant across the  $\Delta M$  and  $\Delta E$  distributions and the usage of the average weight is unacceptable.

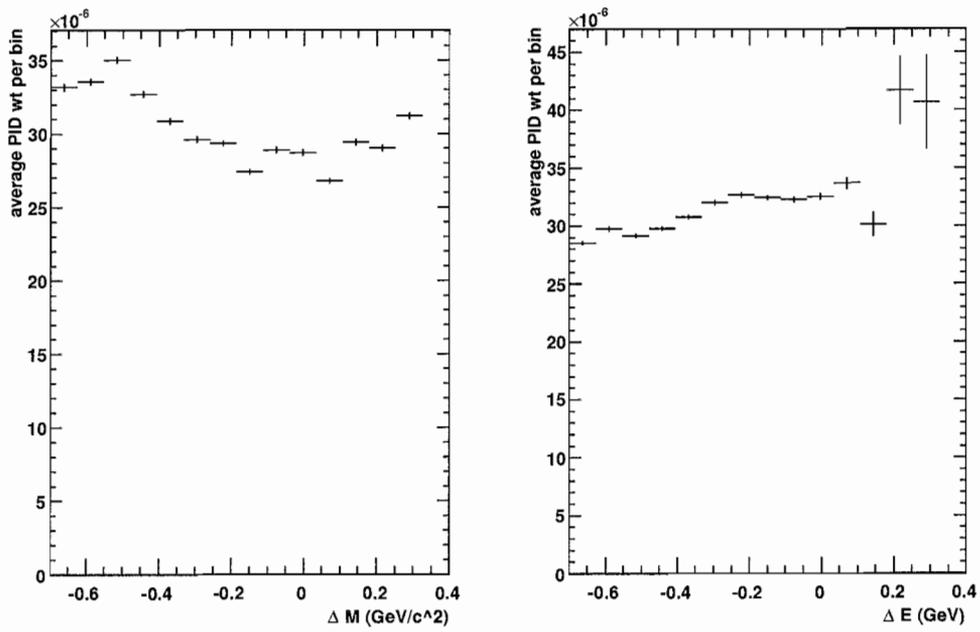


Figure III.8:  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ :  $uds$  background. The histograms show the average PID-weight per bin as a function of  $\Delta M$  (left) and  $\Delta E$  (right).

The two-dimensional  $(\Delta M, \Delta E)$  PDF for the  $uds$  sample is constructed as the product of two one-dimensional PDFs  $(P_{M'}, P_{E'})$ . Since we observe a correlation between  $\Delta M$  and  $\Delta E$  distribution for  $uds$  background, the rotated variables

$$\Delta M' = \cos(\alpha)\Delta M + \sin(\alpha)\Delta E; \quad \Delta E' = -\sin(\alpha)\Delta M + \cos(\alpha)\Delta E \quad (\text{III.25})$$

are used as dependents of each one-dimensional PDF. The angle  $\alpha$  is included into 2-dimensional fit as a free parameter.  $P_{M'}$  is a bifurcated Gaussian and  $P_{E'}$  is given by

$$P_{E'} = \left(1 - \frac{x}{\sqrt{1+x^2}}\right) \cdot (1 + ax + bx^2 + cx^3), \quad (\text{III.26})$$

where  $x = (\Delta E' - \Delta E'_0)/\sigma(E')$ . The values  $\Delta E_0$  and  $\sigma(E)$  are free parameters to be determined from the fit. Therefore fit minimizes the function

$$L = - \sum_i w_i \cdot \log(P_{M'}(\Delta M_i, \Delta E_i; \alpha, \vec{p}_{M'}) \cdot P_{E'}(\Delta M_i, \Delta E_i; \alpha, \vec{p}_{E'})), \quad (\text{III.27})$$

where sum is taken over all points in corresponding sample,  $w_i$  is PID weight of event,  $\vec{p}_{M'}$  and  $\vec{p}_{E'}$  are the parameters of the corresponding PDFs. There are in total nine parameters describing the shape of the  $uds$  PDF. The results of fits to the MC distribution are shown in Figure C.13 in Appendix 3. Although the PDF contains many parameters, the MC statistics are sufficient to constrain all of them.

## High PID-weight Events

The  $uds$  background MC sample contains a small number of high PID-weight events which pass the LB criteria and selection cuts. These high-weight events have particle identification weights which are much greater than the average PID weight for the sample (2-3 orders of magnitude greater). These rare events contain signal-side tracks which are real leptons and thus are have PID efficiencies close to unity. Fits have been done with and without the inclusion of these events, and the resulting background estimations are negligibly different. For plotting purposes, these events have been removed. For actual background estimations, the events have been kept as part of the data sample.

## 7.2 Background from $\tau^+\tau^-$

The two-dimensional PDF for  $\tau^+\tau^-$  is a product of the two one-dimensional PDFs ( $P_{M''}, P_{E''}$ ) for  $\Delta M''$  and  $\Delta E''$  dependents, respectively. The variables  $\Delta M''$  and  $\Delta E''$  are functions of  $\Delta M$  and  $\Delta E$

$$\Delta M'' = \cos(\beta_1)\Delta M + \sin(\beta_1)\Delta E; \quad \Delta E'' = -\sin(\beta_2)\Delta M + \cos(\beta_2)\Delta E, \quad (\text{III.28})$$

but unlike the  $uds$  PDF, they are not required to be perpendicular. Angles  $\beta_1$  and  $\beta_2$  are included in the fit as free parameters. The  $P_{M''}$  PDF is a sum of two Gaussian PDFs with common mean, while  $P_{E''}$  is described by Equation III.26 with  $x = (\Delta E'' - \Delta E''_0)/\sigma(E'')$ . Therefore fit minimizes the function

$$L = - \sum_i w_i \cdot \log(P_{M''}(\Delta M_i, \Delta E_i; \beta_1, \vec{p}_{M''}) \cdot P_{E''}(\Delta M_i, \Delta E_i; \beta_2, \vec{p}_{E''})), \quad (\text{III.29})$$

where sum is taken over all points in the corresponding sample,  $w_i$  is PID weight of event,  $\vec{p}_{M''}$  and  $\vec{p}_{E''}$  are the parameters of the corresponding PDFs. There are in total eleven parameters describing the shape of  $\tau^+\tau^-$  PDF. The results of the fits to the  $\tau^+\tau^-$  distributions are shown in Figure C.14 in Appendix 3.

## 7.3 QED Background

Since the number of Bhabha and  $\mu^+\mu^-$  MC events is smaller than the number of events expected in the data sample, a procedure has been developed to use data control samples to estimate the shape of the QED background in the  $(\Delta M, \Delta E)$  plane. The shapes of the Bhabha and  $\mu^+\mu^-$  backgrounds are actually very similar, and a single PDF is used in each signal channel to parameterize both components. For the final background fit, the PDF extracted from the Bhabha control sample is used in channels where the Bhabha background is dominant ( $e^-e^+e^-, e^-\mu^+\mu^-$ ), while the PDF extracted from the  $\mu^+\mu^-$  control sample is used for  $\mu^-e^+e^-$ . The remaining search channels,  $e^-\mu^+e^-, \mu^-e^+\mu^-$ , and  $\mu^-\mu^+\mu^-$ , have negligible QED backgrounds in the GS.

## Reverse PID Sample

In two of the channels where the data include significant QED background ( $e^-e^+e^-$  and  $e^-\mu^+\mu^-$ ), it is possible to construct an adequate QED control sample by simply looking at events in the grand sideband region that pass all selection cuts except the cut on the 1-prong particle ID (*reverse PID data sample*). In other words, we select events in which the 1-prong track is identified as either eMicroLoose or muNNLoose but otherwise pass the selection cuts.

An analytic PDF for the QED background is then constructed for each signal hypothesis by performing a maximum likelihood fit to the control sample. The QED PDF function  $P_{\text{QED}}$  defined by a product of 1 dimensional distributions  $P'_{M'}$  and  $P'_{E'}$  over  $M' = \cos(\beta)\Delta M + \sin(\beta)\Delta E$  and  $E' = -\sin(\beta)\Delta M + \cos(\beta)\Delta E$  parameters.  $P'_{M'}$  is a third order polynomial in  $M'$  and  $P'_{E'}$  is the Crystal Ball function ( $P_{CB}$ ):

$$P_{CB} = \begin{cases} \exp(-\frac{x^2}{2}) & x > \alpha \\ \frac{(n/\alpha)^n \cdot \exp(-\alpha^2/2)}{n/\alpha - \alpha - x} & x \leq \alpha \end{cases} \quad (\text{III.30})$$

where  $x = (E' - E'_0)/\sigma_{E'}$ , while  $E'_0$  and  $\sigma_{E'}$  are fit parameters which describe the shape of the peak. The rotation angle  $\beta$  is used to account for the observed correlation between  $\Delta M$  and  $\Delta E$  for the QED events. In total, there are six free parameters to describe this unit-normalized 2D PDF.

Fits of this PDF to the Bhabha and di-muon control samples are shown in Figure C.15 in Appendix 3. As can be seen, the control samples have adequate statistics to determine the PDF parameters.

## Alternate QED Sample

For the  $\mu^-e^+e^-$  channel, the reverse PID sample does not have adequate statistics to determine the background shape. An alternate control sample (*alternate QED sample*) is defined by taking preselected data events with a identified muon on the 1-prong side,  $0.5 < m_{1pr} < 2.5$ ,  $p_1^{cms} > 4.8\text{GeV}/c$ , and no PID requirements on the 3-prong side. The 1-prong PID requirement guarantees that this alternate Bhabha sample is independent from any candidate signal events in the  $\mu^-e^+e^-$  data sample. The alternate QED sample is fit with the same PDF as the reverse QED sample.

## 7.4 Final Background Fit

Using the analytic PDFs for the  $uds$ , Bhabha/di-muon, and generic  $\tau^+\tau^-$  backgrounds determined as described above, a final unbinned likelihood fit is performed to the data found in the grand sideband region for each of the signal hypotheses, with the number of sideband events for each of the background classes (yields) as the fit parameters. The results of these fits are shown in Figures III.9-III.14. For some search channels, one of the three background classes has a negligible contribution in the LB. Only background classes with significant contributions are actually included in the final background fit. Table III.8 lists the background contributions to each search channel.

Table III.8: Expected number of background events in the grand sideband (GS) and signal box (SB) after the background fits. By construction, the total number of expected background events in the GS is equal to the number of data events in the GS. The luminosity is  $376 \text{ fb}^{-1}$ .

	$e^-e^+e^-$		$\mu^+e^-e^-$		$\mu^-e^+e^-$		$e^+\mu^-\mu^-$		$e^-\mu^+\mu^-$		$\mu^-\mu^+\mu^-$	
	GS	SB	GS	SB	GS	SB	GS	SB	GS	SB	GS	SB
$uds$	30.5	0.41	18.1	0.25	33.6	0.53	29.2	0.49	28.9	0.41	24	0.29
QED	28.5	0.92	0	0	16.9	0.33	0	0	11.1	0.38	0	0
$\tau^+\tau^-$	0	0	20.9	0.05	15.5	0.03	122.8	0.05	38.0	0.02	92	0.04
Total	59	1.33	39	0.30	66	0.89	152	0.54	78	0.81	116	0.33

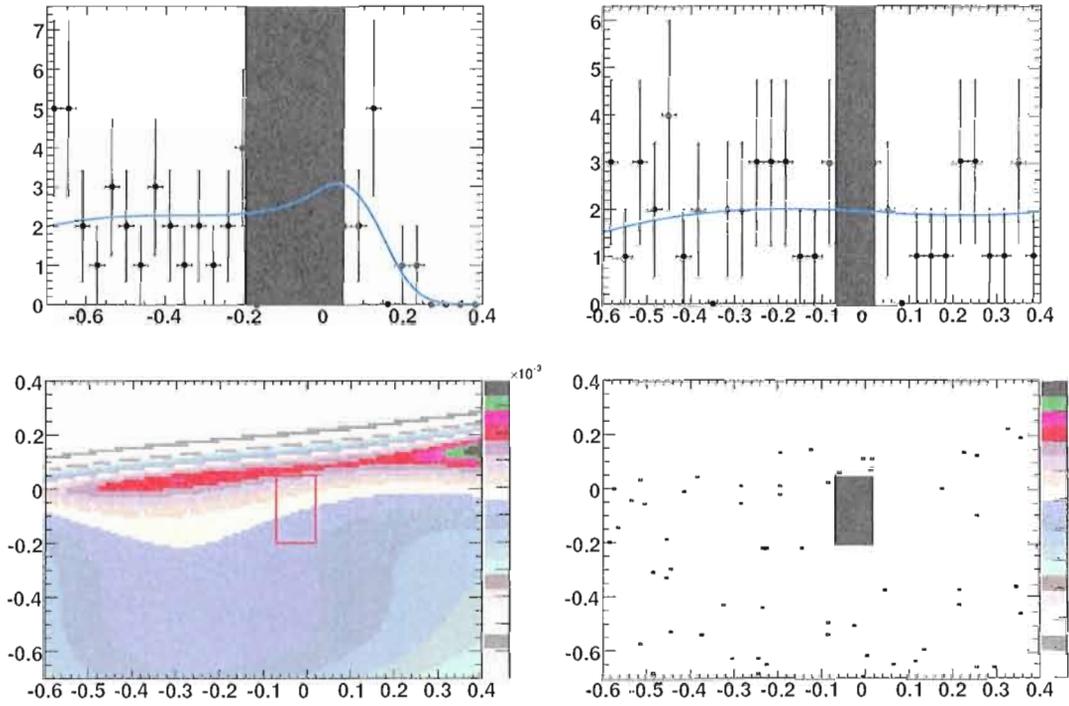


Figure III.9:  $\tau^- \rightarrow e^-e^+e^-$  channel: PDFs with MC-fitted shapes are scaled to data: a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF  $(\Delta M, \Delta E)$  distribution; d) data  $(\Delta M, \Delta E)$  distribution. The filled black boxes and open red box show the signal region (blinded for data).

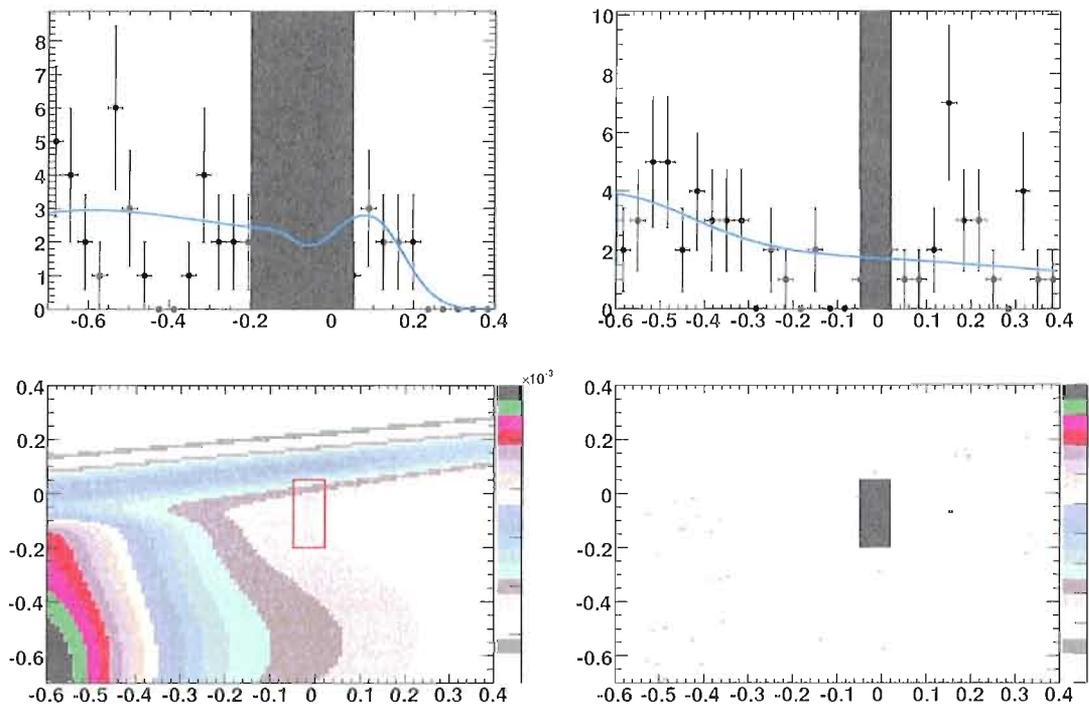


Figure III.10:  $\tau^- \rightarrow \mu^- e^+ e^-$  channel: PDFs with MC-fitted shapes are scaled to data; a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF ( $\Delta M, \Delta E$ ) distribution; d) data ( $\Delta M, \Delta E$ ) distribution. The filled black boxes and open red box show the signal region (blinded for data).

