

**Study of Exclusive Semileptonic B Meson
Decays to Tau Leptons**

Michael A. Mazur

SLAC-R-882

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

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

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

UNIVERSITY OF CALIFORNIA
Santa Barbara

Study of Exclusive Semileptonic B Meson
Decays to Tau Leptons

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

Doctor of Philosophy

in

Physics

by

Michael A. Mazur

Committee in Charge:

Professor Jeffrey Richman, Chair

Professor David Berenstein

Professor Claudio Campagnari

September 2007

The Dissertation of
Michael A. Mazur is approved:

Professor David Berenstein

Professor Claudio Campagnari

Professor Jeffrey Richman, Committee Chairperson

August 2007

Study of Exclusive Semileptonic B Meson Decays to Tau Leptons

Copyright © 2007

by

Michael A. Mazur

Acknowledgements

Thank you...

To Jeff Richman, for putting an idea into my head and for letting me run wild with this analysis, no matter how long it took and how impossible it may have seemed at times.

To Claudio Campagnari, to Harry Nelson, to Yoram Rozen, and to David Stuart for all the time spent discussing physics and everything else that goes along with it. To every teacher I've had along the way.

To everyone at SLAC who took a moment to answer my questions, however naïve they may have been; to everyone at SLAC who came to me with questions, however naïve you may have felt they were. I learned so much from all of you.

To everybody in Pisa, for their infinite patience and all their help throughout the years.

To my friends in the UCSB group, to Tae Min, to Adam, to Steve, to Owen, to Wouter, and to Jeff, it's been a pleasure sharing offices and spending time together.

To my non-UCSB friends, to Stephanie and to Kris, to everybody in the ROB who made my time there a bit more enjoyable, even if you didn't know it.

To the SLAC tours program and to everybody who came to visit the lab, asking me questions and laughing at my jokes — you reminded me that physics can be fun. To Emily and to Barb and to Mika and to Maura and to Melinda and to Kelen and to Ellie, for being such great company.

To my friend Colleen, when I needed it the most, you've been there for me

To Bryan, for all the cribbage, all the Ms. Pac-Man, the endless Chinese food, all for the adventures through the years.

To Mom and Dad, Jonathan and Jessica for putting up with me and my carrots for all these years.

alla vir, per chi sei, per tutto che abbiamo vissuto insieme e per tutto che verrà, per avermi completato, per avermi conosciuto come un pollo, per il modo in cui mi fai ridere, per il 2% che non basta mai. Spero solo di meritarti per sempre.

Curriculum Vitæ

Michael A. Mazur

Education

2002 M.A. University of California, Santa Barbara, CA
1999 M.S. University of Chicago, Chicago, IL
1998 B.S. Worcester Polytechnic Insititute, Worcester, MA

Research

2000 – 2007 Member of the *BABAR* Experiment. Advisor Jeffrey Richman
1998 – 1999 Member of the KTeV Experiment. Advisor Yau Wah

Selected Publications

“Measurement of the Semileptonic Decays $B \rightarrow D\tau^-\bar{\nu}_\tau$ and $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ ”, B. Aubert, et al. (The *BABAR* Collaboration), submitted to the 2007 Europhysics Conference on High Energy Physics.

“Measurement of the Average ϕ Multiplicity in B Meson Decay”, B. Aubert, et al. (The *BABAR* Collaboration), Phys. Rev. D **69** 052005 (2004).

Abstract

Study of Exclusive Semileptonic B Meson Decays to Tau Leptons

Michael A. Mazur

We present the results of a search for four exclusive semileptonic decays $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ in 209 fb⁻¹ of data collected with the *BABAR* detector, corresponding to 232 million $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events. We select events with a $D^{(*)}$ meson and a light lepton (e or μ) recoiling against a fully-reconstructed B meson. We perform a fit to the lepton spectrum and missing mass squared to discriminate signal events from backgrounds, predominantly $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$. A control sample of identified $D^{**}\ell^-\bar{\nu}_\ell$ events is included in the fit to estimate the background contribution from these decays. We measure $\mathcal{B}(B \rightarrow D\tau\nu) = (0.86 \pm 0.24 \pm 0.11 \pm 0.06)\%$ and $\mathcal{B}(B \rightarrow D^*\tau\nu) = (1.62 \pm 0.31 \pm 0.10 \pm 0.05)\%$, where the errors are statistical, systematic, and normalization-mode related, respectively, and where the results are expressed for the \bar{B}^0 lifetime.

Professor Jeffrey Richman
Dissertation Committee Chair

Contents

Acknowledgments	iv
Curriculum Vitæ	vi
Abstract	vii
List of Figures	xiii
List of Tables	xix
1 Introduction	1
1.1 Organization of this Document	2
1.2 Standard Model Predictions	3
1.3 Previous Measurements	5
1.4 Sensitivity to New Physics	7
1.5 Methodology	11
1.6 Notation and Conventions	15
2 Theory of Semileptonic B Decays to τ Leptons	17
2.1 Formalism of Exclusive Semileptonic Decays	18
2.1.1 Decay Kinematics	19
2.1.2 Decay Amplitude	21
2.1.3 Hadronic current for $B \rightarrow D$	23
2.1.4 Hadronic current for $B \rightarrow D^*$	24
2.2 Differential Decay Rate Formulae	25
2.2.1 Differential Decay Rate for $B \rightarrow D\ell^-\bar{\nu}_\ell$	25
2.2.2 Differential Decay Rate for $B \rightarrow D^*\ell^-\bar{\nu}_\ell$	27
2.3 Form Factor Predictions	30

2.3.1	Heavy Quark Effective Theory	32
2.3.2	The CLN Form Factor Model	35
2.4	Kinematic Distributions of Semileptonic Decays	36
2.5	Non-Standard Model Contributions	51
3	The <i>BABAR</i> Experiment	57
3.1	The PEP-II <i>B</i> Factory	57
3.2	The <i>BABAR</i> Detector	60
3.2.1	Silicon Vertex Tracker	61
3.2.2	Drift Chamber	66
3.2.3	Detector of Internally-Reflected Cherenkov Light	70
3.2.4	Electromagnetic Calorimeter	72
3.2.5	Instrumented Flux Return	74
4	Data and MC Samples	75
4.1	<i>BABAR</i> Data Samples	75
4.2	Simulated Event Samples	75
4.3	Semileptonic and Semitauonic Decay Models	77
4.3.1	HQET model	78
4.3.2	ISGW2 Model	79
4.3.3	Goity–Roberts Model	79
4.3.4	$D^{*\ell^-}\bar{\nu}_\ell$ Decay Branching Fractions	80
5	Background Classification	81
5.1	$B \rightarrow X_c \ell^- \bar{\nu}_\ell$ Background	84
5.2	Combinatoric Background	86
5.3	Feed-down and Feed-up	88
6	Event Selection and Optimization	90
6.1	Overview	90
6.2	BSemiExcl Skim	92
6.3	Vcb Reskim	95
6.4	B_{tag} Selection	98
6.5	Charged Particle Identification	102
6.6	Lepton Identification	102
6.6.1	Bremsstrahlung Recovery	103
6.7	Composite Particle Reconstuction	105
6.7.1	π^0 Reconstruction	105
6.7.2	K_S^0 Reconstruction	106
6.7.3	$D^{(*)}$ Meson Reconstruction	106

6.8	Vertexing and Kinematic Fitting	106
6.9	p_{miss}	108
6.10	q^2	108
6.11	Extra Track Veto	109
6.12	E_{extra}	109
6.13	Single Candidate Selection	113
6.14	m_{miss}^2	116
6.15	Signal Efficiencies	116
6.16	Cut-and-Count Event Yields	119
7	Control Samples and MC Validation	123
7.1	$D^{(*)}\ell^{-}\bar{\nu}_{\ell}$ Control Samples	124
7.2	q^2 Control Samples	132
7.3	D^{**} Control Samples	134
8	Fit Procedure	137
8.1	Likelihood Function	139
8.2	Signal and Background Components	140
8.3	p_{ℓ}^* Distribution for Semileptonic and Semitauonic Events	144
8.4	$p_{\ell}^*-m_{\text{miss}}^2$ PDFs for Semileptonic and Semitauonic Events	148
8.5	PDFs from Signal and Semileptonic MC	152
8.5.1	Reweighting $D^{(*)}\tau^{-}\bar{\nu}_{\tau}$ Feed-up into the D^{**} Control Samples	152
8.5.2	Reweighting to $D^{(*)}\tau^{-}\bar{\nu}_{\tau}$ to Simulate $D^{**}\tau^{-}\bar{\nu}_{\tau}$	154
8.5.3	Generically-simulated $D^{**}\ell^{-}\bar{\nu}_{\ell}$ Decays	155
8.6	Fits	159
8.7	$p_{\ell}^*-m_{\text{miss}}^2$ PDFs for Combinatoric Events	192
8.8	Crossfeed Constraints in the Fit	199
8.9	D^* Efficiency Correction and Floating Constraints	200
8.10	Fit Results on Generic Monte Carlo Events	205
9	Fit Validation	214
9.1	Toy Monte Carlo Studies	214
9.2	Validation with Statistically Independent Monte Carlo Samples	218
10	Signal Extraction	220
10.1	Fitting Branching Ratios	220
10.2	Relative Signal Efficiency and Normalization	221
10.3	Combining Results between Charged and Neutral Channels	223
10.4	Validation using the Generic Monte Carlo Sample	224

11	Corrections to the Monte Carlo	226
11.1	D^{**} Modeling	227
11.2	$B \rightarrow D^{(*)}$ Form Factors	227
11.3	Detector Efficiency Corrections	228
11.3.1	Tracking Efficiency	228
11.3.2	PID Efficiency	229
11.3.3	π^0 Efficiency	229
11.3.4	K_S^0 Efficiency	230
11.4	Corrected Crossfeed constraints	230
12	Systematic Uncertainties	232
12.1	Systematic Error Overview	232
12.2	$D^{**}\ell^-\bar{\nu}_\ell$ Modeling	233
12.3	$B \rightarrow D^*$ Form Factors	238
12.4	$B \rightarrow D$ Form Factors	249
12.5	Combinatoric Background Modeling	251
12.6	Monte Carlo Statistics	255
12.6.1	PDF Shapes	256
12.6.2	Crossfeed Constraints	257
12.7	$D^{**}\tau^-\bar{\nu}_\tau$ Abundance	258
12.8	Detector Efficiency	259
12.8.1	Tracking Efficiency	260
12.8.2	PID Efficiency	261
12.8.3	K_S^0 Efficiency	262
12.8.4	Neutral Efficiency	265
12.8.5	Neutral Efficiency and $D^{**}\ell^-\bar{\nu}_\ell$ Feed-down Constraints	267
12.9	m_{miss}^2 Tail Modeling	269
12.10	Bremsstrahlung and Final-State Radiation	273
12.11	Hadronic Daughter Branching Fractions	274
12.12	Tau Branching Fraction	274
13	Unblinded Data	275
13.1	Numerical Results and Fit Projections	275
13.2	Significance of Signals	286
13.3	Stability Checks	294
13.4	Goodness of Fit	295
14	Results and Conclusions	298
	Bibliography	300

Appendices	304
A Definitions of Standard Lists	305
A.1 Raw Lists	305
A.2 Refined Track Lists	306
A.3 Refined Neutral Lists	306
A.4 Electron ID	307
A.5 Muon ID	309
A.5.1 muMicroTight	309
A.5.2 Neural Net Muon ID	311
A.6 Kaon ID	312
A.7 K_S^0 Reconstruction	313
A.8 π^0 Reconstruction	314
A.8.1 pi0VeryLoose	314
A.8.2 Pi0AllDefault	314
B $D^{(*)}$ Candidate Selection	316
B.1 D Candidates	317
B.2 D^* Candidates	317
B.3 $D^{(*)}$ Mass Distributions	319
C Good Track and Neutral Selection	327
C.1 Track Selection	327
C.1.1 Looped Tracks	328
C.1.2 Ghost Tracks	330
C.2 Neutral Selection	333
C.2.1 Unmatched Track Rejection	333

List of Figures

1.1	Standard model Feynman diagram for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$	5
1.2	Measurements of the $b \rightarrow X_c\tau^-\bar{\nu}_\tau$ from LEP experiments	6
1.3	Inclusive $B \rightarrow X\tau^-\bar{\nu}_\tau$ missing energy spectrum from ALEPH	8
1.4	Feynman diagram for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ in a two-Higgs doublet model	9
1.5	Sensitivity of exclusive measurements to a charged Higgs boson	10
1.6	Schematic view of a signal event	12
2.1	Definition of the decay angles θ_ℓ , θ_V , and χ	20
2.2	q^2 distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D\tau^-\bar{\nu}_\tau$	37
2.3	q^2 distributions for $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	38
2.4	Lepton energy distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D\tau^-\bar{\nu}_\tau$	39
2.5	Lepton energy distributions for $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	40
2.6	m_{miss}^2 distribution for $B \rightarrow D\tau^-\bar{\nu}_\tau$	41
2.7	m_{miss}^2 distribution for $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	42
2.8	$\cos\theta_\ell$ distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D\tau^-\bar{\nu}_\tau$	43
2.9	$\cos\theta_\ell$ distributions for $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	44
2.10	Dalitz plot for $B \rightarrow D\ell^-\bar{\nu}_\ell$	45
2.11	Dalitz plot for $B \rightarrow D^*\ell^-\bar{\nu}_\ell$	46
2.12	Dalitz plot for $B \rightarrow D\tau^-\bar{\nu}_\tau$	47
2.13	Dalitz plot for $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	48
2.14	Pseudo-Dalitz plot for $B \rightarrow D\tau^-\bar{\nu}_\tau$	49
2.15	Pseudo-Dalitz plot for $B \rightarrow D^*\tau^-\bar{\nu}_\tau$	50
2.16	Branching ratio of $B \rightarrow D\tau^-\bar{\nu}_\tau$ in a 2HDM	53
2.17	Branching ratio of $B \rightarrow D^*\tau^-\bar{\nu}_\tau$ in a 2HDM	54
2.18	τ polarization in $B \rightarrow D\tau^-\bar{\nu}_\tau$ in a 2HDM	55
2.19	τ polarization in $B \rightarrow D^*\tau^-\bar{\nu}_\tau$ in a 2HDM	56

3.1	Layout of the SLAC linear accelerator, PEP-II storage rings, and <i>BABAR</i> detector	58
3.2	Layout of the <i>BABAR</i> detector, longitudinal cross-section	62
3.3	Layout of the <i>BABAR</i> detector, end view	63
3.4	End view of the SVT	64
3.5	Side view of the SVT	65
3.6	DCH dE/dx for particles of different momenta	69
3.7	Layout of the DIRC detector	71
3.8	Side view of the EMC	73
3.9	Layout of the IFR	74
6.1	m_{ES} and ΔE distributions for the BSemiExcl skim	94
6.2	Optimization of the $ \Delta E $ cut	99
6.3	Distributions of m_{ES} after all cuts, from generic MC	100
6.4	Distributions of ΔE after all cuts, from generic MC	101
6.5	Lepton momentum resolution as a function of polar angle	104
6.6	Optimization of the E_{extra} cut	111
6.7	E_{extra} distributions from generic MC	112
6.8	E_{extra} distributions from signal and semileptonic events	114
6.9	Same as Figure 6.8 with the zero bin suppressed	115
6.10	E_{extra} distributions split according to lepton type	115
6.11	m_{miss}^2 distributions from generic MC	117
6.12	Same as Figure 6.11 but with a logarithmic scale	118
7.1	Distributions of p_ℓ^* used to define the $D^{*0}\ell^-\bar{\nu}_\ell$ control sample	125
7.2	Data–MC comparison of p_ℓ^* in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples	128
7.3	Data–MC comparison of q^2 in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples	129
7.4	Data–MC comparison of E_{extra} in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples	130
7.5	Data–MC comparison of m_{miss}^2 in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples	131
7.6	Distributions of q^2 used to define the q^2 control sample	133
8.1	p_ℓ^* distributions for several components from Monte Carlo	146
8.2	Profile plot of m_{miss}^2 vs p_ℓ^* for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	149
8.3	Profile plot of m_{miss}^2 vs p_ℓ^* for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0$	149
8.4	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	160
8.5	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	161
8.6	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$	162
8.7	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$	163
8.8	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$	164
8.9	Fitted MC distribution of $p_\ell^*-m_{\text{miss}}^2$ for $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$	165

8.10	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^0$	166
8.11	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0$	167
8.12	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0$	168
8.13	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0$	169
8.14	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$	170
8.15	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$	171
8.16	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^{*+}$	172
8.17	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+}$	173
8.18	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^{*+}$	174
8.19	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^+$	175
8.20	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \tau^- \bar{\nu}_\tau \Rightarrow D^+$	176
8.21	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+$	177
8.22	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^+$	178
8.23	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+$	179
8.24	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^{*0} \pi^0$	180
8.25	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^{*0} \pi^0$	181
8.26	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^{*0} \pi^0$	182
8.27	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$	183
8.28	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$	184
8.29	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$	185
8.30	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^{*+} \pi^0$	186
8.31	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^{*+} \pi^0$	187
8.32	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+} \pi^0$	188
8.33	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$	189
8.34	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$	190
8.35	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$	191
8.36	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for charge crossfeed in the signal D channels	194
8.37	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for charge crossfeed in the signal D^* channels	195
8.38	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for other combinatoric back- ground in the signal channels	196
8.39	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for charge crossfeed in the D^{**} control sample	197
8.40	Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$	for other combinatoric back- ground in the D^{**} control sample	198
8.41	m_{miss}^2 projection of the fit to generic MC	in the signal channels	208
8.42	Same as Figure 8.41 (vertical zoom)		209
8.43	Same as Figure 8.41 (vertical zoom, horizontal scale expanded)		210

8.44	m_{miss}^2 projection of the fit to generic MC in the D^{**} control sample	211
8.45	p_ℓ^* projection of the fit to generic MC in the signal channels . .	212
8.46	p_ℓ^* projection of the fit to generic MC in the D^{**} control sample	213
9.1	Toy MC results for $D^0\tau\nu$	216
9.2	Toy MC results for $D^{*0}\tau\nu$	216
9.3	Toy MC results for $D^+\tau\nu$	217
9.4	Toy MC results for $D^{*+}\tau\nu$	217
12.1	Generated branching fraction distributions for D^{**} model reweighting	237
12.2	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$	239
12.3	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^0$	240
12.4	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*+}$	241
12.5	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^+$	242
12.6	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\pi^0$. . .	243
12.7	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^0\pi^0$. . .	244
12.8	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*+}\pi^0$. . .	245
12.9	D^{**} modeling variation for the channel $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^+\pi^0$. . .	246
12.10	Distributions of fit results for D^{**} modeling systematic error .	247
12.11	Number of <i>GoodTracksLoose</i> recoil tracks with $p_T > 200$ MeV/c, comparing signal and normalization	261
12.12	Recoil kaon and pion momentum spectra, comparing signal and normalization	263
12.13	Transverse decay length of K_S^0 candidates, comparing signal and normalization	264
13.1	m_{miss}^2 projection of the unconstrained fit to data, showing the normalization region in the signal channels	276
13.2	m_{miss}^2 projection of the unconstrained fit to data, showing a wider range than Figure 13.1	277
13.3	m_{miss}^2 projection of the unconstrained fit to data, showing the signal region in more detail	278
13.4	m_{miss}^2 projection of the unconstrained fit to data, showing the D^{**} control samples	279
13.5	p_ℓ^* projection of the unconstrained fit to data, showing the signal channels	280
13.6	p_ℓ^* projection of the unconstrained fit to data, showing the D^{**} control samples	283

13.7	p_ℓ^* projection of the unconstrained fit to data, showing the signal channels in the signal-enhanced region $m_{\text{miss}}^2 > 1 \text{ (GeV}/c^2)^2$	284
13.8	p_ℓ^* projection of the unconstrained fit to data, showing the signal channels in the normalization region $m_{\text{miss}}^2 < 1 \text{ (GeV}/c^2)^2$	285
13.9	m_{miss}^2 projection of the R -constrained fit to data, showing the normalization region in the signal channels	286
13.10	m_{miss}^2 projection of the R -constrained fit to data, showing a wider range than Figure 13.9	287
13.11	m_{miss}^2 projection of the R -constrained fit to data, showing the signal region in more detail	288
13.12	m_{miss}^2 projection of the R -constrained fit to data, showing the D^{**} control samples	289
13.13	p_ℓ^* projection of the R -constrained fit to data, showing the signal channels	290
13.14	p_ℓ^* projection of the R -constrained fit to data, showing the D^{**} control samples	291
13.15	p_ℓ^* projection of the R -constrained fit to data, showing the signal channels in the signal-enhanced region $m_{\text{miss}}^2 > 1 \text{ (GeV}/c^2)^2$	292
13.16	p_ℓ^* projection of the R -constrained fit to data, showing the signal channels in the normalization region $m_{\text{miss}}^2 < 1 \text{ (GeV}/c^2)^2$	293
13.17	Distribution of $-\ln \mathcal{L}$ for the unconstrained fit	296
13.18	Distribution of $-\ln \mathcal{L}$ for the R -constrained fit	297
A.1	Efficiency of the PidLHElectrons selector	308
A.2	Efficiency of the muMicroTight selector	311
A.3	Efficiency of the KLHTight selector	313
B.1	Invariant mass of $D^0 \rightarrow K^-\pi^+$ candidates	319
B.2	Invariant mass of $D^0 \rightarrow K^-\pi^+\pi^0$ candidates	320
B.3	Invariant mass of $D^0 \rightarrow K^-\pi^+\pi^-\pi^+$ candidates	320
B.4	Invariant mass of $D^0 \rightarrow K_s^0\pi^+\pi^-$ candidates	321
B.5	Invariant mass of $D^0 \rightarrow K_s^0\pi^+\pi^-\pi^0$ candidates	321
B.6	Invariant mass of $D^+ \rightarrow K^-\pi^+\pi^+$ candidates	322
B.7	Invariant mass of $D^+ \rightarrow K^-\pi^+\pi^+\pi^0$ candidates	322
B.8	Invariant mass of $D^+ \rightarrow K_s^0\pi^+$ candidates	323
B.9	Invariant mass of $D^+ \rightarrow K_s^0\pi^+\pi^0$ candidates	323
B.10	Invariant mass of $D^+ \rightarrow K_s^0\pi^+\pi^-\pi^+$ candidates	324
B.11	D^*-D mass difference of $D^{*0} \rightarrow D^0\pi^0$ candidates	324
B.12	D^*-D mass difference of $D^{*0} \rightarrow D^0\gamma$ candidates	325

B.13	D^*-D mass difference of $D^{*+} \rightarrow D^0\pi^+$ candidates	325
B.14	D^*-D mass difference of $D^{*+} \rightarrow D^+\pi^0$ candidates	326
C.1	$\Delta\theta$ vs. $\Delta\phi$ for same-charge loopers	331
C.2	$\Delta\theta$ vs. $\Delta\phi$ for opposite-charge loopers	332
C.3	$\Delta\theta$ vs. $\Delta\phi$ for ghost tracks	334
C.4	$\Delta\theta$ vs. $\Delta\phi$ for unmatched tracks	337

List of Tables

1.1	Standard model branching fraction predictions for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and $B \rightarrow X_c\tau^-\bar{\nu}_\tau$	4
1.2	Semitauponic branching fractions measured at LEP	6
4.1	Numbers of events and luminosities for various MC event types	78
5.1	Event yields from generic MC before selection	82
6.1	Cumulative cut-by-cut efficiency from Signal MC	120
6.2	Event yields from generic MC after all selection	121
7.1	Composition of the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples in generic MC . .	126
7.2	Composition of the q^2 control samples in generic MC	134
8.1	Components of the signal extraction fit in the signal channels .	141
8.2	Components of the signal extraction fit in the D^{**} control sample channels	142
8.3	p_ℓ^* dependence of the m_{miss}^2 PDF parameterization	151
8.4	Weights used for simulating $D^{**}\tau^-\bar{\nu}_\tau$ with signal MC	156
8.5	Constraints within the signal extraction fit	201
8.6	Feed-down constraints modified by D^* killing	204
8.7	Fit results from 640 fb^{-1} of generic MC	207
9.1	Fit validation from 330 fb^{-1} of generic MC	219
10.1	Normalization branching fractions of $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$	222
10.2	Branching fraction extraction in generic Monte Carlo	224
10.3	Branching fraction extraction with the R -constraint in generic Monte Carlo	225

11.1	Generated and reweighted D^{**} branching fractions	227
11.2	Constraints within the signal extraction fit after all corrections	231
12.1	Contributions to the total systematic error	234
12.2	Ranges of variation for reweighted D^{**} branching fractions	235
12.3	Systematic error estimation for D^{**} modeling	248
12.4	Systematic error estimation for $B \rightarrow D^*$ form factors	250
12.5	Systematic error estimation for $B \rightarrow D$ form factors	251
12.6	Sources and approximate abundances of combinatoric background in generic MC	252
12.7	Branching fractions of D_s^+ and two- and three-body B decays contributing to combinatoric background	253
12.8	Systematic error estimation for combinatoric background modeling	255
12.9	Systematic error estimation for MC statistics on PDF shapes	256
12.10	Systematic error estimation for MC statistics on crossfeed constraints	257
12.11	Systematic error estimation for MC statistics on floating crossfeed constraint parameterization	258
12.12	Systematic error estimation for $D^{**}\tau^-\bar{\nu}_\tau$ systematic error	259
12.13	Systematic error estimation for π^0 crossfeed constraints	267
12.14	Systematic error estimation for $D^{**}\ell^-\bar{\nu}_\ell$ feed-down	270
12.15	Events in the m_{miss}^2 tail of the $D^{(*)}\ell^-\bar{\nu}_\ell$ control sample	271
12.16	Events in the m_{miss}^2 tail of the q^2 control sample	272
12.17	Events in the m_{miss}^2 tail of the two control samples with all channels summed together	272
13.1	Results of the branching fraction fits	280
13.2	Event yields in the unconstrained fit	281
13.3	Event yields in the R -constrained fit	282
13.4	Significances of the signals	294
13.5	Stability checks in the final data sample	295
B.1	D^0 candidate selection criteria	317
B.2	D^+ candidate selection criteria	318
B.3	D^{*0} candidate selection criteria	318
B.4	D^{*+} candidate selection criteria	318

Chapter 1

Introduction

Semileptonic $B \rightarrow X_c \ell^- \bar{\nu}_\ell$ decays with $\ell = e, \mu$ and X_c a charmed meson (either D or D^*) are among the most common and most thoroughly studied B meson decays. Similar processes involving τ leptons, however, have only been observed inclusively — the X_c system has not been specifically identified.

The fact that these exclusive decays have never been observed is rather surprising, given that their branching fractions are expected to be of order 10^{-2} ; B meson decays with branching fractions 10^4 times smaller, such as $B \rightarrow K \ell^+ \ell^-$ [1], have been measured. Experimental study of $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ is difficult because the τ is too short-lived to be observed directly and because at least two (and often three) neutrinos are produced in a signal event. The missing four-momentum carried off by the neutrinos makes reconstruction of signal events very difficult.

Theoretical predictions for these modes in the Standard Model have a good foundation. A great deal is known about the form factors for $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$, both from experimental measurements [2, 3] and from theoretical studies [4]. For $\ell = e, \mu$, where the lepton mass is essentially negligible, there is one form factor for $B \rightarrow D\ell^{-}\bar{\nu}_\ell$ and three for $B \rightarrow D^*\ell^{-}\bar{\nu}_\ell$. However, for $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_\tau$, one additional form factor, due to the large τ mass, enters the amplitude for each mode. These additional form factors, however, can be related to the others using Heavy Quark Effective Theory (HQET) [5].

In this note, we present measurements of the branching fractions of four exclusive B decays:

$$\begin{aligned} B^- &\rightarrow D^0\tau^-\bar{\nu}_\tau & \bar{B}^0 &\rightarrow D^+\tau^-\bar{\nu}_\tau \\ B^- &\rightarrow D^{*0}\tau^-\bar{\nu}_\tau & \bar{B}^0 &\rightarrow D^{*+}\tau^-\bar{\nu}_\tau \end{aligned}$$

as well as two isospin-averaged branching fractions,

$$B \rightarrow D\tau^-\bar{\nu}_\tau \quad \text{and} \quad B \rightarrow D^*\tau^-\bar{\nu}_\tau .$$

1.1 Organization of this Document

In the remainder of this introduction (Chapter 1), we present a summary of current theoretical and experimental studies of $B \rightarrow X_c\tau^-\bar{\nu}_\tau$, and we give

an overview of our analysis. Chapter 2 presents a theoretical treatment of semitauonic B decays. In Chapter 3 we introduce the *BABAR* experiment. Chapter 4 describes the data and Monte Carlo samples used in this analysis, while Chapter 5 describes qualitatively the most important background processes and our overall strategy to minimize their impact on our measurement. Chapter 6 documents our signal selection criteria, and Chapter 7 presents a number of studies we have done to test the validity of our Monte Carlo simulation within the context of this analysis. In Chapter 8 we introduce the fit we use to extract a signal from our event sample, and in Chapter 9 we discuss validation of the fit. In Chapter 10 we show how we extract the branching fractions from the fit results. In Chapters 11 and 12, we discuss corrections to the Monte Carlo and systematic errors. We show the fit to the data in Chapter 13 and summarize our results in Chapter 14.

1.2 Standard Model Predictions

The standard model Feynman diagram for $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_{\tau}$ is shown in Figure 1.1, and predictions for semitauonic decay branching fractions [6, 7] are given in Table 1.1 (these values are largely consistent with older predictions [8, 9]).

Table 1.1: Standard model branching fraction predictions for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and $B \rightarrow X_c\tau\bar{\nu}_\tau$. The predictions for the exclusive \bar{B}^0 modes are taken from [6], where the errors are based on the form factor model used; the predictions for the B^- modes are based on the quoted \bar{B}^0 results, imposing isospin symmetry and using the $B^- - \bar{B}^0$ lifetime ratio in [10]. The prediction for the inclusive mode is taken from Ref. [7], and the dominant uncertainty comes from the experimental uncertainty on the total semileptonic decay rate $B \rightarrow X_c\ell^-\bar{\nu}_\ell$. See Chapter 2 for further details on these calculations and their errors.

Decay Mode	\mathcal{B} (%)
$\bar{B}^0 \rightarrow D^-\tau^-\bar{\nu}_\tau$	0.69 ± 0.04
$\bar{B}^0 \rightarrow D^{*-}\tau^-\bar{\nu}_\tau$	1.41 ± 0.07
$B^- \rightarrow D^0\tau^-\bar{\nu}_\tau$	0.64 ± 0.04
$B^- \rightarrow D^{*0}\tau^-\bar{\nu}_\tau$	1.32 ± 0.07
$B \rightarrow X_c\tau^-\bar{\nu}_\tau$	2.3 ± 0.25

The branching fractions for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ are expected to be about 0.7% and 1.4%, somewhat smaller than the dominant $B \rightarrow X_c\ell\nu$ decays, $B \rightarrow D\ell\nu$ and $B \rightarrow D^*\ell\nu$, which have branching fractions of approximately 2% and 5%, respectively [10]. These would still be, however, among the largest B meson branching fractions, comparable to the dominant hadronic modes $B \rightarrow D^{(*)}\pi$, $D^{(*)}\pi\pi$, and $D^{(*)}\pi\pi\pi$ which have branching fractions between 0.3 and 1.3%, and $B \rightarrow D^{(*)}D_s^{(*)}$, with branching fractions between 0.6 and 1.8%.

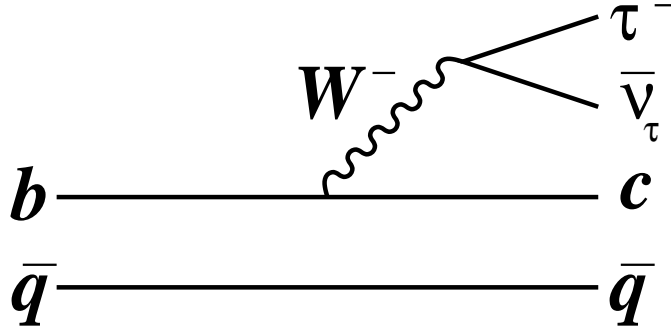


Figure 1.1: Standard model Feynman diagram for $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$.

1.3 Previous Measurements

The exclusive decay modes $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ have not yet been observed. Inclusive measurements of the branching fraction $b \rightarrow X \tau^- \bar{\nu}_\tau$ have been performed by ALEPH [11], L3 [12, 13], and OPAL [14], yielding an average measurement of $(2.48 \pm 0.26)\%$ [10]. These measurements are not inconsistent with the prediction of Falk *et al.* [7], but it must be noted that the calculation is for the inclusive B meson branching fraction, while the LEP experiments measure the inclusive b quark branching fraction; the experimental result also includes contributions from B_s mesons and b baryons. The ALEPH collaboration also measured a submode of the inclusive measurement, $b \rightarrow D^{*\pm} \tau^\mp \nu X$, with a branching fraction $(0.88 \pm 0.31 \pm 0.28)\%$; this is consistent with the predictions of Körner and Schuler [9]. The branching fractions measured by the various LEP experiments are summarized in Table 1.2 and in Figure 1.2.

Table 1.2: Semitauonic branching fractions measured at LEP. The errors given are statistical and systematic, respectively. The two L3 measurements use different experimental techniques.

Decay Mode	Experiment	\mathcal{B} (%)
$b \rightarrow X\tau^-\bar{\nu}_\tau$	ALEPH [11]	$2.43 \pm 0.20 \pm 0.25$
	L3 [12]	$1.7 \pm 0.4 \pm 1.1$
	L3 [13]	$2.4 \pm 0.7 \pm 0.8$
	OPAL [14]	$2.78 \pm 0.18 \pm 0.51$
	PDG average [10]	2.48 ± 0.26
$b \rightarrow D^{*\pm}\tau^\mp\nu X$	ALEPH [11]	$0.88 \pm 0.31 \pm 0.28$

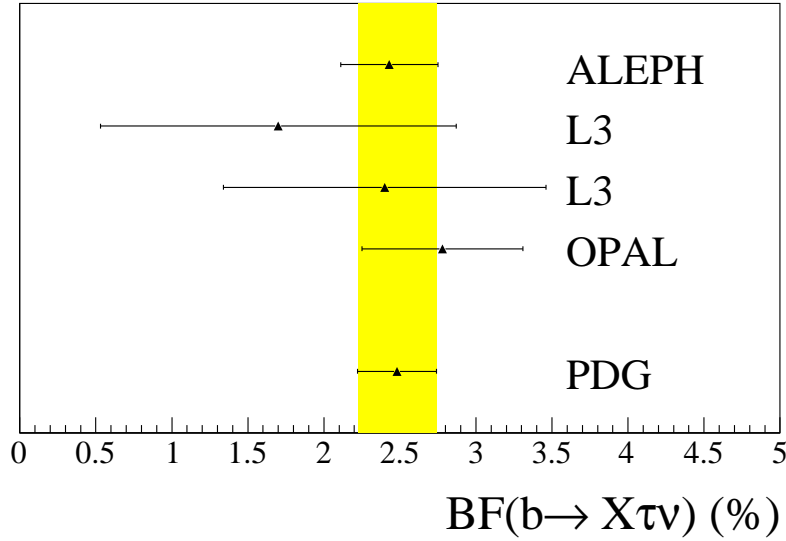


Figure 1.2: Measurements of the $b \rightarrow X_c\tau^-\bar{\nu}_\tau$ from LEP experiments.

These measurements all rely on the jet-like b decays at the Z^0 resonance. Because the two b jets appear back-to-back, the missing energy due to the neutrinos will be spatially correlated with one of the jets. A b -tagging algorithm is used on the opposite jet to distinguish $Z^0 \rightarrow b\bar{b}$ events from other $Z^0 \rightarrow jj$ events. The results of the ALEPH study are shown in Figure 1.3.

1.4 Sensitivity to New Physics

One of the most compelling reasons to search for the exclusive modes is their potential sensitivity to a charged Higgs boson. The tree level $b \rightarrow c\ell^-\bar{\nu}_\ell$ process can be mediated by an H^\pm instead of a W^\pm (as in Figure 1.4), and, since the Higgs coupling is proportional to the particle masses, this effect may be observed in decays involving τ while remaining unobserved in the other semileptonic modes.

Because the H^\pm is a spin-0 particle, the Higgs contribution will be different for the two decays $B \rightarrow D\tau^-\bar{\nu}_\tau$ and $B \rightarrow D^*\tau^-\bar{\nu}_\tau$. The relative proportion of D to D^* may therefore differ (by a potentially large amount) from the SM value. The branching fractions for the two modes depend on the charged Higgs mass, m_H , and on the ratio of vacuum expectation values, $\tan\beta$. In particular, if we consider the dimensionless quantity

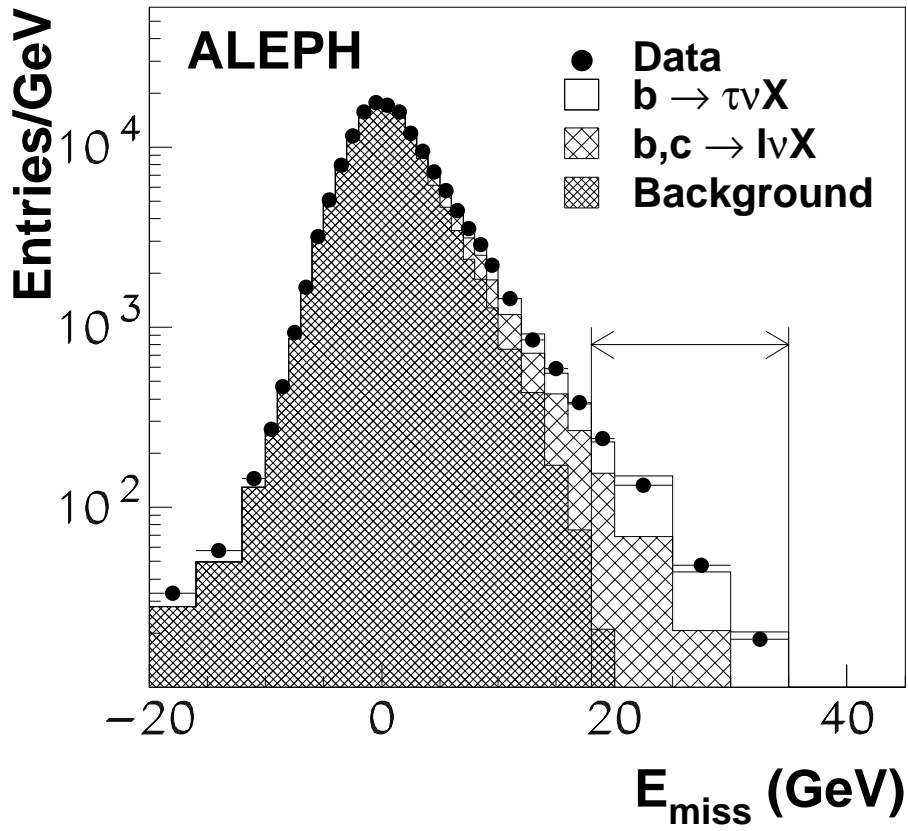


Figure 1.3: Missing energy spectrum for the ALEPH $B \rightarrow X \tau^- \bar{\nu}_\tau$ analysis [11]. Points are data; histograms are simulation. Their E_{miss} signal region is indicated, and is estimated to include 778 signal events out of 2094 total events.

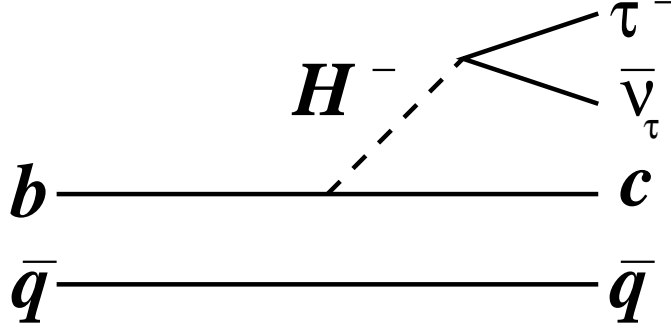


Figure 1.4: Feynman diagram for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ in a two-Higgs doublet model.

$$Q \equiv \frac{m_W}{m_H} \tan \beta ,$$

Miura and Tanaka find that the ratio $\mathcal{B}(B \rightarrow D\tau^-\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D\mu^-\bar{\nu}_\mu)$ depends on Q as seen in Figure 1.5 [15]. The same behavior is seen in the D^* modes and in the inclusive mode, but the effect is smaller in these cases.

Grossman, Haber, and Nir have used the inclusive $b \rightarrow X\tau\nu$ measurement to place a model independent constraint on Q ; they find $Q < 41$ at the 2σ level [16].

The experimental prospects for observing $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ have been discussed by the SuperKEKB Physics Working Group [17]. While their work suggests that such measurements require integrated luminosities greater than 1 ab^{-1} , we believe these conclusions are somewhat pessimistic, and that ob-

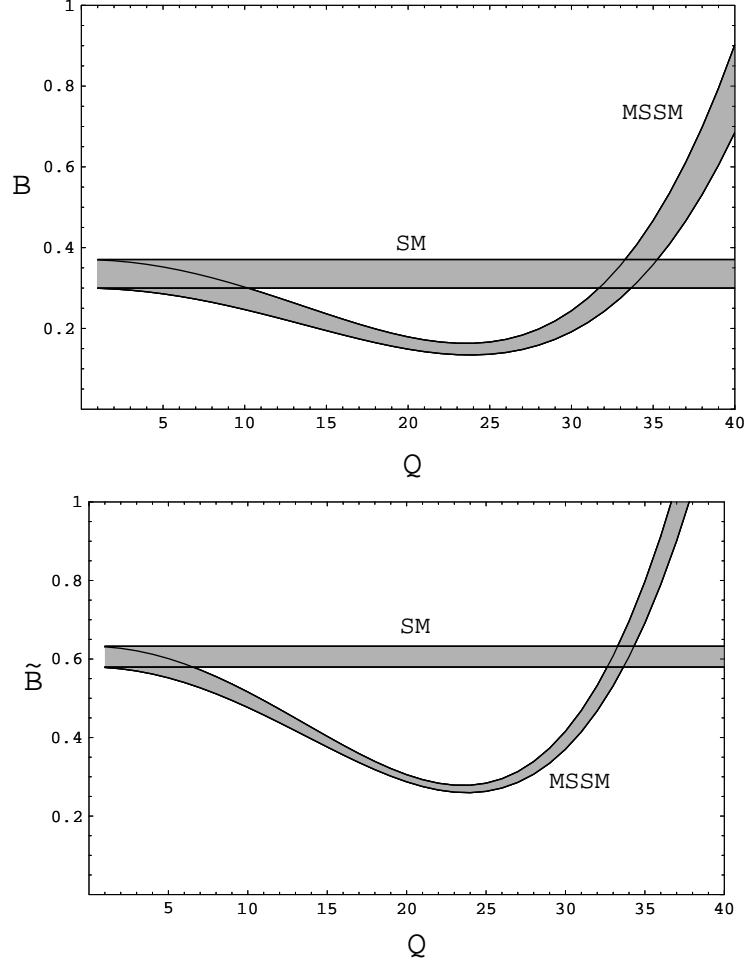


Figure 1.5: $\mathcal{B}(B \rightarrow D\tau^-\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D\mu^-\bar{\nu}_\mu)$ as a function of Q in the MSSM, with the standard model prediction shown as a flat band. The shaded area corresponds to a $\pm 1\sigma$ variation in the experimental measurement of the $B \rightarrow D$ form factor, which is the dominant uncertainty in the calculation. (a) shows the total branching fraction ratio, while in (b), the denominator is only integrated over the phase space region allowed in $B \rightarrow D\tau^-\bar{\nu}_\tau$. Note that, within the present limit on Q , the branching fraction of $B \rightarrow D\tau^-\bar{\nu}_\tau$ may differ from the standard model prediction by up to a factor of two in either direction. Taken from Miura and Tanaka [15].

servation of $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ is possible with B -factory data samples available today.

1.5 Methodology

Because of the large expected background from other $B \rightarrow X_c\ell^-\bar{\nu}_\ell$ decays and the lack of a τ mass peak, we have designed very tight selection criteria. By imposing such constraints on signal candidates, we hope to reduce, as much as possible, the contamination due to misreconstructed or partially reconstructed events.

The key to performing this analysis at *BABAR* is to use the so-called “semiexclusive B reconstruction” technique to fully reconstruct one B meson in a hadronic final state. This B_{tag} sample includes more than 1000 different final states, and yields approximately 2000 fully-reconstructed B mesons per fb^{-1} .

After fully reconstructing the B_{tag} candidate, we reconstruct $D^{(*)}$ and τ candidates from the remaining particles in the event. $D^{(*)}$ mesons are reconstructed in a variety of hadronic modes, while τ candidates are reconstructed in the cleanest channels, the leptonic decays $\tau \rightarrow \ell^-\bar{\nu}_\ell\nu_\tau$, in which only the charged lepton ℓ is actually observed.

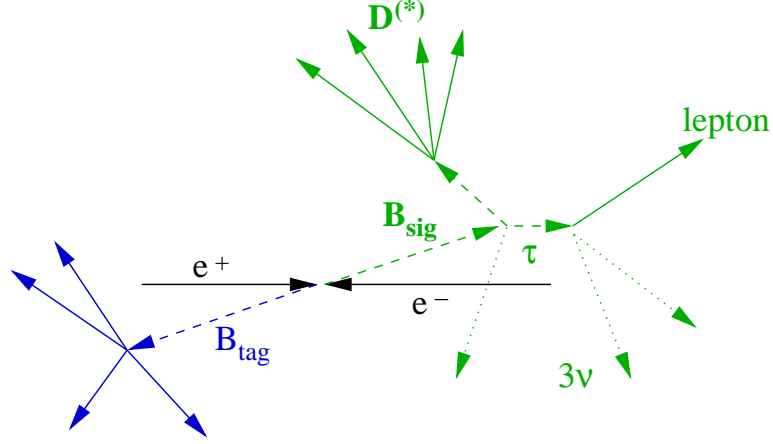


Figure 1.6: Schematic view of a signal event.

A key part of the signature of a signal event is large missing four-momentum due to the neutrinos. Our signal extraction will be based largely on the missing mass squared, defined as

$$m_{\text{miss}}^2 \equiv (p_{\text{miss}})^2 = (p_{\Upsilon} - p_{\text{tag}} - p_{D^{(*)}} - p_{\ell})^2, \quad (1.1)$$

where p_{tag} , $p_{D^{(*)}}$, and p_{ℓ} are the four-momenta of the reconstructed B_{tag} , $D^{(*)}$, and ℓ candidates, respectively. p_{Υ} is the four-momentum of the $\Upsilon(4S)$, and is determined by the colliding beam energies. A schematic view of a signal event is shown in Figure 1.6.

Although working with the B_{tag} sample results in a substantial reduction in the number of events, the fully reconstructed B_{tag} meson provides powerful kinematic constraints. Because the total momentum of the $e^+e^- \rightarrow \Upsilon(4S)$

event is known, we can use the reconstructed four momentum of the B_{tag} meson to infer the four momentum of the signal B meson, whose direction would otherwise be unknown. A second advantage to the B_{tag} sample is that we can account for all of the visible particles in the event, allowing us to reject background events in which particles other than neutrinos may have been lost; these events have large m_{miss}^2 and make signal extraction difficult. We can also reject events in which the total visible charge is not zero; these events have necessarily lost particles and must have been misreconstructed.

While semiexclusive B reconstruction selects a relatively impure sample of B mesons (see, e.g., Figure 6.1), the additional selection on the recoil side of the event cleans up this sample dramatically (see Figure 6.3).

The use of m_{miss}^2 as a signal discriminant has one further advantage: we can simultaneously extract the event yield for the dominant backgrounds $B \rightarrow D^{(*)}e^{-}\bar{\nu}_e$ and $B \rightarrow D^{(*)}\mu^{-}\bar{\nu}_\mu$. These background events only have a single neutrino in the final state, so they will have $m_{\text{miss}}^2 = 0$. This allows us to measure the ratios of branching ratios $R \equiv \mathcal{B}(B \rightarrow D\tau^{-}\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D\ell^{-}\bar{\nu}_\ell)$ and $R^* \equiv \mathcal{B}(B \rightarrow D^*\tau^{-}\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D^*\ell^{-}\bar{\nu}_\ell)$ in such a way that many of the systematic errors cancel. We can then normalize these measurements to the known branching fractions for $B \rightarrow D\ell^{-}\bar{\nu}_\ell$ and $B \rightarrow D^*\ell^{-}\bar{\nu}_\ell$ to extract our final results.

Expected Event Yields

Semiexclusive B reconstruction at *BABAR* has an efficiency of approximately 0.2%. If we assume a signal efficiency of approximately 30% — these are high-multiplicity modes and we will need tight cuts to suppress background — and that the branching fraction of $D^{(*)}$ decays to reconstructible modes is 20%, then we expect, in the Run 1–4 *BABAR* data sample, roughly

$$\begin{aligned} N_{\text{exp}} &= \mathcal{L} \cdot \sigma_{B\bar{B}} \cdot \mathcal{B}(B \rightarrow D\tau\nu) \cdot \mathcal{B}(D^{(*)}) \cdot \mathcal{B}(\tau \rightarrow \ell\nu\bar{\nu}) \cdot \varepsilon(B_{\text{tag}}) \cdot \varepsilon(\text{signal}) \\ &= 200 \text{ fb}^{-1} \cdot 1.1 \text{ nb} \cdot 7 \times 10^{-3} \cdot 0.2 \cdot 0.35 \cdot 2 \times 10^{-3} \cdot 0.3 , \end{aligned} \quad (1.2)$$

which yields 62 events in the $D\tau^-\bar{\nu}_\tau$ modes and, similarly, 96 events in $D^*\tau^-\bar{\nu}_\tau$. If the effective signal-to-background ratio is not much worse than 1:2, then this amounts to a 5σ discovery for both of these modes.