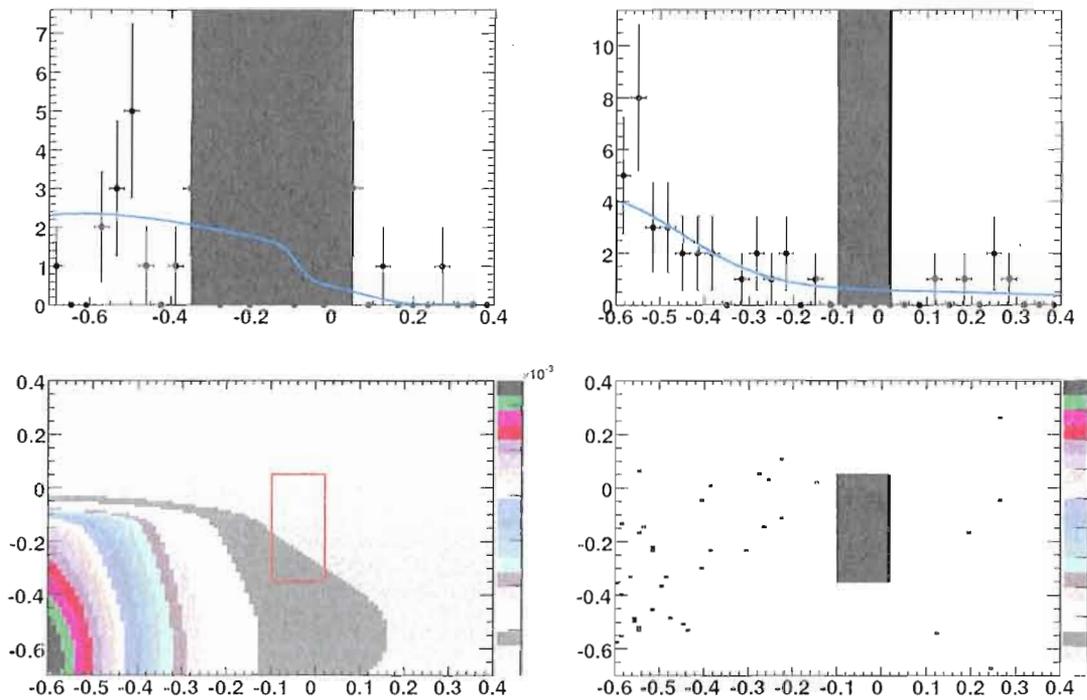


Figure III.11:  $\tau^- \rightarrow \mu^+ e^- e^-$  channel: PDFs with MC-fitted shapes are scaled to data: a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF  $(\Delta M, \Delta E)$  distribution; d) data  $(\Delta M, \Delta E)$  distribution. The filled black boxes and open red box show the signal region (blinded for data).

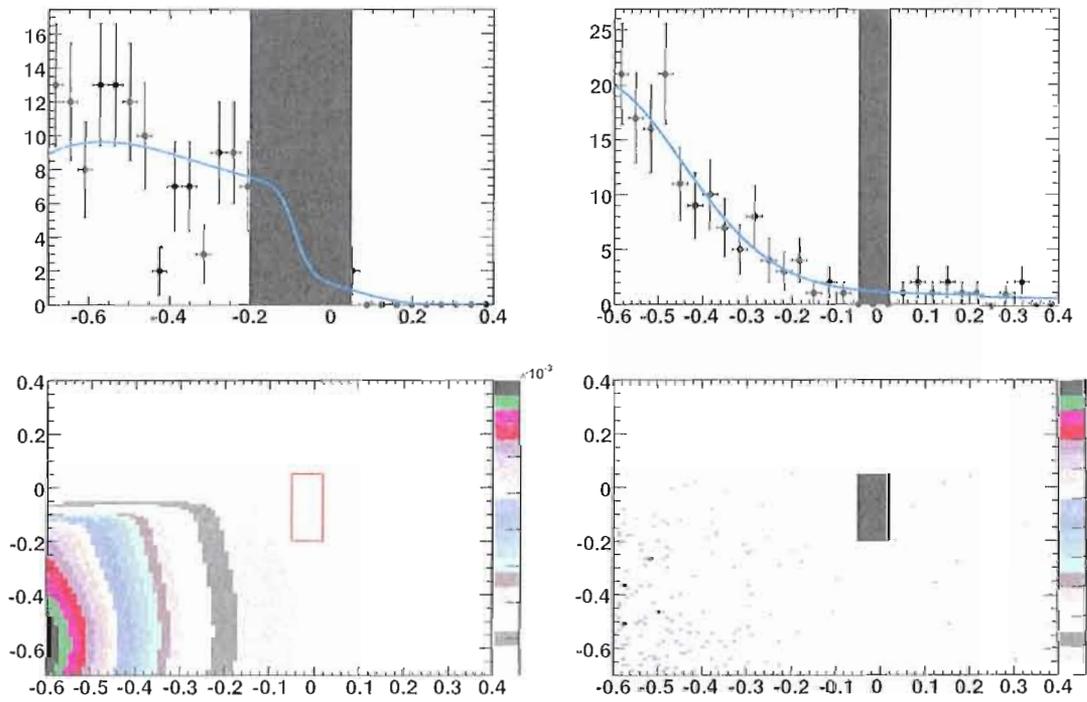


Figure III.12:  $\tau^- \rightarrow e^+\mu^-\mu^-$  channel: PDFs with MC-fitted shapes are scaled to data; a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF ( $\Delta M, \Delta E$ ) distribution; d) data ( $\Delta M, \Delta E$ ) distribution. The filled black boxes and open red box show the signal region (blinded for data).

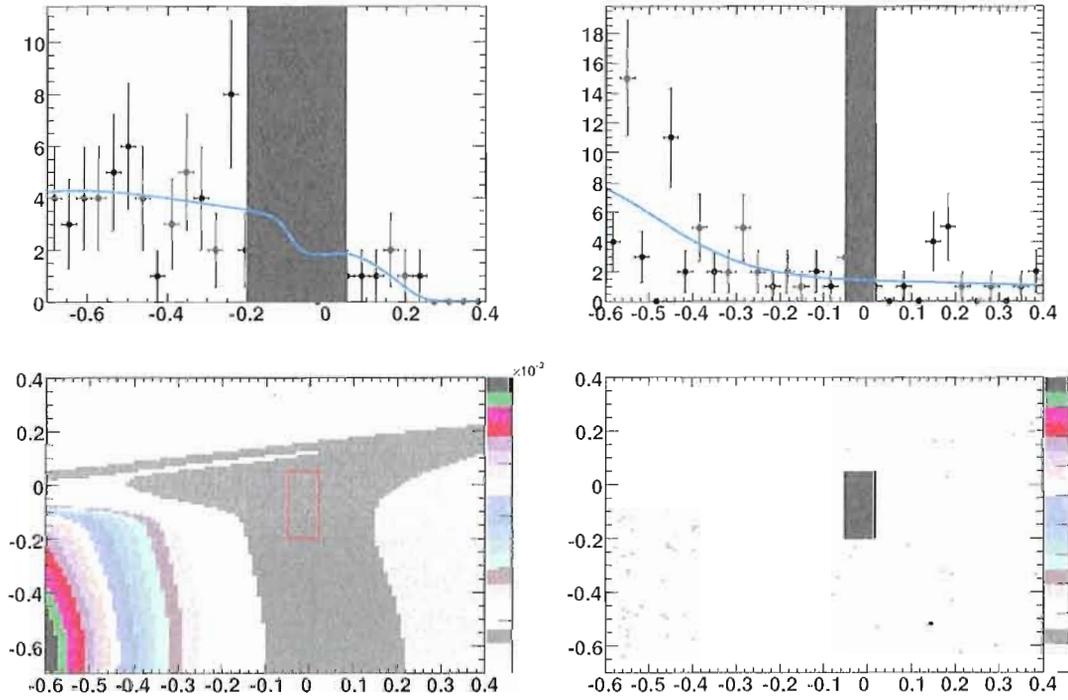


Figure III.13:  $\tau^- \rightarrow e^- \mu^+ \mu^-$  channel: PDFs with MC-fitted shapes are scaled to data: a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF ( $\Delta M, \Delta E$ ) distribution; d) data ( $\Delta M, \Delta E$ ) distribution. The filled black boxes and open red box show the signal region (blinded for data).

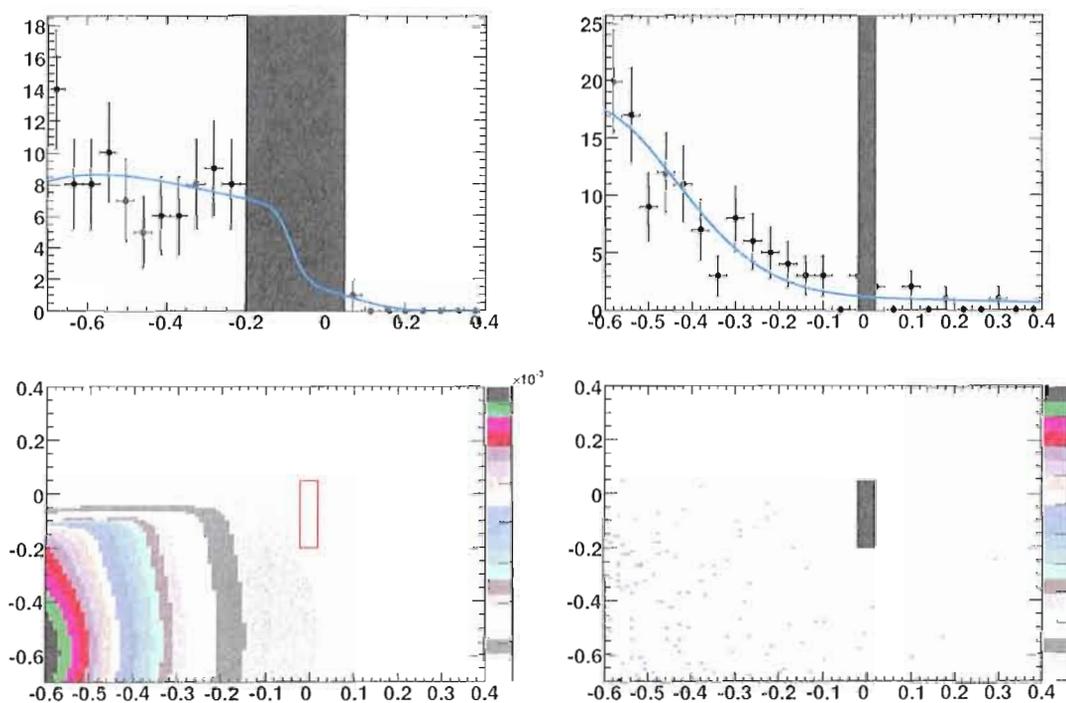


Figure III.14:  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  channel: PDFs with MC-fitted shapes are scaled to data: a)  $\Delta E$  projection of the data (points) and the sum of the background PDFs (curve); b)  $\Delta M$  projection of the data (points) and the sum of the background PDFs (curve); c) PDF ( $\Delta M, \Delta E$ ) distribution; d) data ( $\Delta M, \Delta E$ ) distribution. The filled black boxes and open red box show the signal region (blinded for data).

## 8 Systematic Uncertainties

The systematic uncertainties in this analysis can be divided into three parts: uncertainties related to the signal efficiency, uncertainties related to the background estimate, and uncertainties related to computing the branching fraction (luminosity and  $\tau^+\tau^-$  cross section).

In principle, some sources of uncertainty affect the signal efficiency and background estimate in a correlated way (like tracking efficiency), however these uncertainties have been found to be negligible.

### 8.1 Signal Efficiency

The signal efficiencies are determined from signal MC samples, and hence the efficiency systematics are driven by understanding the deficiencies and uncertainties in the MC modeling.

#### Limited MC Statistics

The absolute uncertainty due to the limited signal MC statistics is calculated using

$$\sigma_{MC} = \sqrt{\frac{\varepsilon(1 - \varepsilon)}{N_{MC}}}, \quad (\text{III.31})$$

where  $N_{MC}$  is the number of events in the initial signal MC samples and  $\varepsilon$  is the total signal selection efficiency. The relative uncertainties range from 0.5-0.8% (depending on selection channel) and are shown in Table III.10. This number does not include the uncertainty due to PID efficiency which is considered separately.

#### Production Model

The signal samples have been produced assuming a flat 3-body phase space decay of the tau lepton. The decision has been made to state this assumption explicitly and assign no additional uncertainty for possible model-dependent structure in the decay.

A Dalitz plot of the selection efficiency for the  $\mu^-\mu^+\mu^-$  channel is shown in Figure III.15. The selection efficiency in the Dalitz plane<sup>4</sup> looks fairly flat. However,

---

<sup>4</sup>The Dalitz plane is defined for sets of three energy-momentum four-vectors,  $p_1, p_2, p_3$ . On one axis is plotted the invariant mass of  $p_1$  and  $p_2$ , and on the other the invariant mass of  $p_1$  and  $p_3$ .

there are notable variations in the efficiency when projected on to each invariant mass spectrum. This is due to the large number of soft muons in the signal MC which get zero weight in the PID selection. Upon removing particle identification from the selection procedure, the signal efficiency is uniform across both of the invariant mass spectra.

## Radiation Modeling

Deficiencies in the description of initial-state (ISR) and final-state (FSR) radiation in the MC can lead directly to errors in the predicted  $(\Delta M, \Delta E)$  distributions. Generator level studies with `KK2f` ISR weights are used to estimate the size of this effect due to missing higher-order corrections. The number of events with 3 signal tracks within the detector acceptance and with invariant mass and energy corresponding to the signal box are compared with and without the  $\mathcal{O}(\alpha^3)$  diagrams, as recommended by the `KK2f` authors. The relative efficiency difference of 0.06% is taken as an estimate of the uncertainty related to the missing higher-order diagrams in the calculation. This negligible uncertainty on the signal efficiency is ignored.

A similar study is done to estimate the uncertainty due to FSR from the outgoing leptons in the decay (generated by `PHOTOS`). The associated systematic uncertainty is also negligible.

## Generic $\tau$ Branching Fraction

The generic decays of the second  $\tau$  in the signal MC are simulated by `TAUOLA` with PDG 2004[65] branching fractions with an additional unitary constraint imposed. The systematic uncertainty related to the branching fraction errors is evaluated as a quadrature sum of the individual branching fraction uncertainties weighted by the relative fraction of selected events in a given signal channel with this generic tau decay mode. The relative systematic uncertainty is 0.9%.

## PID Efficiency

The uncertainty due to particle identification performance for the 3-prong tracks is estimated from the statistical uncertainty of the PID table efficiencies. As a conservative estimate, the relative uncertainty for the event weight is taken to be

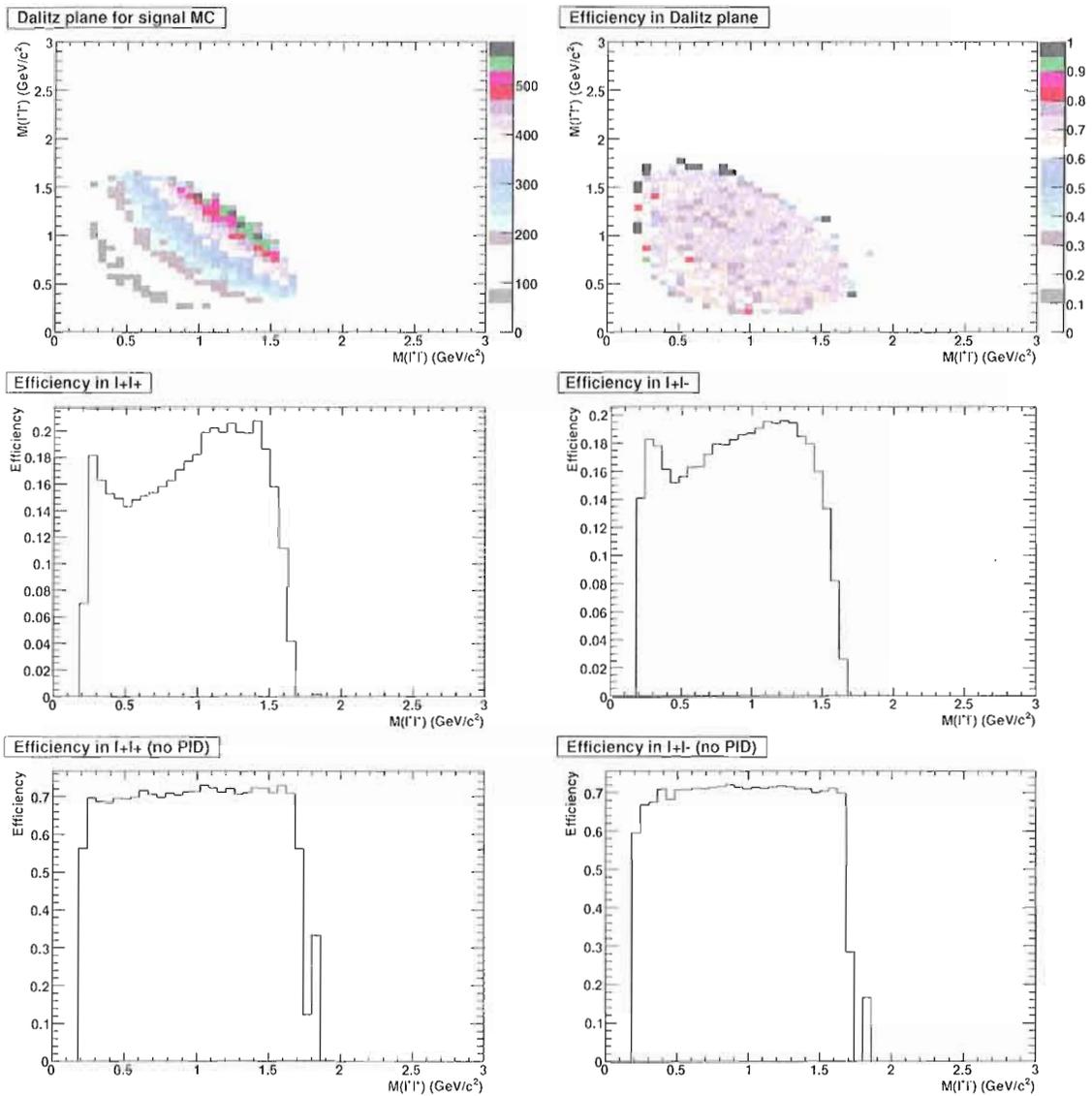


Figure III.15:  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  a) generated MC Dalitz distribution after preselection; b) efficiency to pass all selection except SB as function of Dalitz distribution; c) selection efficiency as a function of invariant mass squared of the pair of same-sign leptons; d) selection efficiency as a function of invariant mass squared of the pair of opposite-sign leptons; e) efficiency for all selection cuts except PID as a function of invariant mass squared of the pair of same-sign leptons; e) efficiency for all selection cuts except PID as a function of invariant mass squared of the pair of opposite-sign leptons;

the quadrature sum of the relative uncertainties of the three-prong track weights. The resulting distribution of event weight uncertainties is significantly asymmetric and has a large tail at high uncertainties. The distribution is integrated from zero up to the value where 68% of the distribution is included. This value is taken as the uncertainty due to particle ID on the 3-prong tracks.

The kaon veto on the 3-prong tracks affects less than 1% of the signal events passing all other selection criteria. It has a negligible effect on the signal efficiency and no uncertainty is assigned. The uncertainty due to the lepton veto on the one-prong track is estimated from the spread around unity of the ratio of MC and data efficiencies, about 1.5% for the electron veto and 6.5% for the muon veto. This uncertainty is added in quadrature with the uncertainty from the 3-prong track PID. Total PID uncertainties for each channel are shown in Table III.10 and range from 1.7% ( $e^-e^+e^-$ ) to 10.7% ( $\mu^-\mu^+\mu^-$ ).

## Tracking Efficiency

Any mismatch between the data and MC tracking efficiency will lead to a bias in the signal efficiency estimate. Internal *BABAR* studies show that the modeling of the single track efficiency in the MC is good to 0.23% per track in low multiplicity events for track momenta  $p_T > 180$  MeV/c. For the few tracks with momentum  $p_T < 180$  MeV/c we conservatively assign a 1.2% uncertainty per track. Using the individual track  $p_T$  values observed, a tracking uncertainty for each event is calculated by simple addition of the individual track uncertainties. This implicitly assumes that the tracking efficiency uncertainties are correlated for all tracks. The total uncertainty is taken to be the mean event uncertainty observed for each signal mode.

The fraction of tracks with  $p_T < 0.18$  GeV/c is approximately 0.4% for all channels. The relative uncertainty on the selection efficiency ranges from 0.99% to 1.01% depending upon the specific channel. Exact values are shown in Table III.10. In principle, this uncertainty is correlated to the background estimate, although due to the way the background rates are fit directly from the data, this correlation is assumed to be negligible.

## Tracking Resolution

If the MC tracking resolution does not match that found in the data, the signal distributions in the  $(\Delta M, \Delta E)$  plane will be incorrect. This affects the number of signal events falling in the signal box and hence the overall signal efficiency. Most studies done of tracking resolution have found that the width of the invariant mass spectra of various control samples are reproduced by the MC simulations to within 5% relative.

To evaluate the efficiency uncertainty due to this level of agreement, an additional smearing of the track momentum was added such that the track resolution was increased by 5%. Assuming that the average track momentum resolution is 0.5%, an additional Gaussian smearing of 0.16% will increase the resolution by 5% relative. This procedure is implemented by adding a random value  $\delta p$  to the momentum magnitude of each track  $p_0$  drawn from a Gaussian with width equal to  $\sigma_{\delta p} = 0.0016p_0$ .

The effect of this additional smearing is to migrate some fraction of the signal events out of the signal box and reduce the efficiency. This effect varies for the different signal channels, giving a reduction ranging from 0.01% to 0.30% relative. Therefore, no uncertainty is assigned for tracking resolution.

## Dependence on Selection Cuts

Most of the uncertainties related to the modeling of the selection cut variables are already accounted for by other systematic uncertainties.

Uncertainties due to PID requirements in the selection are evaluated from the stated PID selector uncertainties.

The uncertainty on the transverse momentum and on the 1-prong momentum and mass distributions are mostly due to the errors in the tracking model, which has also been accounted for explicitly.

## 8.2 Background Estimation

### Fit Uncertainties

Since the data are used directly to evaluate the background level, a primary source of uncertainty in the background estimation comes from the statistical precision of

the background fit to GS data and varies from about 10% for channels with a lot of data in the GS to 36% for channels with only few data events in the GS. These uncertainties are estimated by varying the background yields within their fit errors and refitting for the expected background contribution in the SB. The ratio of the width to the mean of the resulting distribution of SB background contributions is taken as the relative uncertainty due to background yield errors.

Additional systematics come from the choice of background PDF used for the fits. Estimations for this uncertainty are obtained in one of two ways. For channels where the full covariance matrix can be obtained from the fits to all MC and control samples ( $e^-e^+e^-$ ,  $\mu^-e^+e^-$ ,  $e^-\mu^+e^-$ ), the parameters of the background PDFs have been varied according to the error matrix. For each variation, the GS data are refit and a new estimation of the background in the SB is calculated. The relative systematic uncertainty on the background estimate is taken to be the ratio of the width to the mean of the background distribution. For channels where the full covariance matrix for the background fits are not all available ( $\mu^-e^+\mu^-$ ,  $e^-\mu^+\mu^-$ ,  $\mu^-\mu^+\mu^-$ ), the  $(\Delta M, \Delta E)$  background distributions are parameterized by the product of a line in  $\Delta E$  and a line in  $\Delta M$ , and are fit to the MC and control sample distributions. In the same method as before, the sum of these background PDFs is fit to the sideband data events and the expected background in the SB is recalculated. The difference between the expected background from this simplistic model and the expected background from the full parameterization is taken as a conservative estimate of the systematic error due to MC shape modeling.

The uncertainty due to statistical fluctuations in the number of data events in the GS ranges from 8.2% ( $\mu^+e^-e^-$ ) to 17.7% ( $e^+\mu^-\mu^-$ ). The errors on the background estimate are summarized in Table III.10.

To verify that the PDF used fits data, the number of expected and observed events is compared for the boxes neighboring SB as described below.

## Fit Crosscheck

As a cross check we calculate the background level in a set of neighbor boxes and make a comparison with number of events observed there in the data. The neighbor boxes have the twice the size of the signal area. In total 4 boxes are considered: left, bottom, right and top with respect to the signal box. The data events in the neighbor box under consideration are excluded from the background fit in this

cross-check. With a few exceptions, the expected and observed numbers of events agree within the statistical uncertainties (see Table III.9).

Table III.9: The number of expected (left) and observed (right) events in the boxes neighbor to the signal box. Uncertainties on the sum of the expected background for all four boxes are estimated from the uncertainty on the expected background in the SB. Poisson errors are assigned to the sum of the data for all four boxes.

neighbor	left		top		right		bottom		all four together	
$e^-e^+e^-$	3.78	3	1.40	4	3.17	1	3.03	0	$11.4 \pm 2.2$	$8 \pm 2.8$
$\mu^-e^+e^-$	2.45	1	1.50	2	1.71	6	2.02	1	$7.68 \pm 2.3$	$10 \pm 3.2$
$e^-\mu^+e^-$	0.47	3	0.11	0	0.34	0	0.58	2	$1.50 \pm 2.8$	$5 \pm 2.2$
$\mu^-e^+\mu^-$	2.31	1	0.02	0	0.79	3	1.75	0	$4.85 \pm 1.9$	$4 \pm 2.0$
$e^-\mu^+\mu^-$	2.17	2	0.93	1	1.87	1	2.10	3	$7.07 \pm 2.7$	$7 \pm 2.7$
$\mu^-\mu^+\mu^-$	1.07	0	0.17	0	0.56	1	1.01	2	$2.81 \pm 1.6$	$3 \pm 1.7$
Sum	12.25	10	4.11	7	8.44	12	10.49	8	35.31	37

## Two-photon Contribution

Two-photon fusion events can lead to four-fermion final states. These events are characterized by small net transverse momentum in the event ( $p_T^{cms}$ ) and initial leptons flying close to the beam line after scattering. To separate two-photon events from radiative Bhabha and di-muon events, the dependence of the momentum of the 1-prong track ( $p_{1pr}^{cms}$ ) on  $p_T^{cms}$  is studied. From Figure III.16 one can see a large diagonal band due to radiative QED events, as well as a small band at low  $p_T^{cms}$  values due to the two-photon contribution. This contribution is observed in three channels:  $e^-e^+e^-$ ,  $\mu^-e^+e^-$ , and  $e^-\mu^+\mu^-$ . A data control sample with two-photon events is selected from data events passing preselection and PID by making the cuts  $p_{1pr}^{cms} < 4$  and  $p_T^{cms} < 0.2$  GeV/ $c$ . The later cut ensures that the sample is disjoint from the final data sets in the affected channels.

For search channels  $e^-e^+e^-$  and  $e^-\mu^+\mu^-$ , we find that no events from this control sample pass all other selection cuts. For  $\mu^-e^+e^-$ , two events are passed. Most are rejected by the one-prong lepton veto, which passes less than 10% of the two-photon events for all channels. If selection cuts are released, the  $(\Delta M, \Delta E)$  distribution for two-photon data sample looks similar to the  $q\bar{q}$  distribution as one can see from the Figure III.16. Therefore, even if we inappropriately neglect this

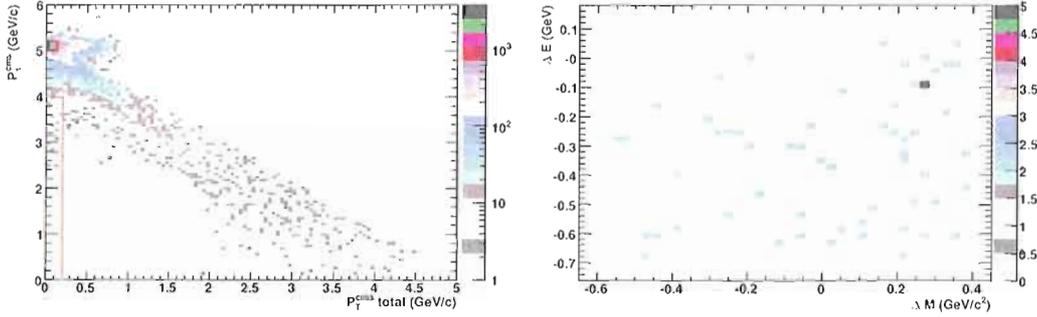


Figure III.16: a) The distribution of  $p_1^{cms}$  versus  $P_T^{cms}$  for the  $\tau^- \rightarrow e^-e^+e^-$  channel after preselection and PID. The z-axis is logarithmic. The red line shows the cuts applied to select the two-photon control sample. b) The  $(\Delta M, \Delta E)$  distribution of these events plotted without the cut on  $P_T^{cms}$ .

type of background, the fit of data will naturally correct for the difference. Thus, we neglect the background uncertainty related to two-photon contribution.

## Tracking Efficiency

Uncertainties in the overall tracking efficiency can be neglected because of the data-driven background estimation which fits the background rates directly from the data observed in the grand sideband region.

## Tracking Resolution

The same smearing study is applied to the  $uds$  and  $\tau^+\tau^-$  background samples as was performed for the signal MC. The relative difference in the accepted background rate in the signal box changes by less than 0.5% for all signal channels.

## Other Unknown Backgrounds

Unknown backgrounds are rather difficult to estimate. Since the backgrounds are already being fit from the data, our procedure probably accommodates any additional unknown background already. Here we assume that none of the backgrounds peak in the signal region and any signature for the peak is due to LFV tau decays. This is true for the standard model tau decays which is verified with generic  $\tau^+\tau^-$  MC sample.

Some backgrounds which are not simulated which we should be able to estimate, however, are the tau decays  $\tau^- \rightarrow \ell^- \ell'^+ \ell'^- \nu_\tau \nu_\ell$ . which have a measured branching fractions of  $(2.8 \pm 1.5) \times 10^{-5}$  ( $e^- e^+ e^- \bar{\nu}_e \nu_\tau$ ) and  $< 3.6 \times 10^{-5}$  ( $\mu^- e^+ e^- \bar{\nu}_\mu \nu_\tau$ ) [68]. For other possible  $\tau^- \rightarrow \ell^- \ell'^+ \ell'^- \nu_\tau \nu_\ell$  modes, the expected branching fraction is close to  $10^{-7}$ .

With  $376 \text{ fb}^{-1}$  of data, we expect a maximum of  $\mathcal{O}(10^3)$  of each of these decays in the data set. The signal total efficiency for pre-selection, PID, and event selection is 12.5% and 10.7% for channels  $e^- e^+ e^-$  and  $\mu^- e^+ e^-$ , respectively. This potentially leaves around 100 events per channel distributed about the LB. However, the SM Feynman diagram for the process includes a virtual photon; therefore the electron-positron pair of the final state have a small invariant mass. The preselection includes at rejection of gamma conversion candidates. Furthermore, channels  $e^- e^+ e^-$  and  $\mu^- e^+ e^-$  also have tighter cuts on the electron-positron invariant mass (see section 6). Given these cuts, and the fact that such SM decays would be distributed similarly to the generic  $\tau$  pair background in the LB, we can safely neglect this background.

### 8.3 Other Systematics

#### Luminosity and $\tau\tau$ Cross Section

The best estimate of the  $\tau$  pair production cross section is  $0.919 \pm 0.003 \text{ nb}$  [69]. Given the 0.9% uncertainty on the luminosity which takes into account the run-by-run variations and the cancellation of the theoretical uncertainty in the product  $\sigma\mathcal{L}$ , the combined uncertainty on the luminosity and the cross section is taken to be 1%.

#### Signal Bias

Since some of the signal events are found outside of the signal box in the grand sideband region, in the case where a signal is found, the background rates predicted by the grand sideband fit will actually be overestimated. If evidence for a signal had been found, a small correction would have been applied to account for this bias.

Table III.10: Systematic uncertainties expressed in relative percent.

	$e^-e^+e^-$	$\mu^-e^+e^-$	$e^-\mu^+e^-$	$\mu^-e^+\mu^-$	$e^-\mu^+\mu^-$	$\mu^-\mu^+\mu^-$
	<b>Uncertainties on the Signal Selection Efficiency</b>					
MC Statistics	0.69	0.73	0.52	0.59	0.73	0.76
Tau BF	0.9	0.9	0.9	0.9	0.9	0.9
PID (3-prong)	1.7	4.1	6.1	8.6	7.1	10.7
PID (1-prong)	1.5	6.5	0	0	1.5	6.5
Tracking Efficiency	1.0	1.0	1.0	1.0	1.0	0.99
Total Uncertainty [%]	2.7	7.9	6.3	8.7	7.4	12.6
	<b>Uncertainties on the Expected Background</b>					
GS fluctuations	12.7	12.4	17.7	8.2	11.1	9.4
Fit to MC	10.6	23.7	179	20.0	12.2	48.1
Fit to data	9.85	13.2	36.2	24.4	20.2	27.1
Total Uncertainty [%]	19.1	29.8	183	39.5	37.8	56.0

## CHAPTER IV

### RESULTS AND CONCLUSIONS

#### 1 Results

All that remains in the  $\tau \rightarrow \ell\ell\ell$  search is to compare the number of observed events in the signal region to the background expectation. This step is referred to as *unblinding* and can only occur after the selection criteria are finalized and all uncertainties are estimated. Let us first recall quantities necessary to place an upper limit.

First, the background expectation in the signal region must be estimated, along with an uncertainty on that estimate. This quantity is calculated separately for each search channel and makes use of both MC and data samples (see Section 7). MC events provide an estimate of the shape of the data distribution in the Large Box, and data events outside the signal region provide an overall normalization. Second, the signal efficiency for each search channel must be calculated. This efficiency takes into account the effects of skimming (Section 6.2), preselection cuts (Section 3), particle identification (Section 4), channel-specific selection criteria (Section 6), and the signal box size (Section 5). Finally, the number of  $\tau^+\tau^-$  pairs produced must be estimated from the luminosity of the data sample, and the  $\tau^+\tau^-$  production cross section for  $e^+e^-$  collisions at 10.58 GeV CM energy. The uncertainties for all three of these quantities are estimated in Section 8.