For comparison, we repeat this exercise for the dominant background modes $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$. The branching fraction for $B \rightarrow D\ell^-\bar{\nu}_\ell$ is larger than that for $B \rightarrow D\tau^-\bar{\nu}_\tau$ by a factor of 3 (for D^* , a factor of 4), and these backgrounds are not reduced by the factor $\mathcal{B}(\tau^- \rightarrow \ell^-\bar{\nu}_\ell\nu_\tau)$. We therefore expect these backgrounds to be about 20 times the size of our signal, roughly 1200 $D\ell^-\bar{\nu}_\ell$ and 2000 $D^*\ell^-\bar{\nu}_\ell$ events (counting both e and μ). Control of these backgrounds is therefore a critical requirement of this measurement.

1.6 Notation and Conventions

- Throughout this paper, the term “semileptonic” will refer exclusively to decays to the light leptons, e and μ ; “semitauonic” will refer to decays involving τ . More generally, the word “lepton” and the symbol ℓ will almost always refer to e and μ . The only exception to this will be Chapter 2, where we will use ℓ to refer to all three charged leptons, and we will explicitly differentiate between the light and heavy leptons where appropriate.
- The term “tagged” will always refer to the B meson identified by the semiexclusive B reconstruction algorithm described below. The other B will be referred to as “signal,” or, when it is explicitly not a signal event, as “recoil.”
- Because we can work in the rest frame of the signal B meson, we denote quantities measured in this system with an asterisk, as in p_ℓ^* . Quantities without an asterisk are measured in the lab frame, and we will explicitly mention any use of the $\Upsilon(4S)$ center-of-mass frame.
- The charge-conjugate mode is always implied for any decay mode mentioned.

- Crossfeed between signal and signal-like modes play an important role in this analysis. Because of this, we need to differentiate between correctly reconstructed events and crossfeed. We will use the notation that $D^{*0}\ell^{-}\bar{\nu}_\ell \Rightarrow D^{*0}$ represents a correctly reconstructed $B \rightarrow D^{*0}\ell^{-}\bar{\nu}_\ell$ event, while $D^{*0}\ell^{-}\bar{\nu}_\ell \Rightarrow D^0$ represents feed-down in which the D^{*0} is misreconstructed as a D^0 .
- Throughout this document, we will use the symbol D^{**} to refer inclusively to both the broad and narrow D^{**} resonances and to nonresonant $D^{(*)}n\pi$ systems.

Chapter 2

Theory of Semileptonic B Decays to τ Leptons

Theoretical investigations of semileptonic B decays to final states with a τ lepton have focused on possible constraints these processes could provide on models with an extended Higgs sector. In such models, a charged Higgs boson in the intermediate state produces an additional amplitude that interferes with the usual W -mediated decay. The large mass of the τ lepton greatly enhances the sensitivity of the decay to the Higgs-mediated amplitude. In contrast to $b \rightarrow s\gamma$ decays, in which the SM decay is forbidden at tree level and sensitivity to new physics occurs via loop diagrams, $B \rightarrow X\tau^-\bar{\nu}_\tau$ decays can have sensitivity to new physics even at tree level.

As a baseline, it is important to understand the SM predictions for semileptonic decays involving a τ lepton. We will briefly introduce the formalism of exclusive semileptonic decays in Section 2.1, paying particular attention to

the details which distinguish semitauonic decays from those involving light leptons. In Section 2.2, we present expressions for the SM differential decay rates for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\ell^-\bar{\nu}_\ell$, and in Section 2.3, we discuss theoretical understanding of the form factors describing the $B \rightarrow D$ and $B \rightarrow D^*$ transitions. In Section 2.4, we use a particular form-factor model to predict the distributions of several important kinematic and event-selection variables, and to compare these distributions between the light leptons and the τ leptons. Finally, in Section 2.5, we discuss possible non-Standard Model contributions to processes involving τ .

2.1 Formalism of Exclusive Semileptonic Decays

In the Standard Model, the leading-order Feynman diagram describing $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ is shown in Figure 1.1. In this diagram, the light antiquark \bar{q} can be either a \bar{u} , in which case the diagram shows a B^- decay, or a \bar{d} , in which case the decaying meson is a \bar{B}^0 , and the $c\bar{q}$ system can hadronize as either a D meson (D^0 or D^+) or a D^* meson (D^{*0} or D^{*+}). The same diagram can also be used to describe semileptonic decays to the light leptons, simply by replacing the $\tau^-\bar{\nu}_\tau$ system with either $e^-\bar{\nu}_e$ or $\mu^-\bar{\nu}_\mu$.

2.1.1 Decay Kinematics

Two independent kinematic variables are used to describe the $B \rightarrow D\ell^-\bar{\nu}_\ell$ decay, while four are needed to describe the $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decay. One of these variables is q^2 , the squared momentum transfer, defined as

$$q^2 \equiv (p_B - p_{D^{(*)}})^2 \equiv (p_\ell + p_{\bar{\nu}})^2, \quad (2.1)$$

which is also equal to the squared mass of the virtual W^- . The remaining variables are decay angles which are defined in Figure 2.1. θ_ℓ is the polar angle between the charged lepton momentum and the direction opposite the B meson momentum, as measured in the rest frame of the virtual W^- , and is used to describe both $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decays. θ_V is the polar angle between the D momentum and the direction opposite the B meson momentum, measured in the rest frame of the D^* meson, while χ is the azimuthal angle between the $W^* \rightarrow \ell^-\bar{\nu}_\ell$ and $D^* \rightarrow D\pi_s$ decay planes,¹ both of which are only defined for the $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decay.

The momentum transfer q^2 is linearly related to another frequently-used variable, w , defined as the product of four-velocities

$$w \equiv v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2m_B m_{D^{(*)}}}. \quad (2.2)$$

¹We use the notation $D^* \rightarrow D\pi_s$ as shorthand here; in D^* decays to $D\gamma$, the definitions of these variables are unchanged.

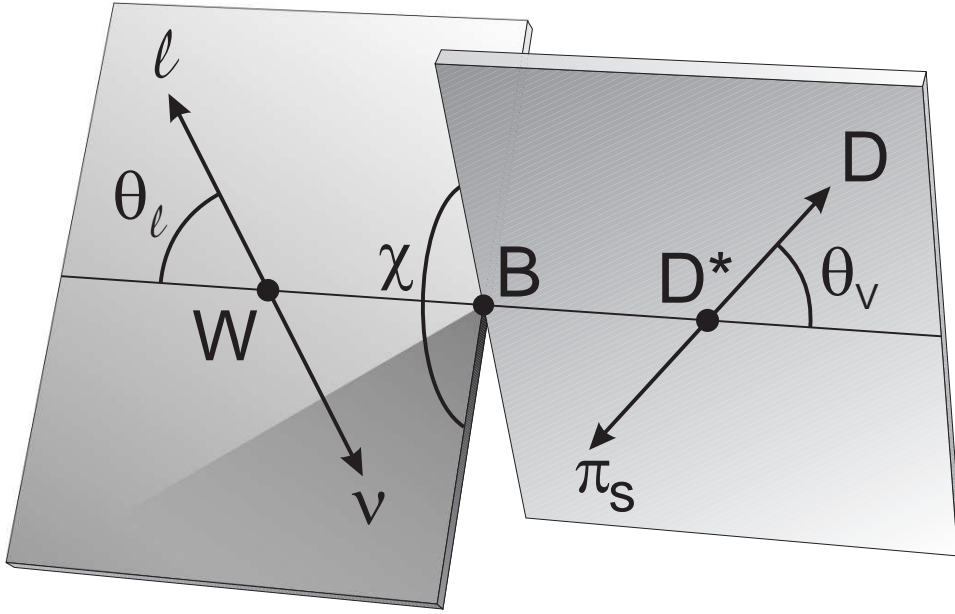


Figure 2.1: Definition of the decay angles θ_ℓ , θ_V , and χ . The angle θ_ℓ is defined for both $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decays, while the other two angles are only defined for the $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ case. The lepton and neutrino are shown back-to-back because they are shown in the W^* rest frame, where the angle θ_ℓ is measured. Similarly, the D and the π are shown in the D^* rest frame, where the angle θ_V is measured. The azimuthal angle χ is measured between the W^* and D^* decay planes, and is illustrated in the rest frame of the B meson.

This variable is particularly useful in HQET expressions for the form factors (see Section 2.3) due to the fact that it is more natural to work with velocities than momenta in HQET.

Another variable related to q^2 and w which is sometimes used is z , defined as

$$z \equiv \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}}. \quad (2.3)$$

The CLN model (also discussed in Section 2.3), which we use to describe both signal $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and normalization $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays, expresses the form factors as functions of z rather than of w . Because the allowed range of z is much smaller than that of w (z varies from zero to ≈ 0.07 for $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays, while w varies from 1 to ≈ 1.6), the Taylor-series expansion of the form factors is more efficient when performed as a function of z with respect to w .

2.1.2 Decay Amplitude

We write the matrix element for the semileptonic decay $M_{Q\bar{q}} \rightarrow X_{q'\bar{q}}\ell^-\bar{\nu}_\ell$ as the product of a hadronic and a leptonic current:

$$\begin{aligned} \mathcal{M}(M_{Q\bar{q}} \rightarrow X_{q'\bar{q}}\ell^-\bar{\nu}_\ell) &= \langle X_{q'\bar{q}} | \bar{q}' \frac{-ig}{2\sqrt{2}} V_{q'Q} \gamma_\mu (1 - \gamma_5) Q | M_{Q\bar{q}} \rangle \\ &\quad \times P^{\mu\nu}(q) \bar{u}_\ell \frac{-ig}{2\sqrt{2}} \gamma_\nu (1 - \gamma_5) v_\ell \end{aligned} \quad (2.4)$$

where the operator Q annihilates the quark Q (or creates \bar{Q}) and the W propagator is given by

$$P^{\mu\nu}(q) = \frac{i(-g^{\mu\nu} + q^\mu q^\nu / M_W^2)}{q^2 - M_W^2} \simeq i \frac{g^{\mu\nu}}{M_W^2}. \quad (2.5)$$

The last expression for the propagator is appropriate when the energies are much less than M_W . We have used the form of the CKM element $V_{q'Q}$ appropriate to the case in which a W^- is emitted (*e.g.*, V_{cb} or V_{ub}); if a W^+ were emitted, as in $c \rightarrow s$ or $c \rightarrow d$, the form of the CKM element is $V_{Qq'}^*$ (*e.g.*, V_{cs}^* or V_{cd}^*), but the absolute magnitude is taken in the end in any case. We therefore have the phenomenological form for the matrix element

$$\mathcal{M}(M_{Q\bar{q}} \rightarrow X_{q'\bar{q}} \ell^- \bar{\nu}_\ell) = -i \frac{G_F}{\sqrt{2}} V_{q'Q} L^\mu H_\mu, \quad (2.6)$$

where $G_F/\sqrt{2} = g^2/8M_W^2$. The leptonic current is exactly known

$$L^\mu = \bar{u}_\ell \gamma^\mu (1 - \gamma_5) v_\nu, \quad (2.7)$$

and the hadronic current is given by

$$H_\mu = \langle X_{q'\bar{q}} | \bar{q}' \gamma_\mu (1 - \gamma_5) Q | M_{Q\bar{q}} \rangle. \quad (2.8)$$

We can use Lorentz invariance to construct the hadronic current from the available four-vectors, which are momenta and spin-polarization vectors. The Lorentz vector or axial-vector quantities thus formed have Lorentz-invariant

coefficients (form factors) that are functions of q^2 . We will consider two main classes of exclusive semileptonic decay: $P \rightarrow P'\ell\nu$, where both P and P' are pseudoscalar particles, and $P \rightarrow V\ell\nu$, where V is a vector particle.

2.1.3 Hadronic current for $B \rightarrow D$

In the case of a $P(Q\bar{q}) \rightarrow P'(q'\bar{q})\ell\nu$ decay, there are only two independent four vectors, which we can take to be $p + p'$ and $q = p - p'$. For these quantum numbers, the hadronic current H^μ has no contribution from the matrix element of the axial-vector current and can be written [5] as

$$\langle P'(p')|V^\mu|P(p)\rangle = f_+(q^2)(p + p')^\mu + f_-(q^2)(p - p')^\mu, \quad (2.9)$$

where $V^\mu = \bar{q}'\gamma^\mu Q$ and where $f_+(q^2)$ and $f_-(q^2)$ are the form factors describing the $P \rightarrow P'$ transition.

For the cases $\ell = e$ and $\ell = \mu$, this expression for the current simplifies, because the terms proportional to q^μ are negligible, both here and for the case $P \rightarrow V\ell\nu$. The reason is that, in the limit $m_\ell \rightarrow 0$, $q^\mu L_\mu = 0$, where L_μ is the lepton current. This means that the decays $B \rightarrow De^-\bar{\nu}_e$ and $B \rightarrow D\mu^-\bar{\nu}_\mu$ can be described by a single form factor, while, for $B \rightarrow D\tau^-\bar{\nu}_\tau$, we must keep the additional form factor.

2.1.4 Hadronic current for $B \rightarrow D^*$

For the process $P(Q\bar{q}) \rightarrow V(q'\bar{q})\ell^-\bar{\nu}_\ell$, where V is a vector meson, each term in the current must be linear in the polarization vector, ε , of the vector meson (actually, the complex conjugate is used since V is an outgoing particle). The most general form for the current is [5, 19]

$$\begin{aligned} \langle V(p', \varepsilon) | V^\mu - A^\mu | P(p) \rangle &= \frac{2i\varepsilon^{\mu\nu\alpha\beta}}{M + m_V} \varepsilon_\nu^* p'_\alpha p_\beta V(q^2) \\ &\quad - (M + m_V) \varepsilon^{*\mu} A_1(q^2) + \frac{\varepsilon^* \cdot q}{M + m_V} (p + p')^\mu A_2(q^2) \\ &\quad + 2m_V \frac{\varepsilon^* \cdot q}{q^2} q^\mu A_3(q^2) - 2m_V \frac{\varepsilon^* \cdot q}{q^2} q^\mu A_0(q^2), \end{aligned} \tag{2.10}$$

where $V^\mu = \bar{q}' \gamma^\mu Q$, $A^\mu = \bar{q}' \gamma^\mu \gamma_5 Q$, and

$$A_3(q^2) = \frac{M + m_V}{2m_V} A_1(q^2) - \frac{M - m_V}{2m_V} A_2(q^2), \tag{2.11}$$

with $A_0(0) = A_3(0)$. The functions $A_0(q^2)$, $A_1(q^2)$, $A_2(q^2)$, and $V(q^2)$ are the independent form factors describing the $P \rightarrow V$ transition. Again, terms proportional to q^μ only play an important role for the case $\ell = \tau$, so that we have effectively three form factors describing the decays $B \rightarrow D^* e^- \bar{\nu}_e$ and $B \rightarrow D^* \mu^- \bar{\nu}_\mu$, and four for the decay $B \rightarrow D^* \tau^- \bar{\nu}_\tau$.

2.2 Differential Decay Rate Formulae

Körner and Schuler [9] have calculated differential decay rates for the processes $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ including lepton mass effects, making these distributions applicable both to the light leptons and to the τ case.

We begin by projecting the form factors onto a helicity basis to obtain a set of helicity form factors H_i . We can then express the hadronic currents in terms of these helicity form factors and can therefore express the decay rate, which goes like $|\mathcal{M}|^2$, as a sum over partial helicity rates \hat{H}_i , each of which is a quadratic function of the H_i .

2.2.1 Differential Decay Rate for $B \rightarrow D\ell^-\bar{\nu}_\ell$

For the $B \rightarrow D$ transition, the helicity form factors are given by

$$H_0(q^2) = \frac{2m_B \cdot p_W^*}{\sqrt{q^2}} f_+(q^2) \quad (2.12)$$

$$H_t(q^2) = \frac{1}{\sqrt{q^2}} [(m_B^2 - m_D^2) f_+(q^2) + q^2 f_-(q^2)] \quad , \quad (2.13)$$

where p_W^* is the momentum of the virtual W^- in the rest frame of the B meson.

The subscripts 0 and t refer to the longitudinal and transverse polarization states of the W^- .

From these helicity factors, we calculate the partial rates

$$\hat{H}_L = |H_0|^2 \quad (2.14)$$

$$\hat{H}_S = 3|H_t|^2 \quad (2.15)$$

$$\hat{H}_{SL} = \Re(H_t H_0^*) , \quad (2.16)$$

where the subscripts L , S , and SL refer to longitudinal, scalar, and scalar-longitudinal interference, and describe the angular dependence of each term in the total decay rate.

The complete decay distribution for $B \rightarrow D\ell^-\bar{\nu}_\ell$ in terms of these \hat{H}_i is

$$\begin{aligned} \frac{d\Gamma}{dq^2 d\cos\theta_\ell} &= \frac{G_F^2}{(2\pi)^3} |V_{cb}|^2 \frac{(q^2 - m_\ell^2)^2 p_W^*}{12m_B^2 q^2} \\ &\times \left\{ \frac{3}{4} \sin^2\theta_\ell \hat{H}_L + \frac{m_\ell^2}{2q^2} \left[\frac{3}{2} \cos^2\theta_\ell \hat{H}_L + \frac{1}{2} \hat{H}_S + 3 \cos\theta_\ell \hat{H}_{SL} \right] \right\} . \quad (2.17) \end{aligned}$$

In the limit $m_\ell \rightarrow 0$, this becomes

$$\frac{d\Gamma}{dq^2 d\cos\theta_\ell} = \frac{1}{4} \frac{G_F^2}{(2\pi)^3} |V_{cb}|^2 (p_W^*)^3 \sin^2\theta_\ell f_+^2(q^2) , \quad (2.18)$$

which is the standard distribution for massless leptons. Note that all of the dependence on f_- appears (within the \hat{H}_S and \hat{H}_{SL} terms) multiplied by $m_\ell^2/2q^2$, which is equivalent to the earlier claim that, for light leptons, the $B \rightarrow D\ell^-\bar{\nu}_\ell$ process is described by only one form factor, f_+ .

2.2.2 Differential Decay Rate for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$

For the $B \rightarrow D^*$ transition, Körner and Schuler use a different normalization convention for the form factors. Their form factors F_i are related to the standard-notation form factors by

$$F_V(q^2) = \frac{-2}{m_B + m_{D^*}} V(q^2) \quad (2.19)$$

$$F_1(q^2) = -(m_B + m_{D^*}) A_1(q^2) \quad (2.20)$$

$$F_2(q^2) = \frac{2}{m_B + m_{D^*}} A_2(q^2) \quad (2.21)$$

$$F_3(q^2) = \frac{1}{m_B + m_{D^*}} A_2(q^2) + \frac{2m_{D^*}}{q^2} (A_3(q^2) - A_0(q^2)) \quad , \quad (2.22)$$

and, in terms of these F_i , the helicity form factors are

$$H_0(q^2) = \frac{1}{2m_{D^*} \sqrt{q^2}} [(m_B^2 - m_{D^*}^2 - q^2) F_1(q^2) + 2m_B^2 (p_W^*)^2 F_2(q^2)] \quad (2.23)$$

$$H_t(q^2) = \frac{m_B p_W^*}{m_{D^*} \sqrt{q^2}} \left[F_1(q^2) + \frac{1}{2} (m_B^2 - m_{D^*}^2 + q^2) F_2(q^2) + q^2 F_3(q^2) \right] \quad (2.24)$$

$$H_\pm(q^2) = F_1(q^2) \pm m_B p_W^* F_V(q^2) \quad . \quad (2.25)$$

With these four helicity form factors, we calculate the partial rates as before. \hat{H}_L , \hat{H}_S , and \hat{H}_{SL} are defined as in Equations 2.14–2.16 (although the definitions of H_0 and H_t have changed), and we now introduce additional

terms:

$$\hat{H}_U = |H_+|^2 + |H_-|^2 \quad (2.26)$$

$$\hat{H}_P = |H_+|^2 - |H_-|^2 \quad (2.27)$$

$$\hat{H}_T = \Re(H_+ H_-^*) \quad (2.28)$$

$$\hat{H}_I = \frac{1}{2} \Re(H_+ H_0^* + H_- H_0^*) \quad (2.29)$$

$$\hat{H}_A = \frac{1}{2} \Re(H_+ H_0^* - H_- H_0^*) \quad (2.30)$$

$$\hat{H}_{ST} = \frac{1}{2} \Re(H_+ H_t^* + H_- H_t^*) , \quad (2.31)$$

where the subscripts describe the angular dependence of each term, and indicate unpolarized-transverse, parity-odd, transverse interference, transverse-longitudinal interference, parity asymmetric, and scalar-transverse interference terms, respectively.

The complete decay distribution for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ in terms of these \hat{H}_i is

$$\begin{aligned}
 \frac{d\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_V d\chi} &= \mathcal{B}(D^* \rightarrow D\pi_s) \frac{G_F^2}{(2\pi)^4} |V_{cb}|^2 \frac{(q^2 - m_\ell^2)^2 p_W^*}{12m_B^2 q^2} \\
 &\times \left\{ \begin{aligned}
 &\frac{3}{8}(1 + \cos^2\theta_\ell) \cdot \frac{3}{4} \sin^2\theta_V \cdot \hat{H}_U \\
 &+ \frac{3}{4} \sin^2\theta_\ell \cdot \frac{3}{2} \cos^2\theta_V \cdot \hat{H}_L \\
 &- \frac{3}{4} \sin^2\theta_\ell \cdot \cos 2\chi \cdot \frac{3}{4} \sin^2\theta_V \cdot \hat{H}_T \\
 &- \frac{9}{16} \sin 2\theta_\ell \cdot \cos\chi \cdot \sin 2\theta_V \cdot \hat{H}_I \\
 &\pm \frac{3}{4} \cos\theta_\ell \cdot \frac{3}{4} \sin^2\theta_V \cdot \hat{H}_P \\
 &\mp \frac{9}{8} \sin\theta_\ell \cdot \cos\chi \cdot \sin 2\theta_V \cdot \hat{H}_A \\
 &+ \frac{3}{4} \sin^2\theta_\ell \cdot \frac{3}{4} \sin^2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_U \\
 &+ \frac{3}{2} \cos^2\theta_\ell \cdot \frac{3}{2} \cos^2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_L \\
 &+ \frac{3}{4} \sin^2\theta_\ell \cdot \cos 2\chi \cdot \frac{3}{4} \sin^2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_T \\
 &+ \frac{9}{8} \sin 2\theta_\ell \cdot \cos\chi \cdot \sin 2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_I \\
 &+ \frac{3}{2} \cos^2\theta_V \cdot \frac{1}{2} \frac{m_\ell^2}{2q^2} \hat{H}_S \\
 &+ 3 \cos\theta_\ell \cdot \frac{3}{2} \cos^2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_{SL} \\
 &+ \frac{9}{4} \sin\theta_\ell \cdot \cos\chi \cdot \sin 2\theta_V \cdot \frac{m_\ell^2}{2q^2} \hat{H}_{ST} \end{aligned} \right\}. \tag{2.32}
 \end{aligned}$$

In the limit $m_\ell \rightarrow 0$, this becomes

$$\begin{aligned}
 \frac{d\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_V d\chi} &= \mathcal{B}(D^* \rightarrow D\pi_s) \frac{3G_F^2 |V_{cb}|^2 p_W^* q^2}{8 (4\pi)^4 m_B^2} \\
 &\times \left\{ \begin{aligned}
 &(1 - \cos\theta_\ell)^2 \sin^2\theta_V |H_+|^2 \\
 &+ (1 + \cos\theta_\ell)^2 \sin^2\theta_V |H_-|^2 \\
 &+ 4 \sin^2\theta_\ell \cos^2\theta_V |H_0|^2 \\
 &- 2 \sin^2\theta_\ell \sin^2\theta_V \cos 2\chi |H_+||H_-| \\
 &- 4 \sin\theta_\ell (1 - \cos\theta_\ell) \sin\theta_V \cos\theta_V \cos\chi |H_+||H_0| \\
 &+ 4 \sin\theta_\ell (1 + \cos\theta_\ell) \sin\theta_V \cos\theta_V \cos\chi |H_-||H_0|
 \end{aligned} \right\}, \tag{2.33}
 \end{aligned}$$

which is the standard formula for massless leptons.

2.3 Form Factor Predictions

The discussion so far has been completely general: the decomposition of the hadronic currents in Equations 2.9 and 2.10 is model-independent, and the differential decay rates in Equations 2.17 and 2.32 follow immediately from this decomposition and the Dirac nature of the fermions. Within the Standard Model, the only unknown quantities are therefore $|V_{cb}|$ and the form factors; $|V_{cb}|$ sets the overall normalization, while the form factors describe

the dynamics of the decay process. Because the form factors parameterize the hadronization of quarks into mesons, a non-perturbative QCD process, calculating them is a difficult problem which has been approached theoretically in a number of different ways.

Calculations of the form factors typically break the problem into two pieces: calculation of the form factors at a fixed value of q^2 , typically $q^2 = q_{\max}^2$, and calculation of the q^2 dependence of the form factors. At $q^2 = q_{\max}^2$, the hadronic system is least disturbed, which makes calculations at this point relatively easy. The variation of the form factors as functions of q^2 is typically assumed to have a very simple form.

One older approach, used in the ISGW model [20], is to assume an exponential distribution for the form factors:

$$f_i(q^2) = C e^{-a(q_{\max}^2 - q^2)} , \quad (2.34)$$

where C is a normalization constant and a is a parameter related to the meson size. Another approach, used in the WSB [19] and KS [9] models is the “nearest-pole dominance” model, in which the form factors are assumed to depend on q^2 like

$$f_i(q^2) = f_i(0) \cdot \left(\frac{m_{\text{pole}}^2}{m_{\text{pole}}^2 - q^2} \right)^n , \quad (2.35)$$

where n is either 1 or 2, depending on the form factor involved, and m_{pole} is the mass of a $Q\bar{q}'$ meson with the quark content and quantum numbers determined by the underlying transition; for the $b \rightarrow c$ transition in $B \rightarrow D^{(*)}$, it would be the mass of the B_c meson, approximately $6.3 \text{ GeV}/c^2$.

2.3.1 Heavy Quark Effective Theory

The development of Heavy Quark Symmetry and Heavy Quark Effective Theory [5] has led to improvements in the precision of form factor predictions.

In the heavy-quark limit $m_b \rightarrow \infty$ and $m_c \rightarrow \infty$, the form factors are easy to calculate. We first introduce the HQET parameterization of the form factors, which are written in terms of four-velocities ($w \equiv v \cdot v'$) rather than momenta ($q \equiv p - p'$). As before, the structure of the currents is completely determined by Lorentz invariance; we have

$$\langle P'(v') | V^\mu | P(v) \rangle = \sqrt{m_B m_D} [h_+(w)(v + v')^\mu + h_-(w)(v - v')^\mu], \quad (2.36)$$

and

$$\begin{aligned} \langle V(v', \varepsilon) | V^\mu - A^\mu | P(v) \rangle &= \sqrt{m_B m_{D^*}} \left[i h_V(v \cdot v') \epsilon^{\mu\nu\alpha\beta} \varepsilon_\nu^* v'_\alpha v_\beta \right. \\ &\quad - h_{A_1}(v \cdot v') \varepsilon^{*\mu} (v \cdot v' + 1) + h_{A_2}(v \cdot v') v^\mu \varepsilon^* \cdot v \\ &\quad \left. + h_{A_3}(v \cdot v') v'^\mu \varepsilon^* \cdot v \right]. \end{aligned} \quad (2.37)$$

The HQET form factors h_i are related to the traditional form factors by [24]

$$Rf_+(q^2) = h_+(w) - \left(\frac{m_B - m_D}{m_B + m_D}\right) h_-(w) \quad (2.38)$$

$$R^{-1}f_-(q^2) = \left[\left(\frac{w+1}{2} - R^{-2}\right) h_+(w) + \left(\frac{w-1}{2} \cdot \frac{m_B + m_D}{m_B - m_D}\right) h_-(w) \right] \frac{m_B^2 - m_D^2}{q^2} \quad (2.39)$$

$$R^* V(q^2) = h_V(w) \quad (2.40)$$

$$R^{*-1}A_1(q^2) = \frac{w+1}{2} h_{A_1}(w) \quad (2.41)$$

$$R^* A_2(q^2) = h_{A_3}(w) + \frac{m_{D^*}}{m_B} h_{A_2}(w) \quad (2.42)$$

$$R^* A_0(q^2) = \frac{m_B(w+1)}{m_B + m_{D^*}} h_{A_1}(w) + \left[\frac{q^2 - m_B^2 + m_{D^*}^2}{2m_{D^*}(m_B + m_{D^*})} \right] \left(h_{A_3}(w) - \frac{m_{D^*}}{m_B} h_{A_2}(w) \right), \quad (2.43)$$

with constants R and R^* are given by

$$R = \frac{2\sqrt{m_B m_D}}{m_B + m_D} \quad \text{and} \quad R^* = \frac{2\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}. \quad (2.44)$$

These form factors satisfy the simple relations in the heavy-quark limit:

$$h_-(w) = h_{A_2}(w) = 0 \quad (2.45)$$

and, at the zero-recoil point $w = 0$ ($q^2 = q_{\text{max}}^2$),

$$h_+(0) = h_V(0) = h_{A_1}(0) = h_{A_3}(0) = 1. \quad (2.46)$$

This can easily be understood by considering the situation in the rest frame of the decaying B meson. The b quark is infinitely massive and so is at rest

in this frame; it decays to an infinitely-massive c quark which, for $w = 0$, is also at rest in the B rest frame. The light degrees of freedom, the \bar{q} and the gluons, see no change in the color field after the $b \rightarrow c$ transition, and their wavefunction remains undisturbed.

In the heavy-quark limit, the w dependence of these form factors is identical, such that these four form factors are all equal, and we write

$$h_+(w) = h_V(w) = h_{A_1}(w) = h_{A_3}(w) = \xi(w) , \quad (2.47)$$

where $\xi(w)$ is called the Isgur-Wise function [25].

When the b and c quark masses become finite, these form factors are no longer equal to $\xi(w)$ or to each other. The Isgur-Wise function is sometimes expanded as a Taylor series about the zero-recoil point; a linear form which has been used extensively is

$$\xi(w) = 1 - \rho^2(w - 1) + \mathcal{O}[(w - 1)^2] . \quad (2.48)$$

HQET is an effective field theory which allows us to calculate deviations from the heavy-quark limit as an expansion in powers of Λ_{QCD}/m_b and Λ_{QCD}/m_c , where Λ_{QCD} is an energy scale typical of QCD processes, and is of the order of 200 MeV.

2.3.2 The CLN Form Factor Model

Caprini, Lellouch, and Neubert (CLN) have calculated the form factors [21] describing $B \rightarrow D^{(*)}$ transitions using dispersion relations. This model of the form factors makes use of heavy-quark symmetry and HQET, and claims to describe the semileptonic form factors with an accuracy better than 2%. We use this model to describe both our signal processes $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and the corresponding normalization processes $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$.

The CLN model calculates one reference form factor, $f_+(q^2)$, as a Taylor series expansion in z (defined in Equation 2.3), where the terms in the expansion take into account all heavy-quark-symmetry-breaking effects. The expansion has the form

$$f_+(q^2) = 1 - 8\rho_+^2 z + (51\rho_+^2 - 10)z^2 - (252\rho_+^2 - 84)z^3, \quad (2.49)$$

where ρ_+^2 is an externally-determined slope parameter describing the form factor. A spin symmetry relates this form factor to $A_1(q^2)$ (the symmetry actually relates $h_+(w)$ and $h_{A_1}(w)$, but these HQET form factors are related to the invariant form factors as above), so that we have, for $A_1(q^2)$,

$$A_1(q^2) = 1 - 8\rho_{A_1}^2 z + (53\rho_{A_1}^2 - 15)z^2 - (231\rho_{A_1}^2 - 91)z^3. \quad (2.50)$$

The remaining form factors can be expressed in terms of these reference form factors as Taylor series expansions in $(w - 1)$, where, again, the terms in the

expansion take into account symmetry-breaking effects. We have

$$f_-(q^2) = f_+(q^2) \left[0.0036 - 0.0068(w-1) + 0.0017(w-1)^2 - 0.0013(w-1)^3 \right] \frac{m_B^2 - m_D^2}{q^2} \quad (2.51)$$

$$V(q^2) = A_1(q^2) \left[R_1 - 0.12(w-1) + 0.05(w-1)^2 \right] \times \frac{(m_B + m_{D^*})^2}{2m_B m_{D^*} (w+1)} \quad (2.52)$$

$$A_2(q^2) = A_1(q^2) \left[R_2 + 0.11(w-1) - 0.06(w-1)^2 \right] \times \frac{(m_B + m_{D^*})^2}{2m_B m_{D^*} (w+1)} \quad (2.53)$$

$$A_0(q^2) = f_+(q^2) \left[1.1548 - 0.2088(w-1) + 0.0032(w-1)^2 - 0.0009(w-1)^3 \right], \quad (2.54)$$

where R_1 and R_2 are parameters which can be determined experimentally.

2.4 Kinematic Distributions of Semileptonic Decays

The form factor parameters ρ_+^2 , $\rho_{A_1}^2$, R_1 , and R_2 in the CLN model have been measured experimentally [44, 3]. Figures 2.2–2.15 show distributions of several kinematic quantities for $B \rightarrow D^{(*)} \ell^- \bar{\nu}_\ell$, both for decays involving light leptons and those involving τ .

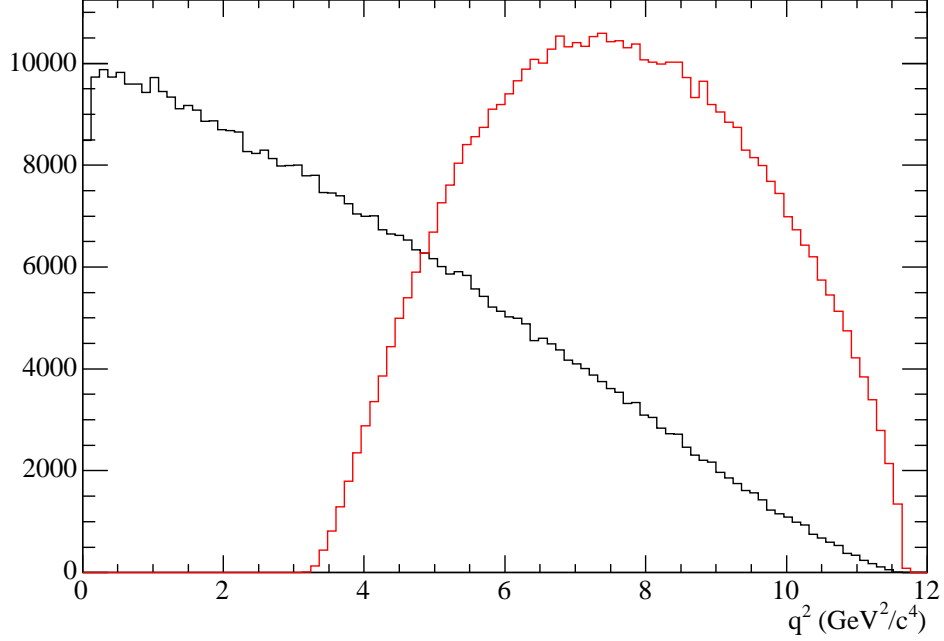


Figure 2.2: q^2 distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ (black) and $B \rightarrow D\tau^-\bar{\nu}_\tau$ (red) in the CLN form factor model with experimentally-measured shape parameters. The distributions are normalized to equal areas.

We see that the q^2 distribution for signal processes is much harder than the distribution for the light leptons, and that, even though the τ lepton is produced with a harder spectrum than the corresponding light-lepton modes, the energy spectrum of the observable secondary lepton is actually much softer.

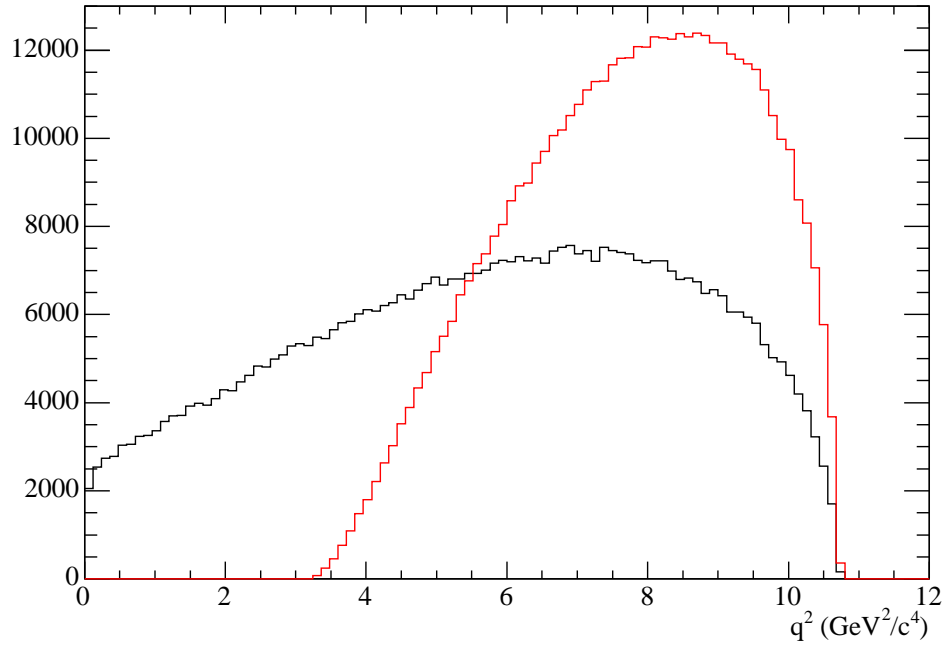


Figure 2.3: q^2 distributions for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ (black) and $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ (red) in the CLN form factor model with experimentally-measured shape parameters. The distributions are normalized to equal areas.

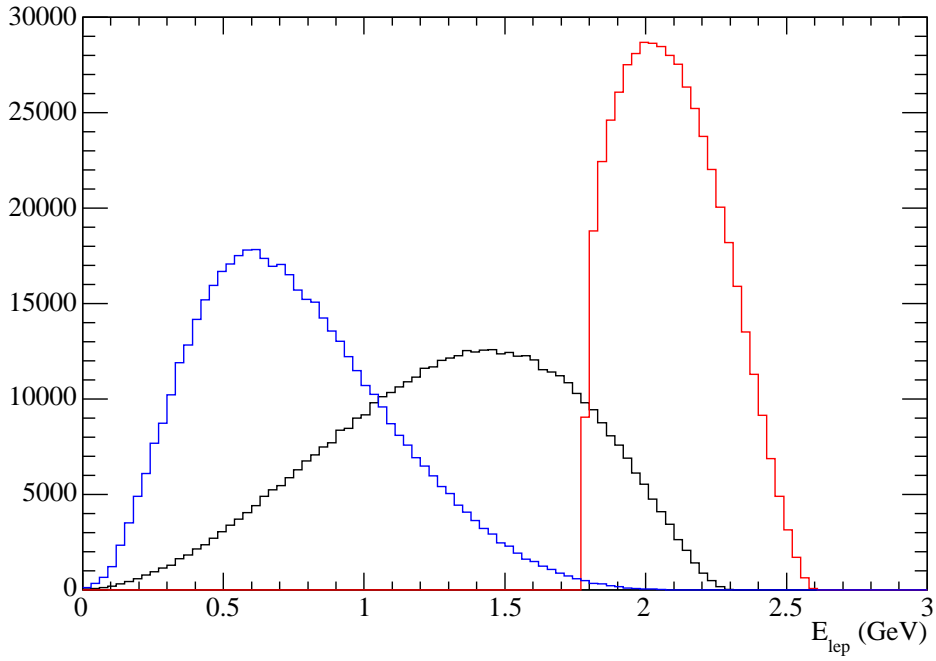


Figure 2.4: Lepton energy distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D\tau^-\bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The three curves show the distributions of the ℓ^- energy in $B \rightarrow D\ell^-\bar{\nu}_\ell$ decays (black), the τ^- energy in $B \rightarrow D\tau^-\bar{\nu}_\tau$ decays (red), and the secondary lepton energy in $B \rightarrow D\tau^-\bar{\nu}_\tau$ decays (blue), all measured in the rest frame of the B meson. The distributions are normalized to equal areas.

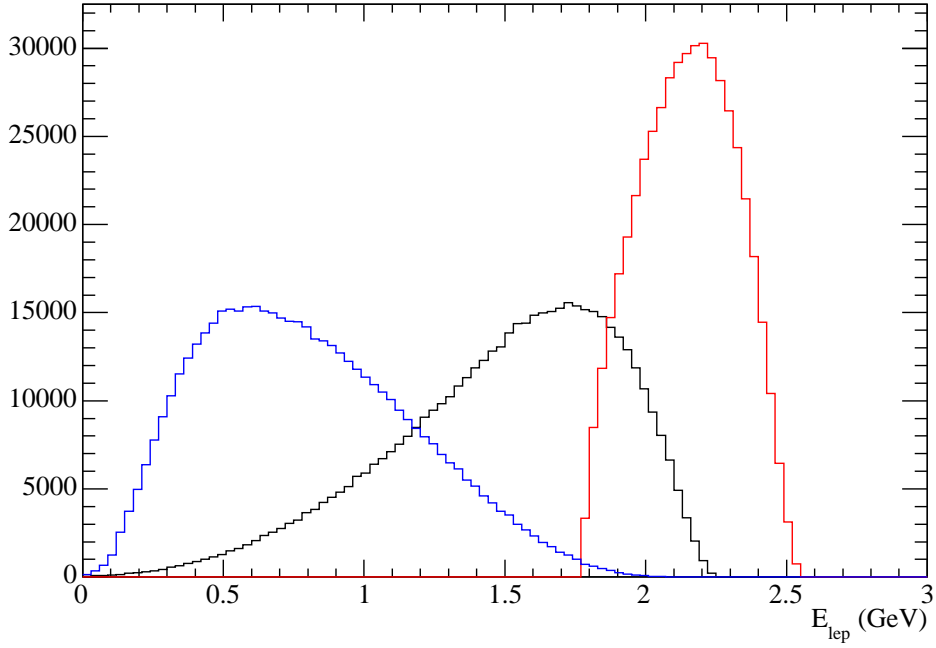


Figure 2.5: Lepton energy distributions for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ and $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The three curves show the distributions of the ℓ^- energy in $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ decays (black), the τ^- energy in $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ decays (red), and the secondary lepton energy in $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ decays (blue), all measured in the rest frame of the B meson. The distributions are normalized to equal areas.

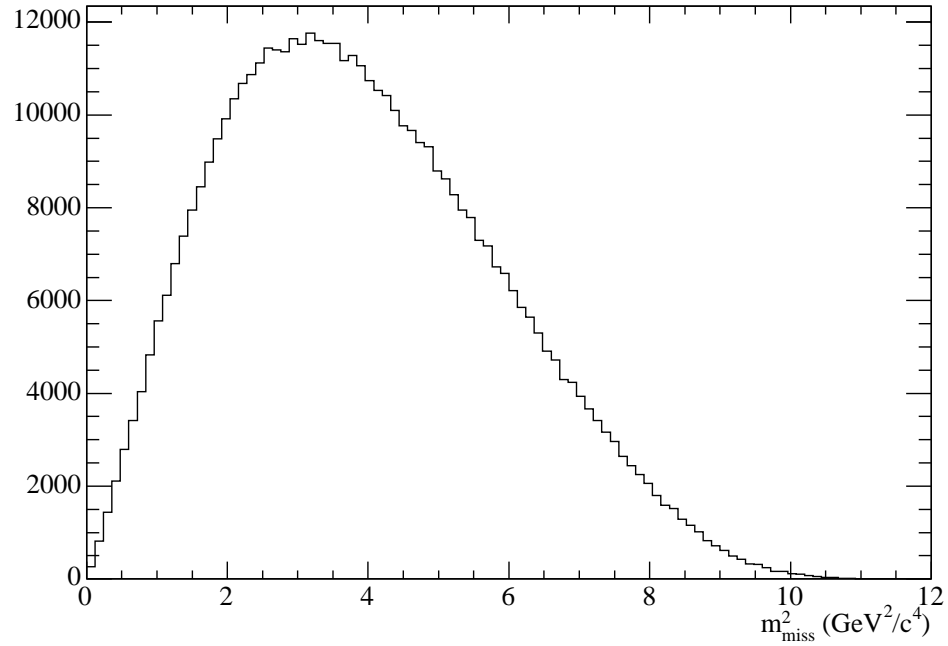


Figure 2.6: m_{miss}^2 distribution for $B \rightarrow D\tau^-\bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The distribution for the normalization mode $B \rightarrow D\ell^-\bar{\nu}_\ell$ is a delta function at zero.

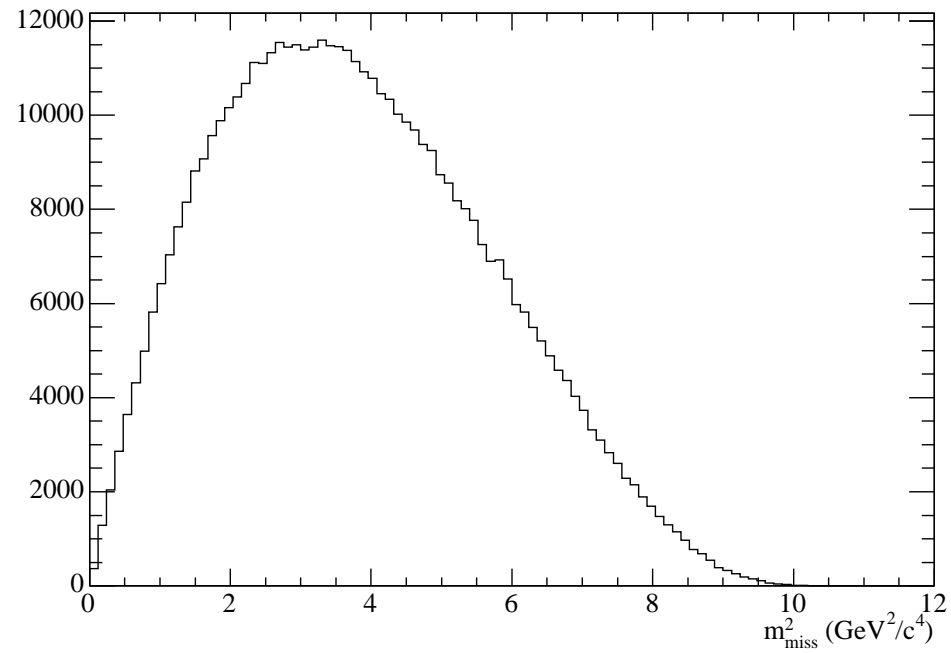


Figure 2.7: m_{miss}^2 distribution for $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The distribution for the normalization mode $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ is a delta function at zero.

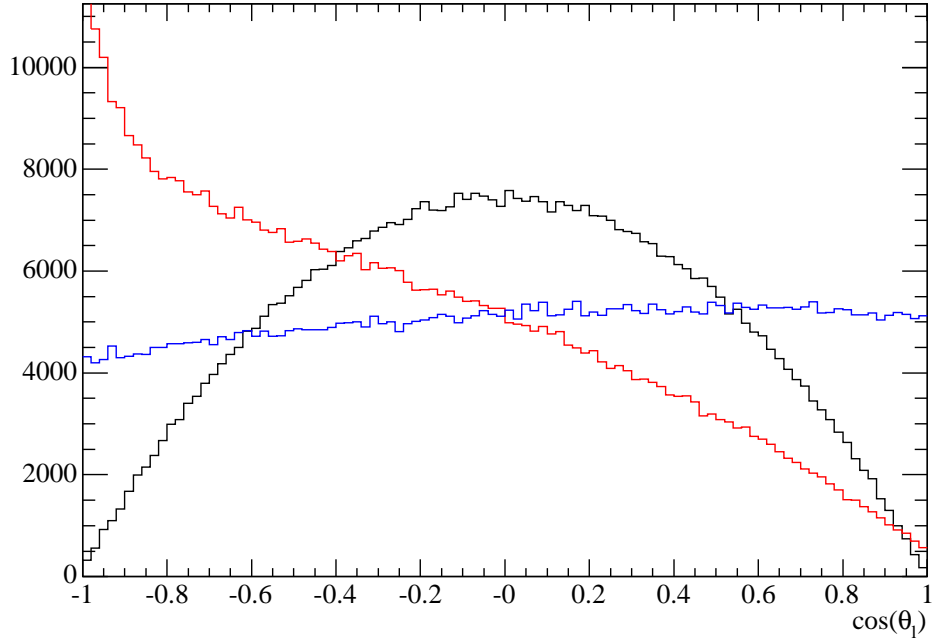


Figure 2.8: $\cos \theta_\ell$ distributions for $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D\tau^-\bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The three curves show the distributions of $\cos \theta_\ell$ in $B \rightarrow D\ell^-\bar{\nu}_\ell$ decays (black), $\cos \theta_\tau$ for the τ^- in $B \rightarrow D\tau^-\bar{\nu}_\tau$ decays (red), and $\cos \theta_\ell$ for the secondary lepton in $B \rightarrow D\tau^-\bar{\nu}_\tau$ decays (blue). Note that the distribution for $B \rightarrow D\ell^-\bar{\nu}_\ell$ is proportional to $\sin^2 \theta_\ell$, as expected from Equation 2.18, and that the $\cos \theta_\ell$ distribution produced in $B \rightarrow D\tau^-\bar{\nu}_\tau$ is largely flattened out when the τ^- decays. The distributions are normalized to equal areas.

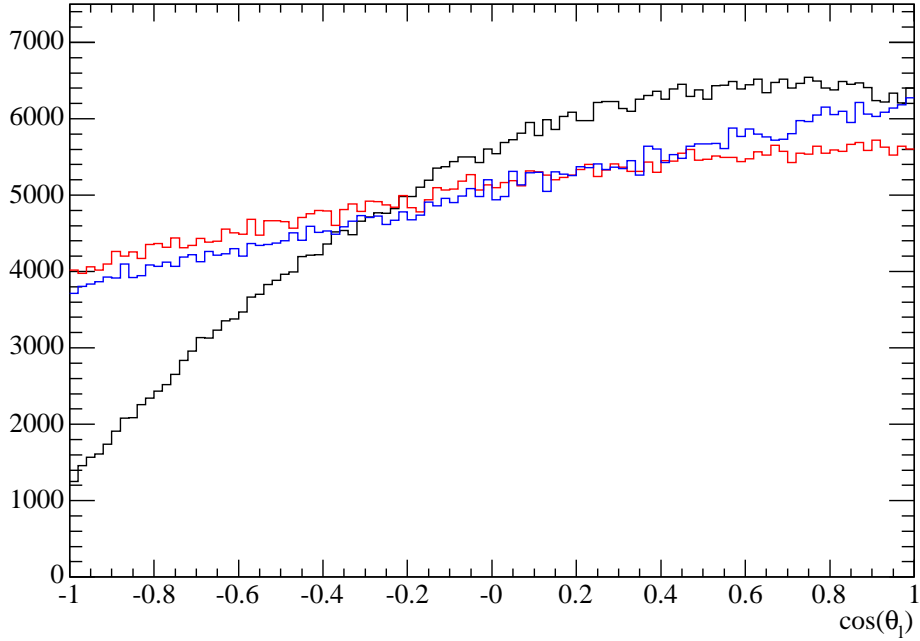


Figure 2.9: $\cos \theta_\ell$ distributions for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ and $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters. The three curves show the distributions of $\cos \theta_\ell$ in $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ decays (black), $\cos \theta_\tau$ for the τ^- in $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ decays (red), and $\cos \theta_\ell$ for the secondary lepton in $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ decays (blue). The distributions are normalized to equal areas.

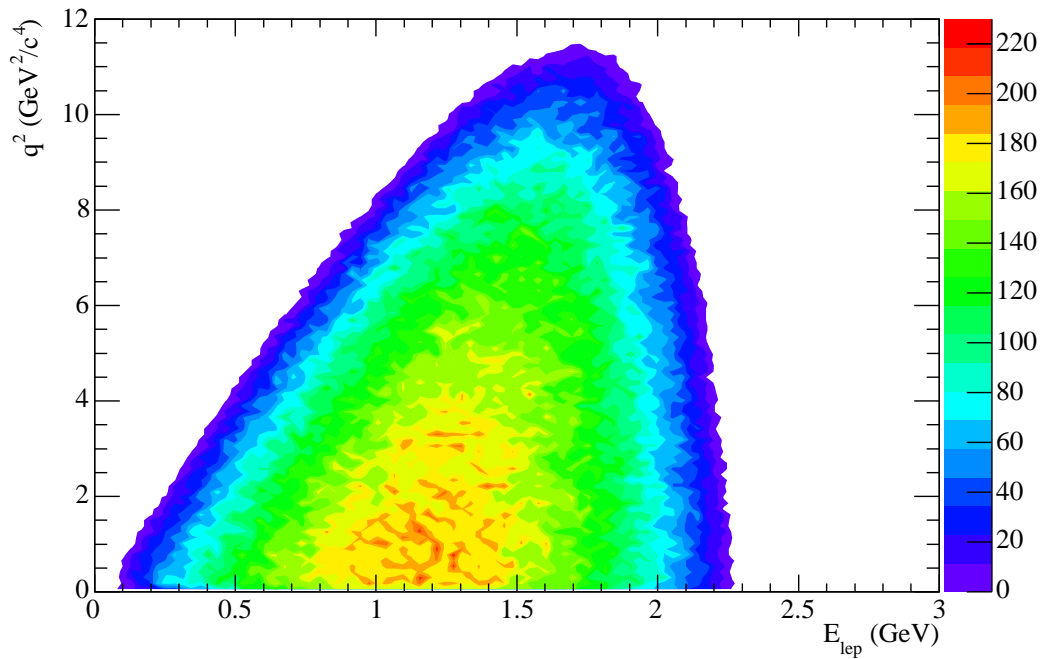


Figure 2.10: Dalitz plot for $B \rightarrow D \ell^- \bar{\nu}_\ell$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_ℓ , where the energy is measured in the rest frame of the B meson.

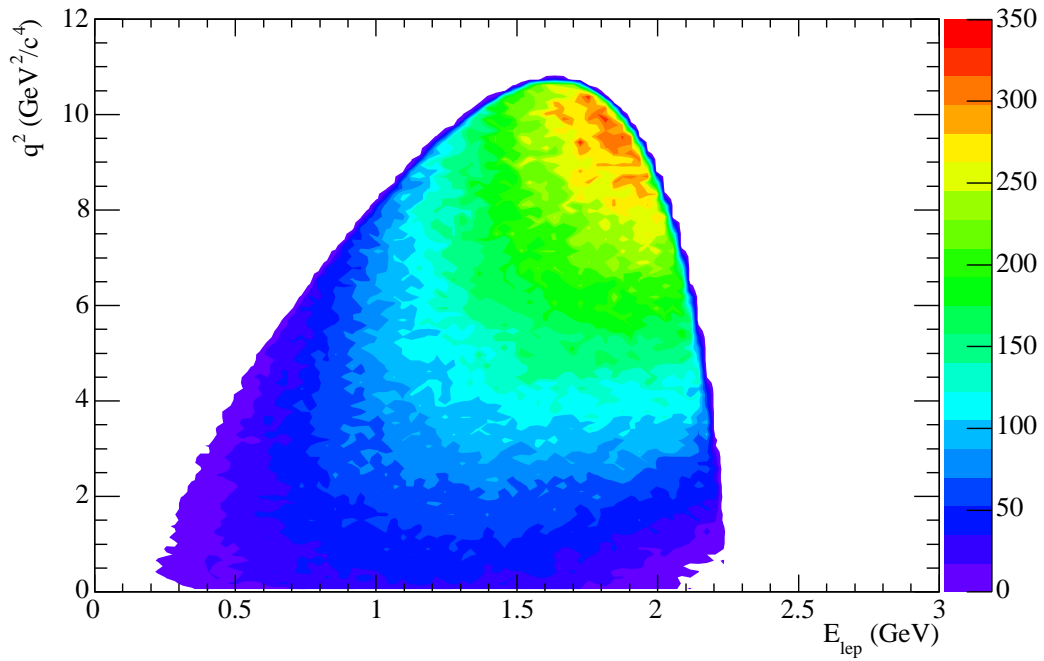


Figure 2.11: Dalitz plot for $B \rightarrow D^* \ell^- \bar{\nu}_\ell$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_ℓ , where the energy is measured in the rest frame of the B meson.

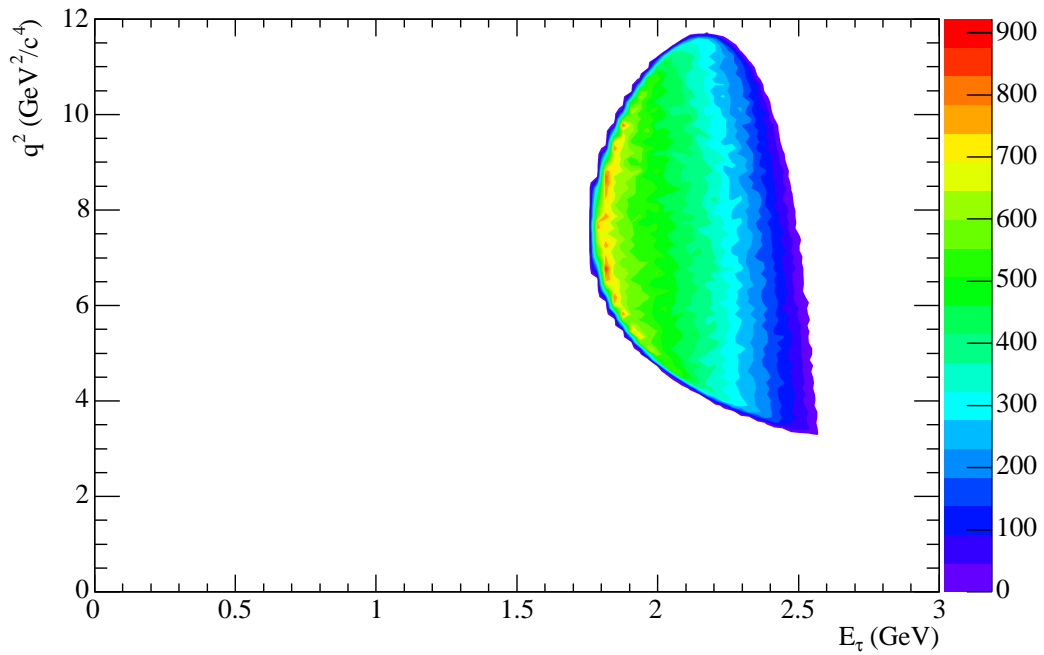


Figure 2.12: Dalitz plot for $B \rightarrow D\tau^-\bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_τ , where the energy is measured in the rest frame of the B meson.

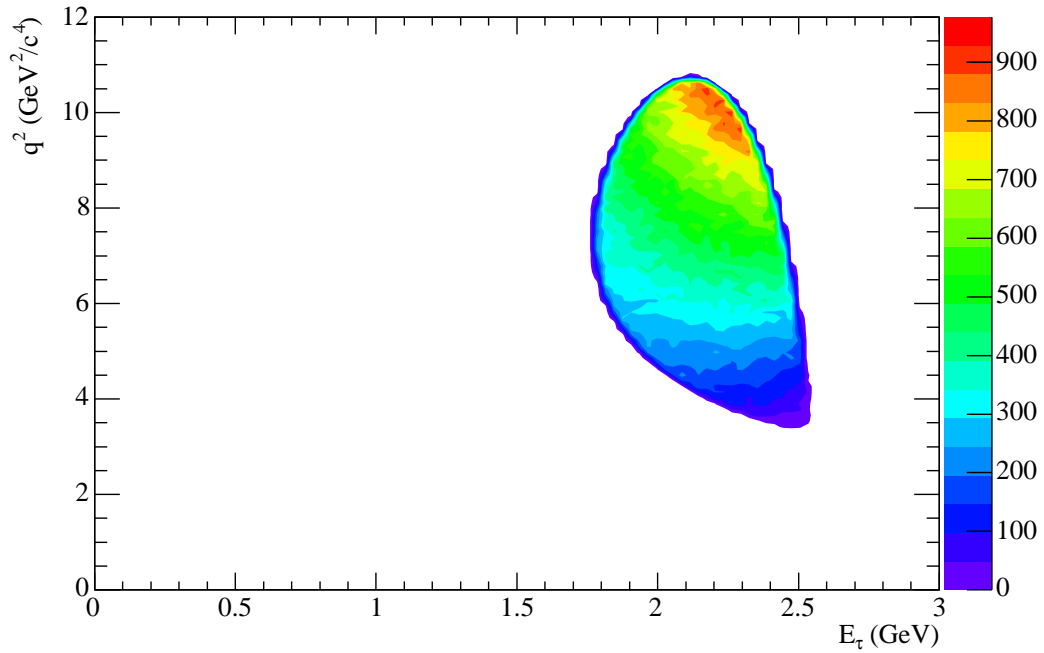


Figure 2.13: Dalitz plot for $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_τ , where the energy is measured in the rest frame of the B meson.

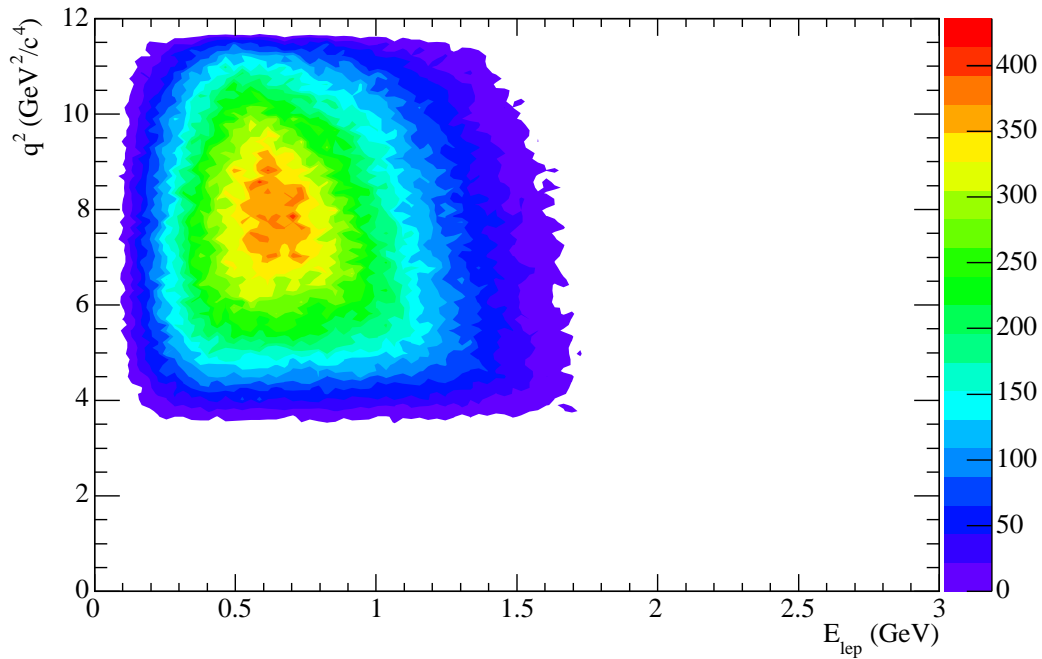


Figure 2.14: Pseudo-Dalitz plot for $B \rightarrow D\tau^-\bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_ℓ , where the energy is that of the secondary lepton and is measured in the rest frame of the B meson.

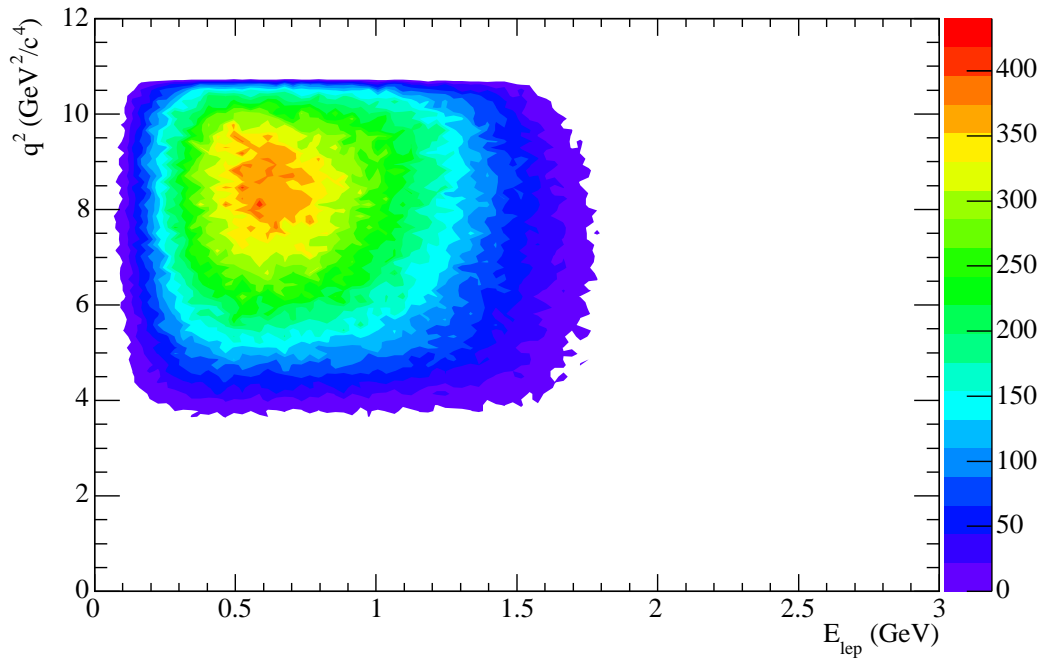


Figure 2.15: Pseudo-Dalitz plot for $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in the CLN form factor model with experimentally-measured shape parameters, showing the distribution of q^2 vs E_ℓ , where the energy is that of the secondary lepton and is measured in the rest frame of the B meson.

2.5 Non-Standard Model Contributions

Many models of physics beyond the Standard Model predict the existence of charged Higgs bosons H^\pm . Figure 1.4 shows a Feynman diagram for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ mediated by H^\pm exchange. As we pointed out in Chapter 1, the fact that the Higgs boson couples to the lepton mass means that that this process, if it occurs, may have observable consequences for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ without affecting $B \rightarrow D^{(*)}e^-\bar{\nu}_e$ or $B \rightarrow D^{(*)}\mu^-\bar{\nu}_\mu$.

The SM includes a single complex Higgs doublet field which produces the neutral scalar Higgs boson H^0 . The simplest class of extensions to the Higgs sector of the SM are known as two-Higgs-doublet models (2HDM), in which a second doublet field is added, giving rise to charged Higgs bosons. In one particular model, the Type-II 2HDM, one of the two Higgs doublets couples to the up-type quarks and generates their masses, while the second doublet couples to the down-type quarks, although more general coupling schemes are also possible. In the Type-II 2HDM, the effects of the charged Higgs boson can be expressed as a function of two new physics parameters: m_H , the mass of the H^\pm , and $\tan\beta \equiv v_2/v_1$, the ratio of the vacuum expectation values of the two Higgs fields.

Supersymmetric models also require the existence of multiple Higgs fields and make predictions similar to 2HDM at tree level, but higher order corrections due to gluino-squark loops and neutralino-squark loops can be significant. These effects can be absorbed into the definition of $\tan \beta$, however, so that results obtained in the Type-II 2HDM are qualitatively still valid [26].

In a Type-II 2HDM, the interaction Lagrangian between the charged Higgs and the quark and lepton fields is given by [15]

$$\mathcal{L}_H = (2\sqrt{2}G_F)^{1/2} \left[\bar{u}_L V_{ij} M_d d_R \tan \beta + \bar{u}_R M_u V_{ij} d_L \cot \beta + \bar{\nu}_L M_\ell \ell_R \tan \beta \right] H^+ , \quad (2.55)$$

where \bar{u} , d , $\bar{\nu}$, and ℓ are the quark and lepton field operators, M_d , M_u , and M_ℓ are diagonal quark and lepton mass matrices, and V_{ij} is the CKM quark mixing matrix. With these couplings, the amplitude of charged Higgs exchange in a $b \rightarrow c \ell^- \bar{\nu}_\ell$ transition is proportional to $m_b m_\ell \tan^2 \beta$, and the effect of the H^\pm is more significant for heavy leptons and for large $\tan \beta$.

Tanaka has studied $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ in a 2HDM [8]. Plots of the branching fractions and τ polarization are shown in Figures 2.16–2.19. These figures show that the branching fraction of $B \rightarrow D \tau^- \bar{\nu}_\tau$ is more sensitive to new physics contributions than that of $B \rightarrow D^* \tau^- \bar{\nu}_\tau$. The τ polarization promises even greater sensitivity, but this quantity is not accessible in our measure-

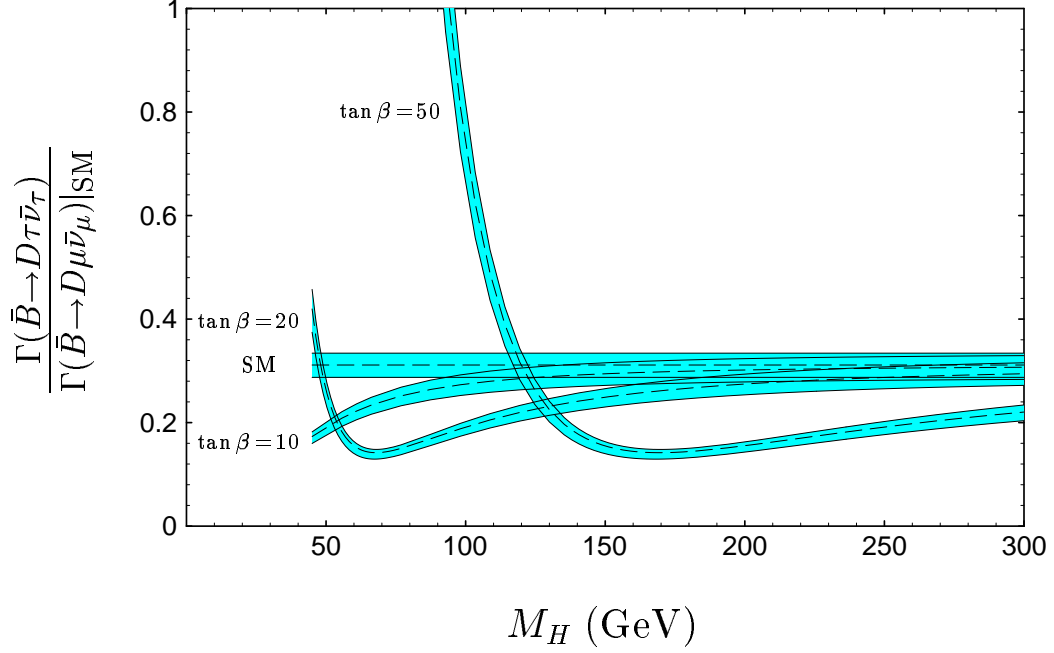


Figure 2.16: Branching ratio $B \rightarrow D\tau^-\bar{\nu}_\tau/B \rightarrow D\mu^-\bar{\nu}_\mu$ in a 2HDM. The branching ratio is shown as a function of the charged Higgs mass for various values of $\tan\beta$, including the Standard Model prediction. The blue bands indicate the uncertainties due to the form factors. Taken from [8].

ment. Chen and Geng [6] make similar predictions, and also find that the lepton forward-backward asymmetry can be very sensitive to a charged Higgs contribution, although this asymmetry is not currently accessible either.

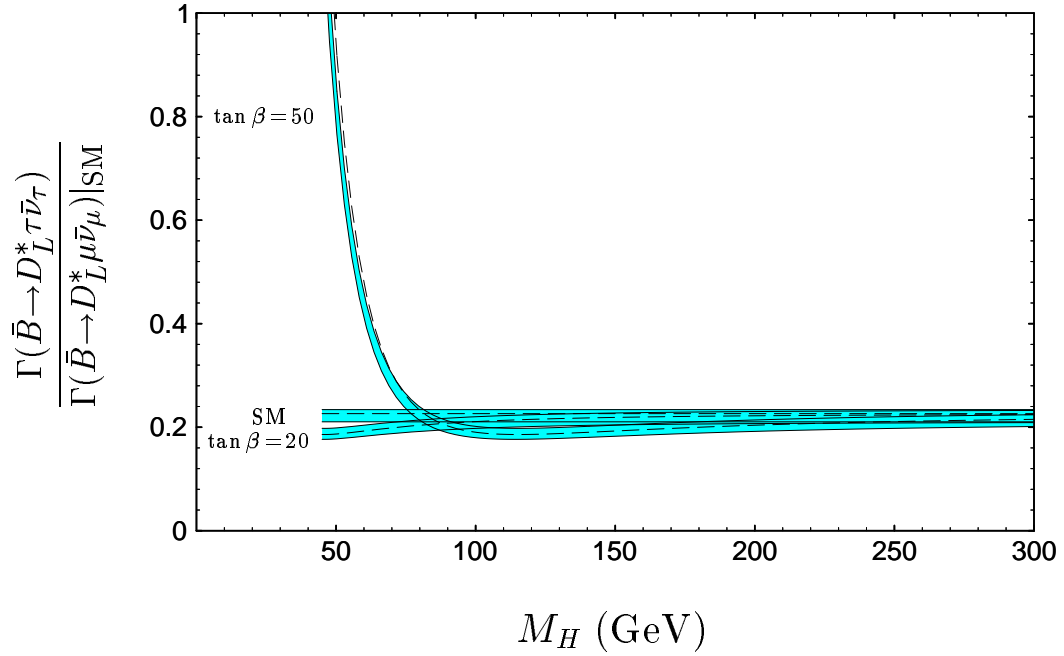


Figure 2.17: Branching ratio $B \rightarrow D^* \tau^- \bar{\nu}_\tau / B \rightarrow D^* \mu^- \bar{\nu}_\mu$ in a 2HDM. The branching ratio is shown only for longitudinally polarized D^* mesons since the Higgs boson does not contribute to the amplitude to the transverse polarization states. The branching ratio is shown as a function of the charged Higgs mass for various values of $\tan \beta$, including the Standard Model prediction. The blue bands indicate the uncertainties due to the form factors. Compared to Figure 2.16, we see that the Higgs contribution is much smaller. Taken from [8].

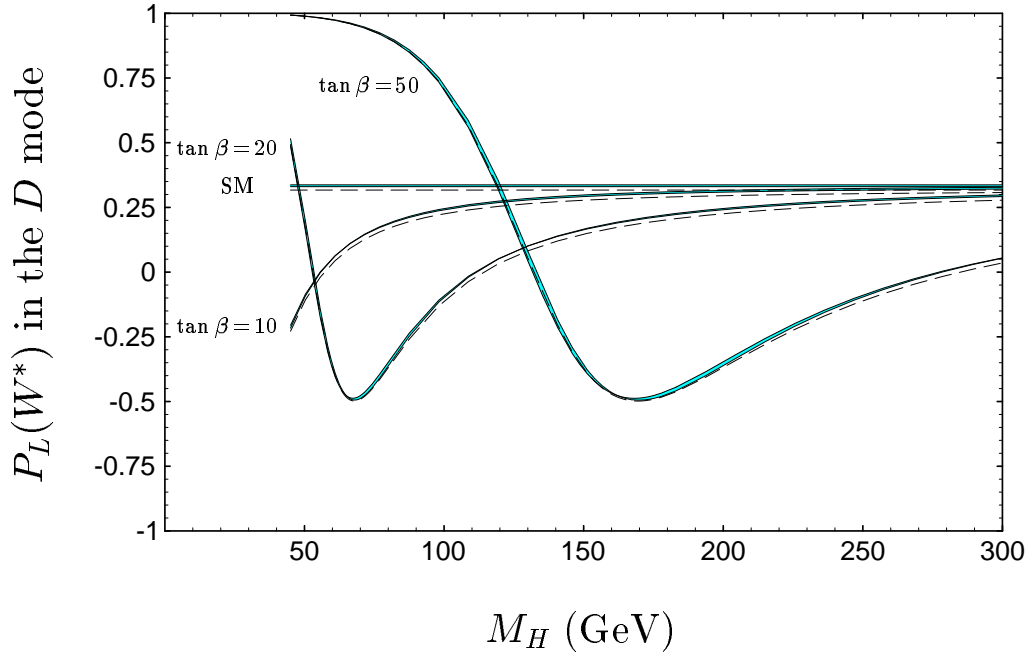


Figure 2.18: τ polarization in $B \rightarrow D\tau^-\bar{\nu}_\tau$ in a 2HDM. The polarization is shown as a function of the charged Higgs mass for various values of $\tan\beta$, including the Standard Model prediction. The blue bands indicate the uncertainties due to the form factors. Compared to Figures 2.16 and 2.17, we see that the polarization can be a much more sensitive probe of new physics than the branching fractions. Taken from [8].

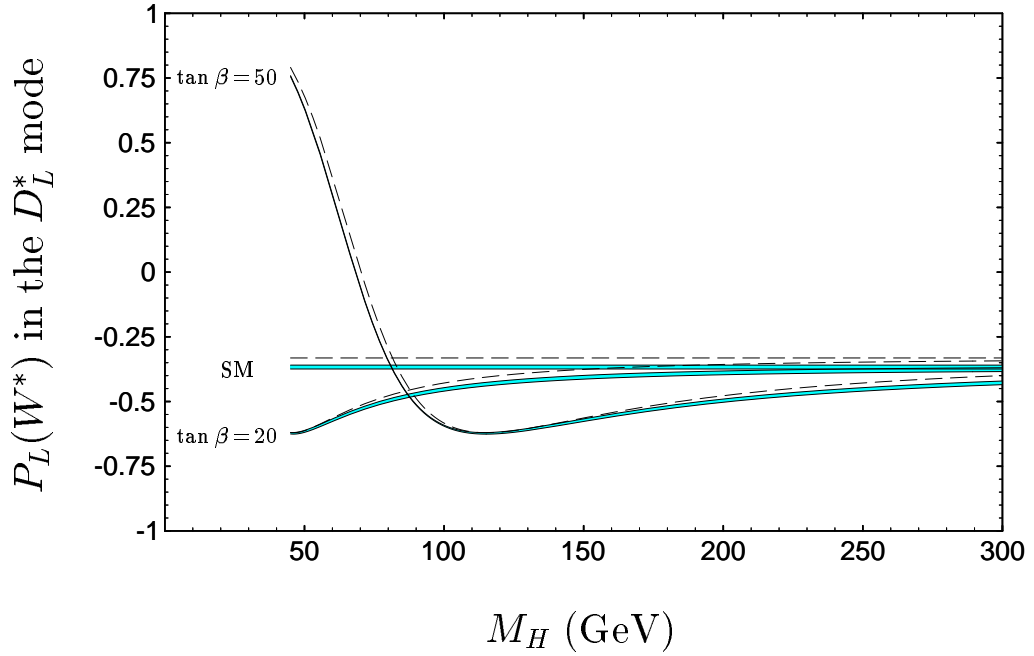


Figure 2.19: τ polarization in $B \rightarrow D^* \tau^- \bar{\nu}_\tau$ in a 2HDM. The polarization is shown as a function of the charged Higgs mass for various values of $\tan \beta$, including the Standard Model prediction. The blue bands indicate the uncertainties due to the form factors. Compared to Figures 2.16 and 2.17, we see that the polarization can be a much more sensitive probe of new physics than the branching fractions. Taken from [8].

Chapter 3

The $B_A B_{AR}$ Experiment

The data used in this dissertation were collected with the *BABAR* detector at the PEP-II asymmetric-energy storage rings located at the Stanford Linear Accelerator Center. The layout of the accelerator and detector is shown schematically in Figure 3.1.

3.1 The PEP-II B Factory

The PEP-II B Factory is an electron-positron collider which typically operates at a center-of-mass energy $\sqrt{s} = 10.58$ GeV. This energy is chosen because it corresponds to the mass of the $\Upsilon(4S)$ resonance, a $b\bar{b}$ bound state which decays almost exclusively to $B\bar{B}$ pairs, either as $B^0\bar{B}^0$ or B^+B^- mesons. The cross section for the $e^+e^- \rightarrow \Upsilon(4S)$ process is approximately 1.1 nb and

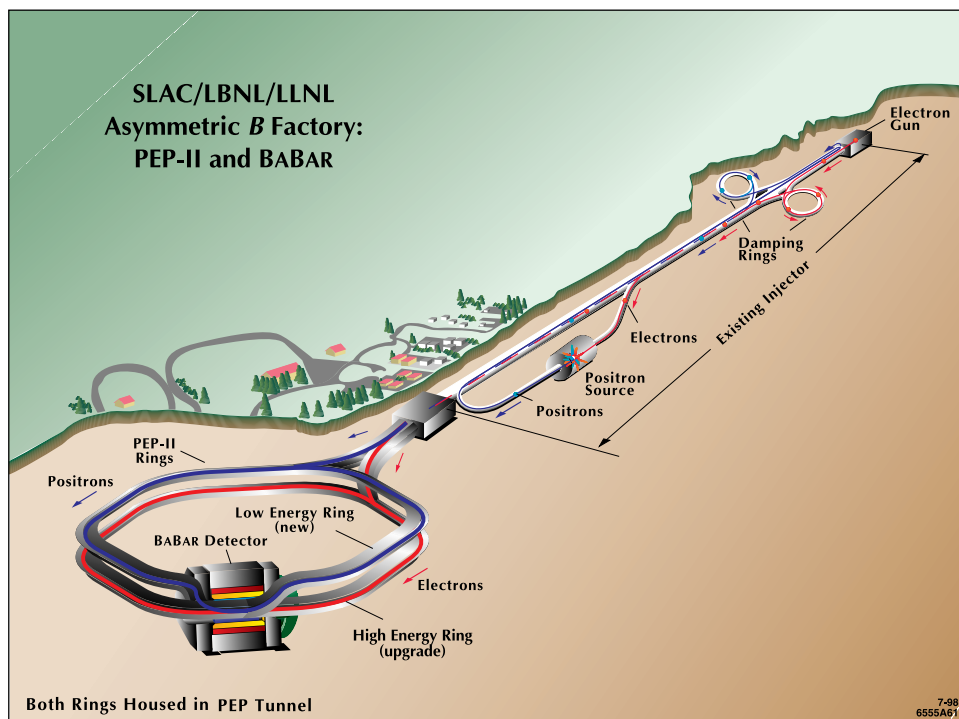


Figure 3.1: Layout of the SLAC linear accelerator, PEP-II storage rings, and BABAR detector.

the collider typically operates at a luminosity of $9 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, so that the B Factory produces approximately ten $B\bar{B}$ pairs per second.

The mass of the $\Upsilon(4S)$ resonance, $10.58 \text{ GeV}/c^2$, is only slightly larger than twice the B meson mass, $2 \times 5.28 \text{ GeV}/c^2 = 10.56 \text{ GeV}/c^2$. As a consequence, B mesons from $\Upsilon(4S) \rightarrow B\bar{B}$ are produced nearly at rest in the center-of-mass frame of the $\Upsilon(4S)$. This allows us to use two powerful kinematic variables to identify reconstructed B mesons:

$$m_{\text{ES}} \equiv \sqrt{s/4 - p_B^2} \quad (3.1)$$

$$\Delta E \equiv E_B - \sqrt{s}/2, \quad (3.2)$$

where p_B and E_B are the momentum and energy, respectively, of a B meson candidate in the $\Upsilon(4S)$ rest frame. For correctly reconstructed B mesons, m_{ES} peaks at the B mass and ΔE peaks at zero.

An additional benefit of working at the $\Upsilon(4S)$ is that there is not enough energy available to produce any hadronic system beyond the $B\bar{B}$ pair. The lightest hadron, the neutral pion, has a mass of $135 \text{ MeV}/c^2$, which is too heavy to be produced via $\Upsilon(4S) \rightarrow B\bar{B}\pi^0$. This means that, in a $B\bar{B}$ event, all observed hadrons must be decay products of one of the two B mesons in the event, a fact which is particularly important in this analysis, where we

will fully reconstruct one of the two B mesons and then associate all of the remaining particles to the second B meson.

PEP-II is an asymmetric collider — the electron beam has an energy of 9 GeV, while the positron beam has an energy of 3.1 GeV. The $\Upsilon(4S)$ mesons are thus produced with a Lorentz boost of approximately 0.55. Since the B mesons are nearly at rest in the $\Upsilon(4S)$ frame, they are produced with a similar boost in the laboratory frame.

3.2 The *B*A*B*A*R* Detector

BABAR is a general purpose particle detector surrounding the interaction region of PEP-II. The detector is nearly hermetic, covering approximately 90% of the solid angle in the e^+e^- center-of-mass frame. Figures 3.2 and 3.3 show the layout of the *BABAR* detector. Five major detector subsystems make up *BABAR*: a silicon vertex tracker (SVT) and a gas-based drift chamber (DCH) for charged track reconstruction, a Cherenkov radiation detector (DIRC) for charged particle identification, a scintillating crystal calorimeter (EMC) for photon reconstruction and electron identification, and an instrumented flux return (IFR) for muon and neutral hadron identification. The SVT, DCH, DIRC, and EMC are located inside the field of a 1.5 T supercon-

ducting solenoid. These systems are described in more detail in the following sections.

Because PEP-II is an asymmetric collider, the *BABAR* detector is itself asymmetric. The region where the two beams collide and where B mesons are produced is off-center in the detector. The EMC has a forward endcap but no backward endcap, and the standoff box for the DIRC and all of the DCH electronics are located in the backward direction. This asymmetry is designed to maximize the acceptance of the detector in the rest frame of the $\Upsilon(4S)$ decay products while minimizing the amount of material in the active detector volume.

3.2.1 Silicon Vertex Tracker

The SVT is a detector designed to precisely reconstruct charged particle trajectories near the interaction region and to measure the position of particle decay vertices. It is the closest subdetector to the interaction region, and consists of five concentric layers of double-sided silicon strip detectors arranged around the beam pipe as shown in Figures 3.4 and 3.5. Layer 1 has a minimum radius of 32 mm, while Layer 4 has a radius of 114–144 mm. The three inner layers are located as close to the beam pipe as practical in order to maximize the vertex resolution. The beam pipe, the SVT support structure, and the SVT

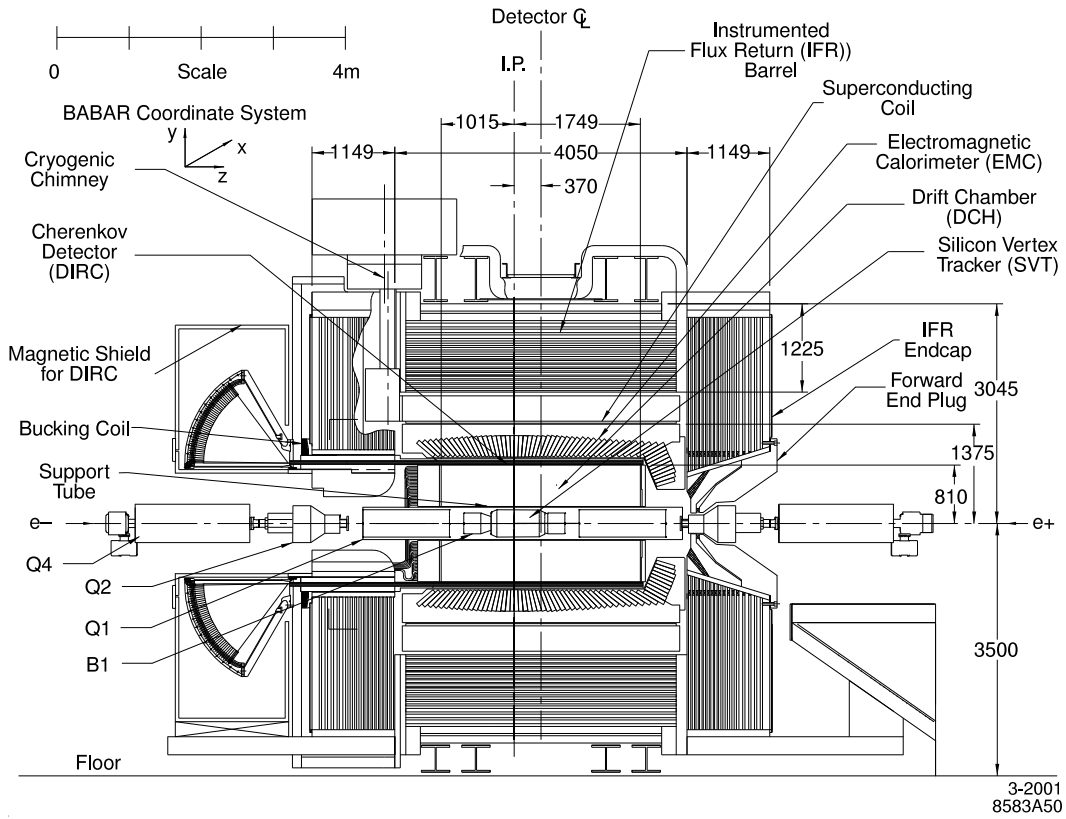


Figure 3.2: Layout of the BABAR detector, longitudinal cross-section.

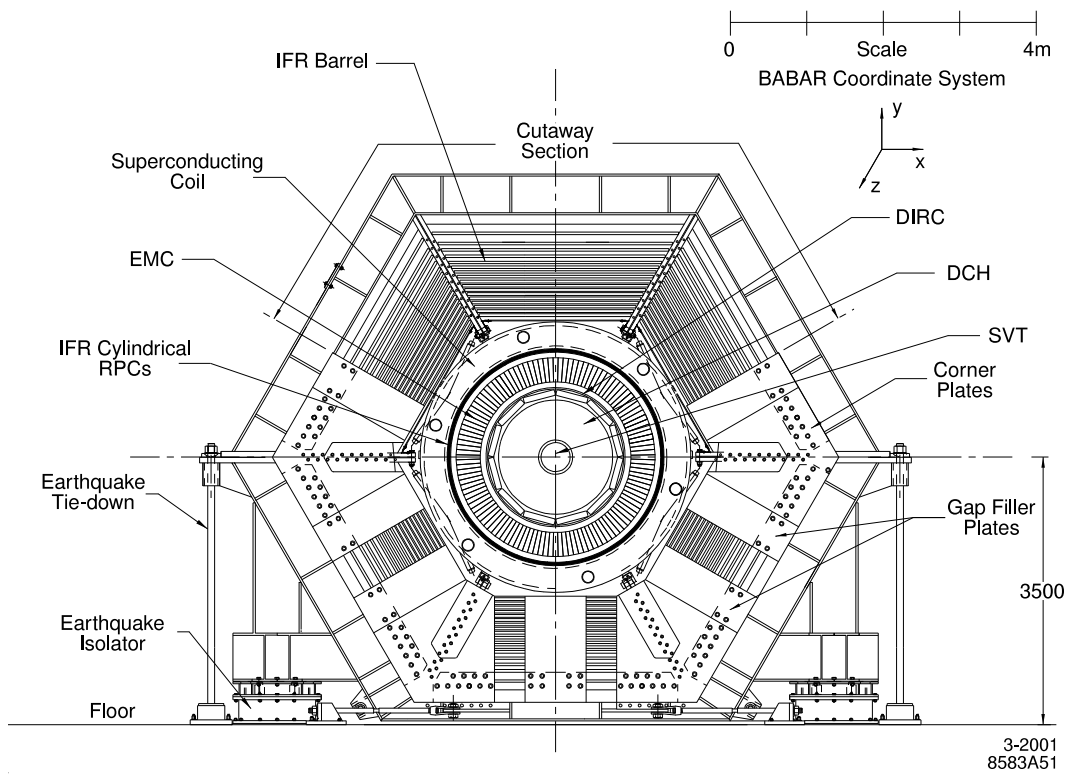


Figure 3.3: Layout of the BABAR detector, end view.

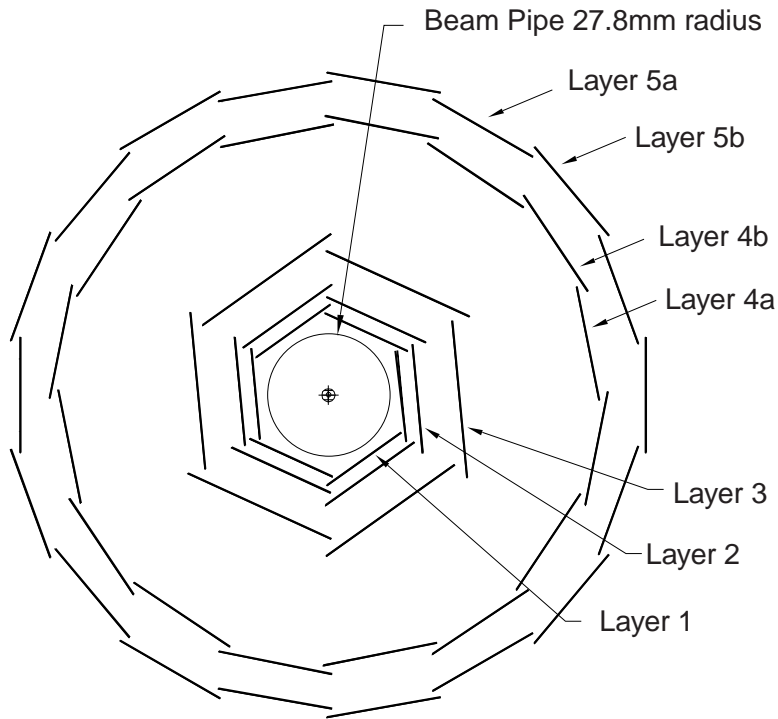


Figure 3.4: End view of the SVT. The beam pipe radius is 27.8 mm. The minimum radius of Layer 1 of the SVT is 32 mm, while the radius of Layer 5 is 114–144 mm.

detectors themselves have all been designed to be as lightweight as possible in order to minimize effects due to multiple scattering.

The SVT is made up of 104 individual 300 μm -thick silicon wafer sensors covering a total area of 0.96 m^2 . Each wafer is a high-resistivity n-type silicon substrate with p^+ and n^+ strips on opposite sides, forming a PIN junction. A voltage of ≈ 40 V is applied across the sensor, fully depleting the silicon. When a charged particle passes through the sensor, it creates electron-hole pairs

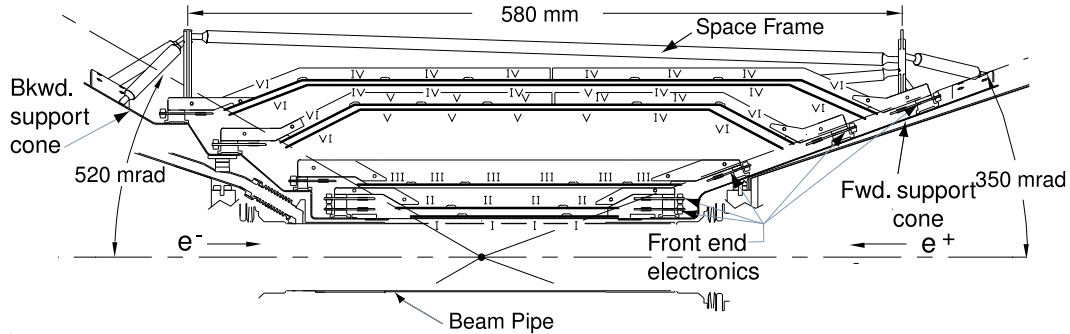


Figure 3.5: Side view of the SVT, showing the five detector layers and the beam pipe. The lower half of the detector is not shown.

which are then collected on the strips, a signal which is read out electronically to provide a measurement of the particle's position. A minimum-ionizing particle typically produces 24,000 electron-hole pairs, and the SVT readout contains about 150,000 channels.

The strips on opposite sides of the sensors run in perpendicular directions: strips parallel to the beam axis measure the azimuthal angle ϕ and strips perpendicular to the beam axis measure the position in z (the direction along the beam axis). The SVT can therefore measure the trajectory of charged particles traversing the detector by recording ϕ and z at each point where the trajectory crosses a silicon wafer. The position resolution on each of these hits depends on the pitch, the separation between readout strips, which ranges from $50 \mu\text{m}$ in the inner layers to $210 \mu\text{m}$ in the outer layers. The measured hit resolution for particles at normal incidence is $10\text{--}15 \mu\text{m}$ in the inner layers

and about $40 \mu\text{m}$ in the outer layers. These hit resolutions result in a typical resolution of the B meson vertex position of $60\text{--}100 \mu\text{m}$, depending on its decay mode.

In addition to providing precise vertex measurements, the SVT is also used for stand-alone track reconstruction of low momentum particles. Charged particles with momentum transverse to the beam axis of $p_T > 100 \text{ MeV}/c$ can be reconstructed with high efficiency in the DCH, but, for lower momenta, the track curls in the magnetic field and does not reach the DCH. In this case, the measurements in the SVT alone are used to reconstruct the track.

The amount of charge deposited in each silicon sensor is related to the ionization energy loss dE/dx , which can be used to measure the mass of the charged particle and to discriminate between various particle hypotheses (e , μ , π , K , or p). While the DCH provides the primary measurement of dE/dx , information from the SVT is used as well to aid in particle identification.

3.2.2 Drift Chamber

The DCH is a gas-filled volume whose primary purpose is to detect the ionization produced by charged particles and to measure the momenta of the particles. The DCH also measures the energy loss dE/dx which is used to identify charged particles, and, for very long-lived particles which decay outside

the SVT (a K_s^0 , for example), the DCH provides a measurement of the decay vertex position.

The chamber is a 280 cm-long cylinder with an inner radius of 23.6 cm and an outer radius of 80.9 cm. The chamber is filled with an 80 : 20 mixture of helium and isobutane (C_4H_{10}). There are nearly 29,000 wires in the chamber, arranged in 7,104 hexagonal cells in 40 cylindrical layers, each parallel to (or nearly so) the chamber axis. Each cell contains a gold-plated tungsten-rhenium “sense” wire surrounded by six gold-plated aluminum “field” wires. The sense wires are held at a potential of $\approx +1900$ V with respect to the field wires. The gas mixture and low-mass field wires were chosen to minimize multiple scattering.

When a charged particle passes through a DCH cell, it ionizes the gas molecules. The resulting electrons are accelerated towards the sense wire, and, as they travel towards the wire, they ionize additional molecules, causing an avalanche of secondary ionization that amplifies the signal. These electrons are collected on the sense wire and the resulting signal is then read out.