After unblinding, the observed number of events in the signal region  $N_{obs}$  is compared to the background expectation  $N_{bgd}$  to test the signal hypothesis in the 6 signal channels. Under the assumption that no evidence for a signal is found, a 90% CL upper limit on each branching fraction can be calculated following the technique detailed in Section 1.3. Table IV.1 shows the final results for each search channel, including the signal efficiency, the expected number of background events in the

signal box, the Feldman-Cousins expected upper limit described in Section 1.5, the number of observed events  $N_{obs}$ , and the upper limit. Figure IV.1 shows the unblinded data distributions in the  $(\Delta M, \Delta E)$  plane, along with regions including 50% and 90% of the signal MC events.

Table IV.1: The total efficiency  $\varepsilon$ , estimated background level in the signal region, expected upper limit, observed number of events in the SB and 90% CL upper limit on  $\mathcal{B}(\tau \rightarrow \ell\ell\ell)$ .

Channel	$\varepsilon, [\%]$	$N_{bkg}$	$\mathcal{B}_{exp}^{UL}$	$N_{obs}$	$\mathcal{B}^{UL}$
$e^-e^+e^-$	$8.9 \pm 0.2$	$1.33 \pm 0.25$	$4.9 \cdot 10^{-8}$	1	$4.3 \cdot 10^{-8}$
$\mu^-e^+e^-$	$8.3 \pm 0.6$	$0.89 \pm 0.27$	$5.0 \cdot 10^{-8}$	2	$8.0 \cdot 10^{-8}$
$e^-\mu^+e^-$	$12.4 \pm 0.8$	$0.30 \pm 0.55$	$2.7 \cdot 10^{-8}$	2	$5.8 \cdot 10^{-8}$
$\mu^-e^+\mu^-$	$8.8 \pm 0.8$	$0.54 \pm 0.21$	$4.6 \cdot 10^{-8}$	1	$5.6 \cdot 10^{-8}$
$e^-\mu^+\mu^-$	$6.2 \pm 0.5$	$0.81 \pm 0.31$	$6.6 \cdot 10^{-8}$	0	$3.7 \cdot 10^{-8}$
$\mu^-\mu^+\mu^-$	$5.5 \pm 0.7$	$0.33 \pm 0.19$	$6.7 \cdot 10^{-8}$	0	$5.3 \cdot 10^{-8}$

## 2 Discussion of Results

In all six search channels, the number of events observed is compatible with the expected background. As expected, the values for  $N_{obs}$  are fluctuations around  $N_{bkd}$ , with the largest upward fluctuation seen in  $\tau^- \rightarrow \mu^+e^-e^-$  where  $N_{bkd} = 0.3 \pm 0.55$  and  $N_{obs} = 2$ , and the largest downward fluctuation seen in  $\tau^- \rightarrow e^-\mu^+\mu^-$ , where  $N_{bkd} = 0.81 \pm 0.31$  and  $N_{obs} = 0$ . Combining all six search channels, we observe a total of 6 events in data, while expecting a total of 4.2 background event. Ignoring the uncertainty on the total background, this observation of 6 events while expecting 4.2 has a Poisson probability of 0.11. Taking into account the uncertainty on the  $N_{bkd}$  values would widen the distribution and mildly increase this probability. Thus, the result can be characterized as somewhat “unlucky” but not suspiciously anomalous.

### 2.1 Implications for Theory

In the absence of Higgs-like couplings, the assumption of lepton universality leads to essentially equal rates for the four lepton-number conserving  $\tau \rightarrow \ell\ell\ell$  decays. While models do exist which predict rates for the other two decays modes [70, 71], they

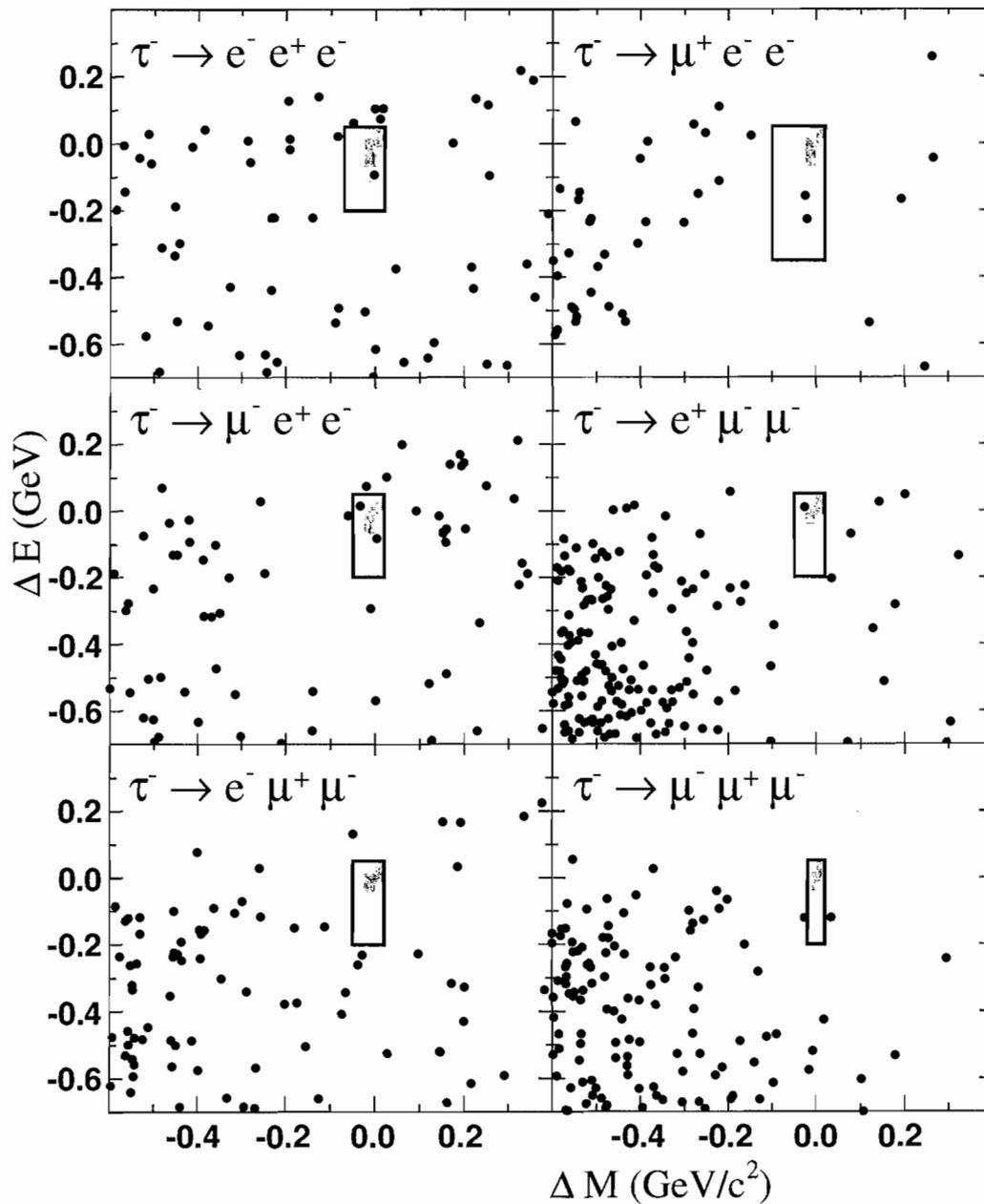


Figure IV.1: Observed data shown as dots in the  $(\Delta M, \Delta E)$  plane and the boundaries of the signal region for each decay mode. The dark and light shading indicates contours containing 50% and 90% of the selected MC signal events, respectively.

are generally much lower than the lepton-number conserving modes. Furthermore, since most models include at least some Higgs contribution, it makes sense to focus on the decay mode most sensitive to these contributions,  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ . The ability to directly constrain new physics models with the  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  result is hampered by the fact that most models predict much higher rates for  $\tau^- \rightarrow \mu^- \gamma$  than for  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  in most areas of parameter space [72, 5, 73]. The situation is remedied slightly by the higher experimental sensitivity to the three body decay (a factor of  $\sim 10$ ).

Two-Higgs Doublet models (2HDM), including minimal supersymmetric models (MSSM), generally have two types of contributions to  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$ . The first type is a subset of  $\tau^- \rightarrow \mu^- \gamma$  in which the photon is off-shell and produces a lepton pair (muons, in our case; rates are similar for electron pairs). This contribution to  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  is naturally suppressed by a factor of  $\sim 100$  relative to  $\tau^- \rightarrow \mu^- \gamma$  [74],

$$\mathcal{B}(\tau \rightarrow 3\mu)_\gamma = \frac{\alpha}{2\pi} \left( \ln \frac{m_\tau^2}{m_\mu^2} - \frac{11}{4} \right) \mathcal{B}(\tau \rightarrow \mu\gamma) \quad (\text{IV.1})$$

except in special cases of fine tuning. The second type of contribution occurs via Higgs-like couplings, as shown in Figure IV.2. In models where the super-particle masses lie above the TeV scale [75], a sizable contribution from the Higgs-mediated channel can lead to ratios like

$$\frac{\mathcal{B}(\tau^- \rightarrow \mu^- \mu^+ \mu^-)}{\mathcal{B}(\tau^- \rightarrow \mu^- \gamma)} > 0.1. \quad (\text{IV.2})$$

With the experimental sensitivity difference noted previously, this means that this sort of new physics could be seen in  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  before  $\tau^- \rightarrow \mu^- \gamma$ . MSSM models where the Higgs contribution is sizable [75, 76, 77] predict the rate for  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  to be

$$\mathcal{B}(\tau^- \rightarrow \mu^- \mu^+ \mu^-) \simeq 1 \times 10^{-7} \left( \frac{\tan\beta}{50} \right)^6 \left( \frac{100\text{GeV}}{m_A} \right)^4, \quad (\text{IV.3})$$

where  $\tan\beta$  is the ratio of the two Higgs doublet VEVs, and  $m_A$  is the mass of the neutral pseudoscalar Higgs particle. In the case of large ( $\sim 50$ )  $\tan\beta$ , the results for  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  presented in this work constraint  $m_A$  to be greater than 100 GeV. Furthermore, no fine-tuned cancellations are required at this point in parameter space to keep the rate for  $\tau^- \rightarrow \mu^- \gamma$  below the experimental limit of  $4.5 \times 10^{-8}$  [78].

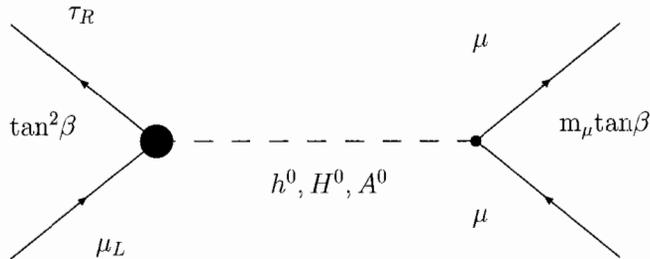


Figure IV.2: Feynman diagram of the leading Higgs-induced contribution to  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  in the MSSM.

### 3 Conclusion

We have used a sample of approximately 350 million  $\tau^+ \tau^-$  pair events recorded at the *BABAR* detector to search for the six lepton flavor violating decays  $\tau \rightarrow \ell \ell \ell$ . In the absence of statistically significant signals for these decays, we have placed upper limits on the branching fractions at the 90% confidence level, using a procedure which takes into account the systematic uncertainties on the signal efficiency, on the number of expected background events, and on the number of  $\tau^+ \tau^-$  pairs produced. The limits on the branching fractions are in the range  $(4 - 8) \times 10^{-8}$ , and improve on the previously published limits by a factor of  $(2 - 5)$  [36, 37].

In Chapter I, we discussed the structure of flavor violation in the interactions of the quarks and leptons. We showed that lepton flavor is essentially conserved in the Standard Model, but that models of new physics provide many options for LFV. We also presented a history of experimental searches for neutrinoless lepton decays. In Chapter II, we presented the *BABAR* experiment. Starting with an overview of the accelerator facilities, we continued with a more detailed look at each of the detector subsystems. The chapter concluded with a discussion of data simulation, triggering, and data processing. In Chapter III, we presented the method by which we actually make the search for  $\tau \rightarrow \ell \ell \ell$  and the statistical procedure that we use to set upper limits on the branching fractions. Chapter IV starts with a presentation of the final results and continues with a discussion of their theoretical implications.

While we were unable to observe the decays  $\tau \rightarrow \ell \ell \ell$ , the limits on the branching fractions that we set with this analysis still constrain theories of physics beyond the Standard Model. And our ability to further constrain these model increases dramatically as limits are pushed into the  $10^{-9}$  range. The PEP-II/*BABAR* facility is

in many ways an ideal  $\tau$  factory. Consequently, the prospects for a Super-B-factory are very exciting. Such an experiment would retain many of the desirable features of *BABAR*, such as the relatively high  $\tau^+\tau^-$  production rate and good separation of the decay products in the detector. With a luminosity 100 times that of *BABAR* and relatively little increase in backgrounds, we can reasonably expect sensitivity to the  $\tau \rightarrow \ell\ell\ell$  decays down to the  $10^{-10}$  range. But the question of what physics we might actually observe at such tiny rates must remain unanswered until such a facility is built.

## APPENDIX A

### TRACK LISTS

#### 1 The CalorClusterNeutral List

The **CalorClusterNeutral** list contains all multi-bump neutral clusters in the EMC, as well as single bumps which are not associated with a charged track.

#### 2 The ChargedTracks List

The **ChargedTracks** list contains all reconstructed tracks with non-zero charge. A pion mass hypothesis is assigned.

#### 3 The GoodTracksVeryLoose List

The **GoodTracksVeryLoose** list contains tracks from the ChargedTracks list for which the following criteria also apply:

- Lab momentum is less than 10 GeV/c.
- Max DOCA (distance of closest approach) in X-Y plane is 1.5 cm.
- Min DOCA in Z is -10 cm.
- Max DOCA in Z is 10 cm.

#### 4 The gammaConversionDefault List

The **gammaConversionDefault** list contains pairs of oppositely-charged tracks from the ChargedTracks list for which the following criteria also apply:

- Max DOCA in X-Y plane is 0.5 cm.

- Max DOCA in Z is 1.0 cm.
- Invariant mass of the two tracks is less than 30 MeV.

## APPENDIX B

### PARTICLE IDENTIFICATION ALGORITHMS

#### 1 The eMicroTight Selector

The tight, cut-based electron selector is called **eMicroTight**. The corresponding PID list contains particle candidates which meet the following criteria:

- dE/dx is in the range [500,1000].
- Minimum of 3 EMC crystals hit.
- The ratio of the EMC energy to the track momentum (E/p) is in the range [0.75,1.3].
- The lateral energy distribution (LAT) is in the range [0,0.6].

$$LAT = \frac{\sum_{i=3}^n E_i r_i^2}{\sum_{i=3}^n E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2} \quad E_1 \geq E_2 \geq \dots E_n, \quad (\text{B.1})$$

where the sum is over all crystals in a shower,  $r_0 = 5\text{cm}$  (the average distance between two crystal frontfaces), and  $r_i$  is the distance between crystal  $i$  and the shower center.

- The shower shape ( $A_{42}$ ) is in the range [-10,10].

$$A_{nm} = \sum_{r_i \leq R_0}^n \frac{E_i}{E} \cdot f_{nm} \left( \frac{r_i}{R_0} \cdot e^{-im\phi_i} \right) \quad (\text{B.2})$$

with  $R_0 = 15\text{ cm}$  and

$$f_{nm}(\rho_i \equiv \frac{r_i}{R_0}) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho_i^{n-2s}}{s! ((n+m)/2 - s)! ((n-m)/2 - s)!} \quad (\text{B.3})$$

with  $n, m \geq 0$ ,  $n - m$  even, and  $m \leq n$ .

## 2 The muNNLoose Selector

The loose, neural-net-based muon selector is called **muNNLoose**. The neural net (NN) consists of:

- One input layer with 8 nodes,
- One hidden layer with 16 nodes,
- One output layer with one node.

The transfer function used in the NN is

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (\text{B.4})$$

with the total input  $x = \sum W_{ij} A_i$ , for weight  $W_{ij}$  and incoming activity  $A_i$ .

The 8 inputs to the NN are the following detector variables normalized to fall in the range  $[0,1]$ :

1. Energy released in the EMC ( $E_{cal}$ )
2. The number of interaction lengths traversed by the track in the detector ( $\lambda_{meas}$ ). This is estimated from the last layer hit by the extrapolated track in the IFR.
3.  $\Delta\lambda = \lambda_{exp} - \lambda_{meas}$ , where  $\lambda_{exp}$  is the expected number of interaction length to be traversed for the track with a muon mass hypothesis.
4. The  $\chi^2/\text{degree of freedom}$  of the IFR hit strips with respect to a third-order polynomial fit of the cluster ( $\chi_{fit}^2$ ).
5. The  $\chi^2/\text{degree of freedom}$  of the IFR hit strips in the cluster with respect to the track extrapolation from the DCH ( $\chi_{mat}^2$ ).
6. The continuity of the track in the IFR ( $T_C$ ).
7. The average multiplicity of hit strips per layer ( $\hat{m}$ ).
8. Standard deviation of the average multiplicity of hit strips per layer ( $\sigma_m$ ).

The **muNNLoose** selector uses multiple kernels for different ranges of momentum, polar angle, and time. The kernels are tuned to provided relatively constant muon identification efficiency.