Position information in the z direction is obtained by placing the wires in 24 of the 40 layers at small angles with respect to the z -axis. These are referred to as “stereo” layers.

The transverse momentum resolution of the DCH is

$$\frac{\sigma_{p_T}}{p_T} = 0.13\% \cdot p_T + 0.45\% , \quad (3.3)$$

where p_T is expressed in GeV/c.

The total ionization in each drift cell is proportional to the energy lost by the charged particle. The Bethe-Bloch equation relates this dE/dx to the relativistic velocity β of the particle:

$$\frac{dE}{dx} = -4\pi r_e^2 m_e c^2 N_A \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right) \right] , \quad (3.4)$$

where r_e and m_e are the classical radius and mass of the electron, N_A is Avogadro's number, Z and A are the atomic number and atomic mass of the ionization medium, T_{\max} is the maximum kinetic energy that can be imparted to a free electron in a single collision, I is the mean excitation energy, and $\delta(\beta\gamma)$ is a density-effect correction. Since the Bethe-Bloch equation depends on velocity and not momentum, dE/dx measurements can be combined with momentum measurements to discriminate between particles with different masses. Figure 3.6 shows dE/dx vs. momentum for charged particles passing through the DCH with the Bethe-Bloch predictions overlaid.

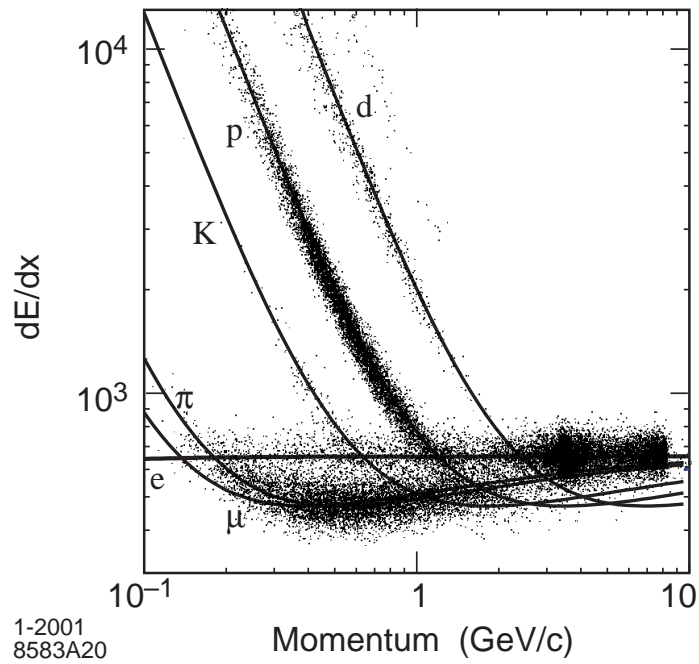


Figure 3.6: DCH dE/dx for particles of different momenta. The Bethe-Bloch predicted curves are overlaid. Note that the measurement of dE/dx provides excellent discrimination between pions and kaons up to about 700 MeV/c.

3.2.3 Detector of Internally-Reflected Cherenkov Light

The DIRC is a charged-particle identification system based on Cherenkov light. When a charged particle passes through a medium faster than the speed of light in that medium, it emits radiation in a conical pattern about the direction of travel, where the opening angle of the cone θ_c is given by

$$\cos \theta_c = \frac{1}{n\beta} \quad (3.5)$$

where n is the index of refraction of the medium and β is the velocity of the particle. Because this angle depends on velocity, we can use this angle to discriminate between particles with different masses by combining this with the momentum measured in the SVT and DCH.

In the DIRC, the Cherenkov radiator consists of 144 synthetic quartz bars with index of refraction $n = 1.473$. The bars, with rectangular cross sections and highly-polished surfaces, serve as light guides for the radiation produced. Some of the light is totally internally reflected, preserving θ_c as the light propagates along the bar since the angle of incidence is preserved when reflecting off flat parallel surfaces.

The layout of the DIRC is shown in Figure 3.7. The light is guided down the quartz bars to the backward end of the detector where it is emitted into the standoff box. Typically 20–65 photons reach the standoff box for each

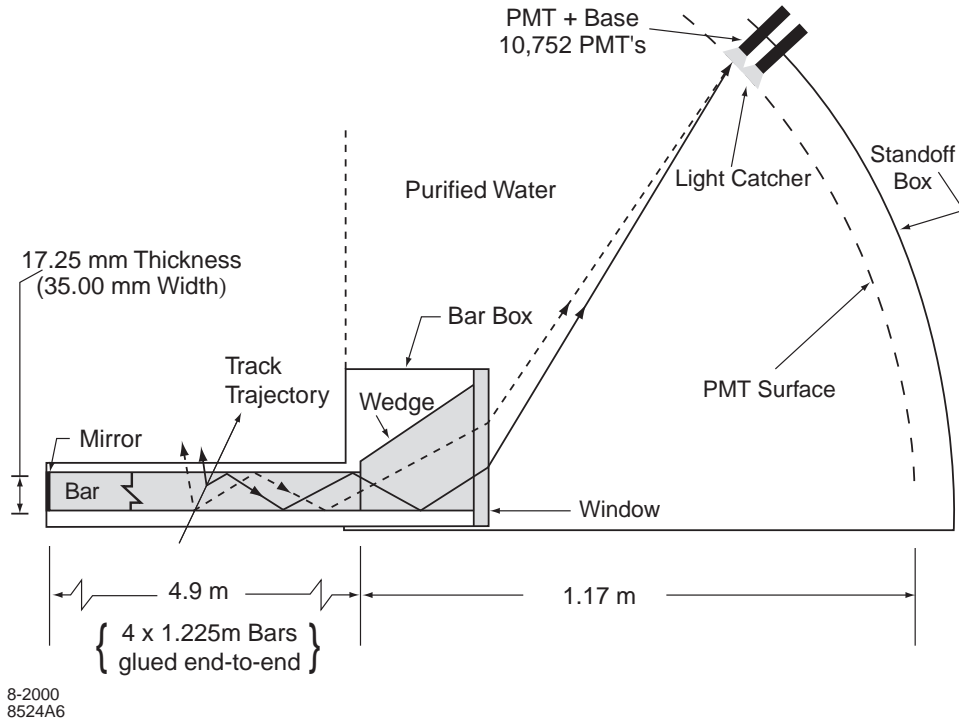


Figure 3.7: Layout of the DIRC detector, showing a charged particle producing Cherenkov light that is guided down quartz bars to the water-filled standoff box where it is recorded by photomultiplier tubes.

charged track. The standoff box is filled with 6000 liters of purified, deionized water ($n = 1.346$), and, as the light from the bars propagates through the water, the photons spread out to form circular arcs. The photons are detected with an array of 10,752 photomultiplier tubes which are mounted on the far end of the standoff box.

By measuring these photons, θ_c can be reconstructed and the particle type determined. The DIRC performance is complementary to particle identifi-

cation using dE/dx in the SVT and DCH — dE/dx can efficiently separate kaons from pions below about 700 MeV/ c , while the DIRC was optimized to perform at higher momenta, providing 4σ kaon–pion separation at momenta up to 3 GeV/ c .

3.2.4 Electromagnetic Calorimeter

The EMC is designed to efficiently detect and measure the energy of electromagnetic showers produced by photons and electrons.

The EMC is made of an array of 6,580 thallium-doped cesium iodide crystals, arranged in 56 azimuthal rings. The detector is symmetric about the beam axis, and the layout is shown in Figure 3.8.

When high-energy photons or electrons arrive at the EMC crystals, an electromagnetic shower develops. High-energy photons are converted into electron-positron pairs ($\gamma \rightarrow e^+e^-$), while high-energy electrons and positrons emit bremsstrahlung photons ($e^\pm \rightarrow e^\pm\gamma$). Both of these processes require a strong external electric field in order to occur — the high- Z nuclei of cesium and iodine provide such a field. The two processes form a feedback loop, so that one incident high-energy particle is quickly converted into a large number of low energy photons and electrons. The energy carried by these shower particles is absorbed by the crystals, which then fluoresce with a characteristic

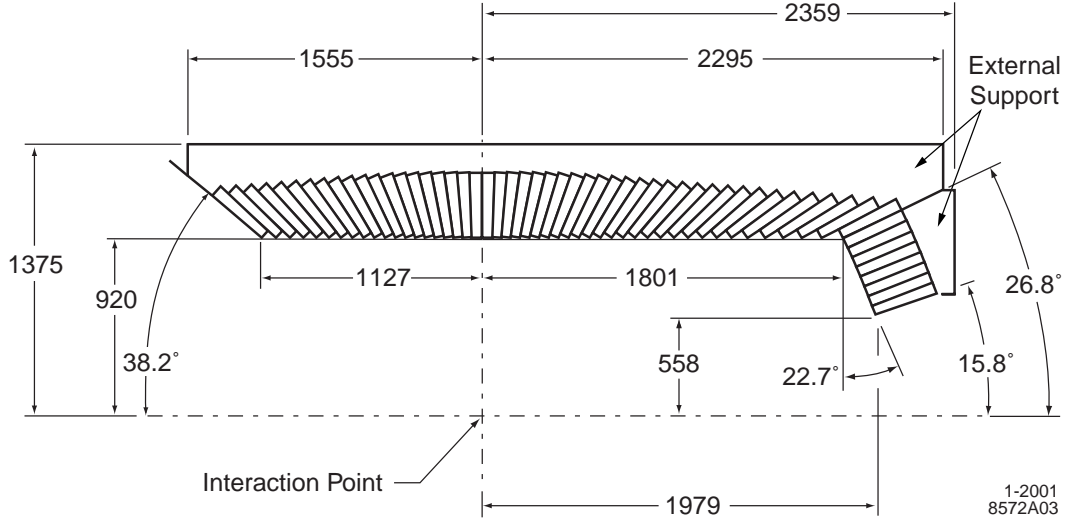


Figure 3.8: Side view of the EMC showing the arrangement of the 56 crystal rings. The lower half of the detector is not shown. All dimensions are in mm.

spectrum. This fluorescence is measured by photodiodes attached to the rear faces of the crystals.

The EMC is over 96% efficient for detecting photons with an energy greater than 20 MeV. The energy resolution of the EMC is found to be

$$\frac{\sigma_E}{E} = \frac{2.3\%}{\sqrt[4]{E}} \oplus 1.9\% , \quad (3.6)$$

where the energy E is expressed in GeV. The first term is primarily due to photon-counting statistics and is dominant at low energies, while the second term arises from uncertainties in scintillation light yield and calibration, and is dominant at higher energies.

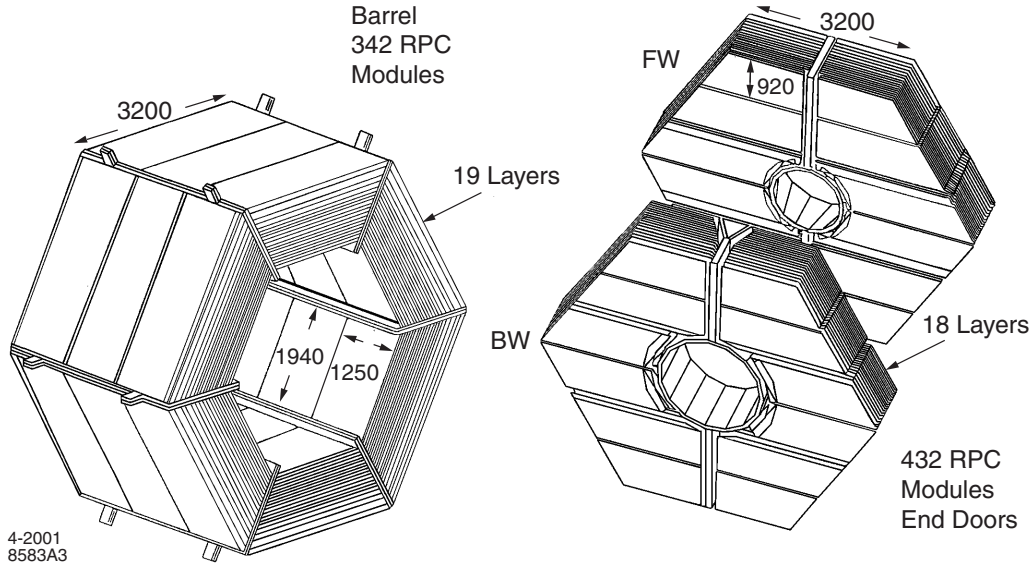


Figure 3.9: Layout of the IFR.

3.2.5 Instrumented Flux Return

The flux of the 1.5 T magnet is returned through the hexagonal IFR surrounding the solenoid coils. Sandwiched between sections of iron are active detectors to identify long-lived, deeply penetrating muons and to detect neutral hadrons such as the K_L^0 . The layout of the IFR is shown in Figure 3.9.

The active detectors are resistive plate chambers consisting of two highly resistive cathodes surrounding a 2 mm-thick chamber filled with a gas mixture of 57% argon, 39% freon 134a, and 4% isobutane. A large voltage (≈ 8 kV) is applied across this gap. When a charged particle passes through the RPC, it ionizes the gas and produces an avalanche similar to what occurs in the DCH.

Chapter 4

Data and MC Samples

4.1 *BABAR* Data Samples

This analysis has been performed on the full Run 1–4 data sample collected by the *BABAR* detector. These data consist of 208.9 fb^{-1} collected at the $\Upsilon(4S)$ resonance, and correspond to 232 million $B\bar{B}$ pairs.

4.2 Simulated Event Samples

We use three types of Monte Carlo simulated events (MC) in this analysis:

- **Generic $B\bar{B}$ events** — in these events, both B mesons decay according to the default values set in the simulation.

- **Signal** $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ events — one B meson decays to a set of final states for which the B_{tag} reconstruction is very efficient.¹ The second B meson decays to a signal mode. In all signal samples, the τ^- decays to $\ell^-\bar{\nu}_\ell\nu_\tau$, and the $D^{(*)}$ decays to a set of reconstructible final states. These final states correspond to those listed in Section 6.3.
- **Semileptonic** One B decays to the same B_{tag} modes as used for the Signal Monte Carlo samples, while the second decays to $D^{(*)}(\ell^-\bar{\nu}_\ell)$. For these events, the label D^{**} refers to a mixture of both the narrow and broad D^{**} resonances and a nonresonant $D^{(*)}\pi$ state, using the same D^{**} mixture found in Generic MC. In all semileptonic samples, the $D^{(*)}$ mesons decay to a set of reconstructible final states, just as in Signal MC.²

The event yield in each of the MC categories is given in Table 4.1. The MC samples are generated to simulate a variety of different conditions, in a luminosity-weighted mixture, so that the samples closely reproduce the time dependence of detector performance and background conditions.

¹The final states used correspond to the low-multiplicity modes $B \rightarrow D^{(*)}\pi^+$, $B \rightarrow D^{(*)}\pi^+\pi^0$, and $B \rightarrow D^{(*)}\pi^+\pi^0\pi^0$, where the light hadronic system also includes resonant contributions from the ρ and a_1 . The D mesons also decay to low-multiplicity (2-, 3-, and 4-body) final states.

²We do not have a dedicated sample of $D^{**}\tau\nu$ MC, although such decays are present in the Generic sample. These events only make up 2–3% of (reconstructed) D^{**} decays, and the effects of this background will be treated as a systematic error later.

MC events are generated and particle decay modeled using the EvtGen package [27], and detector simulation is performed *BABAR*'s GEANT4-based simulation [28].

Note that our signal modes $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_{\tau}$ are included in the default decay table with production rates $\approx 1\%$, so that the generic sample contains both signal and background events in roughly correct (if we assume the SM signal branching fractions) proportions.

We used approximately 640 fb^{-1} of generic MC for designing and optimizing the event selection criteria and for testing the signal-extraction procedure. The remaining 330 fb^{-1} of generic MC was kept as a statistically independent sample for validating the analysis, and was not used during the optimization and fit development. The signal and semileptonic samples were used for developing fit shapes.

4.3 Semileptonic and Semitauonic Decay Models

Three different decay models are used within EvtGen to model B decays to semileptonic and semitauonic final states. The same models are used for these decays within the generic, signal, and semileptonic MC samples. All

Table 4.1: Numbers of events and effective luminosities for various MC event types. All luminosities for signal and semileptonic MC are *very* approximate, due to the uncertainties on the B_{tag} efficiency in MC.

MC event type		N_{evt}	$\mathcal{L}_{\text{eff}}(\text{fb}^{-1})$
Generic	B^+B^-	514 M	940
	$B^0\bar{B}^0$	532 M	970
Signal	$D^0\tau^-\bar{\nu}_\tau$	1868 k	112000
	$D^{*0}\tau^-\bar{\nu}_\tau$	1872 k	66000
	$D^+\tau^-\bar{\nu}_\tau$	934 k	152000
	$D^{*+}(\rightarrow D^0\pi^+)\tau^-\bar{\nu}_\tau$	1872 k	210000
	$D^{*+}(\rightarrow D^+\pi^0)\tau^-\bar{\nu}_\tau$	936 k	300000
Semilep	$D^0\ell^-\bar{\nu}_\ell$	1872 k	15200
	$D^{*0}\ell^-\bar{\nu}_\ell$	1708 k	5300
	$D^{**0}\ell^-\bar{\nu}_\ell$	2978 k	25000
	$D^+\ell^-\bar{\nu}_\ell$	1872 k	42400
	$D^{*+}\ell^-\bar{\nu}_\ell$	1870 k	30600
	$D^{**+}\ell^-\bar{\nu}_\ell$	3744 k	16000

models simulate final state radiation, up to order α_{EM}^2 , using the PHOTOS package [29].

4.3.1 HQET model

$B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decays are modeled using the HQET form factors described in Section 2, where the form factors are parameterized as linear functions of $(w - 1)$. Events generated with this model use the central values of the form factors measured by the CLEO Collaboration [2]:

$$\begin{aligned}
 \rho^2 &= 0.92 \pm 0.16 \\
 R_1 &= 1.18 \pm 0.32 \\
 R_2 &= 0.72 \pm 0.23 .
 \end{aligned}$$

BABAR has recently performed an updated measurement of the $B \rightarrow D^*$ form factors [3] and obtained the values:

$$\begin{aligned}\rho^2 &= 0.79 \pm 0.06 \\ R_1 &= 1.40 \pm 0.06 \\ R_2 &= 0.87 \pm 0.04 ,\end{aligned}$$

which are calculated using the same linear parameterization. The *BABAR* data, however, better support the alternative CLN parameterization described in Section 2, and, in this scheme, the form factor values extracted are:

$$\begin{aligned}\rho^2 &= 1.18 \pm 0.06 \\ R_1 &= 1.42 \pm 0.08 \\ R_2 &= 0.84 \pm 0.04 .\end{aligned}$$

4.3.2 ISGW2 Model

The decays $B \rightarrow D\ell^-\bar{\nu}_\ell$, $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$, and resonant $B \rightarrow D^{**}(\ell^-, \tau^-\bar{\nu})$ are generated using the ISGW2 model of Scora and Isgur [30].

4.3.3 Goity–Roberts Model

The nonresonant decays $B \rightarrow D^{(*)}\pi\ell^-\bar{\nu}_\ell$ are generated using a modified form of the model proposed by Goity and Roberts [31]. The modifications, described in the EvtGen documentation [27], correct errors in the original paper and remove the contribution from $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ from the amplitude so that only the nonresonant piece is modeled here.

4.3.4 $D^{**}\ell^{-}\bar{\nu}_{\ell}$ Decay Branching Fractions

The $B \rightarrow D^{**}\ell^{-}\bar{\nu}_{\ell}$ model in both Generic and Semileptonic MC is a mixture of both resonant and nonresonant decay channels, as described above. The generated branching fractions of the various channels are listed in Table 11.1 in Section 11.1, where we will also discuss reweighting these branching fractions to include more recent measurements and calculations.

Chapter 5

Background Classification

Table 5.1 presents a first look at the signal and background yields in generic MC. We have selected events with B_{tag} , $D^{(*)}$, and ℓ candidates using loose selection criteria (this is the Vcb reskim; see Sec. 6.3), and we tabulate the number of selected events in each $D^{(*)}$ channel. We see that the signal-to-background ratio is approximately 5×10^{-3} in all channels,¹ and that a lot of work must be done before the signal is evident. In this section, we describe the various background sources and our strategies for dealing with them; these strategies will be implemented in Section 6 as our event selection criteria. For comparison, the event yields after the final event selection can be seen in Table 6.2.

¹While this signal-to-background ratio is actually lower than the SM branching fractions of $7\text{--}12 \times 10^{-3}$, the fraction of reconstructible D and τ decays is only about 10%, so a S/B ratio of 5×10^{-3} at this stage is, in fact, slightly better than a random sample of B decays.

Table 5.1: Event yields from 640 fb^{-1} of generic MC before selection. Each row indicates the number of events of a specific generated type which were reconstructed in each of the four signal channels. For the purposes of this table, “combinatoric” background contains all decays other than the 14 specifically listed. “Signal” here includes both correctly identified signal candidates and signal crossfeed background resulting from $D \Rightarrow D^*$ or $D^* \Rightarrow D$, but does not include crossfeed between charged and neutral B decays. Compare to Table 6.2 (Section 6.16), which shows the event yields after the final event selection.

Generated Event Type	Reconstruction Channel			
	D^\pm	$D^{*\pm}$	D^0	D^{*0}
$B^- \rightarrow D^0 e \bar{\nu}_e$	7,184	3,999	18,515	10,254
$B^- \rightarrow D^{*0} e \bar{\nu}_e$	18,467	12,183	52,833	39,216
$B^- \rightarrow D^0 \mu \bar{\nu}_\mu$	6,435	3,767	15,918	9,677
$B^- \rightarrow D^{*0} \mu \bar{\nu}_\mu$	16,288	10,972	44,784	33,789
$B^- \rightarrow D^0 \tau \bar{\nu}_\tau$	1,020	631	2,020	1,279
$B^- \rightarrow D^{*0} \tau \bar{\nu}_\tau$	2,382	1,724	5,267	4,076
$\bar{B}^0 \rightarrow D^+ e \bar{\nu}_e$	13,809	5,787	7,374	5,090
$\bar{B}^0 \rightarrow D^{*+} e \bar{\nu}_e$	46,944	31,355	24,851	17,646
$\bar{B}^0 \rightarrow D^+ \mu \bar{\nu}_\mu$	11,902	5,371	7,617	5,488
$\bar{B}^0 \rightarrow D^{*+} \mu \bar{\nu}_\mu$	41,485	27,695	25,456	18,710
$\bar{B}^0 \rightarrow D^+ \tau \bar{\nu}_\tau$	1,634	785	971	679
$\bar{B}^0 \rightarrow D^{*+} \tau \bar{\nu}_\tau$	5,650	3,726	2,800	2,107
Nonresonant $B \rightarrow D^{(*)} \pi \ell^- \bar{\nu}_\ell$	46,989	27,536	53,057	38,467
Resonant $B \rightarrow D^{**} \ell^- \bar{\nu}_\ell$	58,548	37,676	68,060	50,207
Combinatoric	934,912	712,247	1,058,592	937,939
Total signal	7,284	4,511	7,287	5,355
Total background	1,206,365	880,943	1,380,828	1,169,269
$S/B (10^{-3})$	6.0	5.1	5.3	4.6

Two broad categories of background events are considered in this analysis: $B \rightarrow X_c \ell^- \bar{\nu}_\ell$ background, in which a $B \rightarrow X_c \ell^- \bar{\nu}_\ell$ transition other than $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ recoils against a well-reconstructed B_{tag} , and combinatoric background, which describes all other background sources (this delineation is already implicit in the construction of Table 5.1). Different techniques are required to suppress these backgrounds and to estimate their contamination of the signal. We list the various background categories here and briefly describe our approach to them:

- $B \rightarrow X_c \ell^- \bar{\nu}_\ell$ background, which is composed of
 - **Semileptonic background**, in which a true $D^{(*)} \ell^- \bar{\nu}_\ell$ event is reconstructed, particularly with a non-zero missing mass, and
 - **Feed-down**, in which a soft π or γ is lost from a true D^* or D^{**} , so that the correct final state meson is not reconstructed (also includes $B \rightarrow D^{**} \tau^- \bar{\nu}_\tau$ decays, in which neither the charm meson nor the lepton is fully reconstructed); and
- Combinatoric background, which can be further divided into:
 - **Combinatoric B_{tag} background**, where the B_{tag} candidate is incorrectly reconstructed,

- **Combinatoric $D^{(*)}$ background**, where the signal-side $D^{(*)}$ candidate is incorrectly reconstructed,
- **Fake ℓ background**, where a hadron is misidentified as a lepton, and
- **Combinatoric $\Upsilon(4S)$ background**, where the B_{tag} , $D^{(*)}$, and ℓ candidates are all correctly reconstructed, but the signal hypothesis $\Upsilon(4S) \rightarrow B\bar{B} \rightarrow B_{\text{tag}}D^{(*)}\ell + 3\nu$ is incorrect. This category includes all sources of background not already mentioned, but it is dominated by decays like $B \rightarrow D\bar{D}$ in which one D meson decays leptonically or semileptonically.

5.1 $B \rightarrow X_c\ell^-\bar{\nu}_\ell$ Background

The semileptonic background is our most dangerous background; it has the same final state topology as our signal and a branching fraction roughly twenty times larger. These events can mostly be identified by m_{miss}^2 , but the finite resolution on this measurement means some events will leak into the signal region. We can study this leakage in the simulation, and we will use control samples in data to validate the performance of the MC, giving us some

confidence that we understand the residual $B \rightarrow X_c \ell^- \bar{\nu}_\ell$ background at large m_{miss}^2 .

Feed-down background, such as $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0$, is problematic because, in addition to the fact that we have the wrong charm meson, the lost particle (or particles) tends to increase m_{miss}^2 , causing the candidate to look more signal-like. These events are not particularly problematic if the lost particle is a π^\pm , as these events tend to fail the conservation-of-charge requirement — at least one other particle must be misreconstructed simultaneously in order to do so. If the lost particle is neutral, either a π^0 or γ , we have no such charge conservation constraint, and these backgrounds can be large.

Events which have lost photons outside the EMC acceptance are, at some level, irreducible, and this contribution to the feed-down background must be modeled by the simulation. A large fraction of the feed-down background, however, is due to events in which one or more of the final-state photons interacted in the calorimeter, but we did not reconstruct the correct candidate. These events have neutral clusters remaining after the reconstruction of the signal candidate, and, as such, can be suppressed by cutting on E_{extra} (see Sec. 6.12).

5.2 Combinatoric Background

At the loose event-selection stage described above, events with one or more particles misreconstructed — either the B_{tag} , $D^{(*)}$, or ℓ candidates — dominate our sample. These events will turn out to be relatively easy to suppress, however. Our cuts on particle identification (PID), ΔE and m_{ES} on the tag side, and $m(D)$ (and, for D^* candidates, δm) on the signal side greatly reduce these backgrounds.

Most of the remaining background is eliminated by our “total-containment” requirements — that every particle in the event must be accounted for in our reconstructed event candidate (see Sec. 6.11 and 6.12).

Occasionally, we find events in which both the B_{tag} and the $D^{(*)}$ are misreconstructed, usually due to the swapping of a single particle between the two B mesons. If the momentum of the swapped particle is large enough, the subsequent candidate will fail cuts on ΔE , m_{ES} , or $D^{(*)}$ mass and δm . If, however, the particle has low momentum (the slow pion or γ from D^* or D^{**} decay, for instance), the candidate may pass our event selection criteria (although still suppressed relative to well-reconstructed events). In this case, nothing can be done to further reject these events, so our signal extraction must take them into account.

One of two things can happen, depending on the charge of the particle swapped. If a neutral particle is swapped between the B_{tag} and the $D^{(*)}$, the m_{miss}^2 resolution is worse than it is for correctly reconstructed candidates; the m_{miss}^2 shapes used in the fit take this resolution into account.

If the particle is charged, however, both the B_{tag} candidate and the $D^{(*)}$ candidate are reconstructed with the wrong charge. One component of the fit is dedicated to modeling this charge-crossfeed contribution (Sec. 8.2).

Two-body B Decays

B decays to one charmed meson and one light hadron, such as $B \rightarrow D\pi$ and $B \rightarrow D\rho$ are a significant source of combinatoric $\Upsilon(4S)$ background, usually because a π^\pm track is misidentified as a muon. These decays are easily distinguished from semileptonic and semitauonic decays by their missing momentum; these events have nearly zero missing momentum, while a typical neutrino from a semileptonic event carries off several hundred MeV of momentum.

Most of the remaining combinatoric $\Upsilon(4S)$ background comes from B decays to two charmed mesons, such as $B \rightarrow D^{(*)}D^{(*)}$ and $B \rightarrow D^{(*)}D_s^{(*)+}$. These events produce a fake signal when one of the charmed mesons decays semileptonically, as $D \rightarrow K\ell\nu$, or leptonically, as in $D_s^- \rightarrow \tau^-\bar{\nu}_\tau$ ($\tau^- \rightarrow \ell^-\bar{\nu}_\ell\nu_\tau$).

Events with either $D \rightarrow K^\pm \ell \nu$ or $D \rightarrow K_s^0 \ell^- \bar{\nu}_\ell$ ($K_s^0 \rightarrow \pi^+ \pi^-$) are usually easy to reject because of the additional charged tracks, but events with $D \rightarrow K_L^0 \ell^- \bar{\nu}_\ell$ are much more difficult, as are events with $D_s^- \rightarrow \tau^- \bar{\nu}_\tau$.

Events with either a K_L^0 or multiple neutrinos lost are especially problematic because they will increase m_{miss}^2 and tend to look like our signal.

We suppress these events by cutting on $q^2 \equiv (p_B - p_D^{(*)})^2$, which is equivalent to the momentum of the $D^{(*)}$ in the signal B rest frame. Two-body B decays have very high $D^{(*)}$ momentum and very low q^2 , which distinguish them from our signal.

5.3 Feed-down and Feed-up

Feed-down background occurs when the soft transition particle (π or γ) from a true D^* or D^{**} decay is lost, so that we do not fully reconstruct the X_c system. Because this background is unavoidable, we treat these events specially in our signal extraction (Sec. 8). This is true for signal feed-down ($B \rightarrow D^* \tau^- \bar{\nu}_\tau$ events misreconstructed as $B \rightarrow D \tau^- \bar{\nu}_\tau$) as well as semileptonic feed-down, and, in fact, signal feed-down will be a significant contribution to our total signal yield.

Feed-up background, when a true D meson is combined with a γ or π^0 to create a fake D^* , exists as a subset of the combinatoric D^* background, both for signal and semileptonic events. While this background is much smaller than feed-down (it's easier to lose a particle than to find a fake particle with a specific momentum), it too must be accounted for in the signal extraction.

Chapter 6

Event Selection and Optimization

6.1 Overview

This section presents the event selection for the analysis. Sections 6.2 and 6.3 describe the skimming stages of the analysis. After the skims, we apply cuts on the B_{tag} candidate (Sec. 6.4) and particle identification on the recoil (Sec. 6.5 and 6.6). We reconstruct composite particles on the recoil (Sec. 6.7) and perform a kinematic/vertex fit to the full event candidate (Sec. 6.8). Events are required to have large missing momentum (Sec. 6.9) and large q^2 (Sec. 6.10). As described in Section 5, these cuts are designed to reject combinatoric backgrounds and misidentified particles.

Two cuts are used to reject events with leftover particles visible in the *BABAR* detector, one on the number of charged tracks remaining (Sec. 6.11) and one on the total energy of all neutral clusters remaining (Sec. 6.12).

In events with more than one signal candidate reconstructed, we choose a single candidate as described in Section 6.13.

Many of the selection criteria presented here are expected to have the same efficiency for signal events and for our normalization sample, $B \rightarrow D^{(*)} \ell^- \bar{\nu}_\ell$ events, and so uncertainties in these efficiencies cancel in the ratio. Nevertheless, where possible, distributions of these event-selection variables will be checked in control samples in data; see Section 7.

Most of the selection criteria described here, including the requirements on m_{ES} , particle identification, and invariant masses, have been optimized to maximize $S/\sqrt{S+B}$, given the SM predictions for the signal branching fractions. In optimizing each cut, all of the other cuts have been applied. The ΔE requirement was initially optimized in the same way, but was later made tighter when it became clear that events in the ΔE tails contributed to biases in the signal extraction; the final selection is a compromise between the statistical $S/\sqrt{S+B}$ optimization and the systematic effects due to this bias.

6.2 B SemiExcl Skim

The details of the semiexclusive B reconstruction algorithm are presented in Ref. [32]. For our purposes, it is sufficient to mention that the algorithm extends the standard *BABAR* Composition Tools to search for $B \rightarrow D^{(*)}Y$, where the Y system is a $Q = \pm 1$ combination of π , π^0 , K , and K_s^0 . B candidates are constructed by recursively adding particles to and removing them from the Y system, and are identified by cuts on m_{ES} and ΔE .

The algorithm is semiexclusive in the sense that these hadrons are not constrained to any intermediate resonance states. Because of this, the yield is significantly higher than exclusive B reconstruction. At the same time, however, the impurity — the fraction of events passing the BReco cuts that are not true B mesons — is also much larger.

The purity of a sample of B_{tag} candidates can be measured by fitting the m_{ES} distribution to the sum of a Crystal Ball function [33] representing correctly reconstructed B mesons and an ARGUS function [34] modeling the background. The purity is then extracted as the ratio of the number of events in the signal peak to the total number of events, and is calculated in the signal-enriched range $5.27 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$.

The purity varies widely among different reconstruction modes (there are 1114 modes, corresponding to different $D^{(*)}$ modes and different configurations of the Y^\pm system), and so the measured purity of the various modes can be used as a selection criteria for B_{tag} candidates. Additionally, modes with similar purity are grouped together to define three broad categories: *dirty*, *neat*, and *clean*, with purities in the intervals [0.08–0.5],[0.5–0.8], and [0.8–1.0], respectively.¹ Figure 6.1 shows the m_{ES} and ΔE distributions of B_{tag} candidates within the three categories. This analysis does not treat the three categories separately.

The BSemiExcl skim performs the semiexclusive B reconstruction and stores all B_{tag} candidates passing a minimal purity cut (each mode is required to have a purity greater than 8–10%, see [35]) as persisted candidates within the CM2 framework. This skim is part of the centralized *BABAR* skim production.

¹Technically, these categories are defined by cuts on both the purity and the *integrated purity*, where the latter is the purity summed over all modes in the category, weighted according to their abundances in data. Clean and Neat are defined with the cuts given above applied to the integrated purity; Dirty is required to have an integrated purity less than 50%, but a single-mode purity greater than 8%.²

²Well, to be completely accurate, the minimal single-mode purity allowed for Dirty actually depends on the “seed” mode, the type of D meson reconstructed within the B_{tag} . For D^{*0} seeds, the cut is at 8%, for D^0 and D^- it is at 9%, and for D^{*-} , 10%. More details on this can be found in Ref. [35].

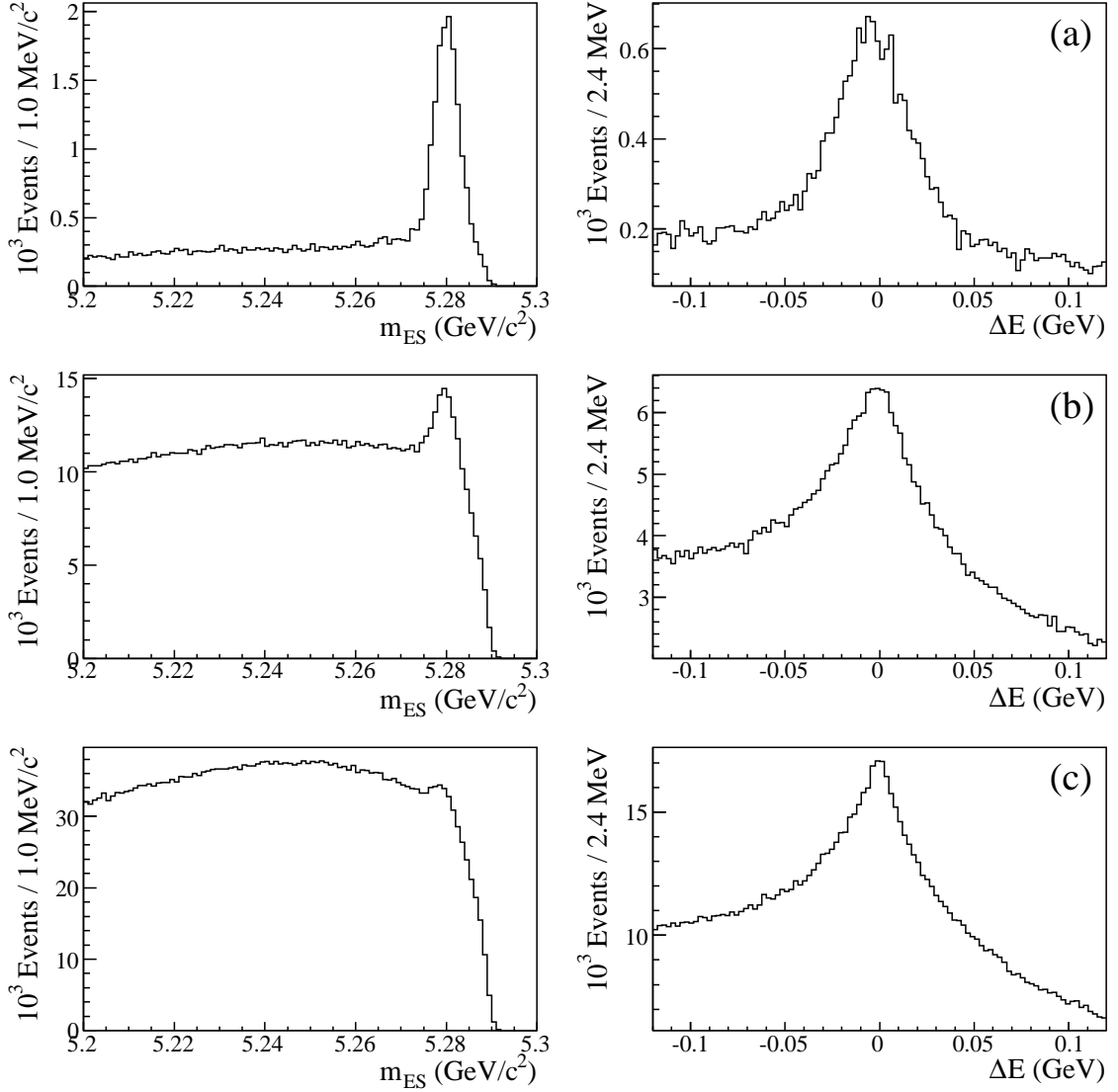


Figure 6.1: m_{ES} and ΔE distributions for the BsemiExcl skim, shown in the (a) Clean, (b) Neat, and (c) Dirty categories as described in the text. The complete Run 1 data sample is shown; m_{ES} plots are shown after $|\Delta E| < 50 \text{ MeV}$, and ΔE plots after $m_{ES} > 5.27 \text{ GeV}/c^2$.

6.3 Vcb Reskim

Events passing the BSEMIExcl skim are then analyzed with a privately-run reskim, the Vcb reskim. This reskim applies two very loose event-selection cuts:

- $N_{\text{ChargedTracks}} < 19$
- $N_{\text{CalorNeutral}} < 26$.

These cuts are designed to minimize CPU usage downstream, and should be completely efficient for well-reconstructed signal events — the highest multiplicity events we select contain 14 charged tracks and 12 neutrals.

We reconstruct D and D^* mesons for the signal-side B meson in the following modes, given with their PDG branching fractions:

Decay mode		$\mathcal{B}(\%)$
D^0	$\rightarrow K\pi$	3.80
	$\rightarrow K\pi\pi^0$	14.1
	$\rightarrow K\pi\pi\pi$	7.72
	$\rightarrow K_s^0\pi\pi$	2.90
	$\rightarrow K_s^0\pi\pi\pi^0$	5.3
Total		31.3
D^+	$\rightarrow K\pi\pi$	9.51
	$\rightarrow K\pi\pi\pi^0$	5.5
	$\rightarrow K_s^0\pi$	1.47
	$\rightarrow K_s^0\pi\pi^0$	7.0
	$\rightarrow K_s^0\pi\pi\pi$	3.11
	$\rightarrow K^+K^-\pi^+$	1.00
Total		24.0
D^{*0}	$\rightarrow D^0\pi^0$	61.9
	$\rightarrow D^0\gamma$	38.1
Total		100.0
D^{*+}	$\rightarrow D^0\pi$	67.7
	$\rightarrow D^+\pi^0$	30.7
Total		98.4

The individual branching fractions listed above do not take into account the $K_s^0 \rightarrow \pi^+\pi^-$ branching fraction, so our total sensitivity to the K_s^0 modes is smaller by a factor of 0.692. The total branching fractions reflect this factor.

During the reskim, we apply only very loose reconstruction cuts:

- π^\pm , K^\pm , K_s^0 , and π^0 candidates are taken from the *ChargedTracks*, *KL-HVeryLoose*, *KsDefault*, and *pi0AllDefault* lists, respectively. These requirements are made tighter after the reskim. The soft π^0 from D^* decay is taken from the *pi0VeryLoose* list, and the soft γ from $D^{*0} \rightarrow D^0\gamma$ is taken from the *CalorNeutral* list and must have a minimum energy of

100 MeV. The detailed requirements for all of the standard lists can be found in Appendix A.

- $m(D)$ within 70 (100) MeV/c^2 of nominal value for modes without (with) a π^0 .
- $130 < m(D^*) - m(D) < 170 \text{ MeV}/c^2$ for D^* mesons.
- All $D^{(*)}$ mesons are constructed using four-vector addition (vertexing will be performed later).

We identify leptons using the *PidLHElectrons*, *muMicroTight*, *muNNTight*, and *muNNTightFakeRate* selectors,³ and we construct $\Upsilon(4S)$ candidates in twelve different final states (three flavor combinations, D and D^* mesons, and two leptons):

$$\Upsilon(4S) \rightarrow B_{\text{tag}}^+ D^{(*)0} \ell^-, B_{\text{tag}}^0 D^{(*)+} \ell^-, \bar{B}_{\text{tag}}^0 D^{(*)+} \ell^- ,$$

where the modes with positively-charged leptons are included in a corresponding manner. The three daughters of the $\Upsilon(4S)$ candidate are required to be mutually nonoverlapping. This $\Upsilon(4S)$ reconstruction allows us to define four analysis **channels**, according to the signal-side charmed meson: D^0 , D^{*0} , D^+ ,

³We use three muon lists during the reskim to maintain flexibility during the subsequent analysis. We select only one muon list (*muMicroTight*) in the final selection, as described below.

and D^{*+} (we combine B^0 and \bar{B}^0 tags to allow for mixed events). Note that, as defined here, each channel includes both electrons and muons. All subsequent steps in the analysis treat the four channels separately.

All $\Upsilon(4S)$ candidates created are persisted with the Vcb reskim, so that the subsequent analysis code does not need to recreate them. As mentioned earlier, the reskim applies, by construction, much looser selection than will be used in the final analysis.

6.4 B_{tag} Selection

The remainder of this Section describes our final event selection applied to the $\Upsilon(4S)$ candidates after the BSemiExcl skim and the Vcb reskim, as described in Section 6.1.

We require B_{tag} candidates to have $|\Delta E| < 72$ MeV, corresponding to $\pm 4\sigma$ about the nominal mean. Figure 6.2 shows the optimization the ΔE cut. From a purely statistical point of view, the optimal point would be between 80 and 100 MeV. We find, however, that for events in the range $72 < \Delta E < 100$ MeV, the B_{tag} candidate is poorly measured, significantly decreasing our m_{miss}^2 resolution. Therefore, we choose a cut which balances the statistical optimization against systematic effects due to the decreased resolution.

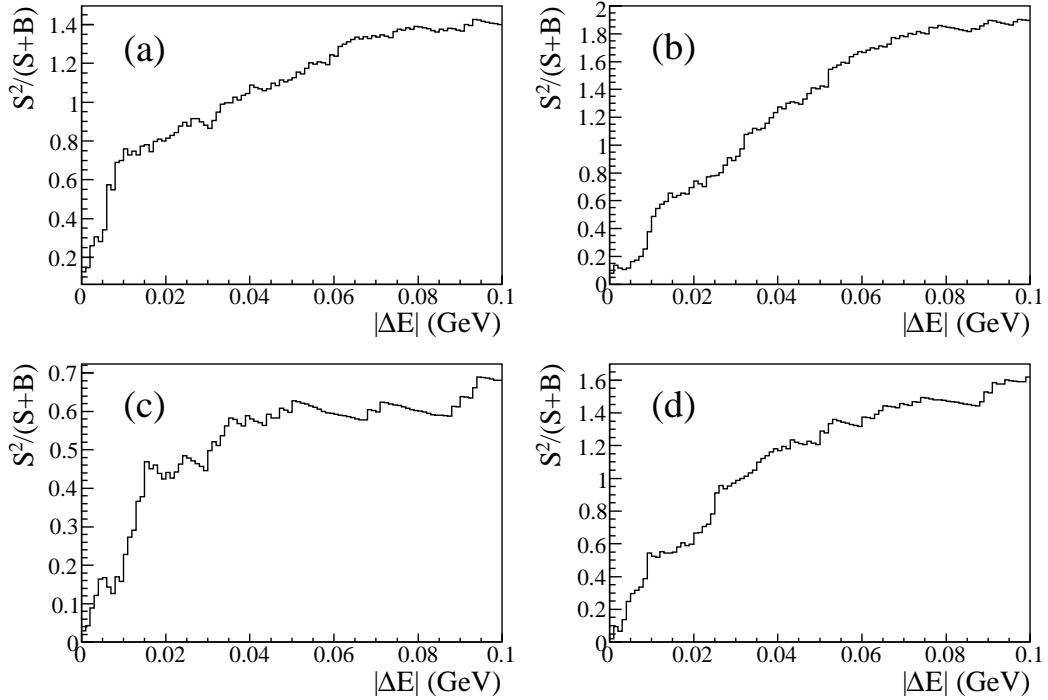


Figure 6.2: Optimization of the $|\Delta E|$ cut, showing $S^2/(S + B)$ in 640 fb^{-1} of generic MC, scaled to the Run 1–4 luminosity, for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels. The bumpiness of the curves is due to the limited number of signal events in generic MC. The statistically optimal cut is quite large, approximately 90 MeV, although the final cut on this variable, taking into consideration systematic effects, is at 72 MeV.

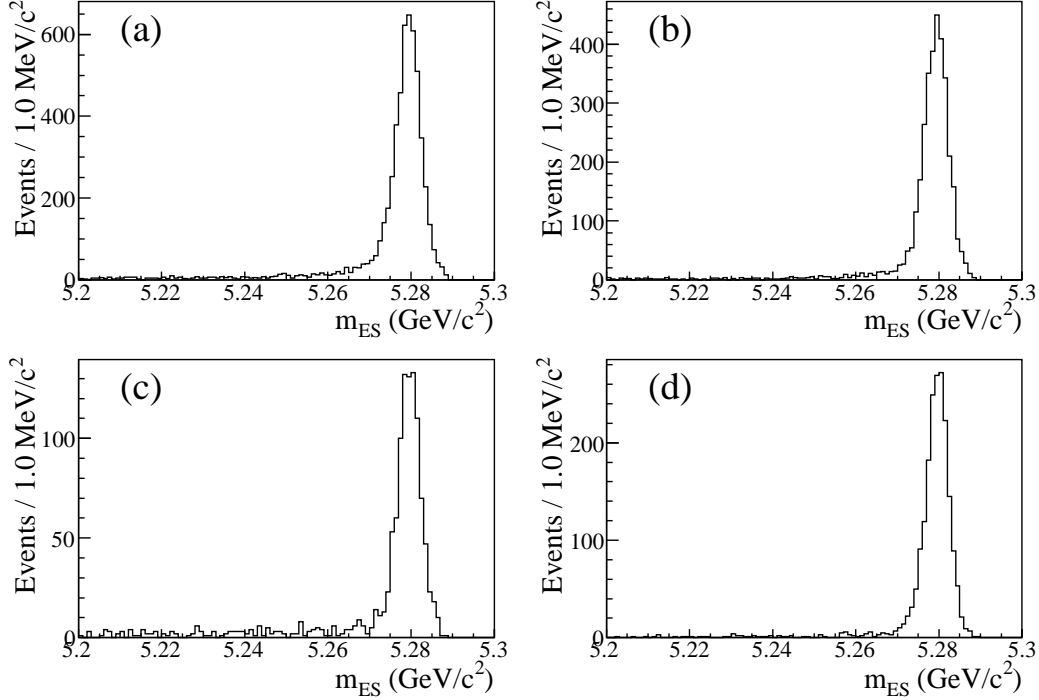


Figure 6.3: Distributions of m_{ES} from 640 fb^{-1} of generic MC, for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels. These distributions are shown after all event selection except the cuts on m_{ES} and m_{miss}^2 . Compared to Figure 6.1, which shows m_{ES} distributions at the skim level, we note that the signal-side selection substantially improves the purity on the tag side.

We further require B_{tag} candidates to satisfy $m_{\text{ES}} > 5.27 \text{ GeV}/c^2$. Figure 6.3 shows distributions of m_{ES} in generic MC after all other event selection has been applied. Compared to Figure 6.1, which shows m_{ES} distributions at the skim level, we note that the signal-side selection substantially improves the purity on the tag side. Figure 6.4 shows the corresponding ΔE distributions, and, again, we note the improvement in purity with respect to the skim.