### 3 The KLHTight Selector

The tight likelihood-ratio-based kaon selector is called **KLHTight**. For each particle candidate, a likelihood is calculated for each particle type. The **KLHTight** list contains particle candidates which fulfill the following conditions on the ratios of the likelihoods:

- $L_K/(L_K + L_\pi) > 0.9$
- $L_K/(L_K + L_p) > 0.2$

Furthermore, particles must not pass the electron likelihood selector to be included in the **KLHTight** list. The likelihoods are calculated from measurements of  $dE/dx$  in the SVT and DCH, and from the Cherenkov angle, the number of Cherenkov photons, and the track quality in the DIRC.

APPENDIX C  
AUXILIARY PLOTS

1 Optimization Plots

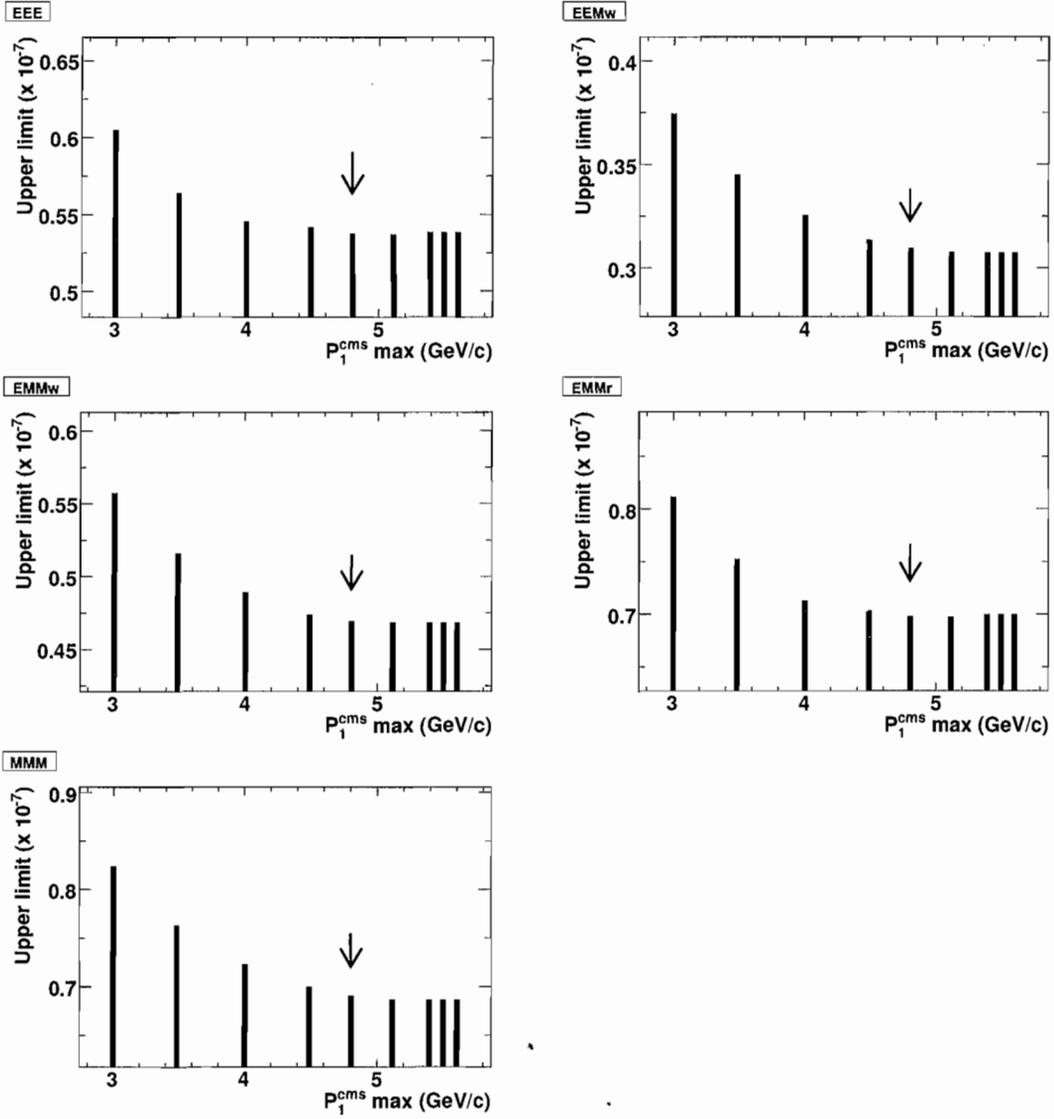


Figure C.1: Expected upper limit on branching fraction as a function of  $p_1^{cms}$ .  $\mu^-e^+e^-$  channel is excluded because the dimuon control sample does not have a requirement on the maximum value of  $p_1^{cms}$  (see section 7). Arrows indicate optimized value for the selection cut.

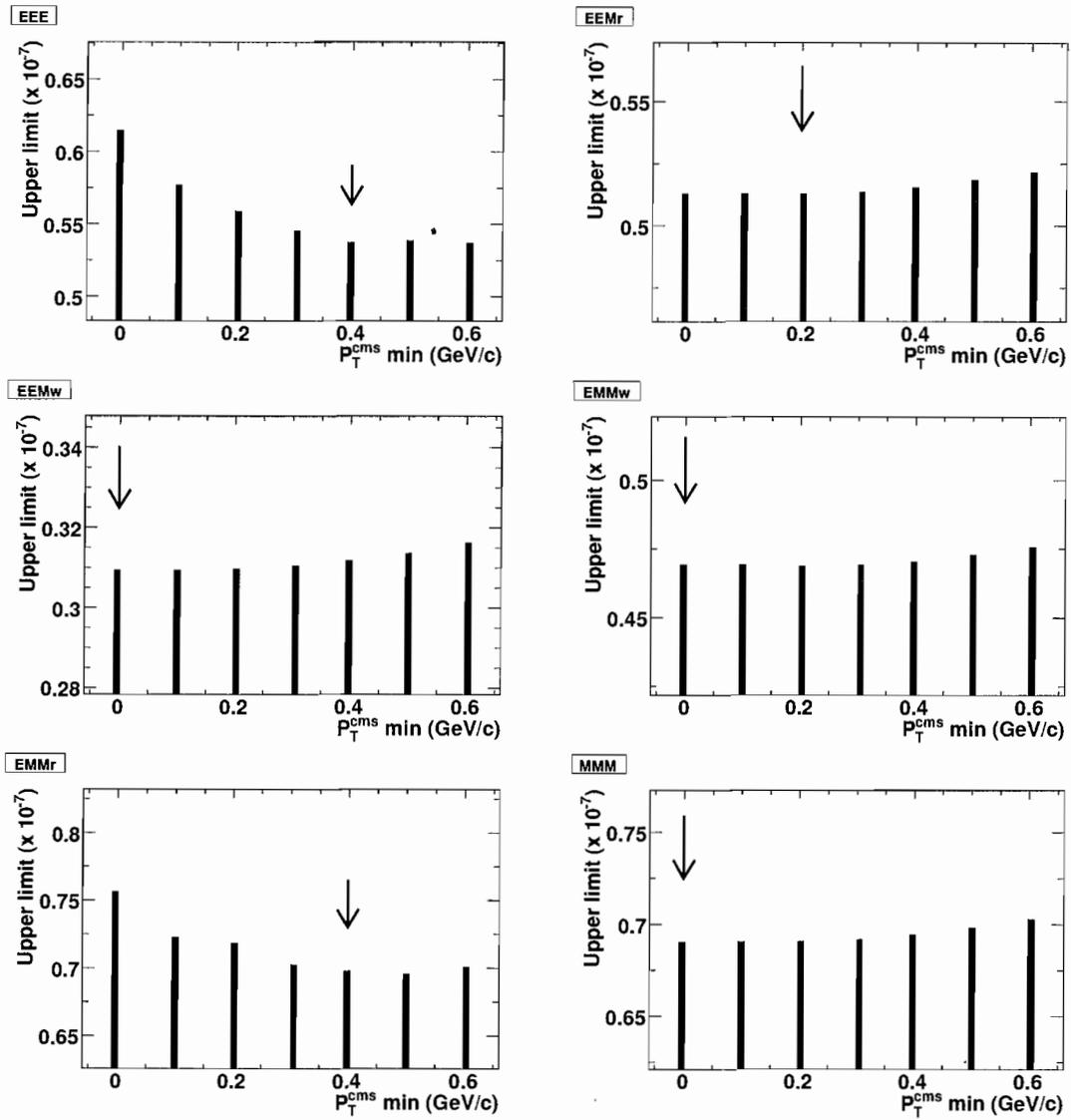


Figure C.2: Expected upper limit on branching fraction as a function of the minimum total transverse momentum  $P_T^{cms}$ . Arrows indicate optimized value for the selection cut.

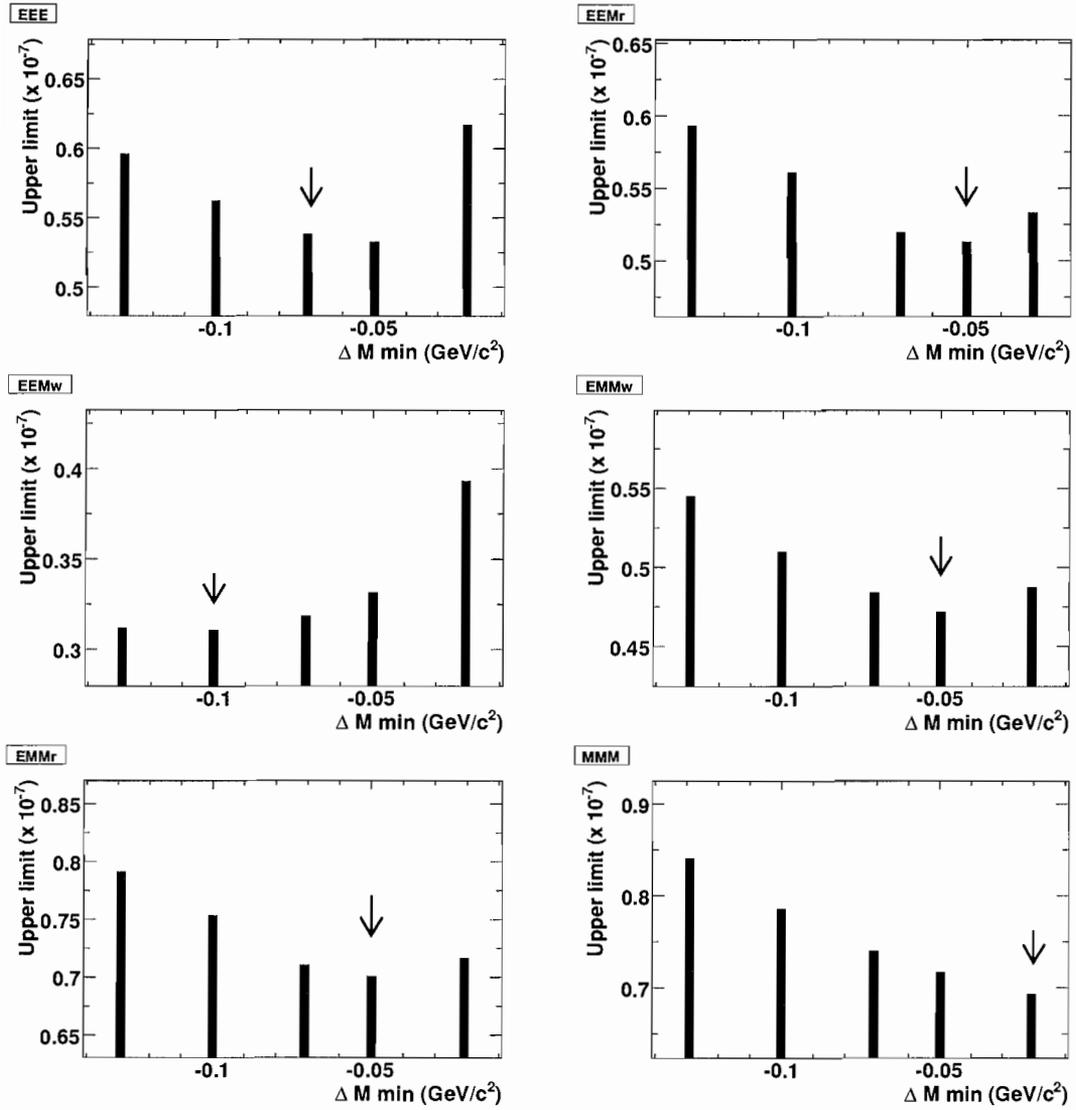


Figure C.3: Expected upper limit on branching fraction as a function of  $\Delta M_{min}^{SB}$ . Arrows indicate optimized value for the selection cut.

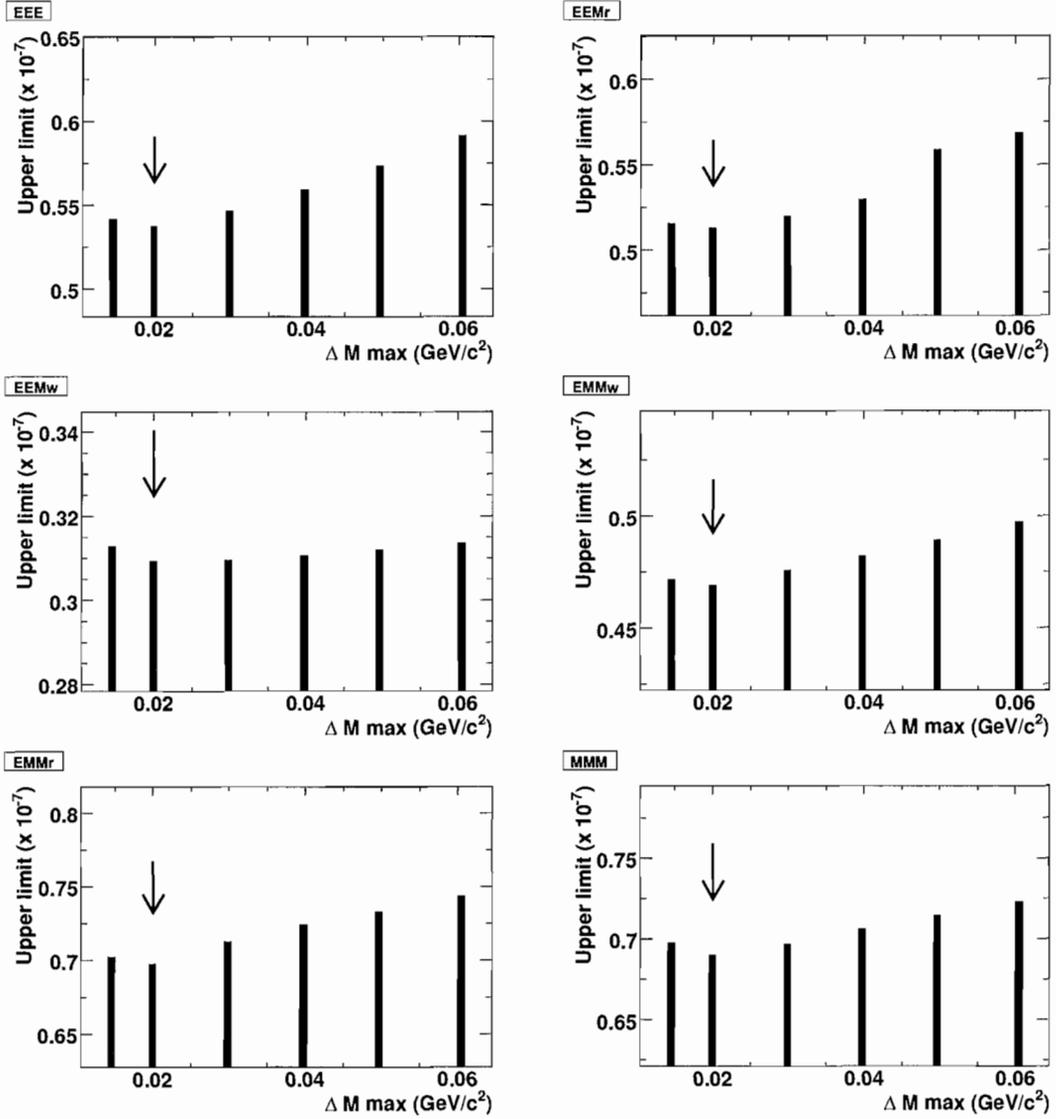


Figure C.4: Expected upper limit on branching fraction as a function of  $\Delta M_{max}^{SB}$ . Arrows indicate optimized value for the selection cut.