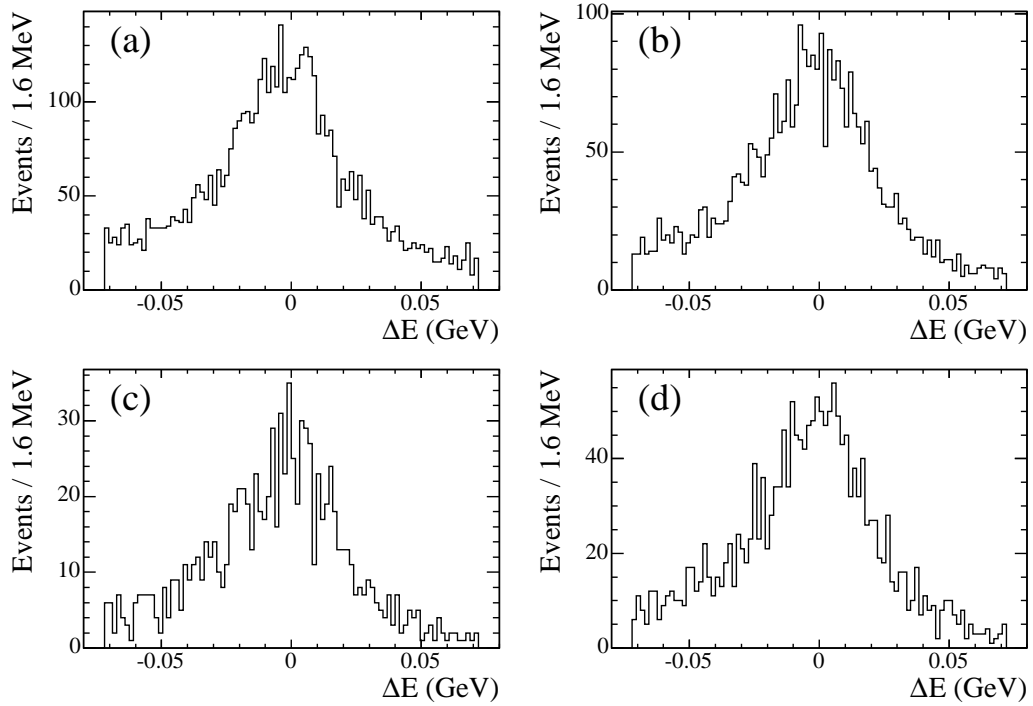


Figure 6.4: Distributions of ΔE from 640 fb^{-1} of generic MC, for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels. These distributions are shown after all event selection except the cut on m_{miss}^2 (this selection includes the cut $|\Delta E| < 72 \text{ MeV}$). Compared to Figure 6.1, which shows ΔE distributions at the skim level, we note that the signal-side selection substantially improves the purity on the tag side.

6.5 Charged Particle Identification

We use the following criteria for particle ID on the recoil:

- We use the *PidLHElectrons* selector for e candidates.
- We use the *muMicroTight* selector for μ candidates.
- We use the *KLHTight* selector for K^\pm candidates.
- We use the *PiLHLoose* selector for π^\pm candidates. This requirement is not applied to the pions identified as $K_s^0 \rightarrow \pi^+\pi^-$ or the soft pion in $D^{*+} \rightarrow D^0\pi^+$.

The definitions of these selectors can be found, along with a brief summary of their performance, in Appendix A.

6.6 Lepton Identification

In addition to the PID selectors, we require that ℓ candidates be found in the *GoodTracksLoose* list and satisfy $0.4 < \theta_\ell < 2.6$ rad. The definition of *GoodTracksLoose* can be found in Appendix A.

These cuts ensure that the momentum of the ℓ candidate is well measured, which is essential to this analysis. Figure 6.5 shows the ℓ momentum resolu-

tion as a function of polar angle; note the large biases near the edges of the acceptance, which are removed by this cut.

Finally, we require $p > 300$ MeV/ c for electron candidates, to ensure that the candidate has a well-measured energy in the EMC; without this cut, PID performance is unreliable.

6.6.1 Bremsstrahlung Recovery

We apply a standard *BABAR* bremsstrahlung recovery algorithm [37] to identify $e^\pm \rightarrow e^\pm\gamma$. At most one γ from *GoodPhotonLoose* (see Appendix A for the definition) is combined with an e^\pm candidate if it satisfies the following criteria:

- $0.0001 < \text{LAT} < 0.8$
- $Z_{42} < 0.25$
- $|\theta_e - \theta_\gamma| < 0.035$ rad
- $\phi_0^{e^-} - 0.05$ rad $< \phi_\gamma < \phi_{\text{cent.}}^{e^-}$ for e^-
- $\phi_{\text{cent.}}^{e^+} < \phi_\gamma < \phi_0^{e^+} + 0.05$ rad for e^+

Here, θ_γ and ϕ_γ are the polar and azimuthal angles of the photon centroid, θ_e and $\phi_{\text{cent.}}^e$ are the angles of the centroid of the cluster associated with the e^\pm

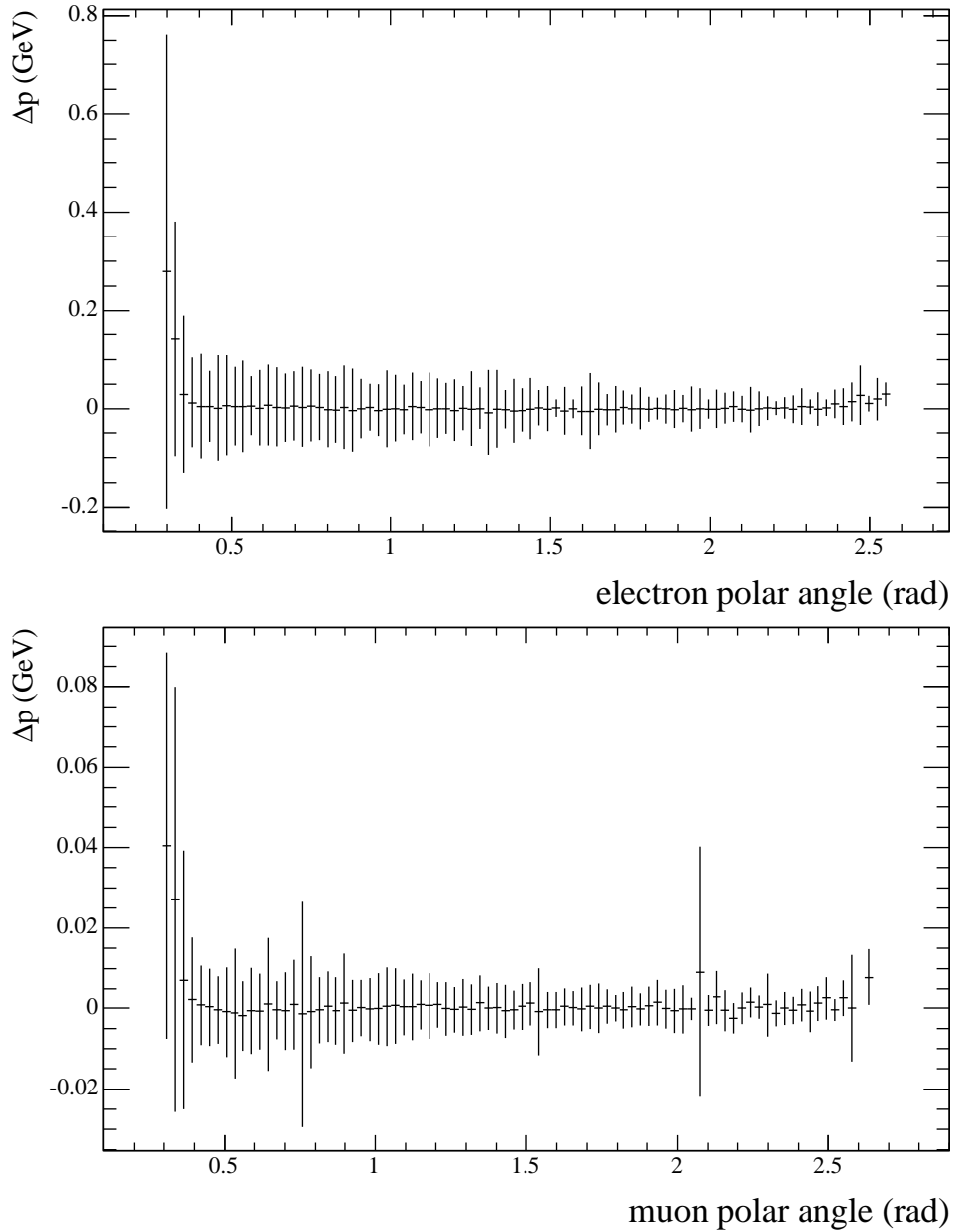


Figure 6.5: Profile plot of lepton momentum resolution as a function of polar angle in radians. Note the large biases in both lepton samples at the edges of the acceptance region; positive numbers here indicate that the measured momentum is lower than the true momentum.

candidate, and ϕ_0^e is the azimuthal angle of the e^\pm candidate measured at the closest approach to the beam spot.

For electrons selected by the bremsstrahlung recovery algorithm, the corrected momentum is used in place of the original momentum throughout the analysis.

6.7 Composite Particle Reconstruction

6.7.1 π^0 Reconstruction

π^0 daughters of D candidates come from the *pi0AllDefault* list, with an additional mass cut of $125 < m_{\gamma\gamma} < 145$ MeV/ c^2 , corresponding to $\pm 3\sigma$ about the fitted mean.

Because the π_s^0 from D^* decay has a much softer momentum spectrum than a typical π^0 from D decay, π_s^0 candidates used for $D^* \rightarrow D\pi_s^0$ come from the *pi0VeryLoose* list. This list is more efficient for low-momentum pions. We do not apply the tight mass window mentioned above because the mass resolution for low-momentum pions is worse than that of a typical π^0 .

The definition of both of these π^0 lists can be found in Appendix A.

6.7.2 K_S^0 Reconstruction

K_S^0 candidates are taken from the *KsDefault* list, with an additional mass cut of $491 < m_{\pi\pi} < 506$ MeV/ c^2 , corresponding to $\pm 3\sigma$ about the fitted mean.

The definition of the *KsDefault* list can be found in Appendix A.

6.7.3 $D^{(*)}$ Meson Reconstruction

For the final analysis, we select D mesons in a subset of the modes reconstructed during the reskim, removing the modes with the lowest sensitivity. In all modes, we fit the D mass distribution and select candidates within 4σ of the fitted mean.

For D^* mesons, the soft particle (either π or γ) is required to have a momentum $p_{\text{soft}}^* < 450$ MeV/ c in the center-of-mass frame of the $\Upsilon(4S)$. We fit the δm distribution and select candidates within 4σ of the fitted mean.

The mass and δm cuts used in this analysis are tabulated in Sections B.1 and B.2.

6.8 Vertexing and Kinematic Fitting

All $\Upsilon(4S)$ candidates are refitted before further consideration. This fit is primarily performed to improve the momentum resolution of the B_{tag} and

signal-side $D^{(*)}$ mesons, which ultimately improves the resolution on m_{miss}^2 . A secondary benefit of the fit is the suppression of combinatoric backgrounds for which the fit often fails to converge.

We use TreeFitter [38] to perform the fit. TreeFitter applies a kinematic constraint on every decay present within the $\Upsilon(4S)$ candidate. A geometric constraint is placed on the D decay vertex, the signal B decay vertex, the B_{tag} decay vertex, and the $\Upsilon(4S)$ decay vertex. The vertex constraint on the signal B meson assumes that the lepton is produced at the B decay vertex; for signal events, this neglects the fact that the τ lepton can travel before decaying, although, given the small boost of τ leptons in B decays,⁴ the effect of ignoring this impact parameter is negligible. The D meson on the signal side is mass constrained, as is the D^* meson if there is one. A beam-spot constraint is applied to the $\Upsilon(4S)$ candidate and to the signal-side D^* meson if there is one. A beam-energy constraint is applied to the $\Upsilon(4S)$ candidate.

⁴In signal events, the τ lepton travels a mean distance of 80μ before decaying. The mean *impact parameter*, the closest distance between the B decay vertex and the ℓ track is smaller still by a factor of $\langle \sin \theta_{\tau\ell} \rangle$, where $\theta_{\tau\ell}$ is the angle between the τ and ℓ flight directions — when the resulting ℓ track is parallel to the τ flight direction, the lepton track points back at the B vertex and the impact parameter is not distinguishable. The vertex resolution in *BABAR* is typically of the order of 100μ , larger than the mean lepton impact parameter, and, given that the D meson also travels a distance of order 100μ , including the τ flight in the fit is not likely to change things. Note that this argument only applies to the one-prong τ decay modes reconstructed here; if three-prong channels were included, the τ vertex becomes reconstructible, not just the impact parameter, possibly providing additional discriminating power and, eventually, enabling studies of the τ polarization, enhancing the New Physics sensitivity.

The complete decay tree is refitted, including π^0 and K_s^0 daughters of $D^{(*)}$ mesons and the B_{tag} , maximizing the sensitivity of the kinematic fit.

The fit is required to converge successfully; no subsequent cut is applied to the χ^2 of the fit.

6.9 p_{miss}

The missing momentum is defined as $p_{\text{miss}} \equiv (p_r - p_{\text{tag}} - p_{D^{(*)}} - p_\ell)$. We require events to satisfy $|p_{\text{miss}}| > 200 \text{ MeV}/c$. This requirement preferentially selects semileptonic and semitauonic events due to the presence of one or more neutrinos, and is particularly effective at removing $B \rightarrow D^{(*)}\pi^\pm$ events where the π^\pm is misidentified as a μ^\pm .

Requiring $|p_{\text{miss}}| > 200 \text{ MeV}/c$ rejects 100% of $B \rightarrow D^{(*)}\pi^\pm$ background while rejecting 0% of $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ signal and 1% of $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ background.

6.10 q^2

We require events to satisfy $q^2 > 4 \text{ (GeV}/c^2)^2$, where $q^2 \equiv (p_r - p_{\text{tag}} - p_{D^{(*)}})^2$. This cut preferentially rejects combinatoric background from two-body

B decays. Because signal events must satisfy $q^2 > m_\tau^2 \approx 3.16 \text{ (GeV}/c^2)^2$, this cut is very efficient, rejecting only 0.5%–1.5% of signal.

6.11 Extra Track Veto

We reject an $\Upsilon(4S)$ candidate if any good track in the event is not used in its reconstruction (see Section C.1 for the definition of “good” tracks). Such candidates must have been misreconstructed, and are likely to have unseen particles, biasing the m_{miss}^2 measurement.

Note that this requirement, coupled with the $\Upsilon(4S)$ channels reconstructed during the Vcb reskim (Sec. 6.3), implies that charge is conserved, *i.e.*, that the total visible event charge is zero.

6.12 E_{extra}

We define $E_{\text{extra}} = \Sigma E_\gamma$, where the sum is taken over all good photons in the event not associated with either the B_{tag} , $D^{(*)}$, or ℓ candidate (see Section C.2 for the definition of “good” photons). A candidate event where a particle has been ignored (a true $D^{*0}\ell^-\bar{\nu}_\ell$ event where the D^0 is reconstructed and the soft π^0 is not, for instance) will tend to have a larger value of E_{extra} than a correctly reconstructed event. Thus, we require events to satisfy $E_{\text{extra}} <$

E_{\max} , where the value of E_{\max} is determined separately for each of the four signal channels.

Figure 6.6 shows the optimization of the E_{extra} cut. Distributions of E_{extra} in generic MC are shown in Figure 6.7, along with the corresponding cuts. The cuts were chosen to optimize $S^2/(S+B)$ and are set at 200 MeV in the D^0 and D^{*0} channels, 150 MeV in the D^+ channel, and 300 MeV in the D^{*+} channel. These values are not very different if we instead optimize S^2/B . As with the ΔE cut, these cuts are somewhat tighter than a purely statistical optimization would require, and were chosen to balance the statistical power against systematic effects due to decreasing m_{miss}^2 resolution and increasing combinatoric background as these cuts are loosened.

Figures 6.8 and 6.9 show distributions of E_{extra} for signal events and semileptonic background events. Note that the distributions look very similar between signal and background, as well as across the four analysis channels. This demonstrates that, in the MC at least, the efficiency of the E_{extra} cut is the same between signal and semileptonic background, so that, when we normalize our measurement to the semileptonic background, this efficiency will cancel. More precisely, the efficiency for a $D^0\tau^-\bar{\nu}_\tau$ event to pass this cut, relative to a $D^0\ell^-\bar{\nu}_\ell$ event is 1.01 ± 0.05 ; for the D^{*0} , D^+ , and D^{*+} channels, the corresponding efficiencies are 1.01 ± 0.05 , 1.00 ± 0.05 , and 1.00 ± 0.05 , respectively.

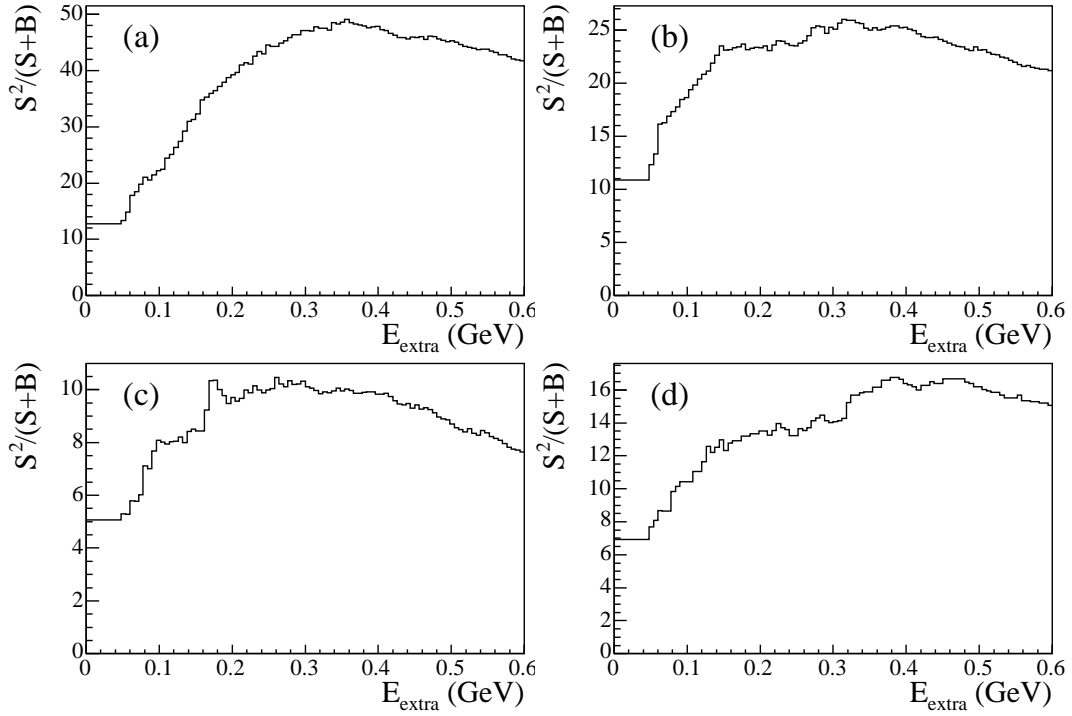


Figure 6.6: Optimization of the E_{extra} cut, showing $S^2/(S + B)$ in 640 fb^{-1} of generic MC, scaled to the Run 1–4 luminosity, for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels.

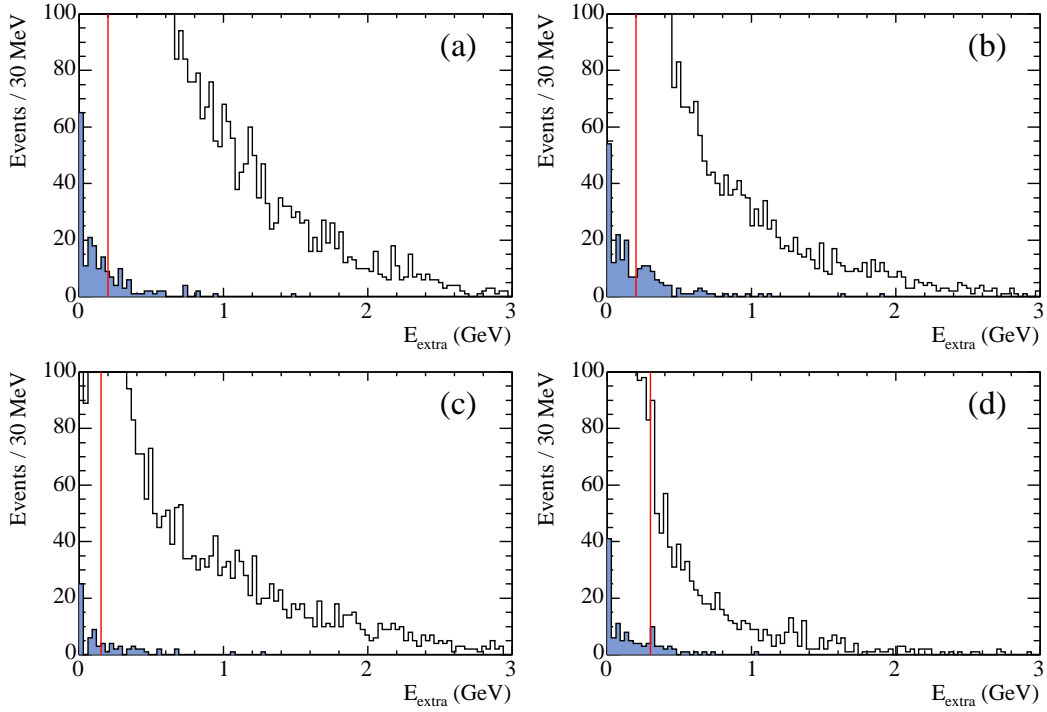


Figure 6.7: E_{extra} distributions from 640 fb^{-1} of generic MC for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels. In all figures, the blue shaded area represents signal $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ events while the unshaded histogram represents background. A red line indicates the cut placement in each mode. All event selection requirements have been applied except E_{extra} and m_{miss}^2 cuts.

Figure 6.10 shows the same distributions, split according to the reconstructed lepton type. Good agreement is seen in both the electron and muon channels.

6.13 Single Candidate Selection

For events with multiple $\Upsilon(4S)$ candidates, we choose the single candidate with the lowest value of E_{extra} . The average number of $\Upsilon(4S)$ candidates per event after the selection described above is approximately 1.9, which, while seemingly large, rate is not especially dangerous. The main reason this rate is so large is that, for every D^{*0} or $D^{*+} \rightarrow D^+\pi^0$ candidate, there also exists a corresponding candidate in the D^0 or D^+ channel in which the only difference is that the soft π^0 or γ is not used in the reconstruction and so is included in E_{extra} . Our arbitration scheme preferentially selects the candidate which uses the greatest fraction of the total visible event, and is therefore least likely to have lost a particle.

The main effect of this algorithm is to preferentially select a candidate in one of the D^* channels when both a D and D^* candidate are present in the event. Because $D^* \Rightarrow D$ feed-down is such a dominant background and $D \Rightarrow D^*$ feed-up is comparatively rare, keeping as many D^* events as possible

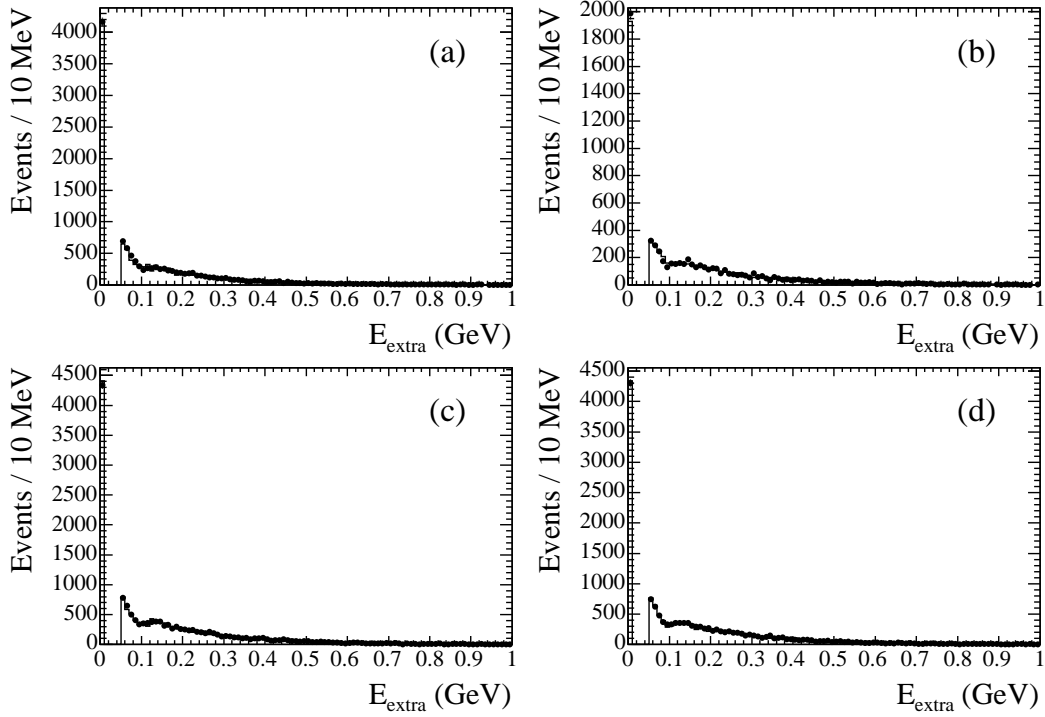


Figure 6.8: E_{extra} distributions from signal and semileptonic MC after all event selection except the cut on E_{extra} . The normalization is arbitrary. In all figures, the distribution from semileptonic MC is shown as a histogram and the distribution from signal MC is shown as points. Shown are: (a) $D^0 \tau^- \bar{\nu}_\tau$ and $D^0 \ell^- \bar{\nu}_\ell$, (b) $D^{*0} \tau^- \bar{\nu}_\tau$ and $D^{*0} \ell^- \bar{\nu}_\ell$, (c) $D^+ \tau^- \bar{\nu}_\tau$ and $D^+ \ell^- \bar{\nu}_\ell$, and (d) $D^{*+} \tau^- \bar{\nu}_\tau$ and $D^{*+} \ell^- \bar{\nu}_\ell$. The gap between the $E_{\text{extra}} = 0$ bin and the remainder of the distribution corresponds to the minimum allowed photon energy, 50 MeV. In all cases, the distributions are the same between signal and semileptonic MC.

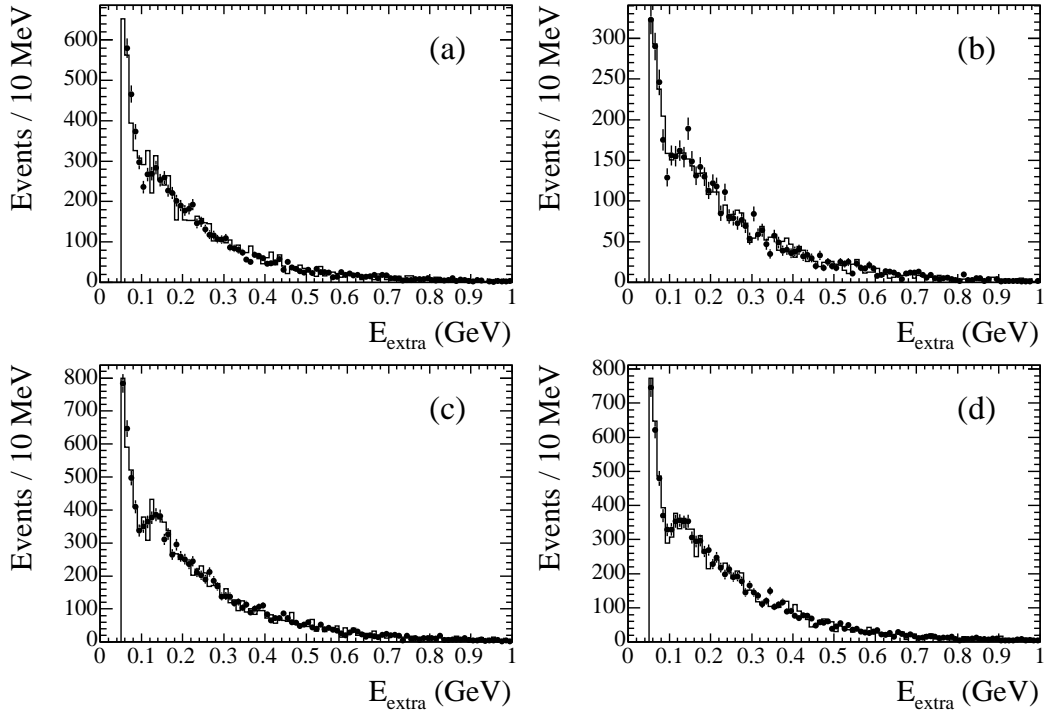


Figure 6.9: Same as Figure 6.8 with the zero bin suppressed.

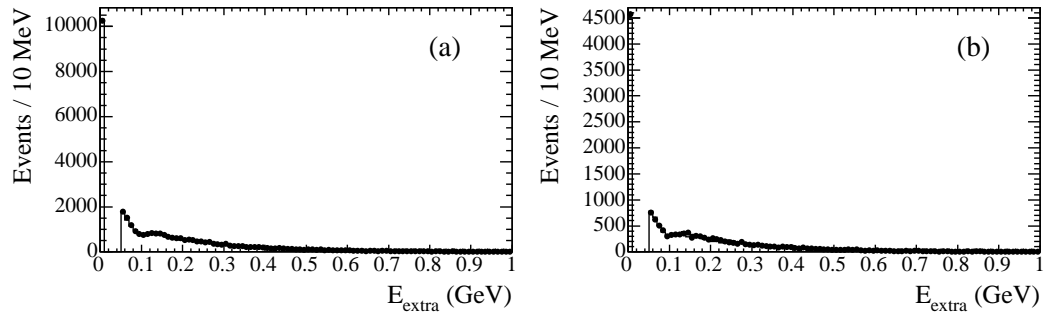


Figure 6.10: E_{extra} distributions from signal and semileptonic MC after all event selection except the cut on E_{extra} . The normalization is arbitrary. The four analysis channels are combined here, and the sample is split according to the reconstructed lepton type: (a) electrons and (b) muons. Again, the signal and normalization channels have indistinguishable distributions.

within the D^* channel is an important part of our event selection; choosing the best candidate based on E_{extra} helps achieve this goal.

6.14 m_{miss}^2

We define $m_{\text{miss}}^2 \equiv (p_r - p_{\text{tag}} - p_{D^{(*)}} - p_\ell)^2$. For the cut-and-count analysis and for optimizing the other cuts, we apply a cut on m_{miss}^2 to reject the dominant semileptonic backgrounds which, because there is only one ν , tends to have $m_{\text{miss}}^2 \approx 0$. This cut will not be used in the final analysis; instead, we perform a fit to the m_{miss}^2 distribution.

Distributions of m_{miss}^2 in the four signal channels are shown in Figure 6.11, along with the corresponding cuts. Figure 6.12 shows the same information on a logarithmic scale. The cuts were chosen to optimize $S^2/(S+B)$ (although, again, optimizing S^2/B is not very different) and are set at 1.6 (GeV/c^2)² in the D^0 channel and 1.8 (GeV/c^2)² in the other three.

6.15 Signal Efficiencies

Table 6.1 shows the efficiency of each cut used in selecting $\Upsilon(4S)$ candidates, taken from Signal MC. These efficiencies are representative of the analysis, but will not be used directly in the signal extraction because of large

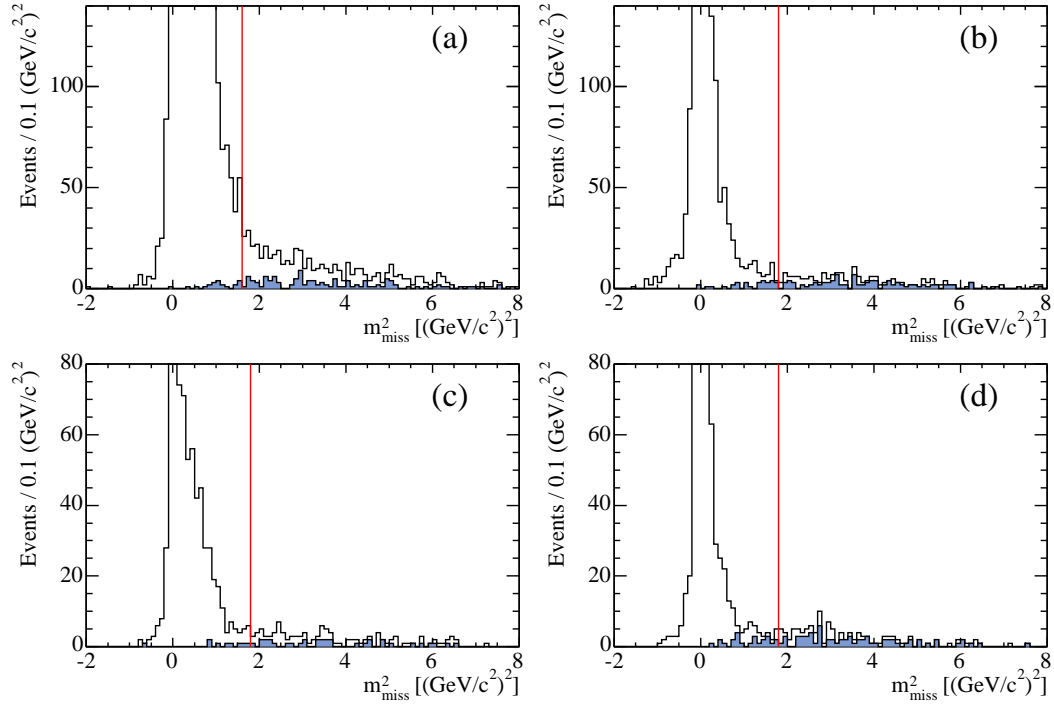


Figure 6.11: m_{miss}^2 distributions from 640 fb^{-1} of generic MC for the (a) D^0 , (b) D^{*0} , (c) D^+ , and (d) D^{*+} channels. In all figures, the blue shaded area represents signal $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ events while the unshaded histogram is background. A red line indicates the cut placement in each mode, which is used for optimization of the other cuts (but not in the final analysis). All event selection requirements have been applied except the m_{miss}^2 cut.

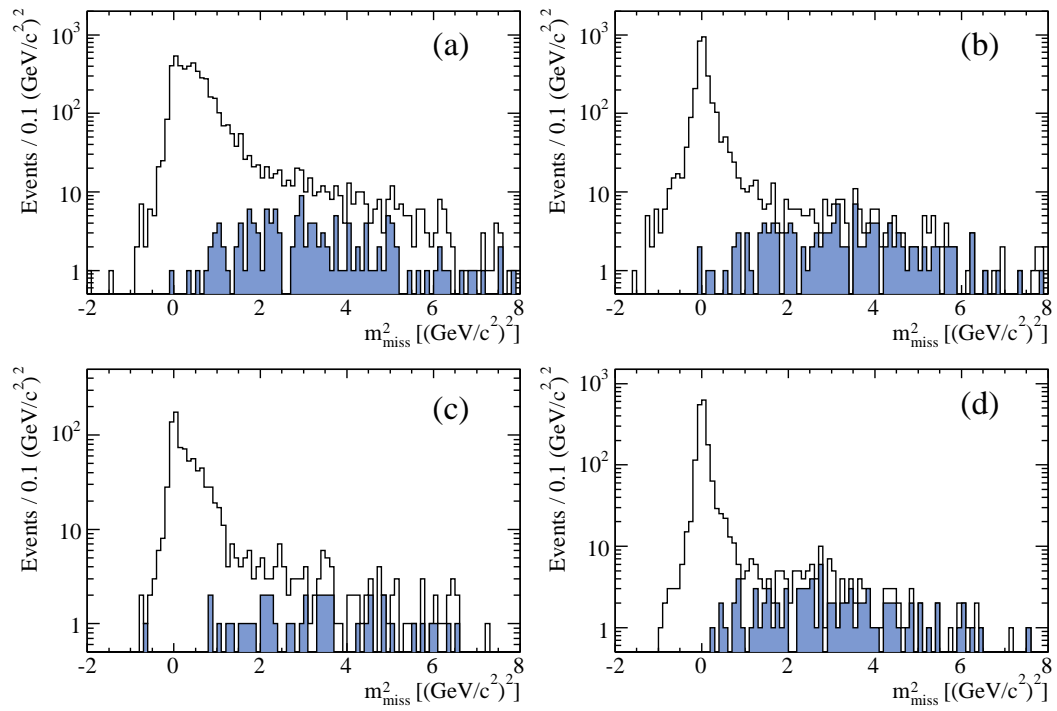


Figure 6.12: Same as Figure 6.11 but with a logarithmic scale

uncertainties in modeling the B_{tag} reconstruction efficiency, as well as several of the signal-side cuts. The relative normalization actually used will be discussed in Section 8. Note also that these tables only give the efficiency for candidates reconstructed in the correct channel; crossfeed is not included in this table, yet it contributes significantly to the signal yield.

6.16 Cut-and-Count Event Yields

As an estimate of our signal sensitivity, we perform a cut-and-count analysis on generic MC. After applying the event selection described above (including cuts on m_{ES} and m_{miss}^2), we count the number of events reconstructed in each of the four analysis channels — the results of this analysis are given in Table 6.2. The MC used here corresponds to approximately 640 fb^{-1} of data, or roughly three times the data luminosity.

We can see that the expected signal-to-background ratio is slightly better than one-to-one in three channels, and slightly worse than one-to-one in D^+ . The main background source is $D^{**}\ell^{-}\bar{\nu}_\ell$, followed by combinatoric background, with very little background from other semileptonic decays. Most of the combinatoric background comes from decays like $B \rightarrow D^{(*)}D^{(*)}$ and

Table 6.1: Cumulative cut-by-cut efficiency from Signal MC. The numbers N_{reco} are the number of events passing the cuts, not the number of $\Upsilon(4S)$ candidates, so that there is no over-counting here. Misreconstructed signal events, however, are included in these numbers, which is why the efficiency for some cuts appears low.

	D^0		D^{*0}	
	$N_{\text{reco}}/936\text{k}$	ε (%)	$N_{\text{reco}}/932\text{k}$	ε (%)
Skim and reskim	34606	3.697	26893	2.886
ΔE	30275	3.235	23180	2.487
ℓ PID and track selection	29834	3.187	22820	2.448
K/π PID, π^0 and K_S^0 mass	22444	2.398	16934	1.817
$m(D)$ and δm	17270	1.845	9452	1.014
Vertex convergence	16188	1.729	8787	0.9428
Extra tracks = 0	15079	1.611	8091	0.8681
Best candidate selection	13587	1.452	7619	0.8175
E_{extra}	8988	0.960	4631	0.4969
	D^+		D^{*+}	
	$N_{\text{reco}}/934\text{k}$	ε (%)	$N_{\text{reco}}/1872\text{k}$	ε (%)
Skim and reskim	53250	5.701	66155	3.534
ΔE	48120	5.152	59416	3.174
ℓ PID and track selection	47460	5.081	58524	3.126
K/π PID, π^0 and K_S^0 mass	33627	3.6	41371	2.21
$m(D)$ and δm	19492	2.087	19801	1.058
Vertex convergence	18582	1.99	17978	0.9604
Extra tracks = 0	17342	1.857	16755	0.895
Best candidate selection	16817	1.801	16163	0.8634
E_{extra}	8876	0.9503	11996	0.6408

Table 6.2: Event yields from 640 fb^{-1} of generic MC after applying all selection requirements in the cut-and-count analysis. Each row indicates the number of events of a specific generated type which were reconstructed in each of the four signal channels. For the purposes of this table, “combinatoric” background contains all decays other than the 14 specifically listed. “Signal” here includes both correctly identified signal candidates and signal crossfeed background, but does not include crossfeed between charged and neutral B decays.

Generated Event Type	Reconstruction Channel			
	D^\pm	$D^{*\pm}$	D^0	D^{*0}
$B^- \rightarrow D^0 e^- \bar{\nu}_e$	0	0	5	1
$B^- \rightarrow D^{*0} e^- \bar{\nu}_e$	2	0	51	11
$B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu$	0	0	1	0
$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	0	0	15	1
$B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$	0	0	120	12
$B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$	1	0	236	102
$\bar{B}^0 \rightarrow D^+ e^- \bar{\nu}_e$	4	0	0	0
$\bar{B}^0 \rightarrow D^{*+} e^- \bar{\nu}_e$	2	9	2	1
$\bar{B}^0 \rightarrow D^+ \mu^- \bar{\nu}_\mu$	3	0	0	0
$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	6	2	2	0
$\bar{B}^0 \rightarrow D^+ \tau^- \bar{\nu}_\tau$	34	4	1	0
$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	19	67	6	2
Nonresonant $B \rightarrow D^{(*)} \pi \ell^- \bar{\nu}_\ell$	16	7	77	11
Resonant $B \rightarrow D^{**} \ell^- \bar{\nu}_\ell$	13	12	75	23
Combinatoric	15	21	50	30
Total signal	53	71	356	124
Total background	62	51	285	80
S/B	0.85	1.39	1.25	1.55

$B \rightarrow D^{(*)} D_s^{(*)+}$, as described in Section 5.2; we will show the composition in more detail in Section 12.5 when we discuss systematic uncertainties.

Chapter 7

Control Samples and MC Validation

The event selection described in Section 6, which involves full-event reconstruction of a high-multiplicity final state, is more complicated than a typical *BABAR* analysis. Further, our selection depends on being able to veto events with extra tracks and extra neutral clusters, quantities for which the reliability of the Monte Carlo simulation has not been thoroughly studied in *BABAR*. We will use two control samples to test the predictions of the MC against the behavior observed in data, the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample and the q^2 control sample.

We also use a control sample to constrain the D^{**} contamination within our signal region. While the dominant backgrounds in this analysis are from $D^{(*)}\ell^{-}\bar{\nu}_\ell$ events, these events are confined to the region $m_{\text{miss}}^2 \approx 0$, and additionally, these processes are understood quite well, so uncertainties associ-

ated with their kinematics should be quite small. For $D^{**}\ell^{-}\bar{\nu}_\ell$ events, however, the situation is different. To pass our event selection, a $D^{**}\ell^{-}\bar{\nu}_\ell$ event must have lost at least one particle — the soft transition particle(s) from $D^{**} \rightarrow D^{(*)}\pi(\pi\dots)$ — and so must have $m_{\text{miss}}^2 > 0$. Furthermore, neither the decay kinematics nor the individual branching fractions for $B \rightarrow D_1\ell^{-}\bar{\nu}_\ell$, $B \rightarrow D_2^*\ell^{-}\bar{\nu}_\ell$, $B \rightarrow D_0^*\ell^{-}\bar{\nu}_\ell$, $B \rightarrow D_1^*\ell^{-}\bar{\nu}_\ell$, and nonresonant $B \rightarrow D^{(*)}\pi\ell^{-}\bar{\nu}_\ell$ are well known, so any MC prediction of D^{**} contamination is susceptible to large systematic uncertainties. We will instead use a control sample in data to determine this background.

7.1 $D^{(*)}\ell^{-}\bar{\nu}_\ell$ Control Samples

We take advantage of the fact that our dominant background, $D^{(*)}\ell^{-}\bar{\nu}_\ell$, has a different lepton momentum spectrum than our signal to define the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample. Figure 7.1 shows the p_ℓ^* spectrum for generic MC events passing all selection for the D^{*0} channel. We see that there is almost no signal above 1.5 GeV/c, because the lepton in this case is the secondary lepton from $\tau^- \rightarrow \ell^{-}\bar{\nu}_\ell\nu_\tau$ and is typically low momentum (see also Section 2 for more details about the p_ℓ^* spectrum in $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and in $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$). Table 7.1 shows the composition of generic MC events in the region $p_\ell^* > 1.5$ GeV/c,

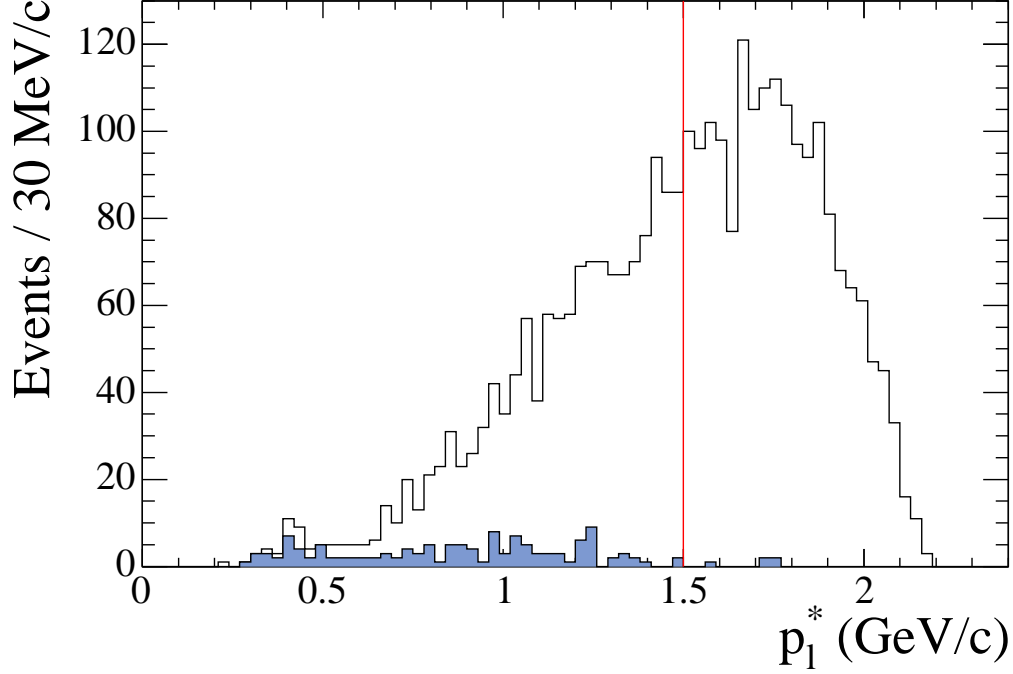


Figure 7.1: Distributions of p_ℓ^* from 640 fb^{-1} of generic Monte Carlo, used to define the $D^{*0}\ell^-\bar{\nu}_\ell$ control sample. The p_ℓ^* distribution is shown for D^{*0} candidates passing all selection. True $D^{*0}\tau^-\bar{\nu}_\tau$ events are shaded in blue. The $D^{(*)}\ell^-\bar{\nu}_\ell$ control sample is defined as $p_\ell^* > 1.5 \text{ GeV}/c$.

where we can see that this selects samples which are approximately 95% pure $D^{(*)}\ell^-\bar{\nu}_\ell$.

We define four control samples, one corresponding to each signal channel, by requiring $p_\ell^* > 1.5 \text{ GeV}/c$. The $D^{*+}\ell^-\bar{\nu}_\ell$ control sample looks qualitatively like the $D^{*0}\ell^-\bar{\nu}_\ell$ sample, except for a lower event yield (due to the lower B_{tag} efficiency for neutral B mesons); the $D\ell^-\bar{\nu}_\ell$ control samples, however, contain both $D\ell\nu$ events and a significant contribution from $D^*\ell^-\bar{\nu}_\ell$ feed-down. As this

Table 7.1: Composition of the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control samples in generic MC. Here, “charge crossfeed” refers to true $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$ events in which both the B_{tag} and the signal $D^{(*)}$ are reconstructed with the wrong charge; “combinatoric” refers to all other decays. We see that the control samples are approximately 95% pure $D^{(*)}\ell^{-}\bar{\nu}_\ell$.

Component	Composition (%)			
	D^0	D^{*0}	D^+	D^{*+}
$B \rightarrow D\ell^{-}\bar{\nu}_\ell$	23.1	2.2	45.6	0.4
$B \rightarrow D^*\ell^{-}\bar{\nu}_\ell$	70.8	94.5	48.8	96.7
$B \rightarrow D^{**}\ell^{-}\bar{\nu}_\ell$	2.9	1.4	2.3	1.4
$B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_\tau$	0.6	0.5	0.8	0.8
Charge crossfeed	1.9	1.1	0.9	0.1
Combinatoric	0.7	0.3	1.6	0.6

is also the case for the signal region in p_ℓ^* , these control samples will provide a valuable test of the Monte Carlo modeling of the kinematic variables in feed-down events.

Figures 7.2–7.5 show data–MC comparisons for several important distributions within the four $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control samples. We have examined a large number of variables in order to have a comprehensive picture of the performance of the Monte Carlo and to eliminate, as much as possible, the chance that a data–MC discrepancy in some region of phase space could bias the measurement. We have examined several broad categories of variables, including event-selection variables (E_{extra} , B_{tag} multiplicity and purity, kinematic fit

probability), kinematic variables (p_ℓ^* , q^2 , p_{miss}), m_{miss}^2 resolution, and soft π^0/γ efficiency modeling.

In each plot, the histogram is the MC prediction and the points with error bars are data. All distributions are sideband-subtracted, using m_{ES} for the B_{tag} candidate, to account for the fact that the MC does not correctly model the B_{tag} efficiency.¹ In all figures, the MC prediction is scaled to the event yield in data.

All of these figures show excellent agreement between the data and the MC. In particular, note that the $D\ell^-\bar{\nu}_\ell$ plots in Figure 7.5 show both a peak at $m_{\text{miss}}^2 = 0$ corresponding to true $D\ell\nu$ events, as well as a shoulder at larger values of m_{miss}^2 , corresponding to $D^*\ell^-\bar{\nu}_\ell$ feed-down, and that the data–MC agreement in this case is quite good. The p_ℓ^* dependence of this shape is also seen to be modeled well, as are the shapes for electrons and muons separately.

We have also examined variables related to D^* reconstruction. We see good agreement in the observed momentum spectrum of the soft π/γ and the angular distributions of the π/γ , leading us to conclude that the MC does a good job modeling the soft π/γ efficiency. We also see that the MC is able to predict the m_{miss}^2 spectrum when the soft π/γ from D^* decay is lost, which

¹The background to be subtracted typically accounts for 0.4–1.0% of the events in the signal range, defined as $5.27 < m_{\text{ES}} < 5.29$ GeV/ c^2 .

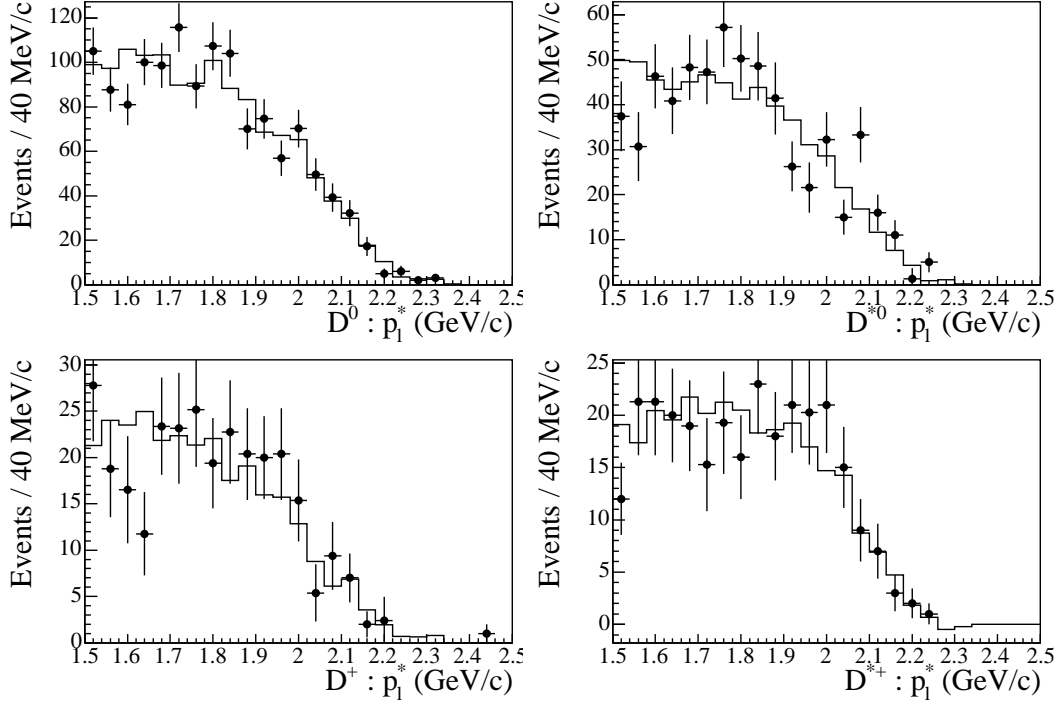


Figure 7.2: Data–MC comparison of p_ℓ^* in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples. Points with error bars correspond to the full Run 1–4 data sample, and the histogram corresponds to generic MC. Both the data and MC are shown after an m_{ES} sideband subtraction and the MC histogram is normalized to the number of events in the data, as described in the text.

gives us additional confidence that the MC is reliable in the large m_{miss}^2 region where our analysis is most sensitive.

In Section 12, we will revisit this control sample to help estimate systematic uncertainties in our m_{miss}^2 modeling.

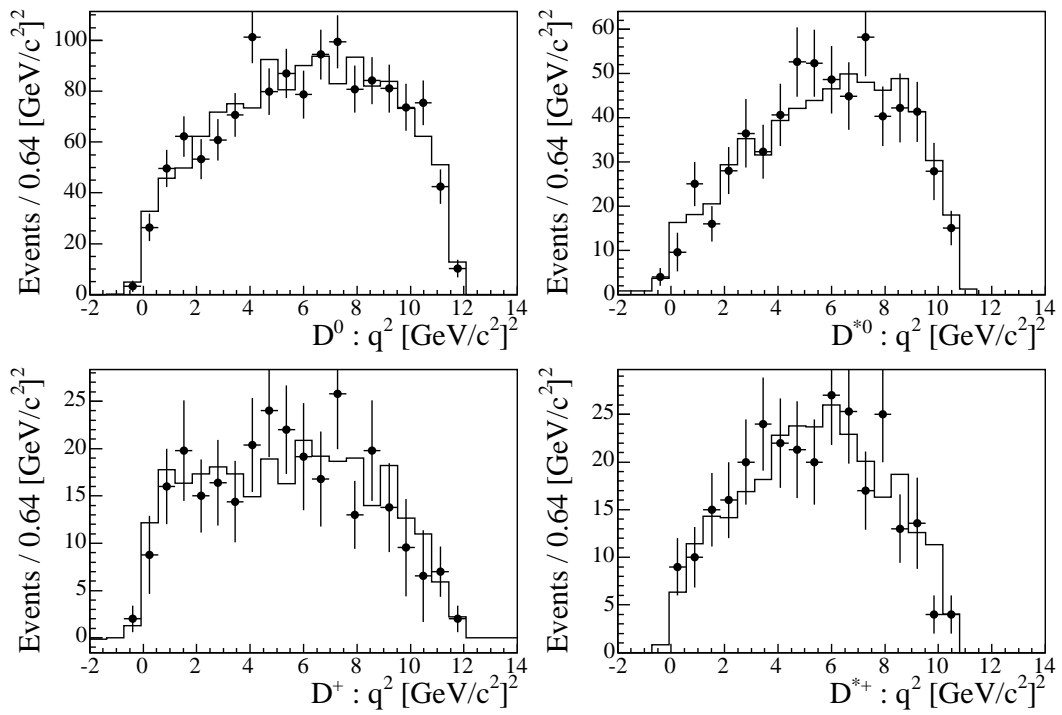


Figure 7.3: Data–MC comparison of q^2 in the $D^{(*)}\ell^-\bar{\nu}_\ell$ control samples

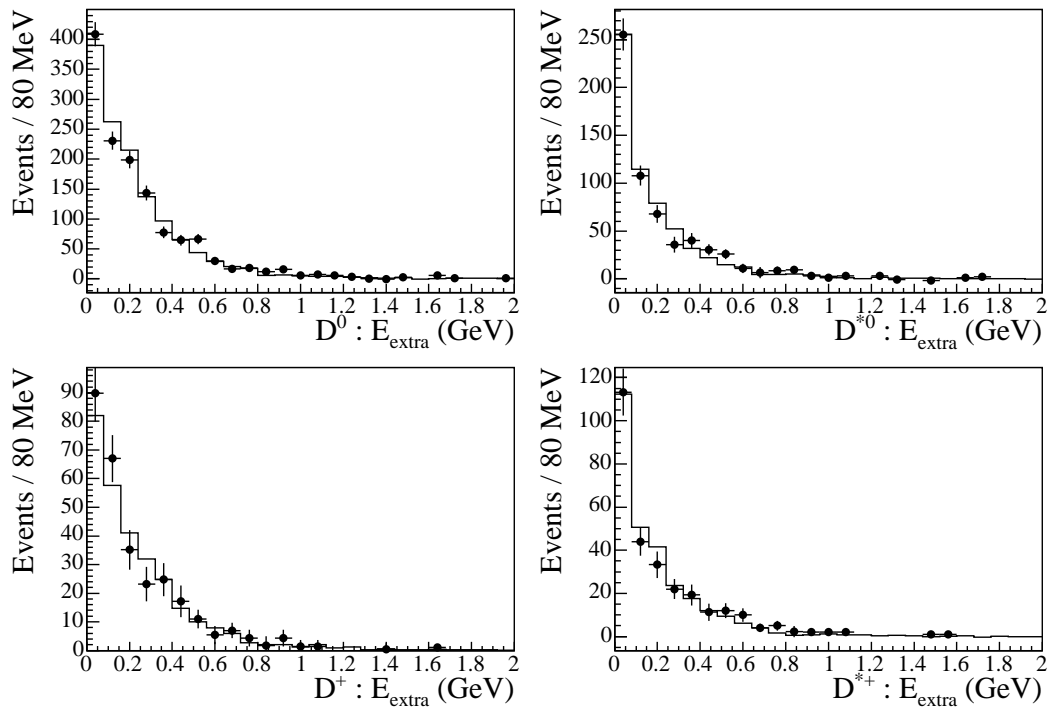


Figure 7.4: Data–MC comparison of E_{extra} in the $D^{(*)}\ell^{-}\bar{\nu}_{\ell}$ control samples

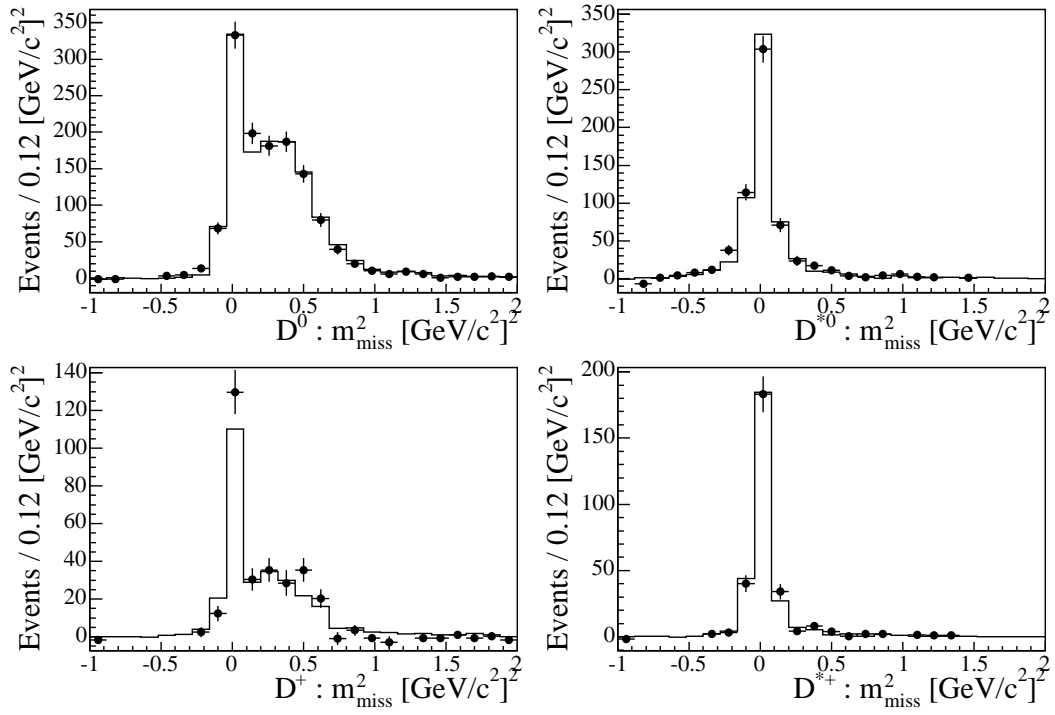


Figure 7.5: Data–MC comparison of m_{miss}^2 in the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control samples

7.2 q^2 Control Samples

The main shortcoming of the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample is that it requires a cut on p_ℓ^* ; we would like to test the MC in a region of lepton momentum which is more signal-like. To do this, we look at a second control sample, defined using $q^2 \equiv (p_{B_{\text{sig}}} - p_{D^{(*)}})^2$. For $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_\tau$ events, conservation of momentum requires $q^2 > m_\tau^2 \sim 3.2 \text{ (GeV}/c^2\text{)}^2$, so we may select a signal-depleted sample of events by cutting on q^2 . Figure 7.6 shows the q^2 distribution for events passing all selection criteria (except the cut on q^2 itself) for the D^0 channel. We see that there is almost no signal below $q^2 = 5 \text{ (GeV}/c^2\text{)}^2$.

As before, we define four control samples, one corresponding to each signal channel, by requiring $q^2 < 5 \text{ (GeV}/c^2\text{)}^2$. Table 7.2 shows the composition of generic MC events in these control samples. Again, these samples are dominated by $D^{(*)}\ell^{-}\bar{\nu}_\ell$ events (approximately 90%), although the contamination from $D^{**}\ell^{-}\bar{\nu}_\ell$, signal, and combinatoric background is larger.

As with the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control samples, we see excellent agreement between data and MC, and we point out that these two control samples are largely statistically independent. Approximately 50% of the events in these samples are common to the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample.

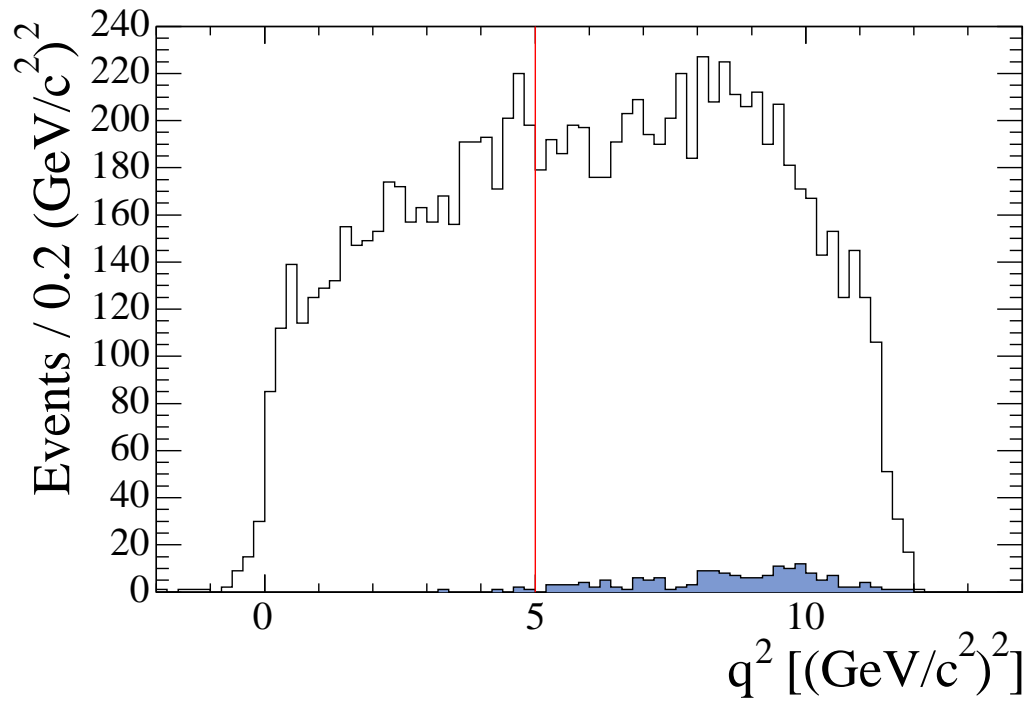


Figure 7.6: Distributions of q^2 from 640 fb^{-1} of generic Monte Carlo, used to define the q^2 control sample. The q^2 distribution is shown for D^0 candidates passing all selection except for the cut on q^2 . True $D^0 \tau^- \bar{\nu}_\tau$ events are shaded in blue. The q^2 control sample is defined as $q^2 < 5(\text{GeV}/c^2)^2$.

Table 7.2: Composition of the q^2 control samples in generic MC. Here, “charge crossfeed” refers to true $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$ events in which both the B_{tag} and the signal $D^{(*)}$ are reconstructed with the wrong charge; “combinatoric” refers to all other decays. We see that the control samples are approximately 90% pure $D^{(*)}\ell^{-}\bar{\nu}_\ell$.

Component	Composition (%)			
	D^0	D^{*0}	D^+	D^{*+}
$B \rightarrow D\ell^{-}\bar{\nu}_\ell$	43.3	3.7	61.6	0.2
$B \rightarrow D^*\ell^{-}\bar{\nu}_\ell$	47.2	86.4	25.2	89.6
$B \rightarrow D^{**}\ell^{-}\bar{\nu}_\ell$	2.8	3.2	3.6	3.1
$B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_\tau$	2.1	2.8	3.2	3.7
Charge crossfeed	1.5	0.9	1.2	0.2
Combinatoric	3.1	3.2	5.2	3.2

7.3 D^{**} Control Samples

According to Monte Carlo studies, the dominant source (70%) of $D^{**}\ell^{-}\bar{\nu}_\ell$ background in the signal region is events in which a π^0 was lost, either from $D^{**} \rightarrow D^{(*)}\pi^0$ or from nonresonant $B \rightarrow D^{(*)}\pi^0\ell^{-}\bar{\nu}_\ell$. Unfortunately, D^{**} transitions involving a π^0 have never been observed, nor have the nonresonant modes, so any background estimation is based on a number of assumptions which necessarily lead to large uncertainties. On the other hand, events in which one or two π^\pm are lost are highly suppressed in our analysis, due to the requirements on extra tracks and charge-correlation between the $D^{(*)}$ meson and the B_{tag} , and are expected to contribute less than one event to the D^{**} background yield in data. Note that the only 1/3 of the *produced* D^{**} mixture

in the Monte Carlo decays to $D^{(*)}\pi^0$, both for resonant and nonresonant decays; the fact that 70% of *reconstructed* decays involve a π^0 shows how efficient we are at rejecting D^{**} backgrounds involving charged pions.

Because our background is dominated by events with lost π^0 mesons, we will use events in which we find extra π^0 mesons to define $D^{**}\ell^-\bar{\nu}_\ell$ control samples, and to help constrain the amount of D^{**} background in our signal region. We are only dependent on the *BABAR* MC for the π^0 efficiency. Assuming this is known, we may use the control sample in data to determine the feed-down into our signal region without needing to know either the production branching fractions $B \rightarrow D^{**}\ell^-\bar{\nu}_\ell$, the $D^{**} \rightarrow D^{(*)}\pi^0$ daughter branching fractions, or the dynamical details of either process.

To select these samples, we apply the event selection criteria described above, omitting the cut on E_{extra} . We then search the event for extra π^0 mesons, which are π^0 candidates from the *pi0VeryLoose* list (see Appendix A) that do not overlap the $\Upsilon(4S)$ candidate.²

The extra π^0 is required to satisfy:

- $125 < m_{\gamma\gamma} < 145 \text{ MeV}/c^2$.
- $p_\pi^0 > 400 \text{ MeV}/c$ for all four channels

²We use the term overlap here in the *BABAR* Beta sense: neither the π^0 itself nor either of its two γ daughters may be used in reconstructing the $\Upsilon(4S)$.

- E_{extra,π^0} , the remaining neutral energy, not counting the extra π^0 , is required to satisfy $E_{\text{extra},\pi^0} < 500$ MeV for all four channels

For events with multiple extra π^0 candidates, we select the π^0 with the largest energy. These selection criteria were optimized to select a sample of $D^{**}\ell^{-}\bar{\nu}_\ell$ events with minimal contamination from combinatoric background.

In generic MC, these criteria select event samples that are composed of 60–80% pure $D^{**}\ell^{-}\bar{\nu}_\ell$ (of which, more than 90% involve true D^{**} transitions with π^0 mesons), $\sim 10\%$ feed-up from $D\ell^{-}\bar{\nu}_\ell$ and $D^*\ell^{-}\bar{\nu}_\ell$, and the remainder combinatoric background.

The event yields in both MC and data are too small to allow the type of statistical comparison done in the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ and q^2 control samples. Therefore, rather than perform the same MC validation of kinematic distributions and *BABAR* performance, we will use the $D^{**}\ell^{-}\bar{\nu}_\ell$ control sample directly within the signal fit (see Section 8). The yield of D^{**} events in the control sample is of the same order as the yield in the signal region, and we will use the measurement in the control sample to constrain the yield in the signal region (see Section 8.8).

Chapter 8

Fit Procedure

We extract our signal by performing an extended unbinned maximum likelihood fit to the m_{miss}^2 and p_ℓ^* joint distribution. In doing this, we do not apply the cut on m_{miss}^2 described earlier. The fit is performed simultaneously in eight samples: the four analysis channels D^0 , D^{*0} , D^+ , and D^{*+} , and the four corresponding channels in the D^{**} control sample, $D^0\pi^0$, $D^{*0}\pi^0$, $D^+\pi^0$, and $D^{*+}\pi^0$.

Most of the probability distribution function (PDF) shape parameters are fixed to Monte Carlo-derived values, as described below. All of the combinatoric background yields, both for charge crossfeed and “other” combinatoric backgrounds, are fixed to the expected value as predicted by the MC in the four signal channels; the charge crossfeed background yields are also fixed in the D^{**} control samples. The only free parameters in the fit are the yields of signal and semileptonic backgrounds and “other” combinatoric backgrounds

in the D^{**} control samples, and parameters describing the $D^* \Rightarrow D$ feed-down rates. We are able to float the combinatoric background yields in the D^{**} control samples because the broad shape in m_{miss}^2 is sufficiently different from the peaking D^{**} “signal” shape in these channels to be distinguishable.

In Section 8.1, we introduce the likelihood function for the fit, and in Section 8.2, we define each of the signal and background components that are used to model the eight channels in the fit. In Section 8.3, we introduce a model for the lepton spectrum, and, in Section 8.4, we show how this one-dimensional model is extended to create a two-dimensional $m_{\text{miss}}^2-p_\ell^*$ PDF. We document our use of the signal and semileptonic MC samples in Section 8.5, and, in Section 8.6, we show projections of the resulting fits to these MC samples. Our Monte Carlo samples for combinatoric background, along with the resulting PDFs, are presented in Section 8.7. We apply a number of constraints on the fit, which are described in Sections 8.8 and 8.9. Finally, we put all of the pieces together and perform a complete fit to the generic MC; this fit is shown in Section 8.10.

8.1 Likelihood Function

The complete extended likelihood function is given by:

$$\begin{aligned}
 \mathcal{L} &= \prod_{i=1}^8 \mathcal{L}_i \\
 &= \prod_{i=1}^8 \left[\frac{e^{-\mu_i} \mu_i^{N_i^{\text{evt}}}}{N_i^{\text{evt}}!} \times \prod_{k=1}^{N_i^{\text{evt}}} \left(\sum_{j=1}^{n_i^c} \frac{N_{ij}}{\mu_i} \mathcal{P}_{ij}(p_\ell^*, m_{\text{miss}}^2) \right) \right] \\
 &= \prod_{i=1}^8 \left[e^{-\mu_i} \mu_i^{N_i^{\text{evt}}} \times \prod_{k=1}^{N_i^{\text{evt}}} \left(\sum_{j=1}^{n_i^c} \frac{N_{ij}}{\mu_i} \mathcal{P}_{ij}(p_\ell^*, m_{\text{miss}}^2) \right) \right]. \quad (8.1)
 \end{aligned}$$

Here, the index i runs over the eight channels. j runs over the various signal and background components within sample i ; these components are described in detail below, and there are n_i^c such components in channel i (in each of the four signal channels, $n_i^c = 7$, while in each of the four D^{**} channels, $n_i^c = 5$). The index k runs over each event reconstructed in channel i , from 1 to N_i^{evt} , and the reconstruction variables p_ℓ^* , and m_{miss}^2 are all obviously independently measured for each event k . The only free parameters within the fit are the N_{ij} terms, the number of events in the i^{th} channel and the j^{th} component, and two parameters describing $D^* \Rightarrow D$ feed-down (see Section 8.9; the rest of the shape parameters describing the various \mathcal{P}_{ij} are fixed in fits to Monte Carlo samples. There are 48 yield parameters in total (7×4 signal channels $+ 5 \times 4$ D^{**} channels), although, due to the constraints on the fit and the fixed combinatoric backgrounds, the total number of free parameters is only 18. The

number of expected events in the i^{th} channel, μ_i , is defined as $\mu_i \equiv \sum_{j=1}^{n_i^c} N_{ij}$. The factor $N_i^{\text{evt}}!$ in the second expression is independent of the fit variables and so can safely be discarded.

8.2 Signal and Background Components

As mentioned above, each of the four signal channels is fit to the sum of seven components, and each of the four D^{**} control sample channels is fit to the sum of five components. The signal channel components are listed in Table 8.1 along with their approximate abundances in generic MC; an analagous breakdown of the D^{**} control sample channels is shown in Table 8.2.

In each of the four signal channels, there are two signal components, one representing correctly-reconstructed signal events, and the second representing signal feed-up ($D \Rightarrow D^*$) or feed-down ($D^* \Rightarrow D$), depending on the channel. Likewise, there are two $D^{(*)}\ell^{-}\bar{\nu}_\ell$ components, one correctly reconstructed and one feed-up or feed-down. In each signal channel, there is a component describing the feed-down of $D^{**}\ell^{-}\bar{\nu}_\ell$ events, including the small contribution from $D^{**}\tau^{-}\bar{\nu}_\tau$.¹

¹As noted in Section 4, we do not have dedicated signal MC samples for $D^{**}\tau^{-}\bar{\nu}_\tau$. In Section 8.5, we discuss our approach to modeling this background.

Table 8.1: Components of the signal extraction fit in the signal channels, and their approximate abundances in generic MC. The structure of the fit is identical between the B^- and \bar{B}^0 channels. There are seven components in each of the four signal channels. The last column gives the figures where the fitted $m_{\text{miss}}^2 - p_\ell^*$ distributions, as described in the following sections, can be found.

Channel	Source	Abundance in generic MC (%)	Figure
D^{*0}	$D^{*0} \tau^- \bar{\nu}_\tau$ signal	5	8.4
	$D^0 \tau^- \bar{\nu}_\tau$ signal feed-up	0.5	8.5
	$D^{*0} \ell^- \bar{\nu}_\ell$ background	90	8.6
	$D^0 \ell^- \bar{\nu}_\ell$ feed-up	2	8.7
	$D^{**} (\ell^- / \tau^-) \bar{\nu}$ feed-down	3	8.8
	Combinatoric: charge crossfeed	0.5	8.37
	Other combinatoric background	1	8.38
D^0	$D^0 \tau^- \bar{\nu}_\tau$ signal	4	8.9
	$D^{*0} \tau^- \bar{\nu}_\tau$ signal feed-down	3	8.10
	$D^0 \ell^- \bar{\nu}_\ell$ background	35	8.11
	$D^{*0} \ell^- \bar{\nu}_\ell$ feed-down	50	8.12
	$D^{**} \ell^- \bar{\nu}_\ell$ feed-down	5	8.13
	Combinatoric: charge crossfeed	2	8.36
	Other combinatoric background	2	8.38
D^{*-}	$D^{*-} \tau^- \bar{\nu}_\tau$ signal	5	8.14
	$D^- \tau^- \bar{\nu}_\tau$ signal feed-up	0.5	8.15
	$D^{*-} \ell^- \bar{\nu}_\ell$ background	90	8.16
	$D^- \ell^- \bar{\nu}_\ell$ feed-up	2	8.17
	$D^{**} \ell^- \bar{\nu}_\ell$ feed-down	3	8.18
	Combinatoric: charge crossfeed	0.5	8.37
	Other combinatoric background	1	8.38
D^-	$D^- \tau^- \bar{\nu}_\tau$ signal	4	8.19
	$D^{*-} \tau^- \bar{\nu}_\tau$ signal feed-down	3	8.20
	$D^- \ell^- \bar{\nu}_\ell$ background	35	8.21
	$D^{*-} \ell^- \bar{\nu}_\ell$ feed-down	50	8.22
	$D^{**} \ell^- \bar{\nu}_\ell$ feed-down	5	8.23
	Combinatoric: charge crossfeed	2	8.36
	Other combinatoric background	2	8.38

Table 8.2: Components of the signal extraction fit in the D^{**} control sample channels, and their approximate abundances in generic MC. The structure of the fit is identical between the B^- and \bar{B}^0 channels. There are five components in each of the four D^{**} control sample channels. The last column gives the figures where the fitted $m_{\text{miss}}^2 - p_\ell^*$ distributions, as described in the following sections, can be found.

Channel	Source	Abundance in generic MC (%)	Figure
$D^{*0}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	65	8.24
	$D^{*0}\ell^-\bar{\nu}_\ell$ feed-up	18	8.25
	$D^0\ell^-\bar{\nu}_\ell$ feed-up	2	8.26
	Combinatoric: charge crossfeed	1	8.39
	Other combinatoric background	15	8.40
$D^0\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	75	8.27
	$D^{*0}\ell^-\bar{\nu}_\ell$ feed-up	7	8.28
	$D^0\ell^-\bar{\nu}_\ell$ feed-up	3	8.29
	Combinatoric: charge crossfeed	1	8.39
	Other combinatoric background	13	8.40
$D^{*-}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	65	8.30
	$D^{*-}\ell^-\bar{\nu}_\ell$ feed-up	18	8.31
	$D^-\ell^-\bar{\nu}_\ell$ feed-up	2	8.32
	Combinatoric: charge crossfeed	1	8.39
	Other combinatoric background	15	8.40
$D^-\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	75	8.33
	$D^{*-}\ell^-\bar{\nu}_\ell$ feed-up	7	8.34
	$D^-\ell^-\bar{\nu}_\ell$ feed-up	3	8.35
	Combinatoric: charge crossfeed	1	8.39
	Other combinatoric background	13	8.40

We use two components in the fit to model combinatoric backgrounds. The first, charge crossfeed background, consists of $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$ events in which both the B_{tag} and the signal $D^{(*)}$ are misreconstructed with the wrong charge, due to the swapping of a charged pion between the two B mesons in the event. The second component is used to model all remaining combinatoric background sources, including events with fake leptons or misreconstructed B_{tag} candidates. The precise composition of this background component will be discussed in more detail in Section 12.5. The main reason for this distinction is that the charge crossfeed events peak in m_{miss}^2 —the tag and recoil sides have swapped particles, but the total momentum is still close to the correct value. The m_{miss}^2 distribution for the remaining combinatoric events is very broad by comparison (see Figures 8.36, 8.37, and 8.38).

The four D^{**} control sample channels are structured similarly. In each, there is a “signal” component modeling correctly-reconstructed $B \rightarrow D^{**}\ell^{-}\bar{\nu}_\ell$ decays.² Two components are used to model the feed-up, one from $D\ell^{-}\bar{\nu}_\ell$ and the other from $D^*\ell^{-}\bar{\nu}_\ell$ decays (the small feed-up contributions from $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_\tau$ decays are included in these components, as described in Section 8.5).

As in the signal channels, we use two components to model the combinatoric

²The lack of $D^{**}\tau^{-}\bar{\nu}_\tau$ MC affects the D^{**} control samples in the same way as the feed-down into the signal channels. This will be discussed in Section 8.5 as well.

background in the D^{**} control samples, one for charge crossfeed events and a second for all other sources of combinatoric background.

8.3 p_ℓ^* Distribution for Semileptonic and Semitauonic Events

The functional form of the two-dimensional $p_\ell^*-m_{\text{miss}}^2$ PDF is complicated by the fact that the two variables are, in some cases, highly correlated. We will see, however, that we are able to factorize the problem somewhat, by modeling the m_{miss}^2 dimension as a purely “resolution” dimension,³ with a resolution function that depends on p_ℓ^* . This factorization is motivated by the two-dimensional behavior seen in the next Section (Figures 8.2 and 8.3, for example). Note also that this choice does not imply any loss of generality; it simplifies the form of the p_ℓ^* model, while incorporating all of the two-dimensional correlations into the m_{miss}^2 resolution function.

³For channels in which the only missing particle is a single neutrino, the true m_{miss}^2 spectrum is a delta function centered at zero (assuming massless neutrinos), and the observed m_{miss}^2 distribution is a pure resolution function. For channels with multiple missing particles, the observed m_{miss}^2 distribution is the convolution of the physical m_{miss}^2 spectrum with our detector resolution. Our model of the m_{miss}^2 distribution will be used to describe both of these physical cases.

We present first a model of the lepton spectrum. In Section 8.4, we show how we extend the one-dimensional spectrum to create a two-dimensional PDF.

Figure 8.1 shows the p_ℓ^* distribution for one signal component, $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$, and three semileptonic components, $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$, $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0$, and $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0$, taken from signal and semileptonic MC. These distributions are representative of the spectra in all components, which have qualitatively similar features: a spectrum that rises from zero at $p_\ell^* \sim 200$ MeV/ c to a broad maximum, and then falls back to zero at a kinematic endpoint between 2 and 2.3 GeV/ c .

We use a generalized form of a Gaussian to model the p_ℓ^* distribution. The Gaussian distribution,

$$\mathcal{G}(p_\ell^*) \propto \exp\left(-\left|\frac{p_\ell^* - p_0}{2\sigma}\right|^2\right), \quad (8.2)$$

has the same general properties as our distributions — it rises smoothly from zero to a peak and then falls smoothly back to zero again. Here, p_0 represents the value of p_ℓ^* for which \mathcal{G} peaks and σ represents the width of the Gaussian distribution. This gross agreement is not enough, however, so we modify the form slightly to allow for greater flexibility. We therefore define a modified

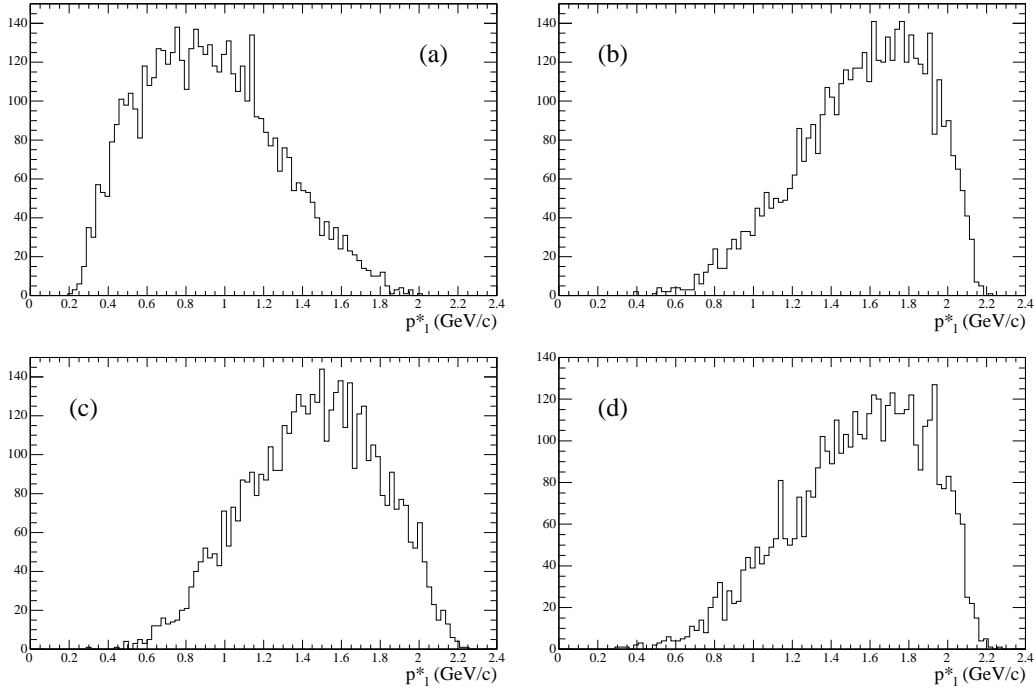


Figure 8.1: p_ℓ^* distributions for (a) $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$, (b) $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^{*0}$, (c) $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0$, and (d) $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0$, taken from signal and semileptonic MC. In (a), ℓ refers, as always, to the secondary lepton from $\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\tau$, which is why the spectrum is clearly softer than other modes. Note, however, that the shapes in all cases are qualitatively similar.

Gaussian function,

$$\mathcal{H}(p_\ell^*) \propto \exp \left(- \left| \frac{p_\ell^* - p_0}{\sigma(p_\ell^*)} \right|^{\nu(p_\ell^*)} \right), \quad (8.3)$$

where, for convenience, we have absorbed the constant factor of 2 into the definition of $\sigma(p_\ell^*)$.

By allowing the width and exponent of the Gaussian to be functions of p_ℓ^* , we are able to describe a much greater variety of shapes. Specifically, we parameterize $\nu(p_\ell^*)$ as a linear function,

$$\nu(p_\ell^*) = \nu_L + \frac{\nu_H - \nu_L}{2.4 \text{ GeV}/c} \cdot p_\ell^* \quad (8.4)$$

where ν_L and ν_H are the values of the exponential term at the low and high endpoints of p_ℓ^* , fixed at zero and 2.4 GeV/c, respectively. Similarly, we parameterize $\sigma(p_\ell^*)$ as a bilinear function,

$$\sigma(p_\ell^*) = \begin{cases} \sigma_L + \frac{\sigma_0 - \sigma_L}{p_0} \cdot p_\ell^* & p_\ell^* < p_0 \\ \sigma_0 + \frac{\sigma_H - \sigma_0}{2.4 \text{ GeV}/c - p_0} \cdot (p_\ell^* - p_0) & p_\ell^* > p_0 \end{cases}, \quad (8.5)$$

where σ_L , σ_0 , and σ_H represent the width of the Gaussian at $p_\ell^* = 0$, $p_\ell^* = p_0$, and $p_\ell^* = 2.4 \text{ GeV}/c$, respectively. Even though this parameterization is discontinuous at the point $p_\ell^* = p_0$, the resulting function $\mathcal{H}(p_\ell^*)$ remains smooth since the numerator in the exponent, $(p_\ell^* - p_0)$, goes to zero at the same point.

The p_ℓ^* parameterization therefore has six free parameters: p_0 , the peak; ν_L and ν_H , describing the exponential term; σ_L , σ_0 , and σ_H , describing the

width. When performing fits using this PDF, we integrate \mathcal{H} numerically to compute the normalization.

8.4 $p_\ell^* - m_{\text{miss}}^2$ PDFs for Semileptonic and Semitauonic Events

We begin constructing a complete two-dimensional PDF by looking at the dependence of m_{miss}^2 on p_ℓ^* ; profile plots for two modes, $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$ and $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^0$, are shown in Figures 8.2 and 8.3. Both the mean value and spread of m_{miss}^2 are seen to decrease as p_ℓ^* increases; this is typical of most PDFs, although, for components in which only a single neutrino is missed (*i.e.* $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^{*0}$), the m_{miss}^2 distribution is centered at zero, more-or-less independent of p_ℓ^* .

We now construct two complete two-dimensional PDFs, $\mathcal{P}_1(p_\ell^*, m_{\text{miss}}^2)$ and $\mathcal{P}_2(p_\ell^*, m_{\text{miss}}^2)$. We will use the simpler PDF, \mathcal{P}_1 , to model most of the semileptonic and semitauonic fit components (22 out of 32). For the remaining ten components, however, the more complicated parameterization \mathcal{P}_2 is necessary to adequately describe the m_{miss}^2 tail. Eight of these components are the ones in which the only missing particle is a single neutrino,

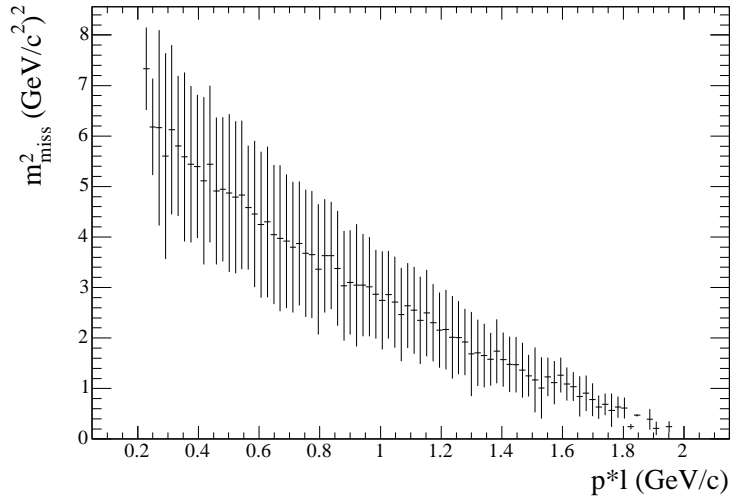


Figure 8.2: Profile plot of m^2_{miss} vs p^*_ℓ for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$, showing a decrease in the both mean value and spread in m^2_{miss} as p^*_ℓ increases.

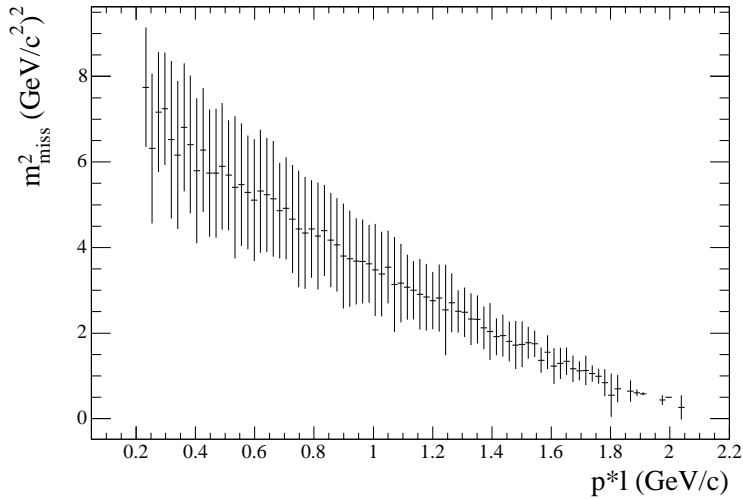


Figure 8.3: Profile plot of m^2_{miss} vs p^*_ℓ for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0$, showing a decrease in the both mean value and spread in m^2_{miss} as p^*_ℓ increases.

$$\begin{aligned}
 D^{*0} \ell^- \bar{\nu}_\ell &\Rightarrow D^{*0} & D^0 \ell^- \bar{\nu}_\ell &\Rightarrow D^0 \\
 D^{*+} \ell^- \bar{\nu}_\ell &\Rightarrow D^{*+} & D^+ \ell^- \bar{\nu}_\ell &\Rightarrow D^+ \\
 D^{**} \ell^- \bar{\nu}_\ell &\Rightarrow D^{*0} \pi^0 & D^{**} \ell^- \bar{\nu}_\ell &\Rightarrow D^0 \pi^0 \\
 D^{**} \ell^- \bar{\nu}_\ell &\Rightarrow D^{*+} \pi^0 & D^{**} \ell^- \bar{\nu}_\ell &\Rightarrow D^+ \pi^0 ,
 \end{aligned}$$

and the remaining two are components in which a single neutrino and a soft π^0 or γ are missing,

$$D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0 \quad D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^+ .$$

Using the model of the lepton spectrum $\mathcal{H}(p_\ell^*)$ introduced above, we construct the PDFs as:

$$\begin{aligned}
 \mathcal{P}_1(p_\ell^*, m_{\text{miss}}^2) &\propto \mathcal{H}(p_\ell^*) \times \\
 &\quad [f_1(p_\ell^*) \mathcal{G}_1(p_\ell^*; m_{\text{miss}}^2) + (1 - f_1(p_\ell^*)) \mathcal{G}_2(p_\ell^*; m_{\text{miss}}^2)]
 \end{aligned}$$

and

$$\begin{aligned}
 \mathcal{P}_2(p_\ell^*, m_{\text{miss}}^2) &\propto \mathcal{H}(p_\ell^*) \times \\
 &\quad [f_1(p_\ell^*) \mathcal{G}_1(p_\ell^*; m_{\text{miss}}^2) + f_2(p_\ell^*) \mathcal{BG}(p_\ell^*; m_{\text{miss}}^2) + \\
 &\quad (1 - f_1(p_\ell^*) - f_2(p_\ell^*)) \mathcal{G}_2(p_\ell^*; m_{\text{miss}}^2)] .
 \end{aligned}$$

Here, the functions \mathcal{G}_i and \mathcal{BG} are Gaussians and Bifurcated Gaussians (Gaussian with different σ parameters on either side of the mean), respectively; all are functions of m_{miss}^2 , with parameters dependent on p_ℓ^* .

Table 8.3: p_ℓ^* dependence of the m_{miss}^2 PDF parameterizaion. The form of f_2 is chosen to allow the \mathcal{BG} term to contribute at low p_ℓ^* , but to drive this term rapidly to zero as p_ℓ^* increases. The form of σ_H is chosen to allow for a long tail towards high m_{miss}^2 at low p_ℓ^* , but to drive this term rapidly to zero as p_ℓ^* increases (note that there is no problem having σ approach zero since the amplitude of this term goes to zero as well; the result is finite and well-behaved).

Function	Parameter	Dependence on p_ℓ^*	Number of free parameters
\mathcal{G}_i	mean	quadratic	3
\mathcal{G}_i	σ	linear	2
\mathcal{BG}	mean	constant	1
\mathcal{BG}	σ_L	constant	1
\mathcal{BG}	σ_H	$\sigma_{H0} \cdot \left[1 - \left(\frac{p_\ell^*}{2.4 \text{ GeV}/c} \right)^\phi \right]$	2
\mathcal{P}_i	f_1	linear	2
\mathcal{P}_2	f_2	$F \cdot \left(\frac{2.4 \text{ GeV}/c - p_\ell^*}{2.2 \text{ GeV}/c} \right)^\theta$	2

The p_ℓ^* dependence of the various parameters of \mathcal{G}_i and \mathcal{BG} is listed in Table 8.3. The total number of free parameters for \mathcal{P}_1 is eighteen: six for $\mathcal{H}(p_\ell^*)$, five each for \mathcal{G}_1 and \mathcal{G}_2 , and two for f_1 . The total number of free parameters for \mathcal{P}_2 is 24: six for $\mathcal{H}(p_\ell^*)$, five each for \mathcal{G}_1 and \mathcal{G}_2 , four for \mathcal{BG} , and two each for f_1 and f_2 .

8.5 PDFs from Signal and Semileptonic MC

We extract each shape from a fit to a MC sample. For most shapes, we simply select true $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ events from signal and semileptonic MC reconstructed in one of the eight channels, as appropriate. For the $D^{**}\ell^-\bar{\nu}_\ell$ PDFs, however, we fit a weighted mixture of two or more MC samples, as described in Sections 8.5.1, 8.5.2, and 8.5.3.

These shapes are appropriate for subsequent fitting of *BABAR* MC event samples. For fitting to data, we need to correct these shapes for known data-MC differences. To do this, we apply event-by-event reweighting to the MC samples during the shape fits; this reweighting is described in Section 11.

8.5.1 Reweighting $D^{(*)}\tau^-\bar{\nu}_\tau$ Feed-up into the D^{**} Control Samples

Each of the D^{**} control samples contain two feed-up components, one from true $D\ell^-\bar{\nu}_\ell$ events and one from $D^*\ell^-\bar{\nu}_\ell$. We define this component to include feed-up from both $D^{(*)}\ell^-\bar{\nu}_\ell$ events with light leptons and feed-up from $D^{(*)}\tau^-\bar{\nu}_\tau$ events.

We generate fit samples for these PDFs by mixing MC events from signal $D^{(*)}\tau^-\bar{\nu}_\tau$ and semileptonic $D^{(*)}\ell^-\bar{\nu}_\ell$ decays. In these samples, the signal events

receive weights

$$\begin{aligned}
 w_{D\tau^-\bar{\nu}_\tau \Rightarrow D\pi^0} &= R_{\text{SM}} \cdot \mathcal{B}(\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\tau) = 0.061 \\
 w_{D^*\tau^-\bar{\nu}_\tau \Rightarrow D^*\pi^0} &= R_{\text{SM}}^* \cdot \mathcal{B}(\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\tau) = 0.041 \ , \quad (8.6)
 \end{aligned}$$

where $R_{\text{SM}}^{(*)}$ is the standard model expectation of the ratio of branching ratios $\mathcal{B}(B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell)$. Feed-up events from the light leptons enter the MC sample without reweighting,

$$w_{D^{(*)}\ell^-\bar{\nu}_\ell \Rightarrow D^{(*)}\pi^0} = 1 \ . \quad (8.7)$$

Note that signal events are only expected to contribute 4–6% of the feed-up (assuming equal efficiencies for feed-up for signal and semileptonic events), which itself makes up less than 20% of the $D^*\pi^0$ control samples and less than 10% of the $D\pi^0$ control samples. If we assume the standard model signal branching fractions are correct, we expect to see 1.8 ± 0.6 signal feed-up events in all four D^{**} control samples together in the Run 1–4 data sample.

Note that this procedure explicitly introduces an assumption about the branching fractions of the signal modes. Fortunately, this assumption is only relevant at higher order — it only affects the feed-up into the D^{**} control samples — and it applies to a small number, so this is not expected to dramatically bias our signal extraction.

One possible way to reduce this bias, if necessary, would be to iterate the analysis. We use the standard model assumptions to make a first measurement, then we take our first results as an improved input to the reweighting procedure and repeat the analysis. This procedure would quickly converge to an unbiased measurement.

A second possible solution would be to introduce terms in the fit dedicated to modeling $D^{(*)}\tau^-\bar{\nu}_\tau$ feed-up, instead of combining these terms with $D^{(*)}\ell^-\bar{\nu}_\ell$. We would introduce feed-up constraints on these terms as we do for other feed-up components (see Section 8.8), completely removing any input assumptions on the branching fractions. It is difficult, however, to imagine introducing components in the fit in which the expected number of events in data is less than one event per channel.

8.5.2 Reweighting to $D^{(*)}\tau^-\bar{\nu}_\tau$ to Simulate $D^{**}\tau^-\bar{\nu}_\tau$

All of the D^{**} components in the fit (both in the D^{**} control samples and in the signal channels) are defined to model the contribution from light-lepton $B \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ decays as well as $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ decays. We expect, based on the current *BABAR* MC model, that 3.5–4.5% of all reconstructed D^{**} decays are $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$, corresponding to approximately 14 events over all eight channels. Again, because of the small number of expected events, we choose

to include all $D^{**}(\ell^-/\tau^-)\bar{\nu}$ decays in a single model, rather than introducing separate $D^{**}\tau^-\bar{\nu}_\tau$ components.