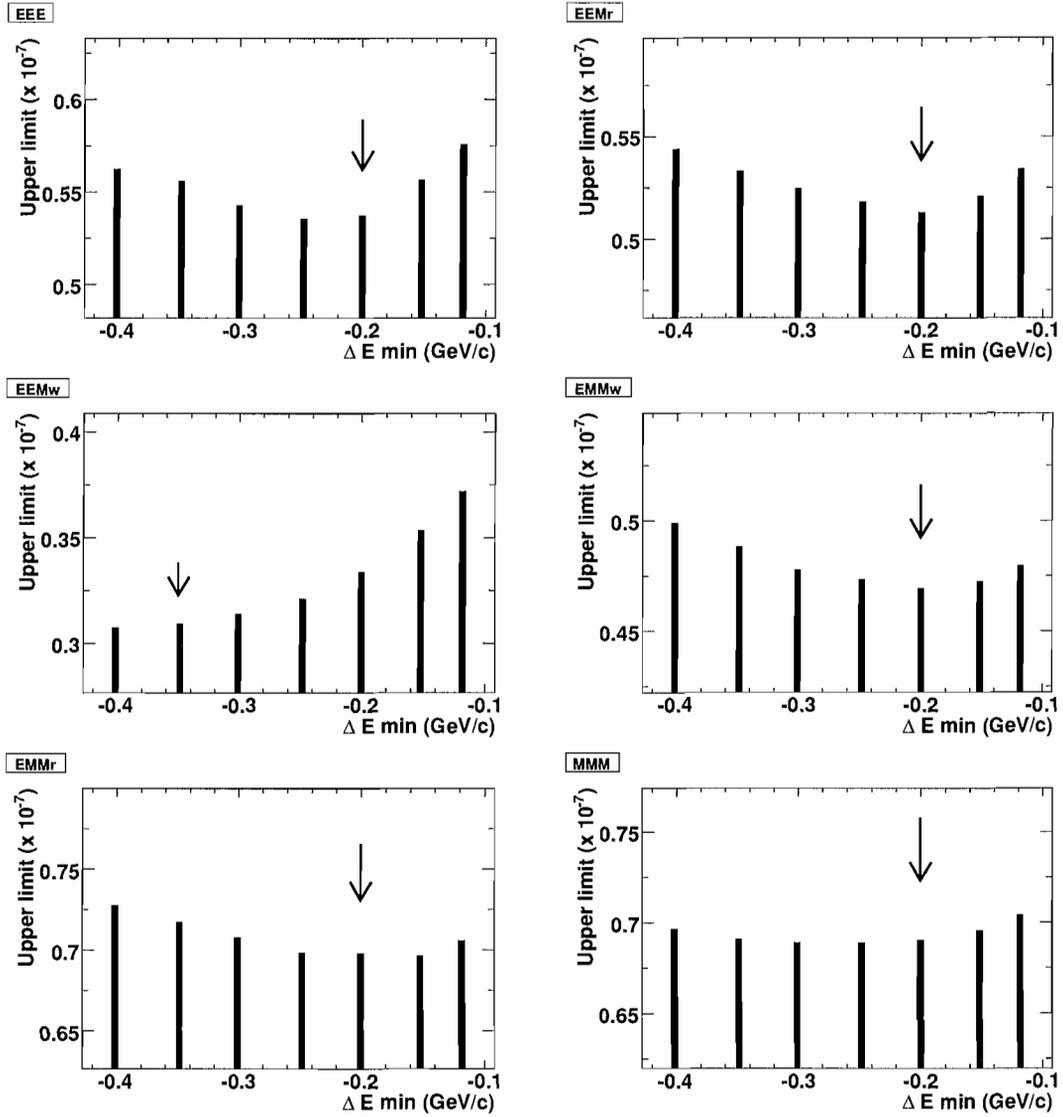


Figure C.5: Expected upper limit on branching fraction as a function of  $\Delta E_{min}^{SB}$ . Arrows indicate optimized value for the selection cut.

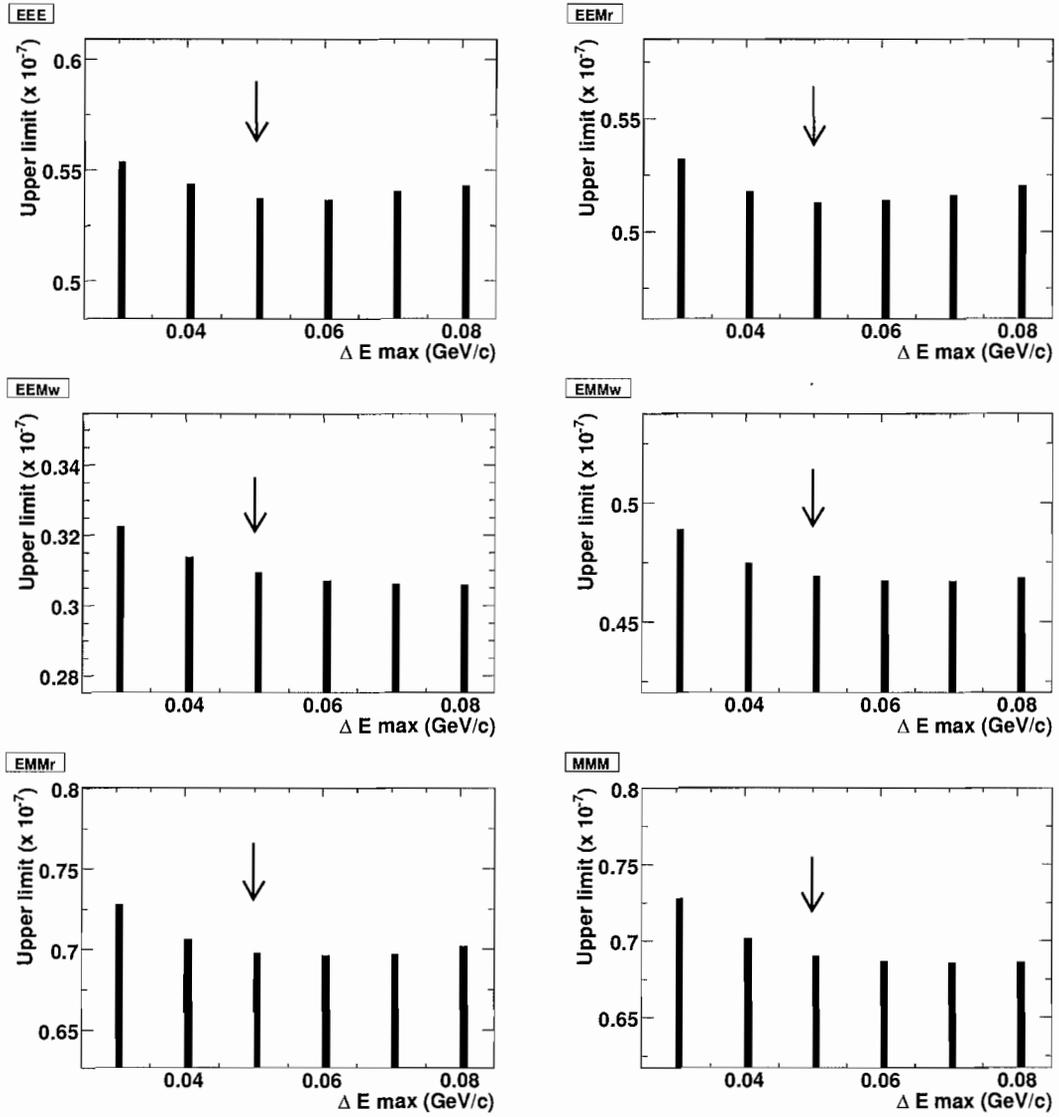


Figure C.6: Expected upper limit on branching fraction as a function of  $\Delta E_{max}^{SB}$ . Arrows indicate optimized value for the selection cut.

## 2 N-1 Plots

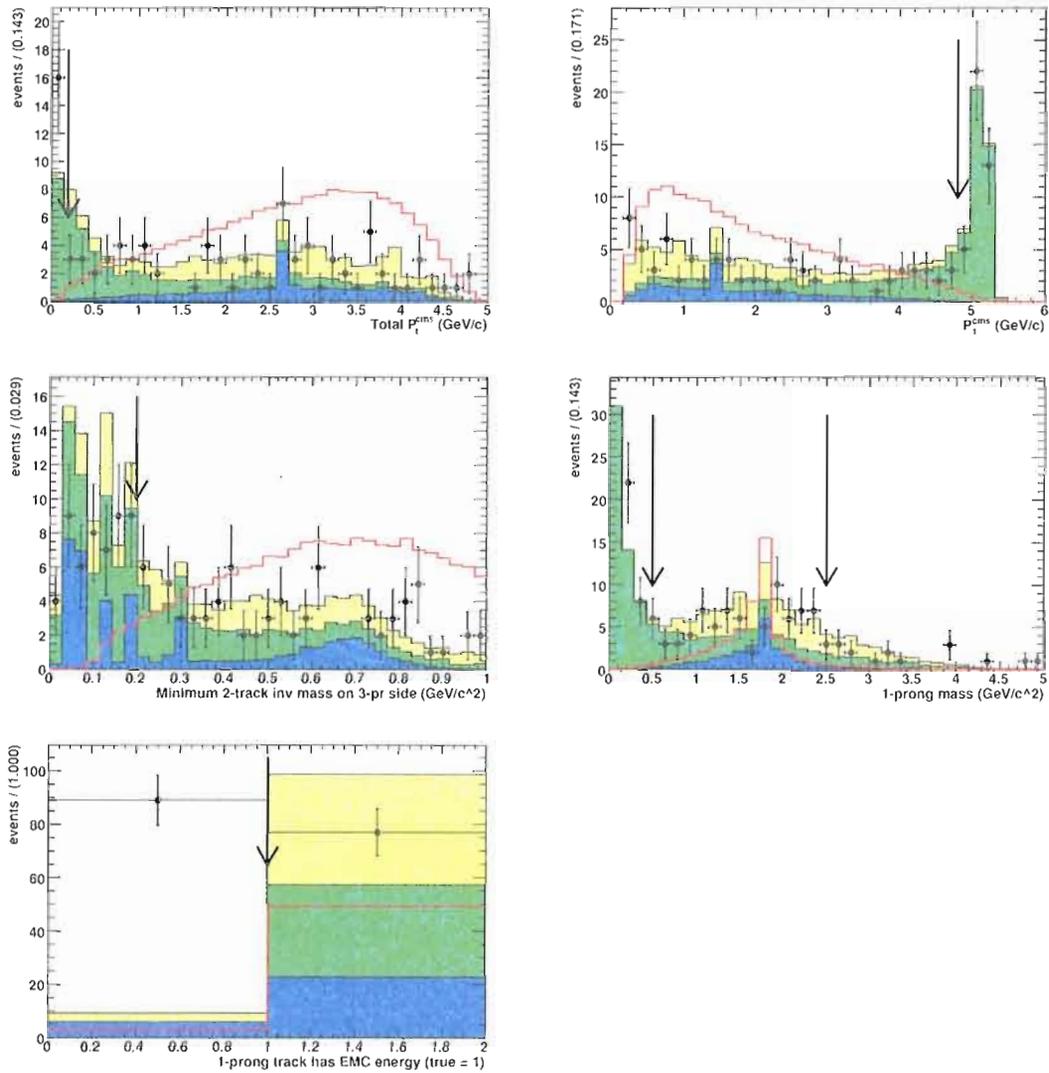


Figure C.7:  $\tau^- \rightarrow e^-e^+e^-$  channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+\tau^-$  background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

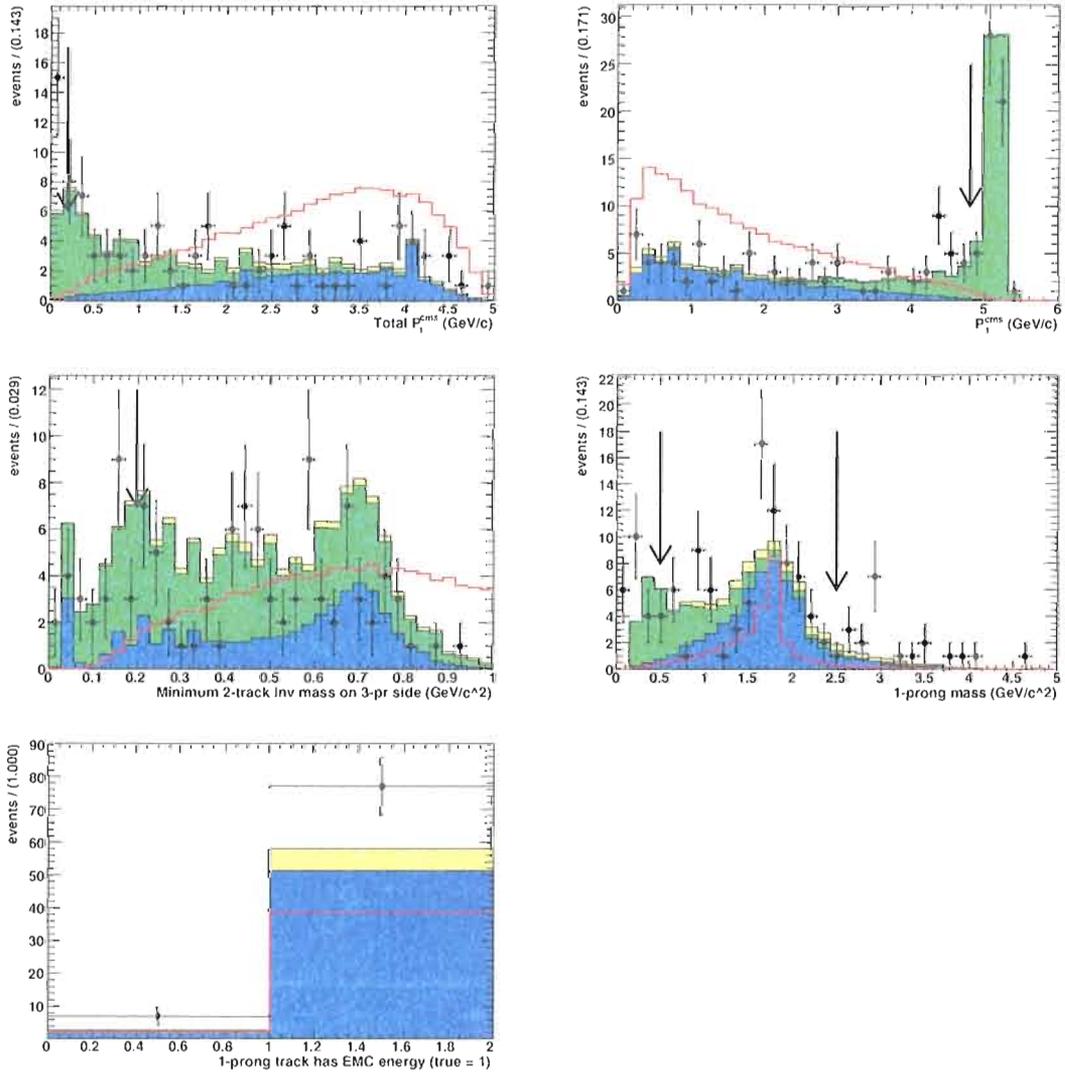


Figure C.8:  $\tau^- \rightarrow \mu^- e^+ e^-$  channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+ \tau^-$  background level, the green histogram shows the expected dimuon background level, and the yellow histogram shows the expected  $uds$  background level. All normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

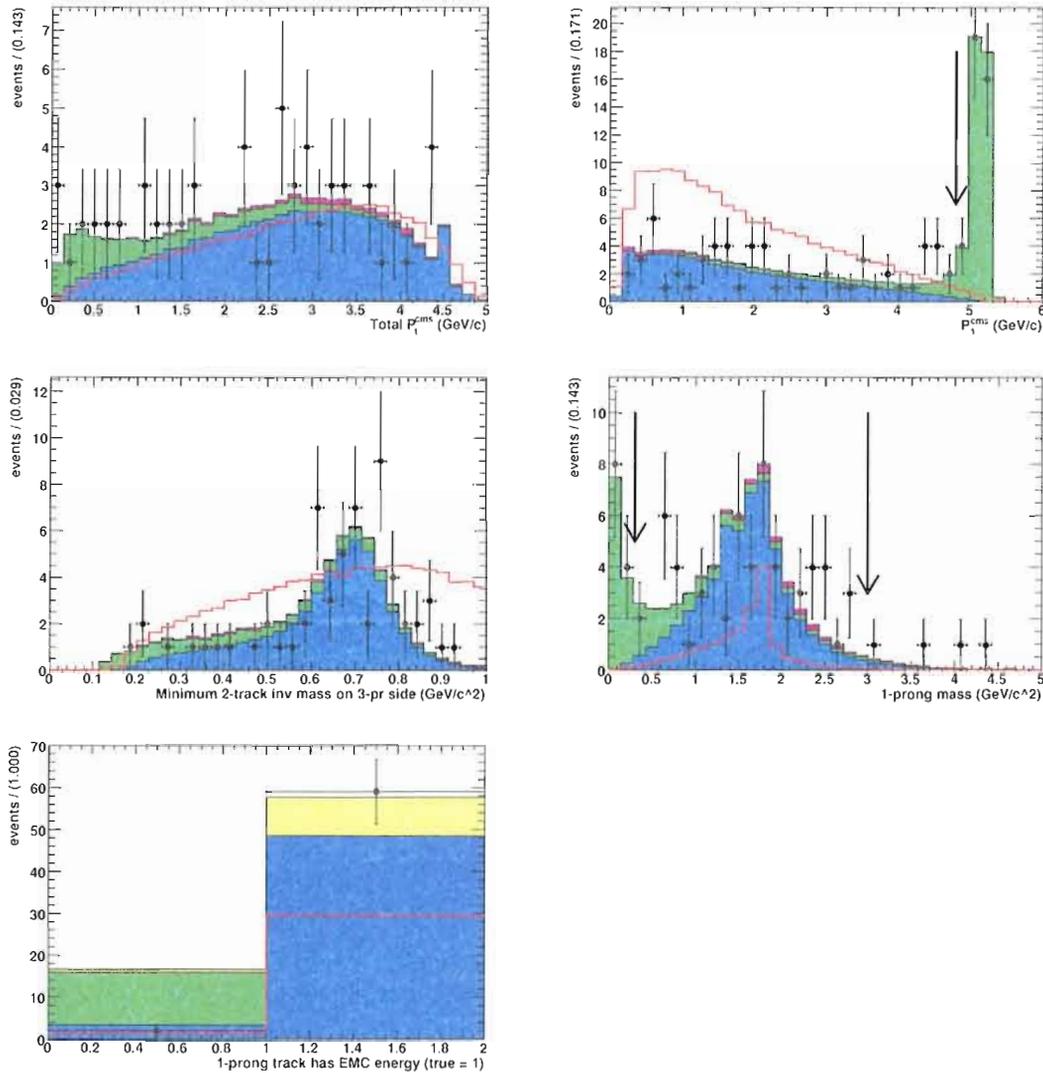


Figure C.9:  $\tau^- \rightarrow \mu^+ e^- e^-$  channel a) total transverse momentum; b) 1-prong momentum; c) min2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+\tau^-$  background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

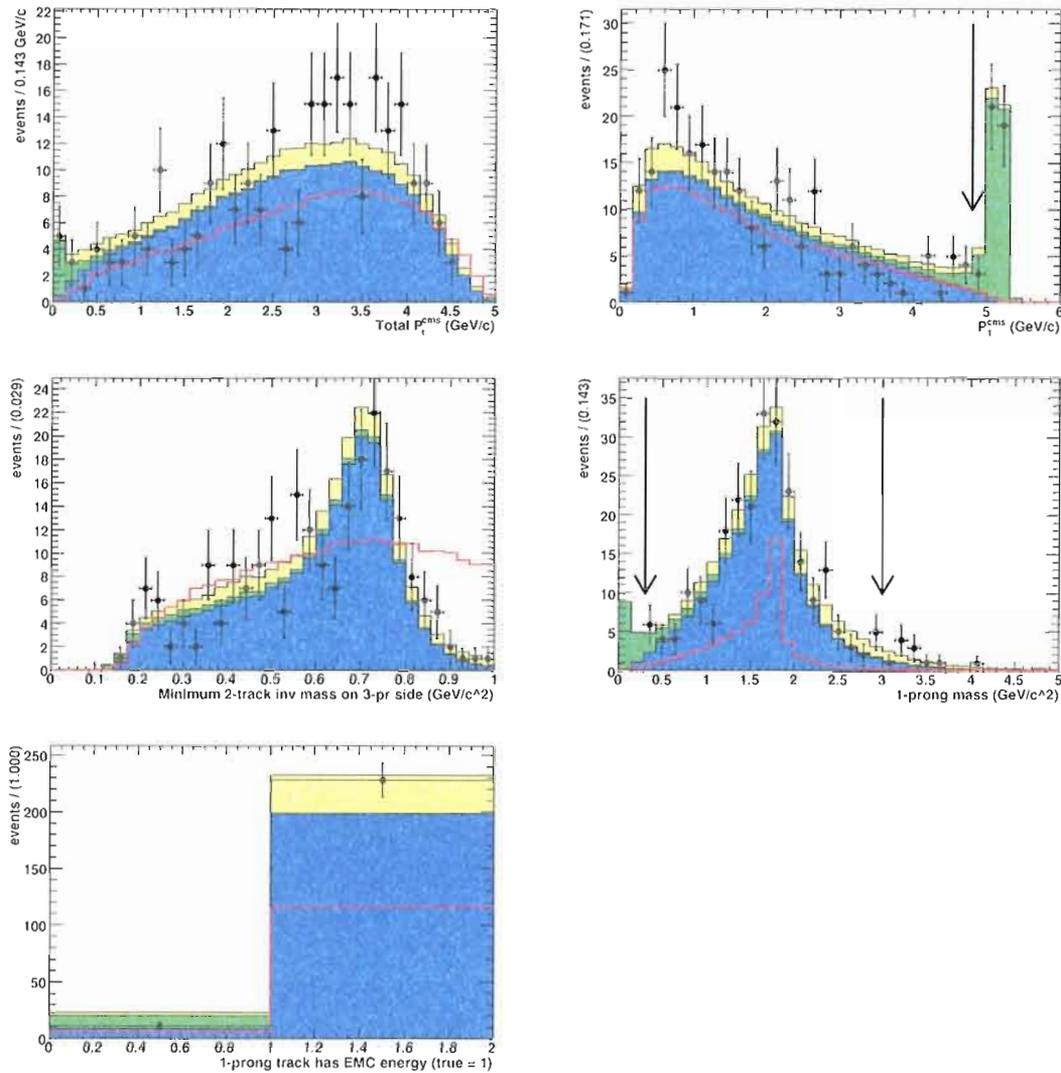


Figure C.10:  $\tau^- \rightarrow e^+ \mu^- \mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) min 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+ \tau^-$  background level, the green histogram shows the expected Bhabha background level, and the yellow histogram shows the expected  $uds$  background level. all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

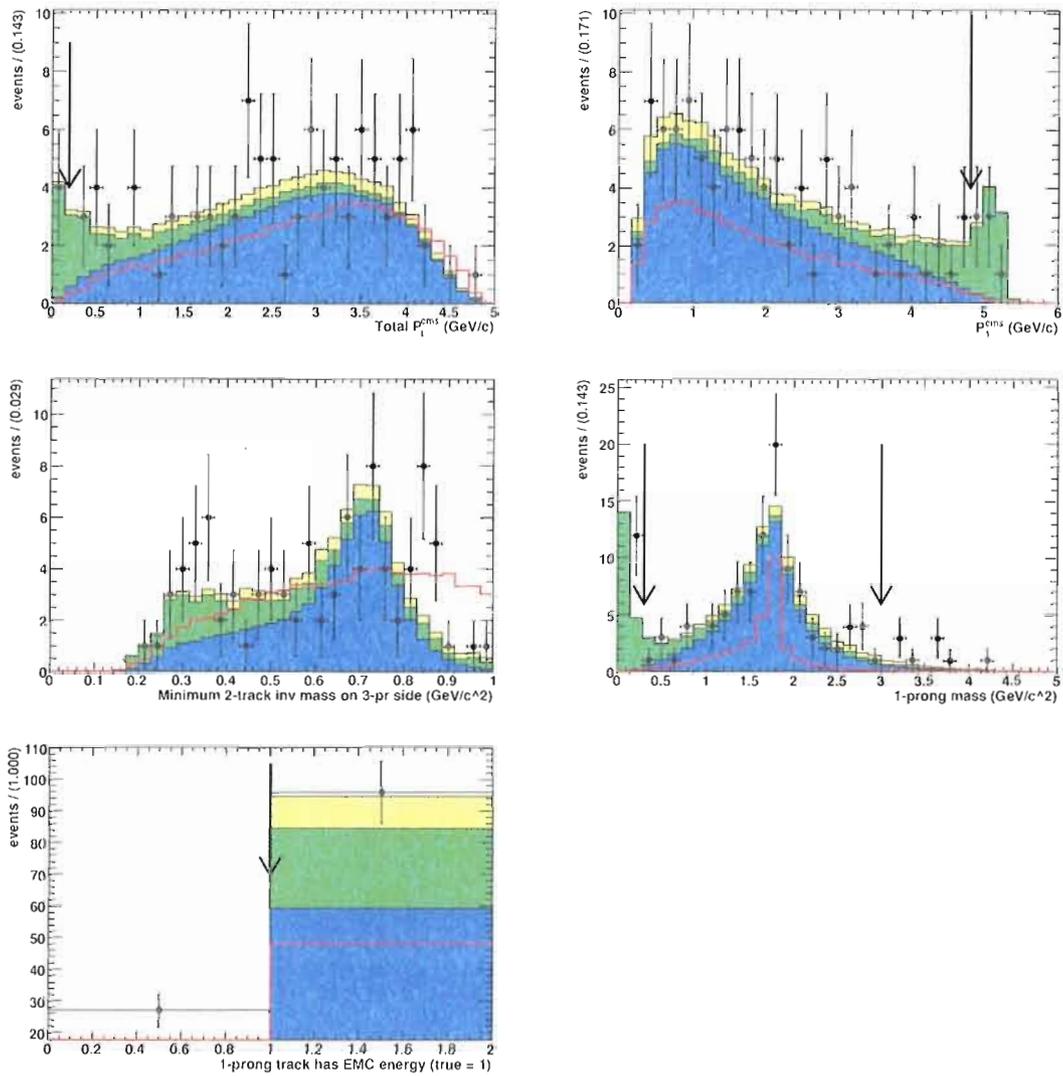


Figure C.11:  $\tau^- \rightarrow e^- \mu^+ \mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) minimum 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The green histogram shows the expected Bhabha background level and the yellow histogram shows the expected  $uds$  background level, all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

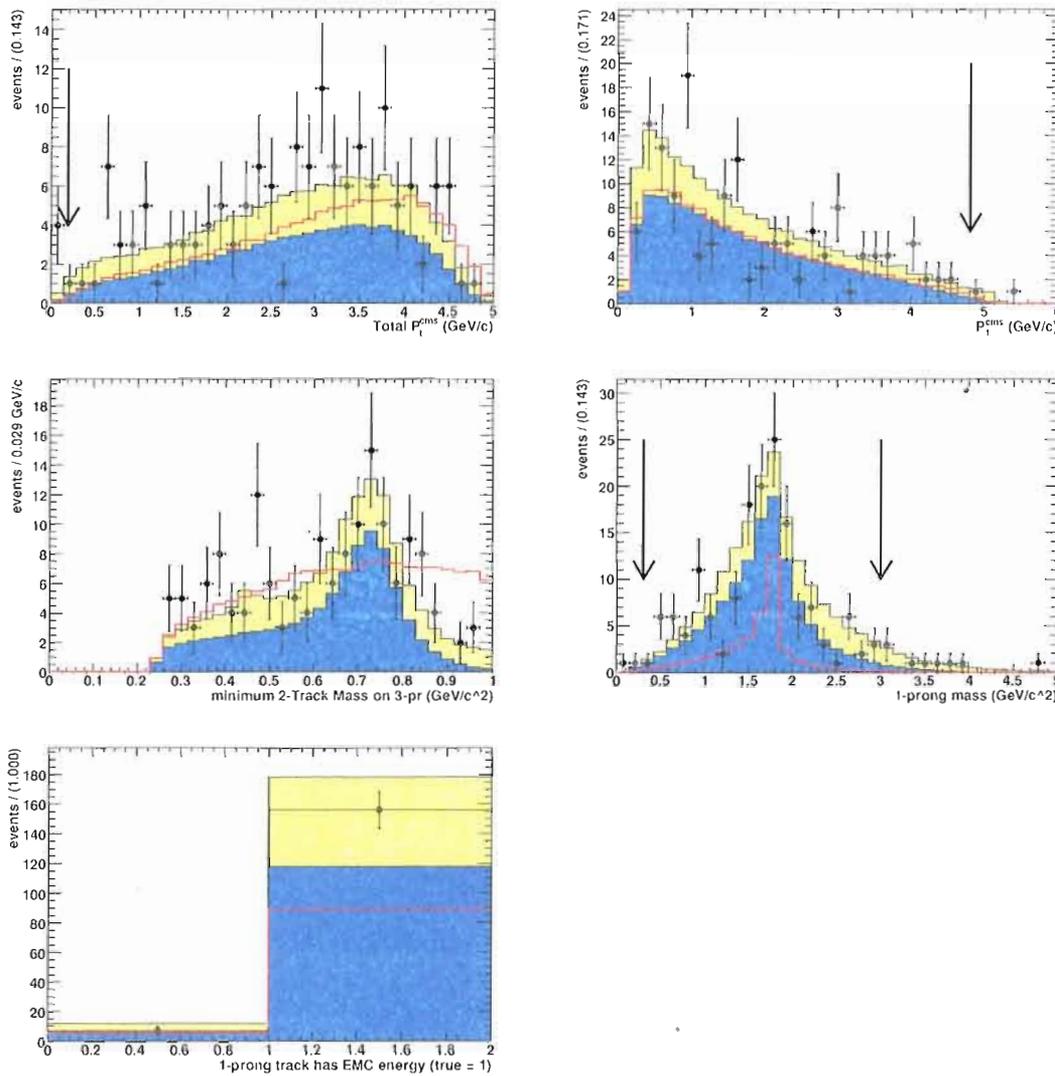


Figure C.12:  $\tau^- \rightarrow \mu^- \mu^+ \mu^-$  channel a) total transverse momentum; b) 1-prong momentum; c) minimum 2-track mass; d) 1-prong mass; e) (bool) one-prong track has EMC energy. The points show the data distributions for events in the grand sideband region with all other cuts applied. The blue histogram shows the expected  $\tau^+ \tau^-$  background level and the yellow histogram shows the expected  $uds$  background level. all normalized by the background fits with all selection cuts applied. Arrow(s) indicate the chosen cut value(s). For comparison, the red curve shows the MC signal distribution for the large box with arbitrary normalization.

### 3 Plots of Background Fits

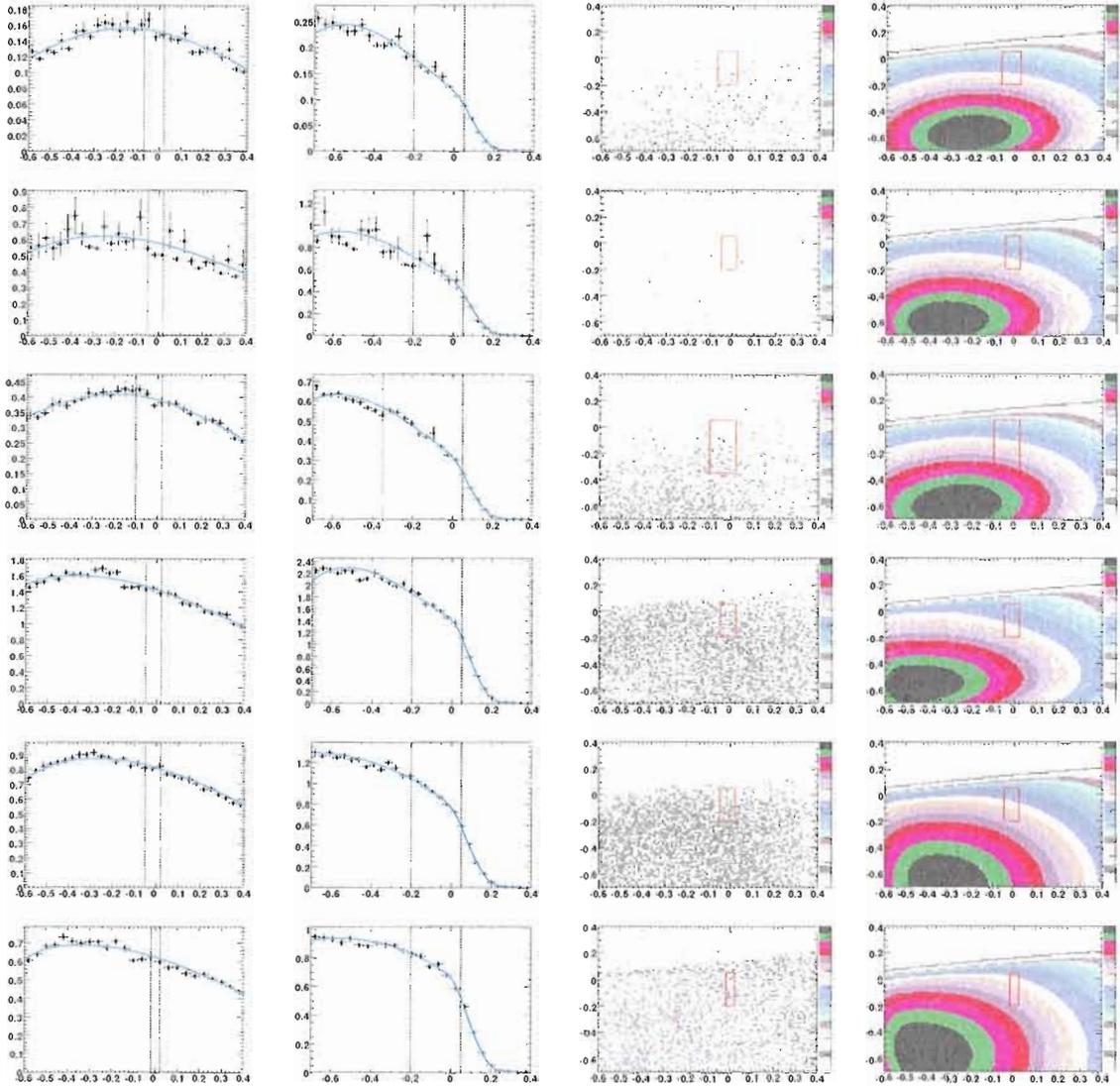


Figure C.13:  $uds$  background: Fit of the  $uds$  ( $\Delta M, \Delta E$ ) MC distribution with PDF described in the text. column 1)  $\Delta M$  projection of the MC distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the MC distribution (points) and the PDF (curve); column 3) MC  $(\Delta M, \Delta E)$  distribution; column 4) MC PDF  $(\Delta M, \Delta E)$  distribution; row 1)  $e^- e^+ e^-$ ; row 2)  $\mu^- e^+ e^-$ ; row 3)  $e^- \mu^+ e^-$ ; row 4)  $\mu^- e^+ \mu^-$ ; row 5)  $e^- \mu^+ \mu^-$ ; row 6)  $\mu^- \mu^+ \mu^-$ .  $\Delta M$  is plotted in  $(\text{GeV}/c^2)$  and  $\Delta E$  is plotted in  $(\text{GeV})$ .