Because we do not have dedicated $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ Monte Carlo samples, we use our signal MC for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ to simulate $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ decays. Doing this, we assume that the p_ℓ^* and m_{miss}^2 distributions of $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ events are dominated by the $\tau^- \rightarrow \ell^-\bar{\nu}_\ell\nu_\tau$ decay, and are not extremely sensitive to the quantum numbers of the D^{**} system.

We generate fit samples for these PDFs by mixing MC events from signal $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ and semileptonic $B \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ decays. The semileptonic events receive weights of 1.0, while the weights for signal events are chosen so that the resulting abundances of signal events (separately for $D\tau^-\bar{\nu}_\tau$ and $D^*\tau^-\bar{\nu}_\tau$) match the abundances seen in generic MC. These weights are shown in Table 8.4.

8.5.3 Generically-simulated $D^{**}\ell^-\bar{\nu}_\ell$ Decays

A non-negligible fraction of $D^{**}\ell^-\bar{\nu}_\ell$ events reconstructed in the four signal channels in generic MC is found to come from “unsimulated” D meson decays, that is, D decays which are not present in the signal MC samples. The D mesons are therefore misreconstructed; the most common decay chain is $B^- \rightarrow D^{**0}\ell^-\bar{\nu}_\ell \rightarrow D^+\pi^-\ell^-\bar{\nu}_\ell$, where the true D^+ meson loses *two* charged pions and

Table 8.4: Weights used for simulating $D^{**}\tau^-\bar{\nu}_\tau$ with signal MC. The weights used for simulating the D^{**} control sample are much larger than those used in the signal channels; this is due to the lower number of reconstructed events feeding up into the control samples.

Signal process	Used to simulate	w
$D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^0$	0.0017
$D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^0$	0.0035
$D^0\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	0.05
$D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$	0.008
$D^+\tau^-\bar{\nu}_\tau \Rightarrow D^+$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^+$	0.0017
$D^{*+}\tau^-\bar{\nu}_\tau \Rightarrow D^+$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^+$	0.0037
$D^+\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}$	0.062
$D^{*+}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}$	0.0013
$D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^0\pi^0$	1.0
$D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^0\pi^0$	1.0
$D^0\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\pi^0$	1.0
$D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\pi^0$	0.6
$D^+\tau^-\bar{\nu}_\tau \Rightarrow D^+\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^+\pi^0$	0.38
$D^{*+}\tau^-\bar{\nu}_\tau \Rightarrow D^+\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^+\pi^0$	0.35
$D^+\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}\pi^0$	1.0
$D^{*+}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}\pi^0$	$D^{**}\tau^-\bar{\nu}_\tau \Rightarrow D^{*+}\pi^0$	0.35

a fake D^0 meson is reconstructed by combining the π^- from D^{**} decay with the remainder of the D^+ .

Note that, in most cases, when a D meson loses particles, the invariant mass of the remaining particles is far below the true D mass, and this candidate will correctly be rejected by our event selection. The only way for this candidate to pass our selection and fake a D is by combining with another particle in the event (with approximately the same momentum as the lost particles); because of our total-event reconstruction, however, there are typically no additional particles left in the event for this combinatoric process to take place. Only in D^{**} decays, where the transition pion from $D^{**} \rightarrow D^{(*)}\pi$ provides the additional particle, is this process likely to occur. Indeed, less than 1% of $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ events in generic MC are generated with the D decaying in an unsimulated mode; these events are evenly distributed across all channels (as we will see, this is not the case in $D^{**}\ell^-\bar{\nu}_\ell$ events), and the m_{miss}^2 resolution in these events is comparable to that of the “normal” modes.

For $D^{**}\ell^-\bar{\nu}_\ell$ events reconstructed in the signal channels, at least one (sometimes two or three) of the transition particles has not been reconstructed, making it possible to misreconstruct an unsimulated D mode as described above. In generic MC, unsimulated D decays account for between 10% (in the D^{*+} channel) and 40% (in the D^+ channel) of reconstructed $D^{**}\ell^-\bar{\nu}_\ell$ de-

cays. In the D^{**} control sample channels, however, where an additional π^0 has been reconstructed, the combinatoric possibilities are greatly reduced, and we find that unsimulated decays account for only 7% of the $D^{**}\ell^-\bar{\nu}_\ell$ events in the control sample. The asymmetry between these numbers means that the D^{**} feed-down constraints (see Section 8.8) are significantly different between the generic and signal MC samples, and, in order to correctly calculate the constraints, these unsimulated modes need to be accounted for.

We account for these events by creating hybrid MC samples. In each of the eight analysis channels, we select $B \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ events from generic MC in which the D decays to a mode not present in the signal MC sample. We add these events to the corresponding signal MC samples. In the four signal channels, we assign a weight to the unsimulated events from generic MC by normalizing to the semileptonic decays $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$:

$$w_{\text{unsim}} = \frac{N_{D^{(*)}\ell^-\bar{\nu}_\ell, \text{signalMC}}}{N_{D^{(*)}\ell^-\bar{\nu}_\ell, \text{genericMC}}} . \quad (8.8)$$

In the four D^{**} control samples, we use the same weights as the corresponding signal channel.

Note that this procedure, adding a small number of events with large relative weights, is conceptually different from what is done in Sections 8.5.1 and 8.5.2, where we add large numbers of events with small weights. In this case,

the small number of events taken from generic MC give rise to large *apparent* fluctuations in these data samples; when we fit these samples, the resulting PDF (and its errors) correctly handle these fluctuations. These fluctuations will be seen in the fits in Figures 8.8, 8.13, 8.18, 8.23, 8.24, 8.27, 8.30, and 8.33.

8.6 Fits

We fit the signal and semileptonic MC samples described above to determine the 32 PDFs, as described in Section 8.2. Each of these shape fits is performed as an unbinned maximum likelihood fit to the appropriate MC sample.

One-dimensional projections of these fits are shown in Figures 8.4–8.35 for each of the 32 PDFs. In each figure, the full m_{miss}^2 and p_ℓ^* projections are shown, along with m_{miss}^2 projections in four slices of p_ℓ^* to demonstrate the correlation between the variables.

$$D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$$

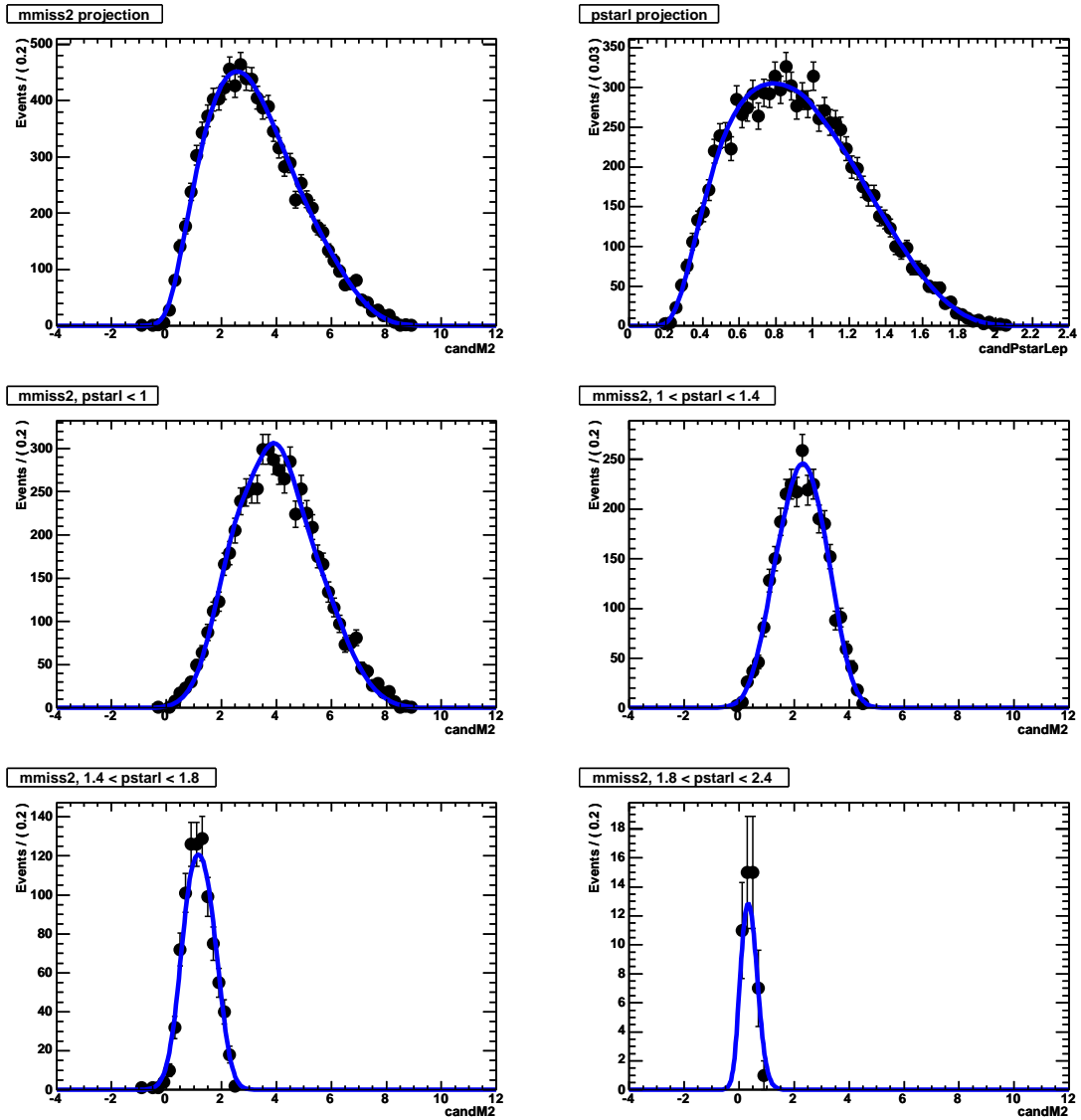


Figure 8.4: Fitted MC distribution of $p_l^* - m_{\text{miss}}^2$ for $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$.

$$D^0 \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$$

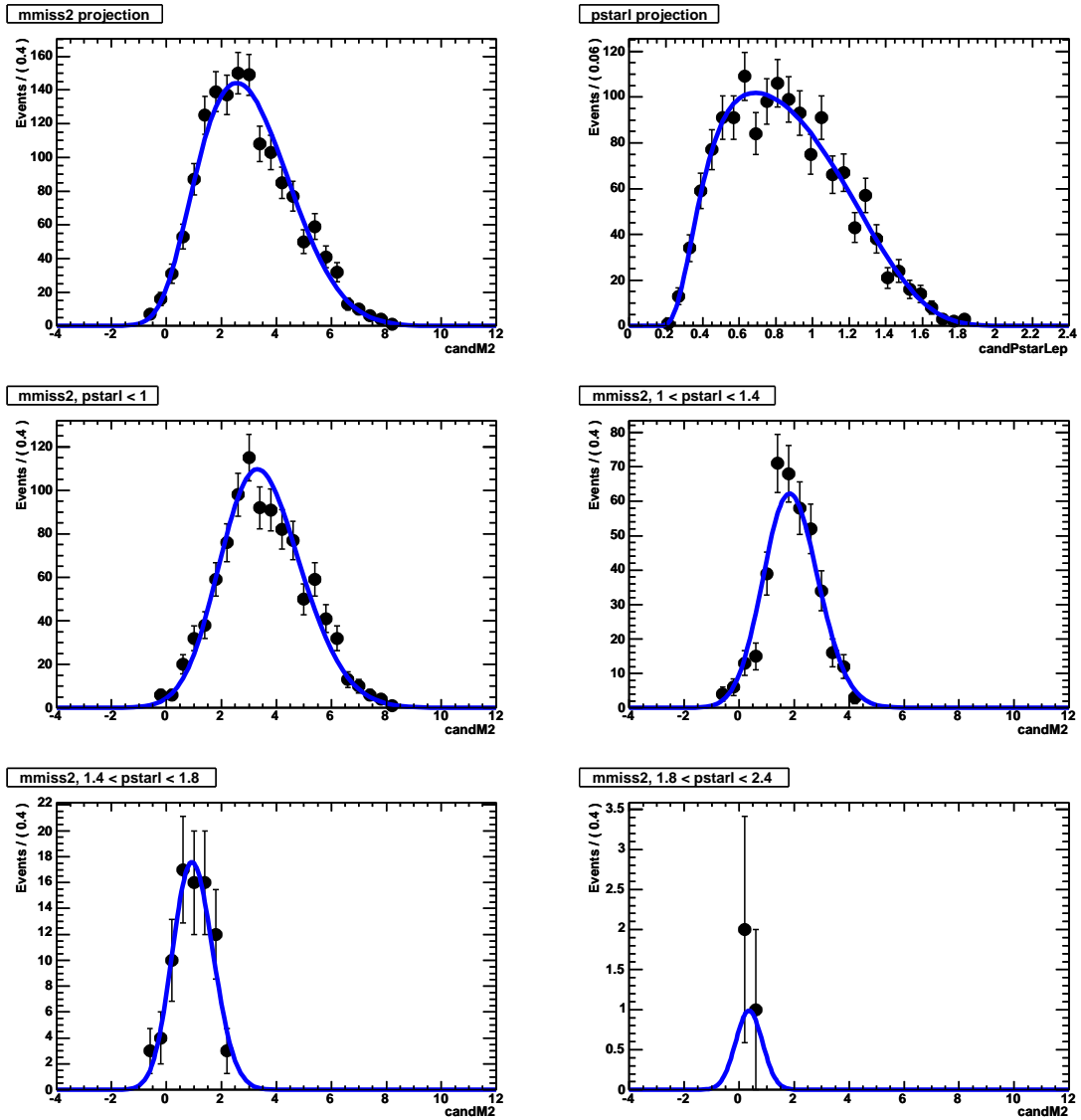


Figure 8.5: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \tau^- \bar{\nu}_\tau \Rightarrow D^{*0}$.

$$D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0}$$

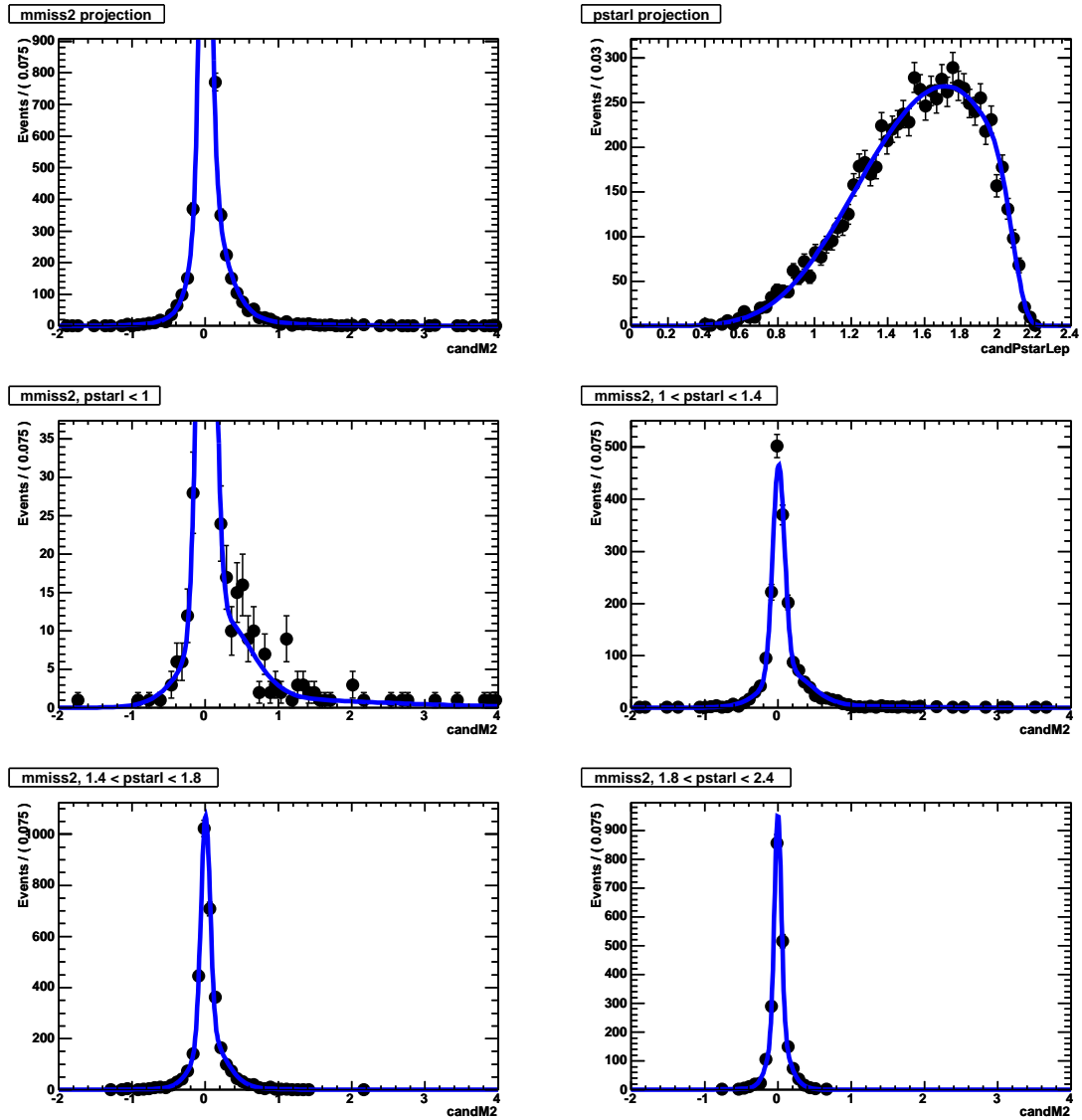


Figure 8.6: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0}$.

$$D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^{*0}$$

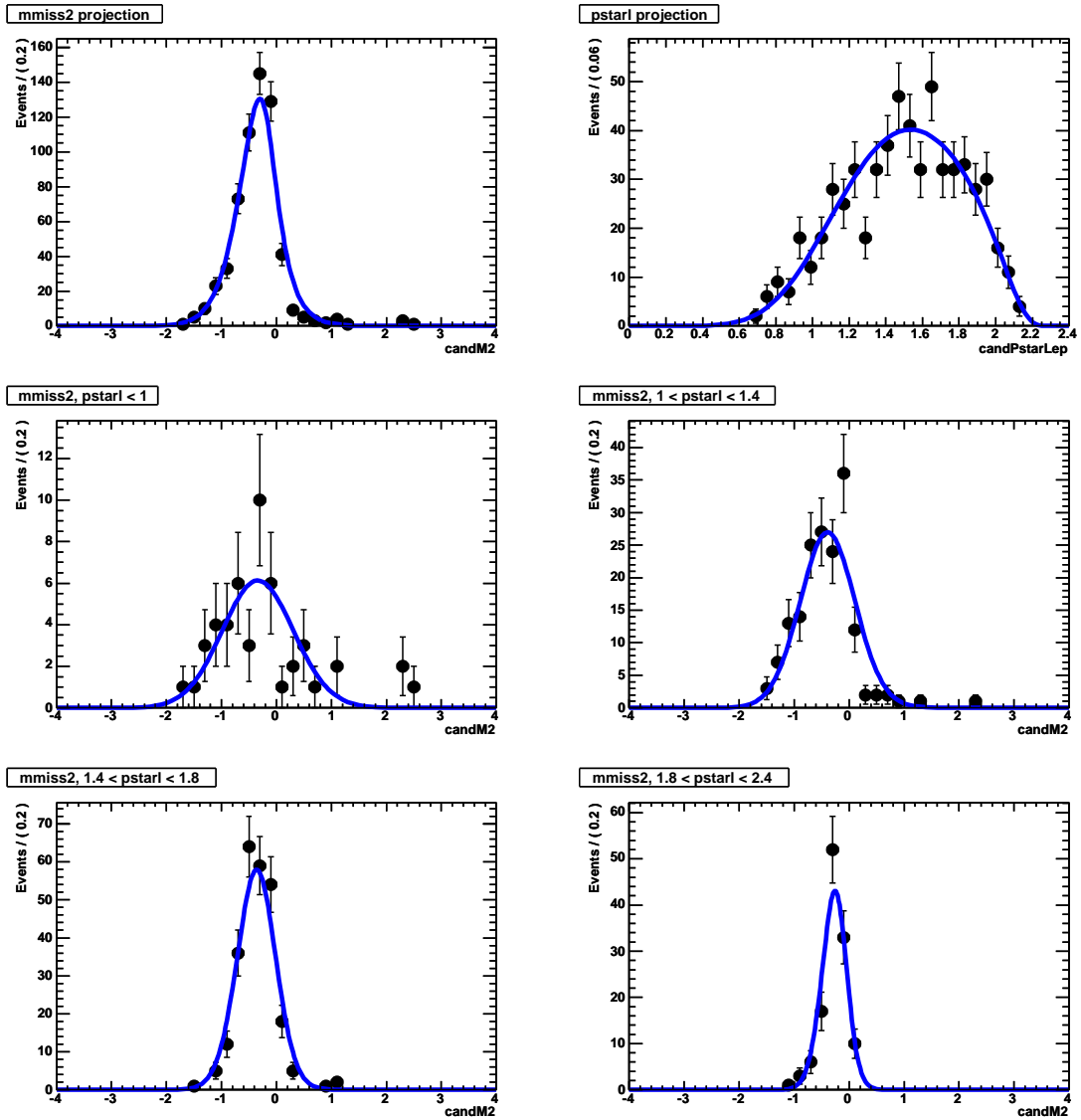


Figure 8.7: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^{*0}$.

$$D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0}$$

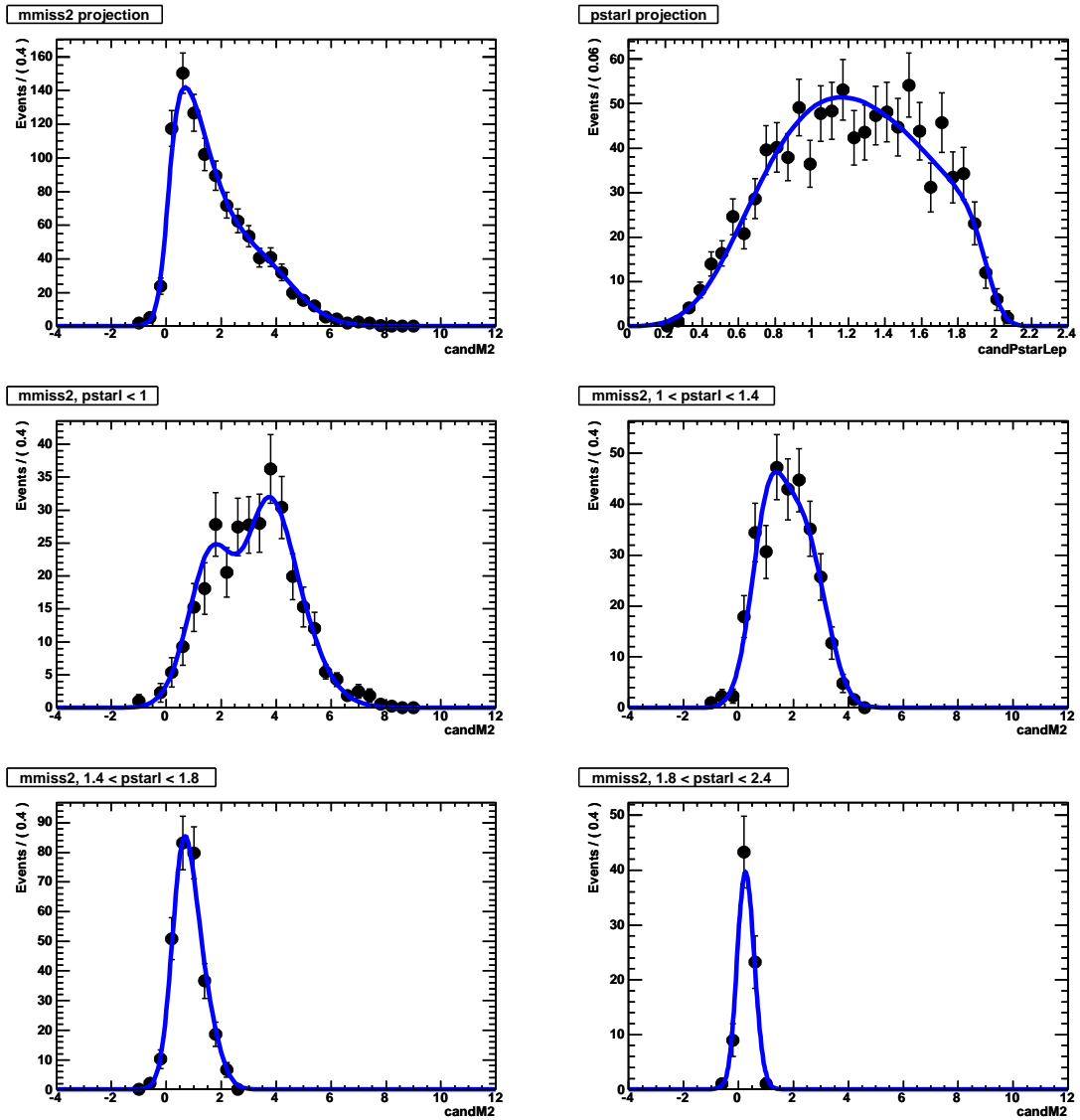


Figure 8.8: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0}$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^0 \tau^- \bar{\nu}_\tau \Rightarrow D^0$$

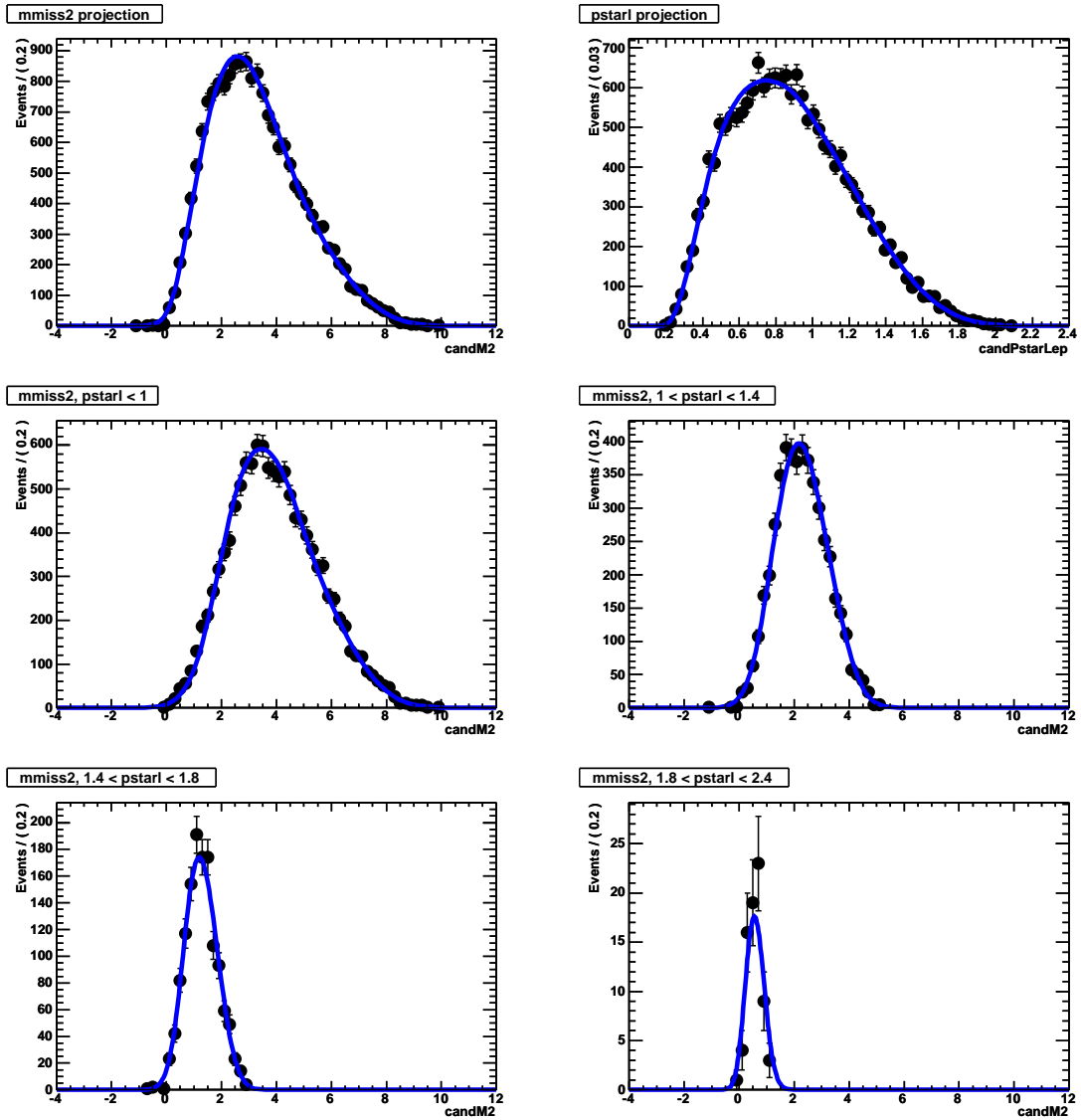


Figure 8.9: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \tau^- \bar{\nu}_\tau \Rightarrow D^0$.

$$D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^0$$

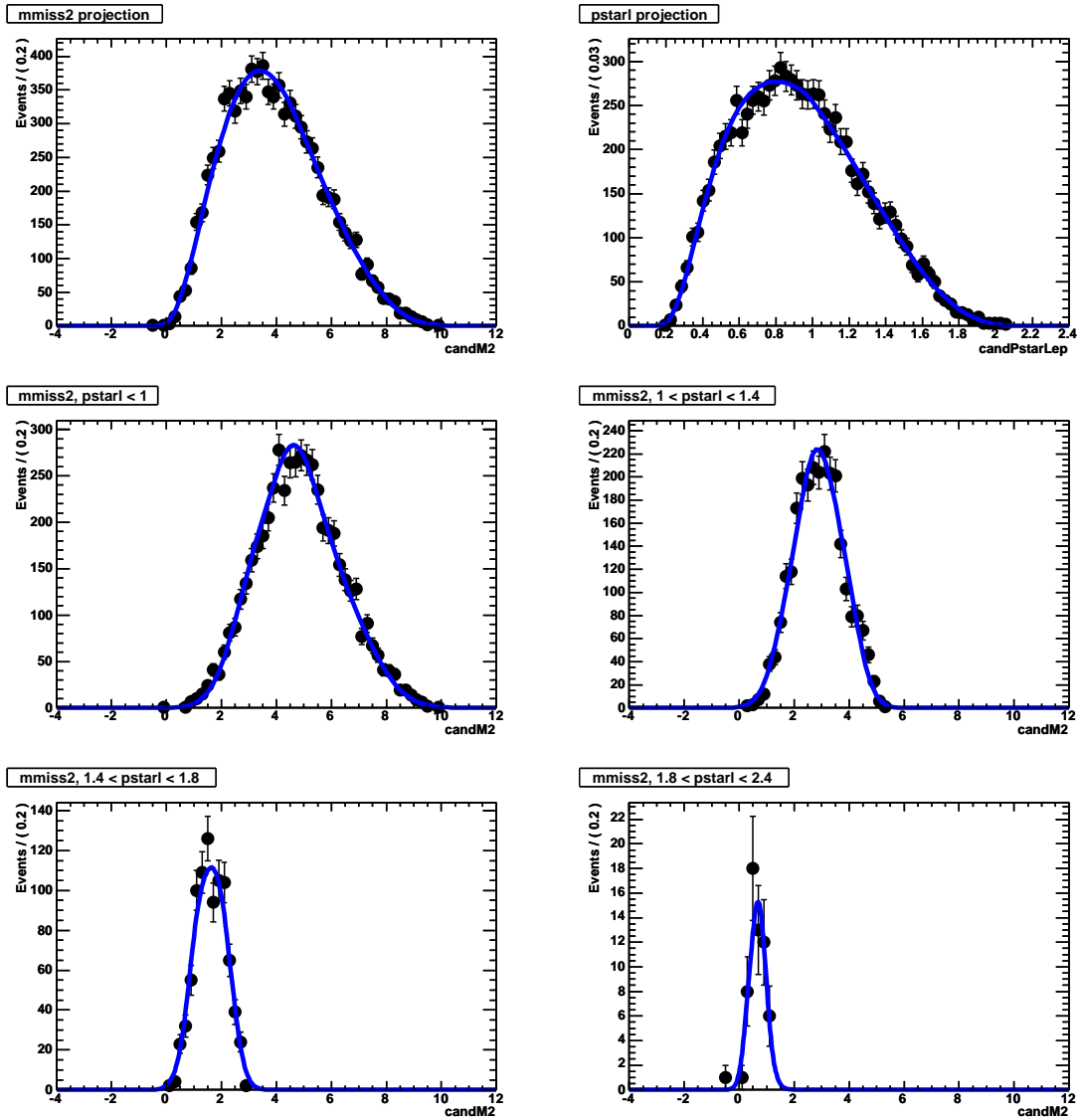


Figure 8.10: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{*0} \tau^- \bar{\nu}_\tau \Rightarrow D^0$.

$$D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0$$

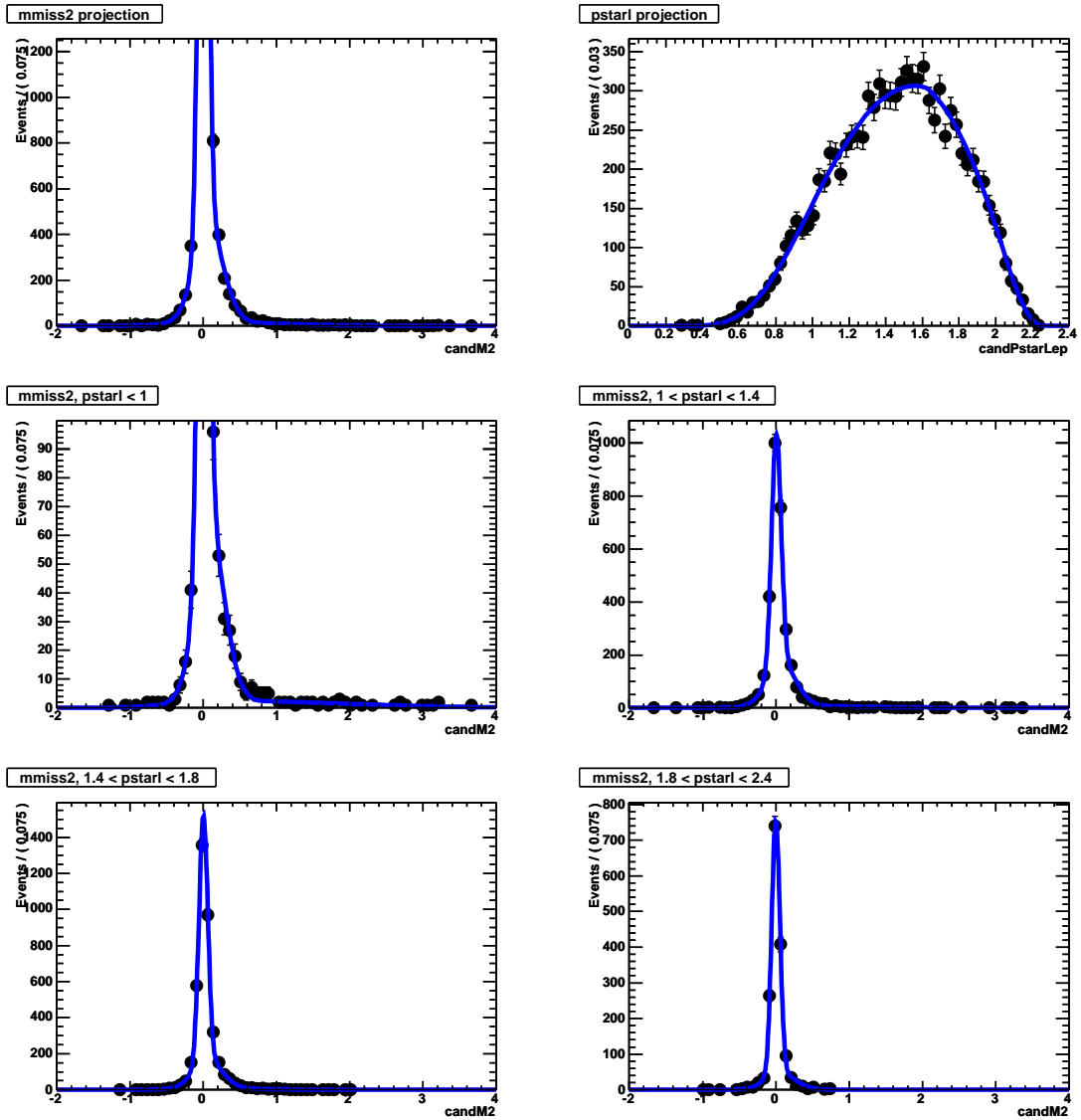


Figure 8.11: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0$.

$$D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0$$

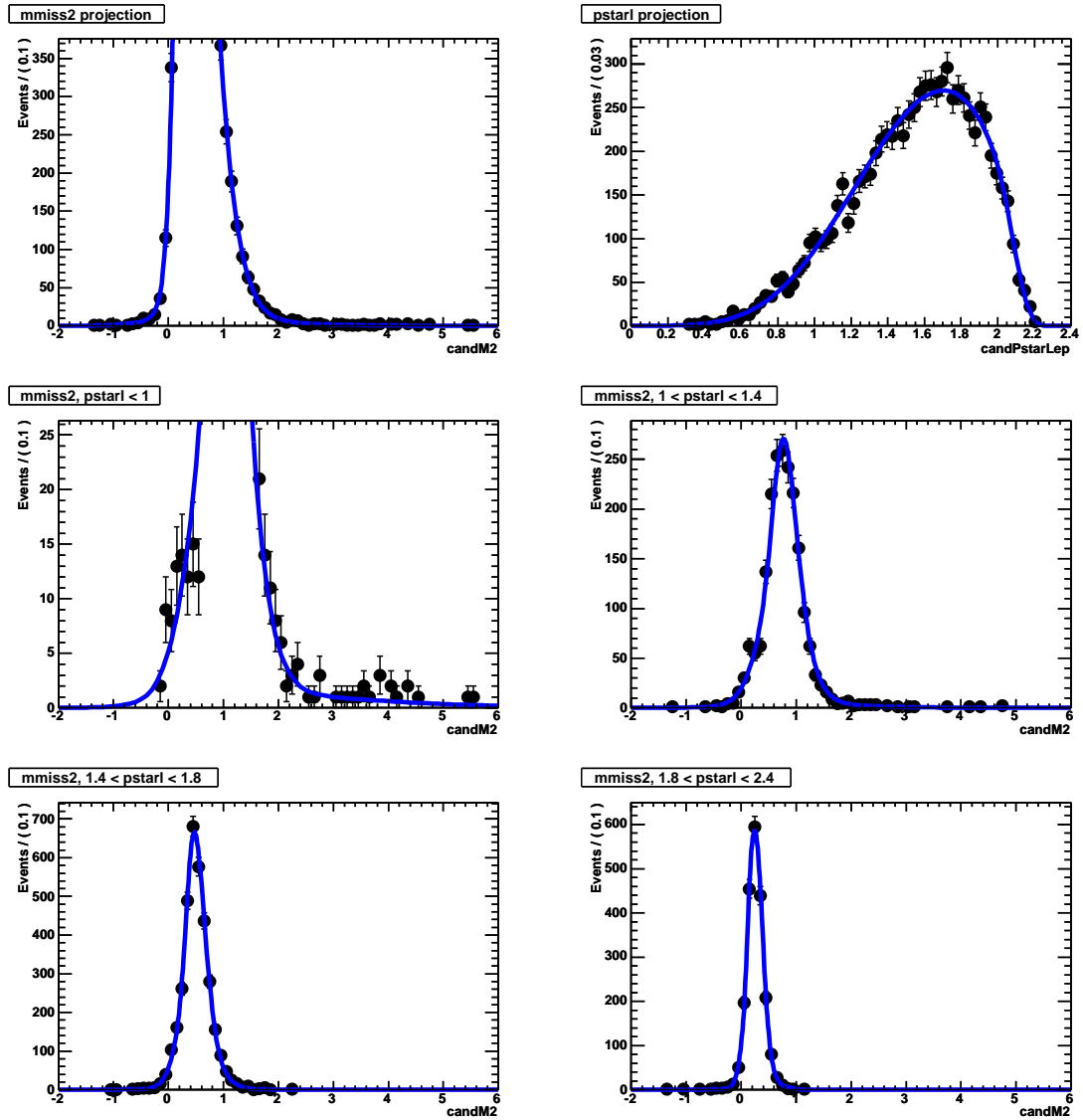


Figure 8.12: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{*0} \ell^- \bar{\nu}_\ell \Rightarrow D^0$.

$$D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0$$

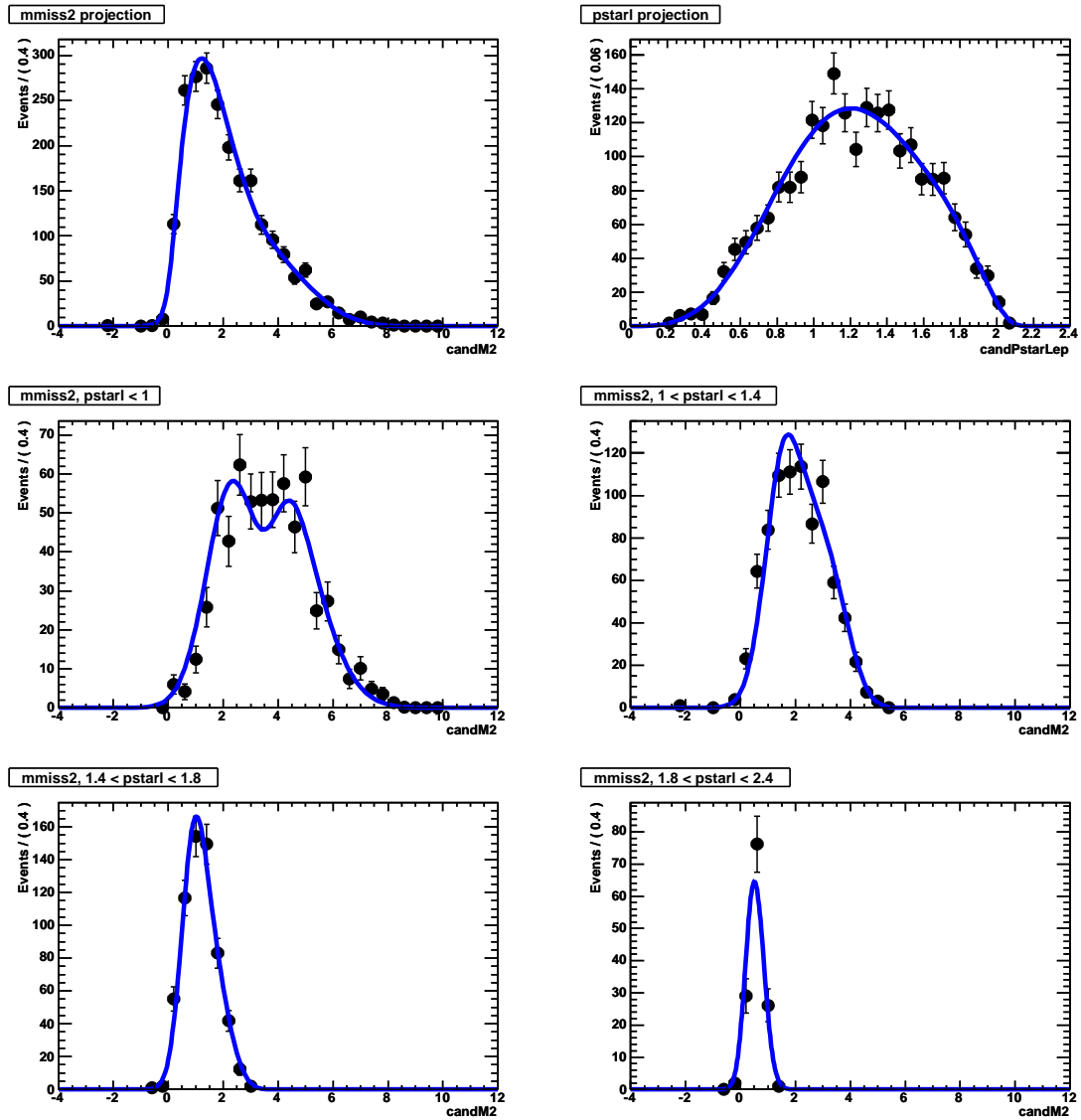


Figure 8.13: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{*+} \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$$

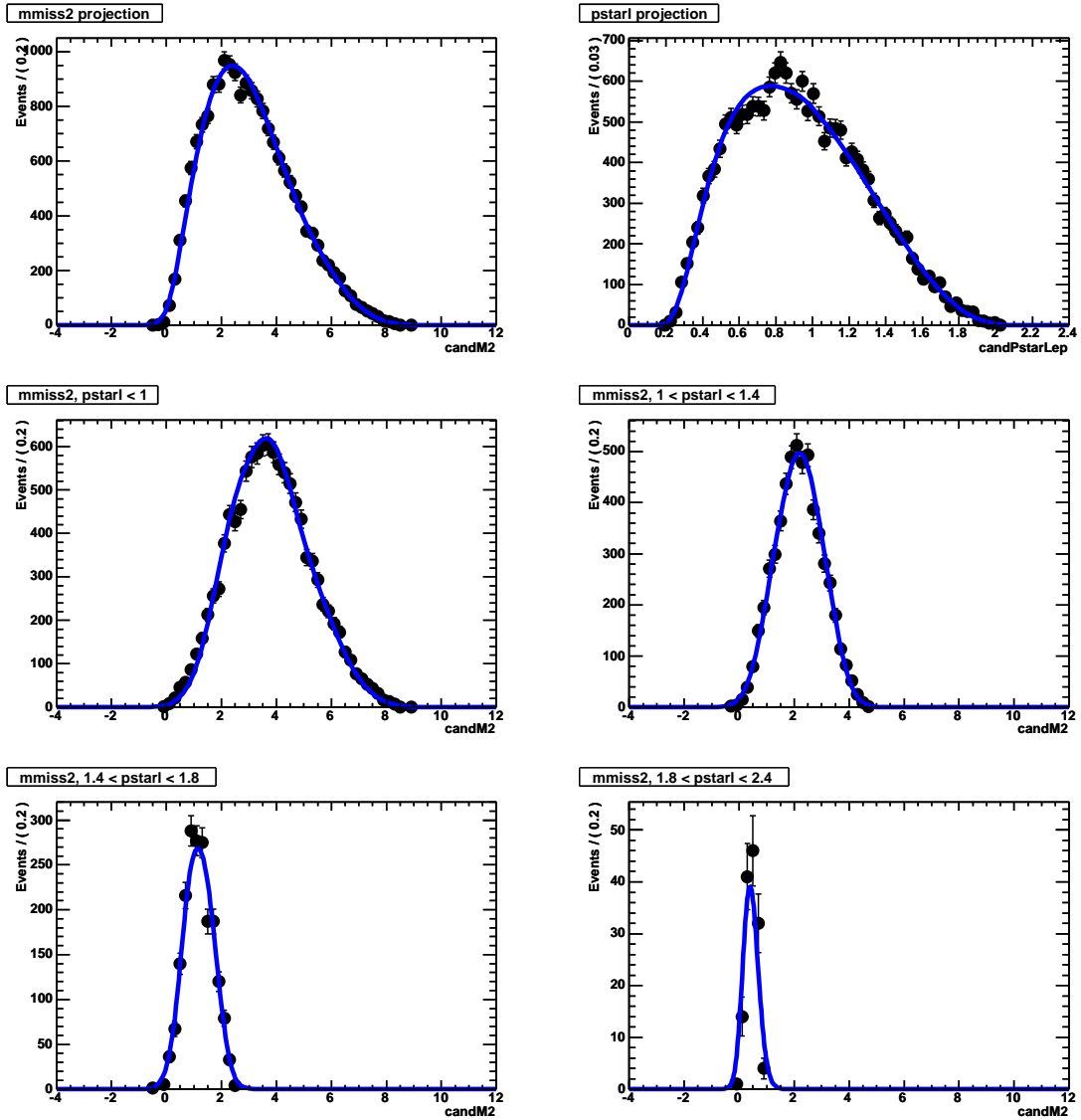


Figure 8.14: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{*+} \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$.