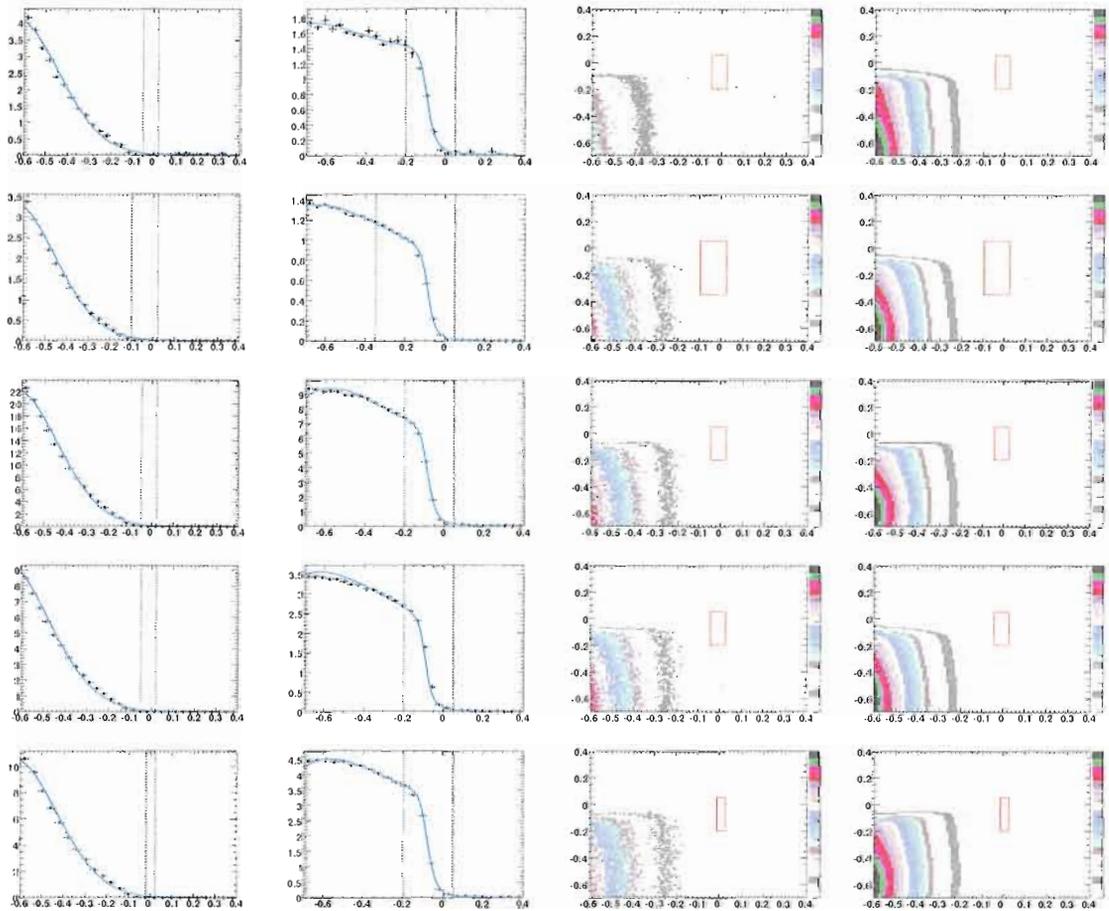


Figure C.14:  $\tau^+\tau^-$  background: Fit of the  $\tau^+\tau^-$  ( $\Delta M, \Delta E$ ) MC distribution with PDF described in the text. column 1)  $\Delta M$  projection of the MC distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the MC distribution (points) and the PDF (curve); column 3) MC ( $\Delta M, \Delta E$ ) distribution; column 4) MC PDF ( $\Delta M, \Delta E$ ) distribution: row 1)  $\mu^-e^+e^-$ ; row 2)  $e^-\mu^+e^-$ ; row 3)  $\mu^-e^+\mu^-$ ; row 4)  $e^-\mu^+\mu^-$ ; row 5)  $\mu^-\mu^+\mu^-$ .  $\Delta M$  is plotted in ( $\text{GeV}/c^2$ ) and  $\Delta E$  is plotted in ( $\text{GeV}$ ). Only channels with significant  $\tau^+\tau^-$  contributions in the LB are shown.

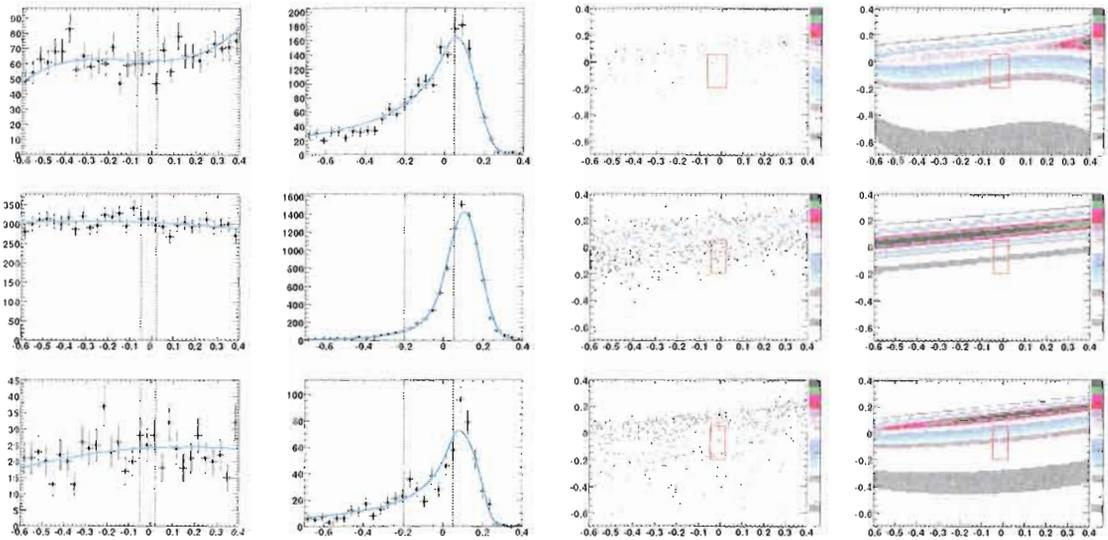


Figure C.15: Bhabha and di-muon backgrounds: Fit of the Bhabha and di-muon  $(\Delta M, \Delta E)$  MC distributions with PDF described in the text. column 1)  $\Delta M$  projection of the control sample distribution (points) and the PDF (curve); column 2)  $\Delta E$  projection of the control sample distribution (points) and the PDF (curve); column 3) control sample  $(\Delta M, \Delta E)$  distribution; column 4) PDF  $(\Delta M, \Delta E)$  distribution; row 1)  $e^-e^+e^-$ ; row 2)  $\mu^-e^+e^-$ ; row 3)  $e^-\mu^+\mu^-$ .  $\Delta M$  is plotted in  $(\text{GeV}/c^2)$  and  $\Delta E$  is plotted in  $(\text{GeV})$ . Only channels with significant QED contributions in the LB are shown.

## REFERENCES

- [1] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev.* **D2**, 1285 (1970).
- [2] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967).
- [3] K. Ichikawa, M. Fukugita, and M. Kawasaki, *Phys. Rev.* **D71**, 043001 (2005), [astro-ph/0409768](#).
- [4] X.-Y. Pham, *Eur. Phys. J.* **C8**, 513 (1999), [hep-ph/9810484](#).
- [5] P. Paradisi, *JHEP* **08**, 047 (2006), [hep-ph/0601100](#).
- [6] E. P. Hincks and B. Pontecorvo, *Phys. Rev.* **73**, 257 (1948).
- [7] T. O'Keefe, M. Rigby, and J. Wormald, *Proc. Phys. Soc. (London)* **173**, 951 (1959).
- [8] H. F. Davis, A. Roberts, and T. F. Zipf, *Phys. Rev. Lett.* **2**, 211 (1959).
- [9] D. Berley, J. Lee, and M. Bardon, *Phys. Rev. Lett.* **2**, 357 (1959).
- [10] J. Ashkin, T. Fazzini, G. Fidecaro, N. H. Lipman, A. W. Merrison, and J. Paul, *Nuovo Cimento* **14**, 1266 (1959).
- [11] S. Frankel, V. Hagopian, J. Halpern, and A. L. Whetstone, *Phys. Rev.* **118**, 589 (1960).
- [12] D. Bartlett, S. Devons, and A. M. Sachs, *Phys. Rev. Lett.* **8**, 120 (1962).
- [13] S. Frankel, J. Halpern, L. Holloway, W. Wales, M. Yearian, O. Chamberlain, A. Lemonick, and F. M. Pipkin, *Phys. Rev. Lett.* **8**, 123 (1962).
- [14] J. Lee and N. P. Samios, *Phys. Rev. Lett.* **3**, 55 (1959).

- [15] A. Babaev, M. Balats, V. Kaftanov, L. Landsberg, V. Lyubimov, and O. Yu, Soviet Phys.-JETP **16**, 1397 (1963).
- [16] S. Parker, H. L. Anderson, and C. Rey, Phys. Rev. **133**, B768 (1964).
- [17] H. P. Povel et al., Phys. Lett. **B72**, 183 (1977).
- [18] A. van der Schaaf et al., Nucl. Phys. **A340**, 249 (1980).
- [19] U. Bellgardt et al. (SINDRUM), Nucl. Phys. **B299**, 1 (1988).
- [20] R. D. Bolton et al., Phys. Rev. **D38**, 2077 (1988).
- [21] M. L. Brooks et al. (MEGA), Phys. Rev. Lett. **83**, 1521 (1999), hep-ex/9905013.
- [22] M. L. Perl et al., Phys. Rev. Lett. **35**, 1489 (1975).
- [23] K. G. Hayes et al., Phys. Rev. **D25**, 2869 (1982).
- [24] H. Albrecht et al. (ARGUS), Phys. Lett. **B185**, 228 (1987).
- [25] S. Keh et al. (Crystal Ball), Phys. Lett. **B212**, 123 (1988).
- [26] T. J. V. Bowcock et al. (CLEO), Phys. Rev. **D41**, 805 (1990).
- [27] H. Albrecht et al. (ARGUS), Z. Phys. **C55**, 179 (1992).
- [28] J. E. Bartelt et al. (CLEO), Phys. Rev. Lett. **73**, 1890 (1994).
- [29] D. W. Bliss et al. (CLEO), Phys. Rev. **D57**, 5903 (1998), hep-ex/9712010.
- [30] P. Abreu et al. (DELPHI), Phys. Lett. **B359**, 411 (1995).
- [31] A. Bean et al. (CLEO), Phys. Rev. Lett. **70**, 138 (1993).
- [32] K. W. Edwards et al. (CLEO), Phys. Rev. **D55**, 3919 (1997).
- [33] G. Bonvicini et al. (CLEO), Phys. Rev. Lett. **79**, 1221 (1997), hep-ex/9704010.
- [34] R. Godang et al. (CLEO), Phys. Rev. **D59**, 091303 (1999), hep-ex/9902005.
- [35] S. Chen et al. (CLEO), Phys. Rev. **D66**, 071101 (2002), hep-ex/0208019.

- [36] B. Aubert et al. (BABAR), Phys. Rev. Lett. **92**, 121801 (2004),  
hep-ex/0312027.
- [37] Y. Yusa et al. (Belle), Phys. Lett. **B589**, 103 (2004), hep-ex/0403039.
- [38] K. Abe et al. (Belle), Phys. Rev. Lett. **92**, 171802 (2004), hep-ex/0310029.
- [39] Y. Enari et al. (Belle), Phys. Lett. **B622**, 218 (2005), hep-ex/0503041.
- [40] B. Aubert et al. (BABAR), Phys. Rev. Lett. **95**, 041802 (2005),  
hep-ex/0502032.
- [41] B. Aubert et al. (BaBar), Phys. Rev. Lett. **95**, 191801 (2005), hep-ex/0506066.
- [42] Y. Miyazaki et al. (BELLE), Phys. Lett. **B632**, 51 (2006), hep-ex/0508044.
- [43] Y. Miyazaki et al. (BELLE), Phys. Lett. **B639**, 159 (2006), hep-ex/0605025.
- [44] B. Aubert et al. (BABAR), Phys. Rev. Lett. **96**, 041801 (2006),  
hep-ex/0508012.
- [45] Y. Yusa et al. (BELLE), Phys. Lett. **B640**, 138 (2006), hep-ex/0603036.
- [46] B. Aubert et al. (BABAR), Phys. Rev. Lett. **98**, 061803 (2007),  
hep-ex/0610067.
- [47] Y. Miyazaki et al. (BELLE), Phys. Lett. **B648**, 341 (2007), hep-ex/0703009.
- [48] B. Aubert et al. (BABAR), Phys. Rev. Lett. **99**, 251803 (2007), 0708.3650.
- [49] B. Aubert et al. (BABAR), Nucl. Instrum. Meth. **A479**, 1 (2002),  
hep-ex/0105044.
- [50] S. Jadach, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **130**, 260  
(2000), hep-ph/9912214.
- [51] S. Jadach, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **79**, 503  
(1994).
- [52] S. Jadach and Z. Was, Comput. Phys. Commun. **85**, 453 (1995).
- [53] D. J. Lange, Nucl. Instrum. Meth. **A462**, 152 (2001).

- [54] T. Sjostrand (1995), [hep-ph/9508391](#).
- [55] Z. Was, Nucl. Phys. Proc. Suppl. **98**, 96 (2001), [hep-ph/0011305](#).
- [56] W. M. Yao et al. (Particle Data Group), J. Phys. **G33**, 1 (2006).
- [57] E. Barberio and Z. Was, Comput. Phys. Commun. **79**, 291 (1994).
- [58] S. Agostinelli et al. (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003).
- [59] K. Kinoshita, Nucl. Instrum. Meth. **A276**, 242 (1989).
- [60] G. Cowan, *Statistical Data Analysis* (Oxford University Press, 1998).
- [61] J. Neyman, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences **236**, 333 (1937), ISSN 00804614, URL <http://www.jstor.org/stable/91337>.
- [62] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998), [physics/9711021](#).
- [63] R. D. Cousins and V. L. Highland, Nucl. Instrum. Meth. **A320**, 331 (1992).
- [64] R. Barlow, Comput. Phys. Commun. **149**, 97 (2002), [hep-ex/0203002](#).
- [65] S. Eidelman et al. (Particle Data Group), Phys. Lett. **B592**, 1 (2004).
- [66] e. Harrison, P. F. and e. Quinn, Helen R. (BABAR) (1998), 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.
- [67] B. F. L. Ward, S. Jadach, and Z. Was, Nucl. Phys. Proc. Suppl. **116**, 73 (2003), [hep-ph/0211132](#).
- [68] M. S. Alam et al. (CLEO), Phys. Rev. Lett. **76**, 2637 (1996).
- [69] S. Banerjee, B. Pietrzyk, J. M. Roney, and Z. Was, Phys. Rev. **D77**, 054012 (2008), [0706.3235](#).
- [70] M. Blanke, A. J. Buras, B. Duling, A. Poschenrieder, and C. Tarantino, JHEP **05**, 013 (2007), [hep-ph/0702136](#).

- [71] C.-X. Yue, S. Zhao, and W. Ma, Nucl. Phys. **B784**, 36 (2007), 0706.0232.
- [72] E. Arganda and M. J. Herrero, Phys. Rev. **D73**, 055003 (2006), hep-ph/0510405.
- [73] J. R. Ellis, J. Hisano, M. Raidal, and Y. Shimizu, Phys. Rev. **D66**, 115013 (2002), hep-ph/0206110.
- [74] A. Dedes, J. R. Ellis, and M. Raidal, Phys. Lett. **B549**, 159 (2002), hep-ph/0209207.
- [75] A. Brignole and A. Rossi, Phys. Lett. **B566**, 217 (2003), hep-ph/0304081.
- [76] K. S. Babu and C. Kolda, Phys. Rev. Lett. **89**, 241802 (2002), hep-ph/0206310.
- [77] A. Brignole and A. Rossi, Nucl. Phys. **B701**, 3 (2004), hep-ph/0404211.
- [78] K. Hayasaka et al. (Belle) (2007), 0705.0650.