$$D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$$

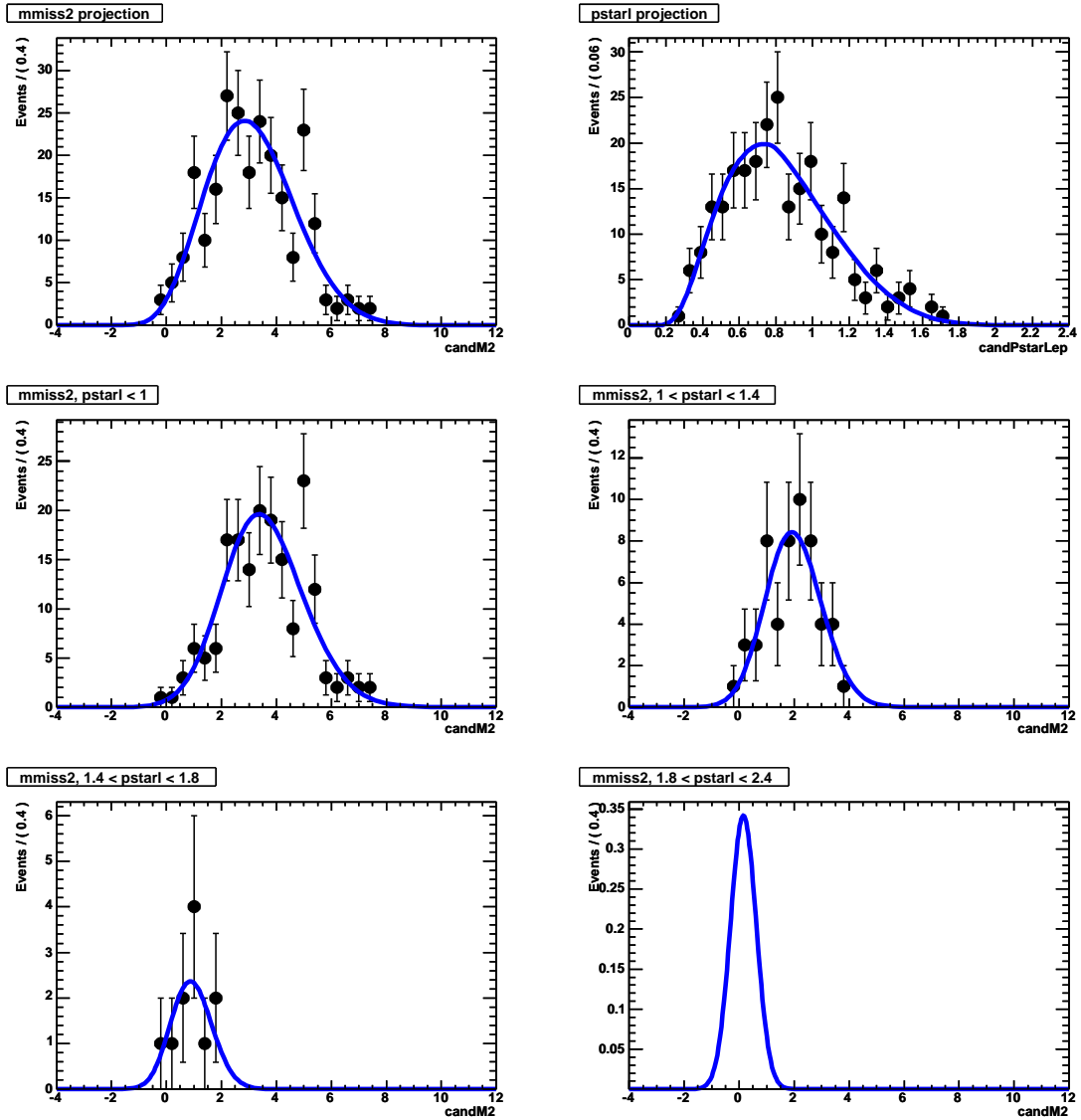


Figure 8.15: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^{*+}$.

$$D^{*+}\ell^{-}\bar{\nu}_{\ell} \Rightarrow D^{*+}$$

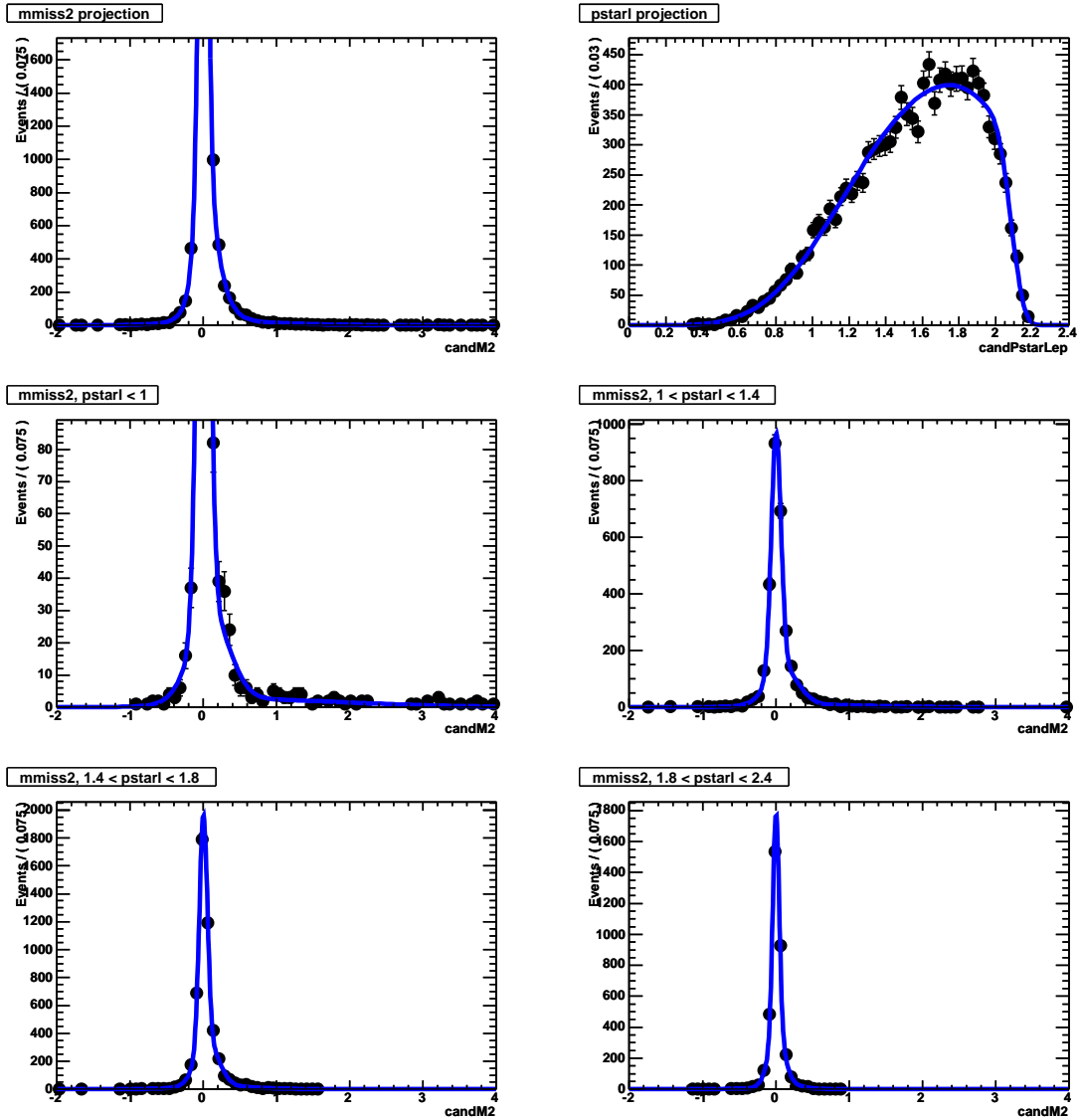


Figure 8.16: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*+}\ell^{-}\bar{\nu}_{\ell} \Rightarrow D^{*+}$.

$$D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+}$$

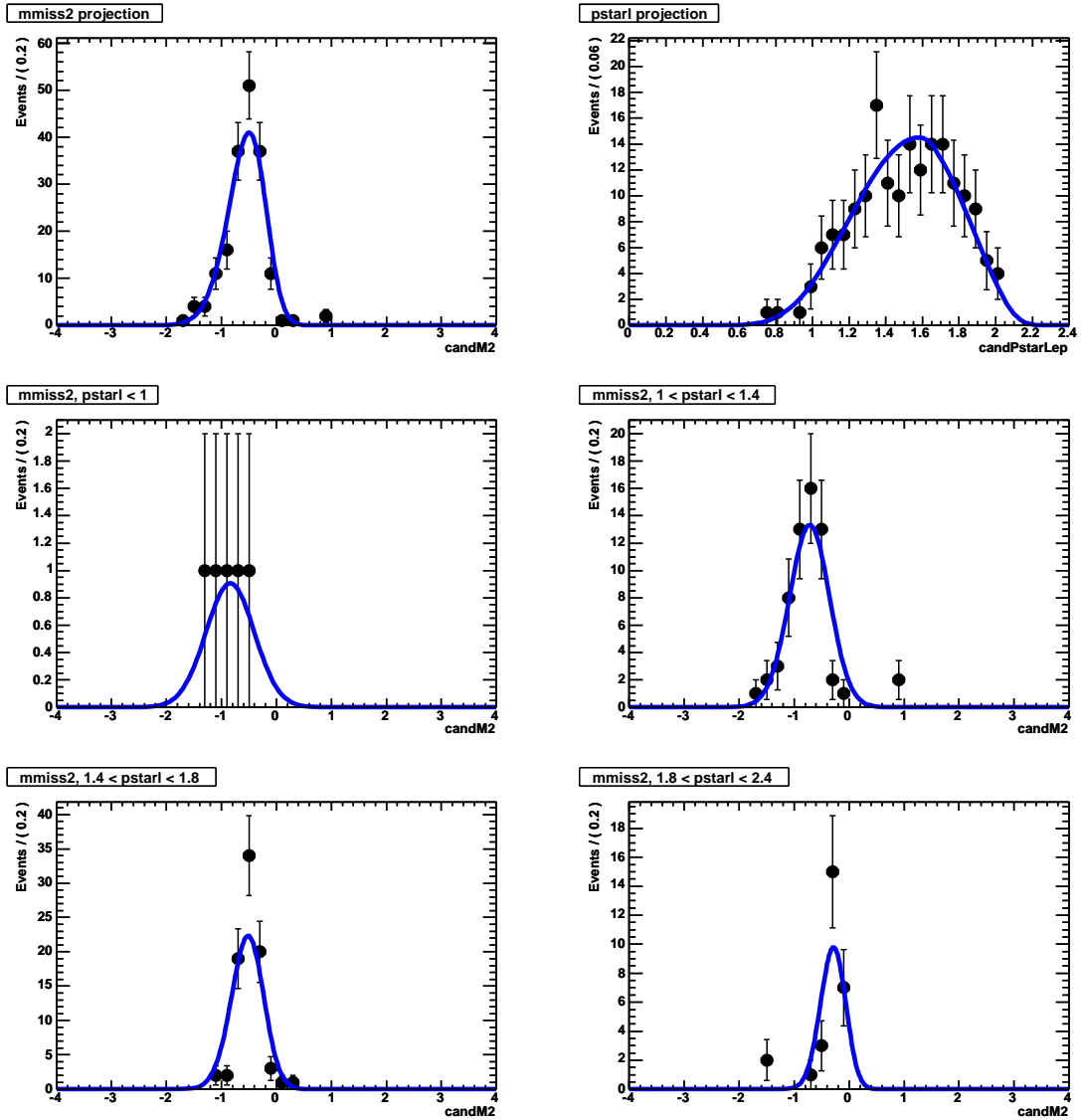


Figure 8.17: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+}$.

$$D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+}$$

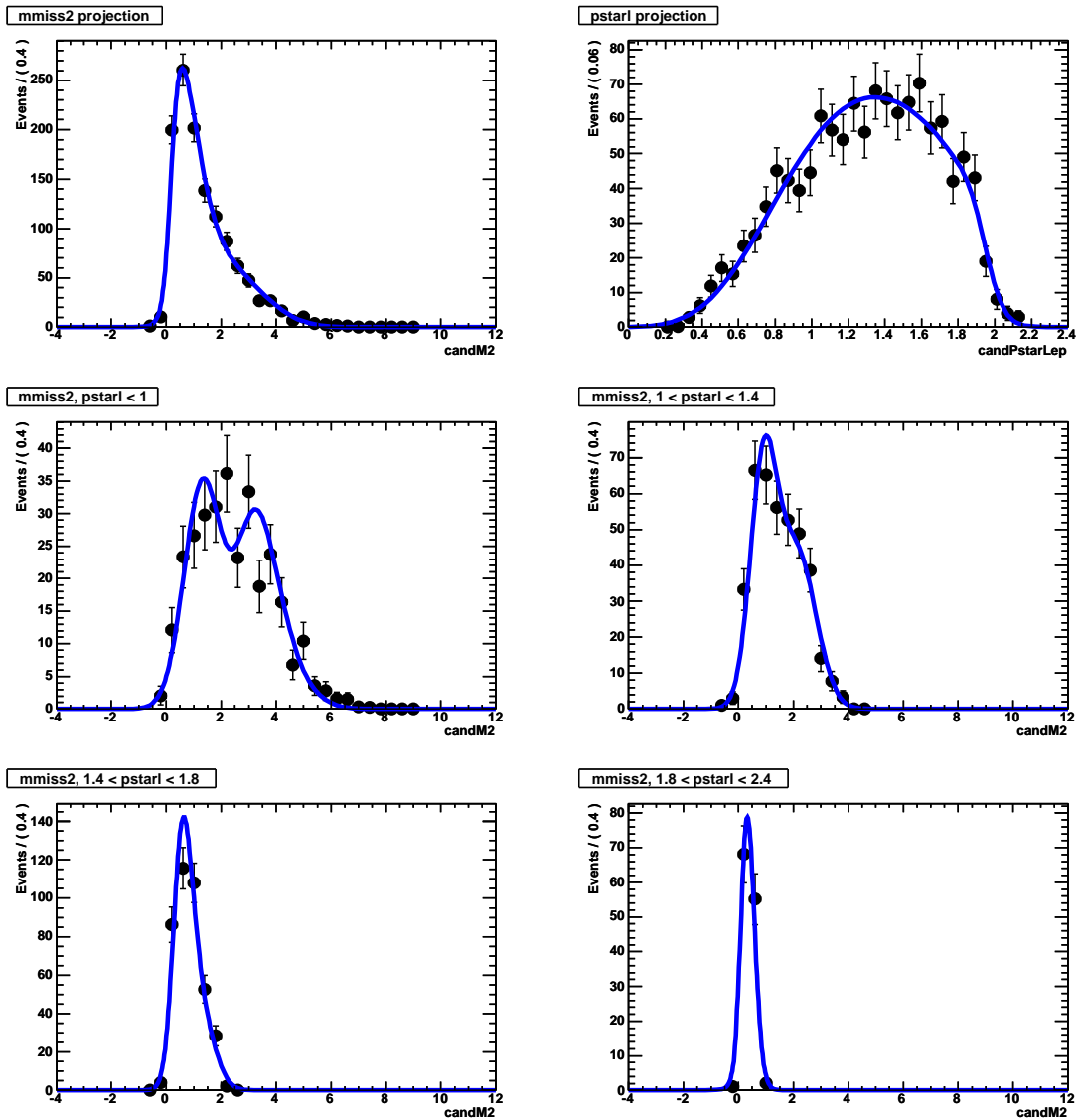


Figure 8.18: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+}$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^+$$

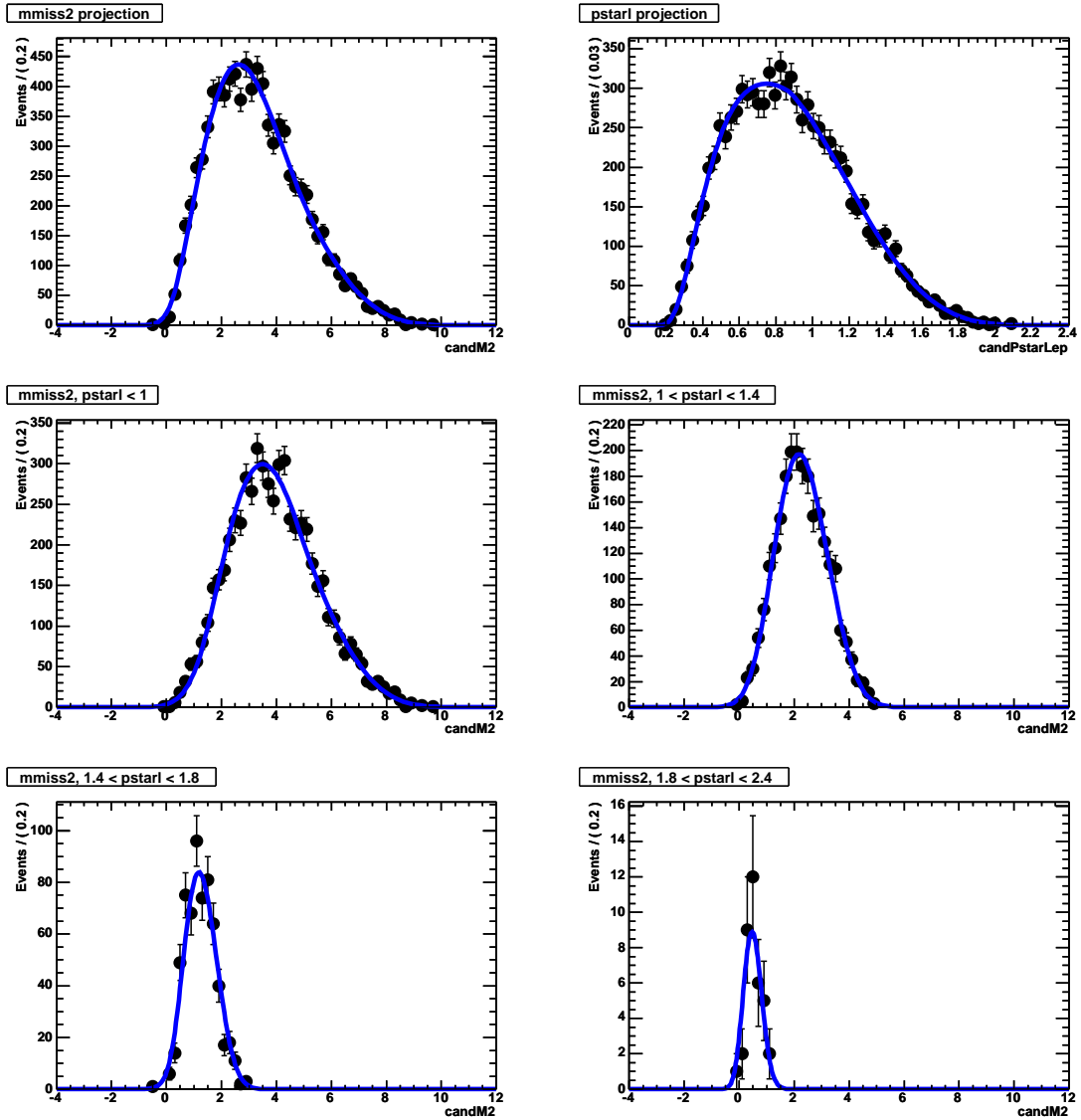


Figure 8.19: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \tau^- \bar{\nu}_\tau \Rightarrow D^+$.

$$D^{*+}\tau^{-}\bar{\nu}_{\tau}\Rightarrow D^{+}$$

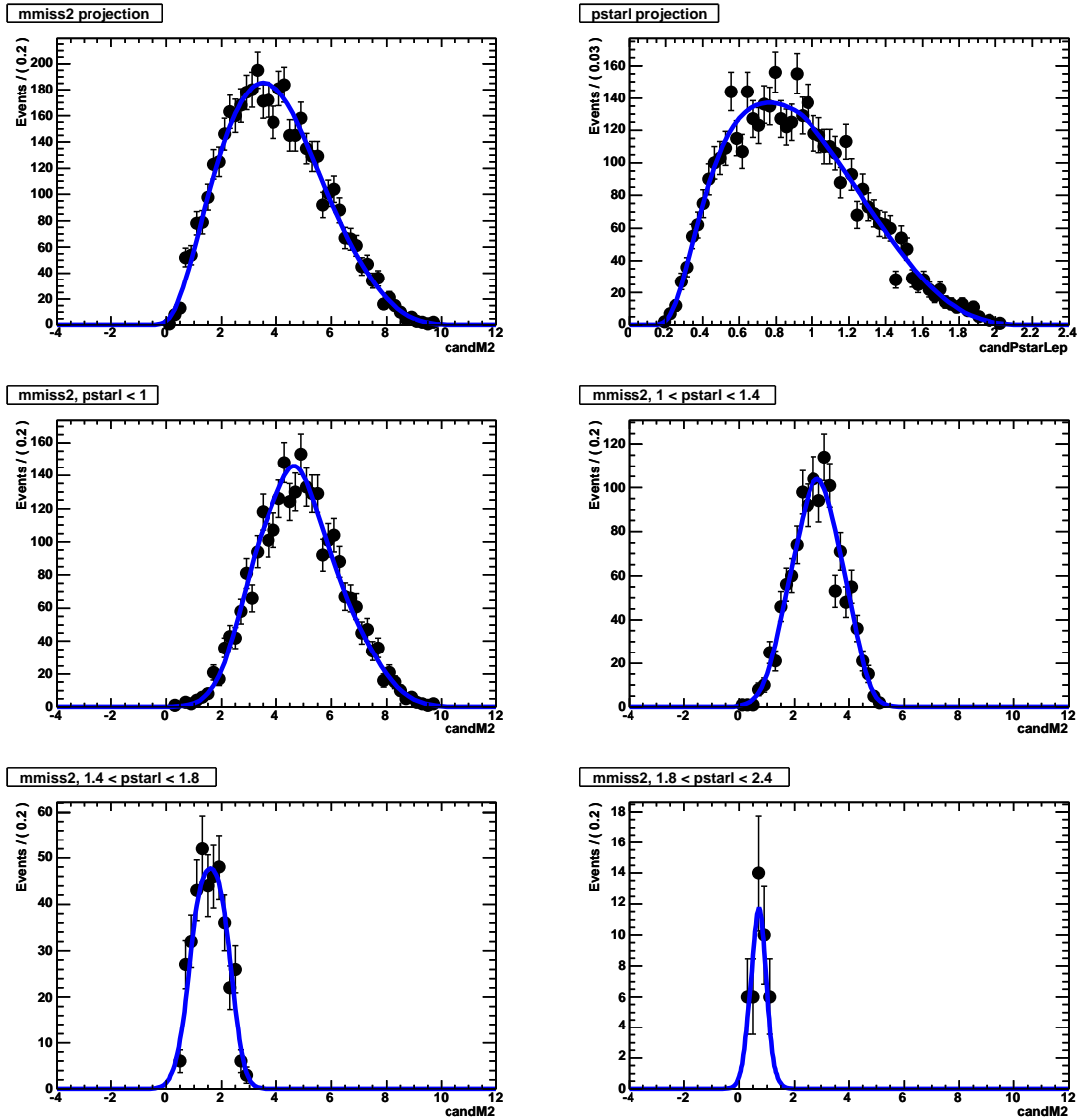


Figure 8.20: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*+}\tau^{-}\bar{\nu}_{\tau}\Rightarrow D^{+}$.

$$D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+$$

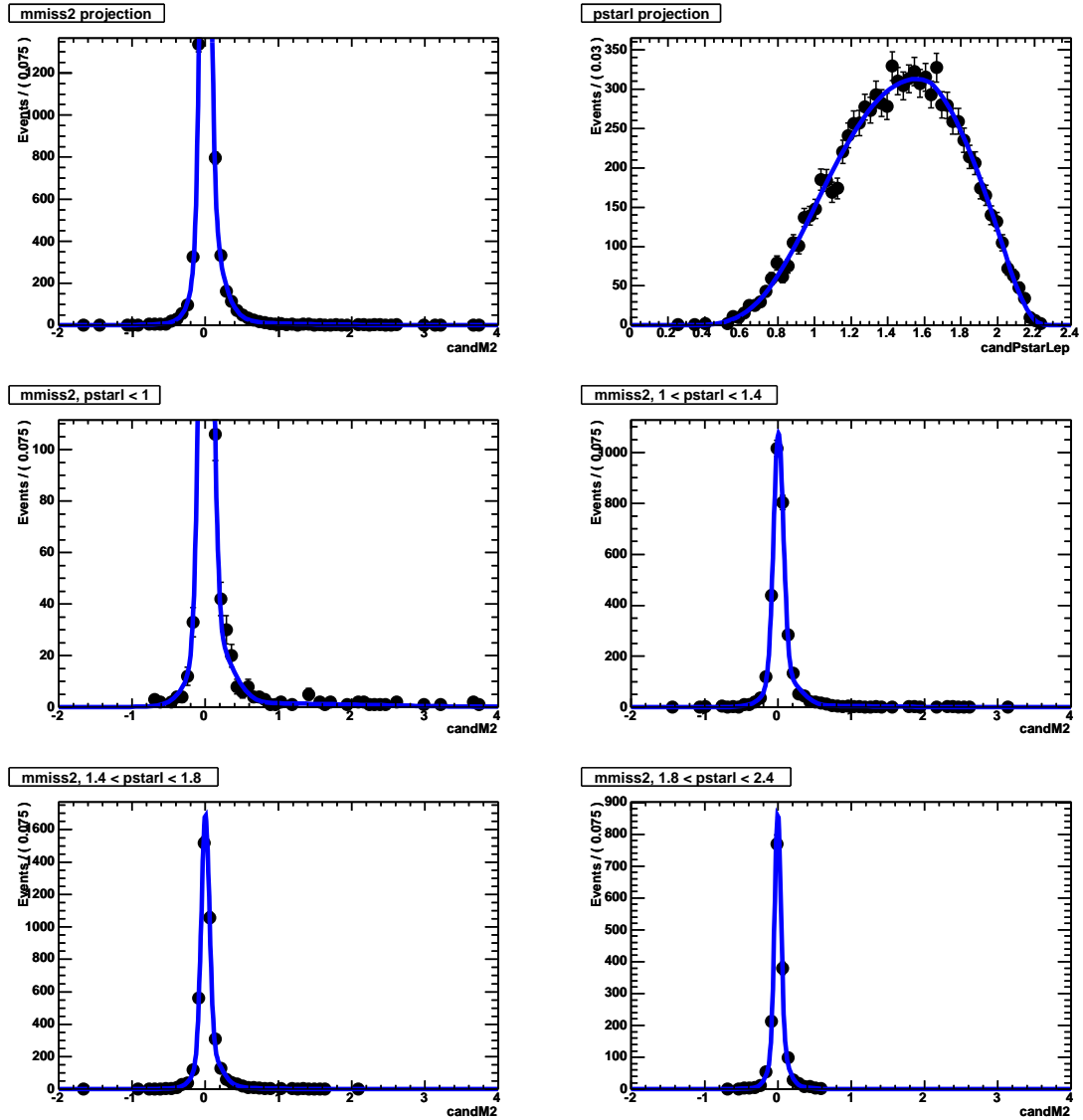


Figure 8.21: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+$.

$$D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^+$$

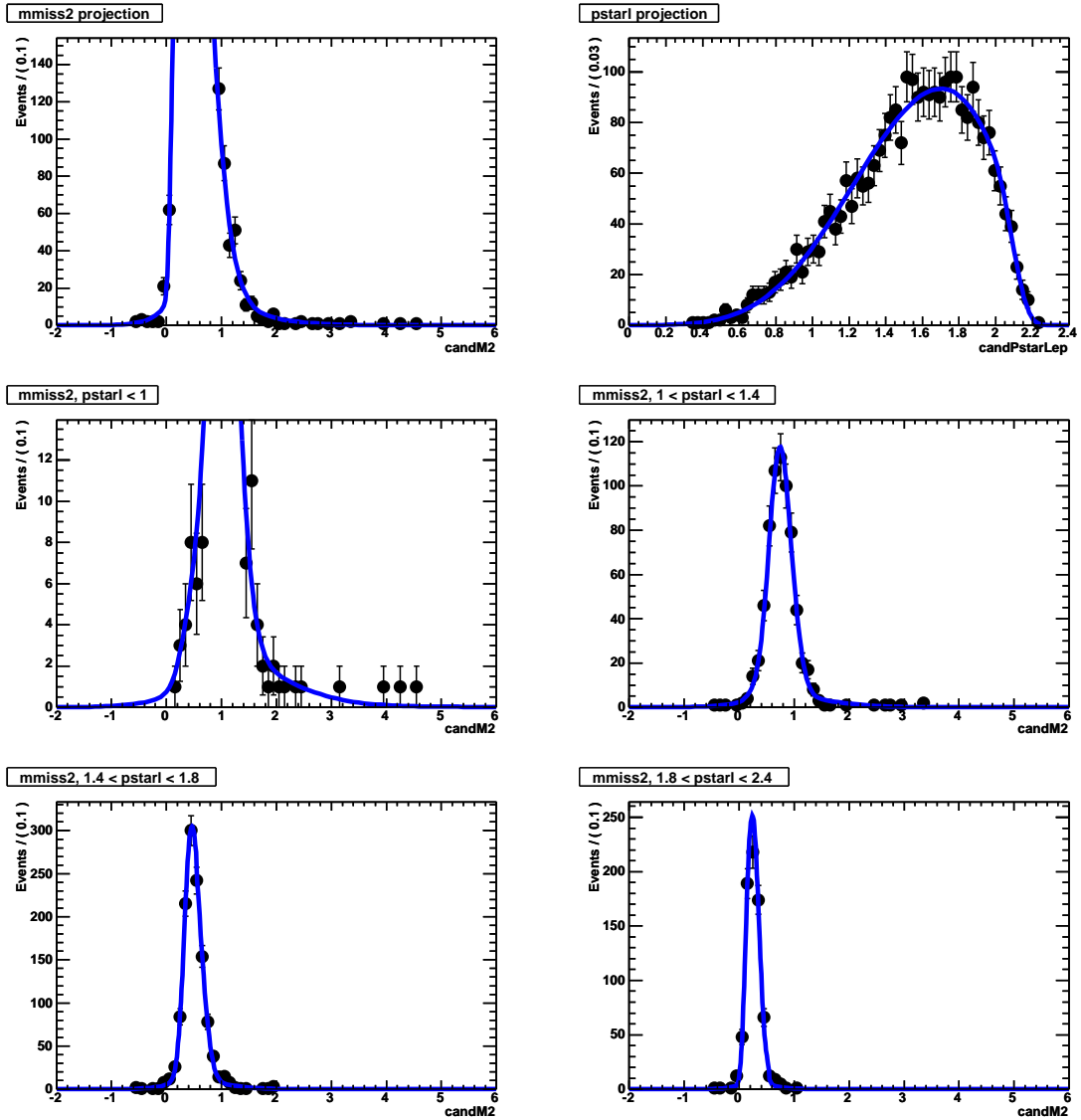


Figure 8.22: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*+} \ell^- \bar{\nu}_\ell \Rightarrow D^+$.

$$D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+$$

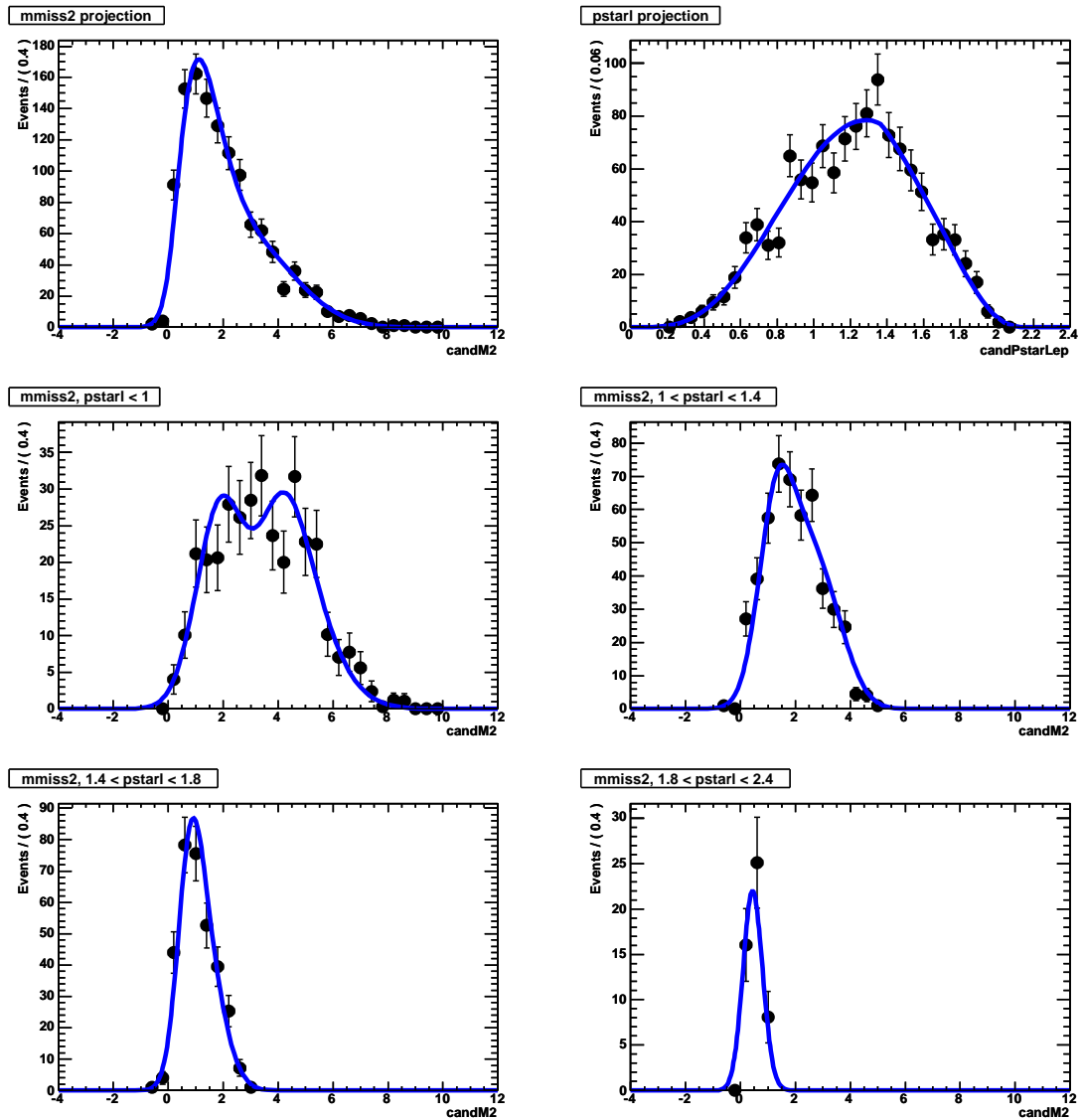


Figure 8.23: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0} \pi^0$$

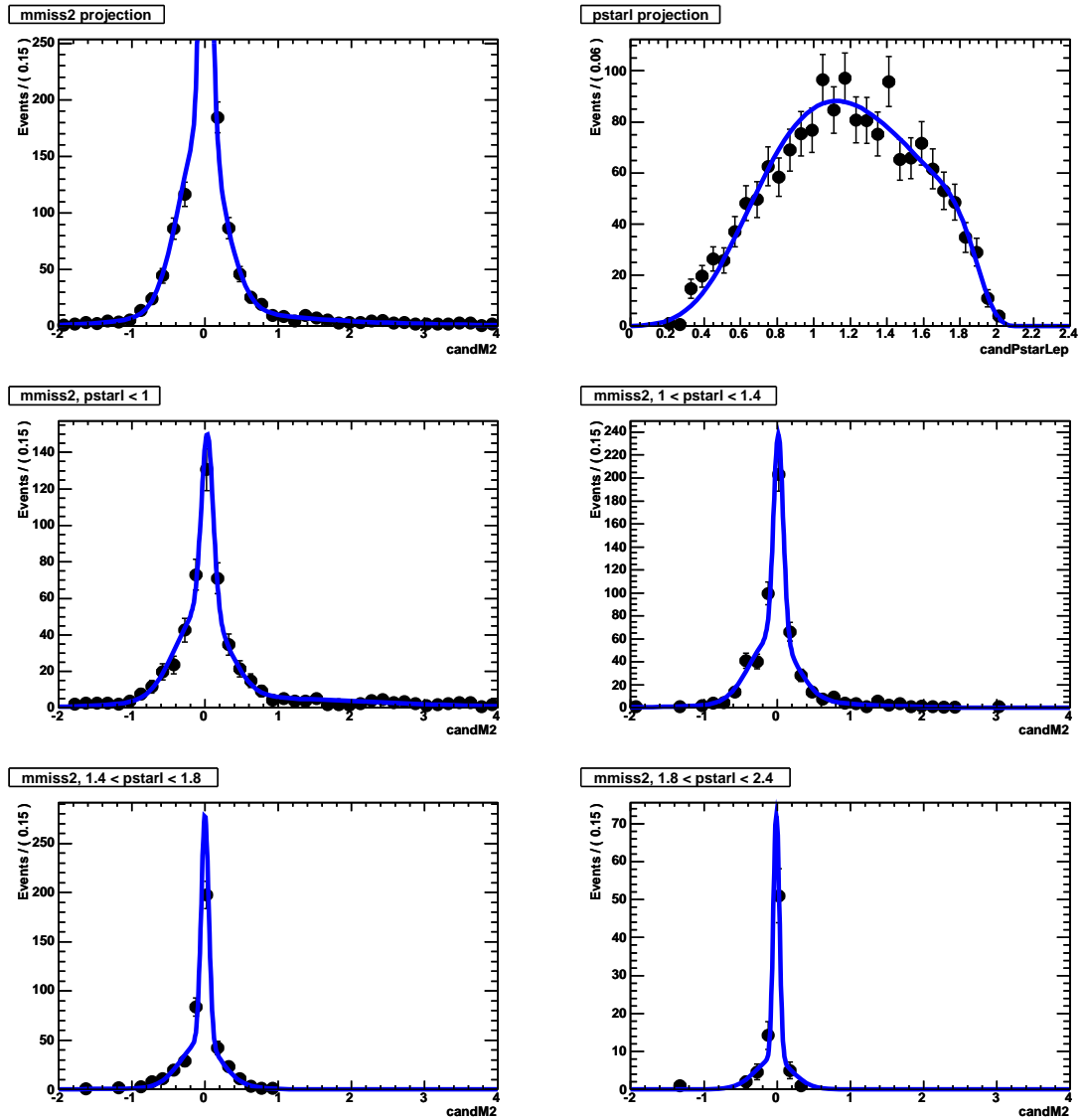


Figure 8.24: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0} \pi^0$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0} \pi^0$$

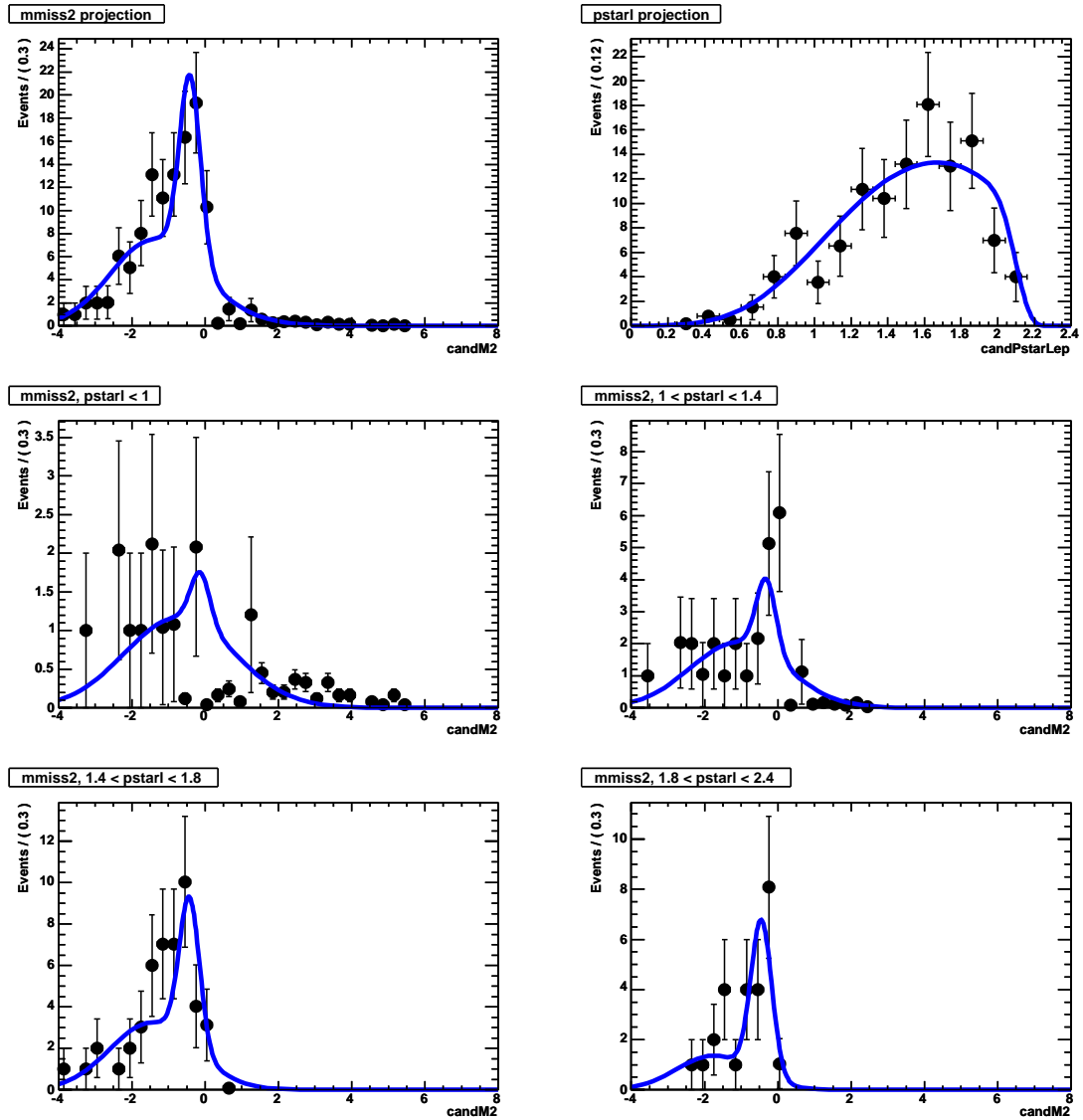


Figure 8.25: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*0} \pi^0$.

$$D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^{*0} \pi^0$$

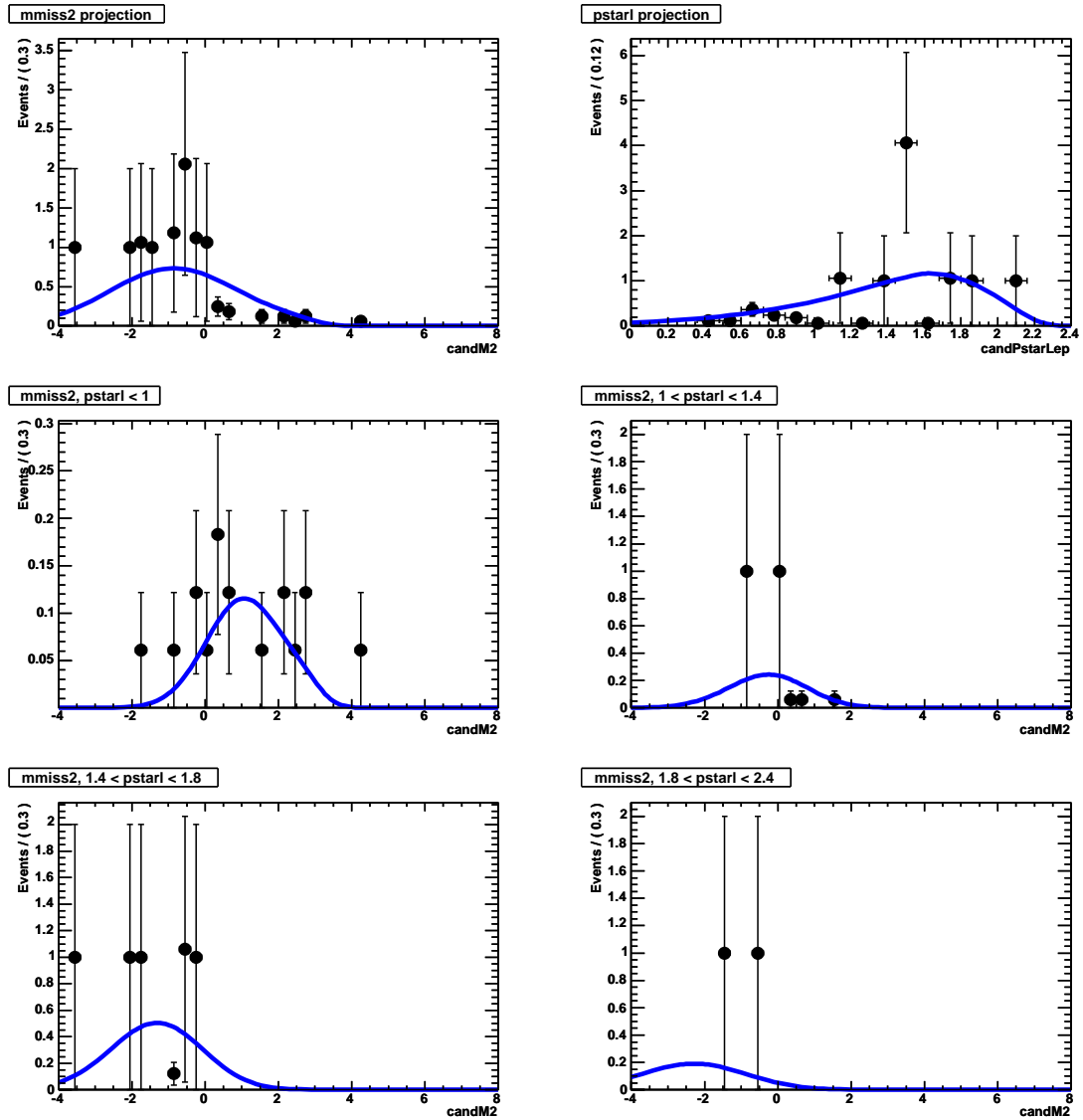


Figure 8.26: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^{*0} \pi^0$.

$$D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$$

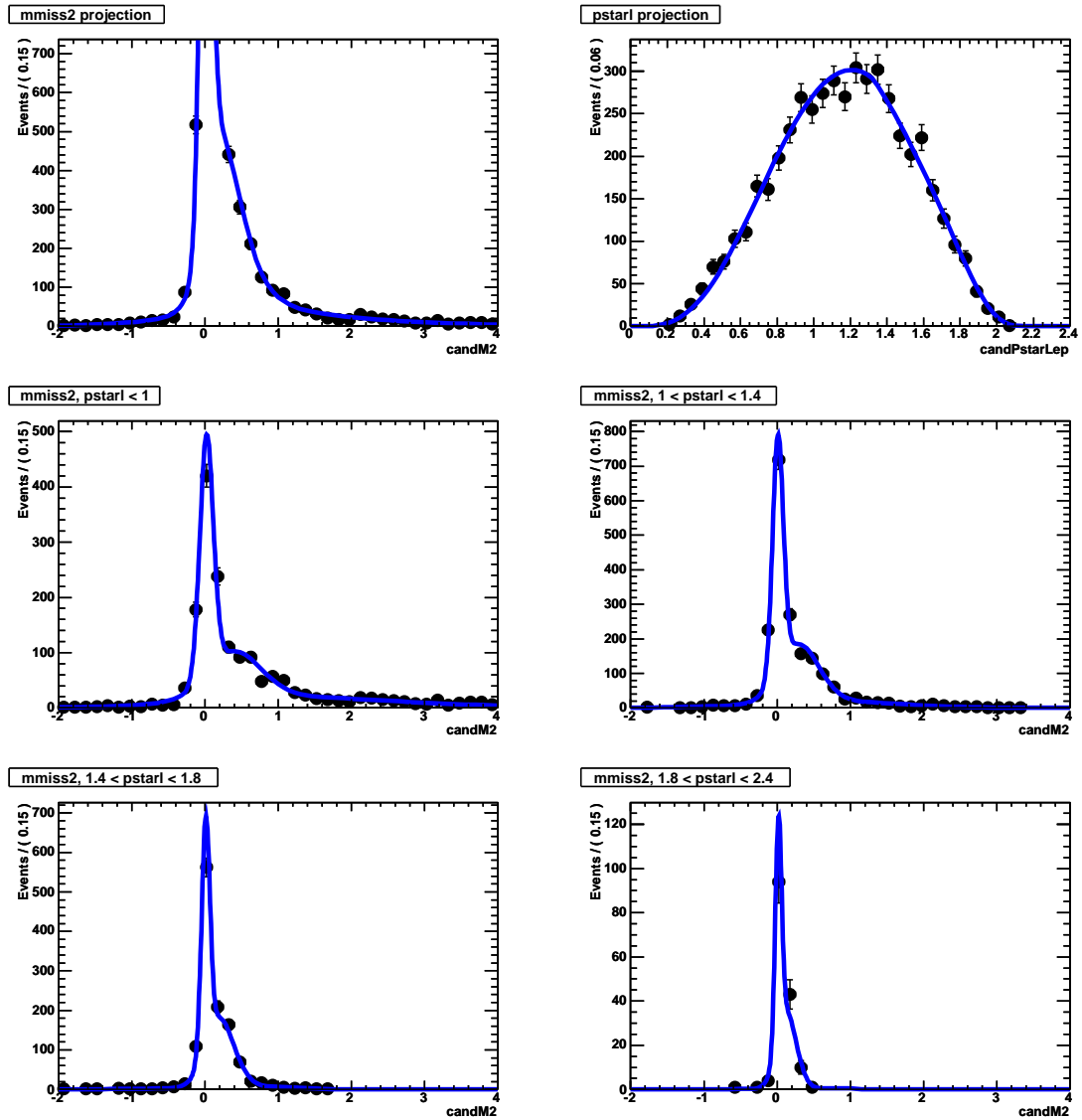


Figure 8.27: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^0 \pi^0$$

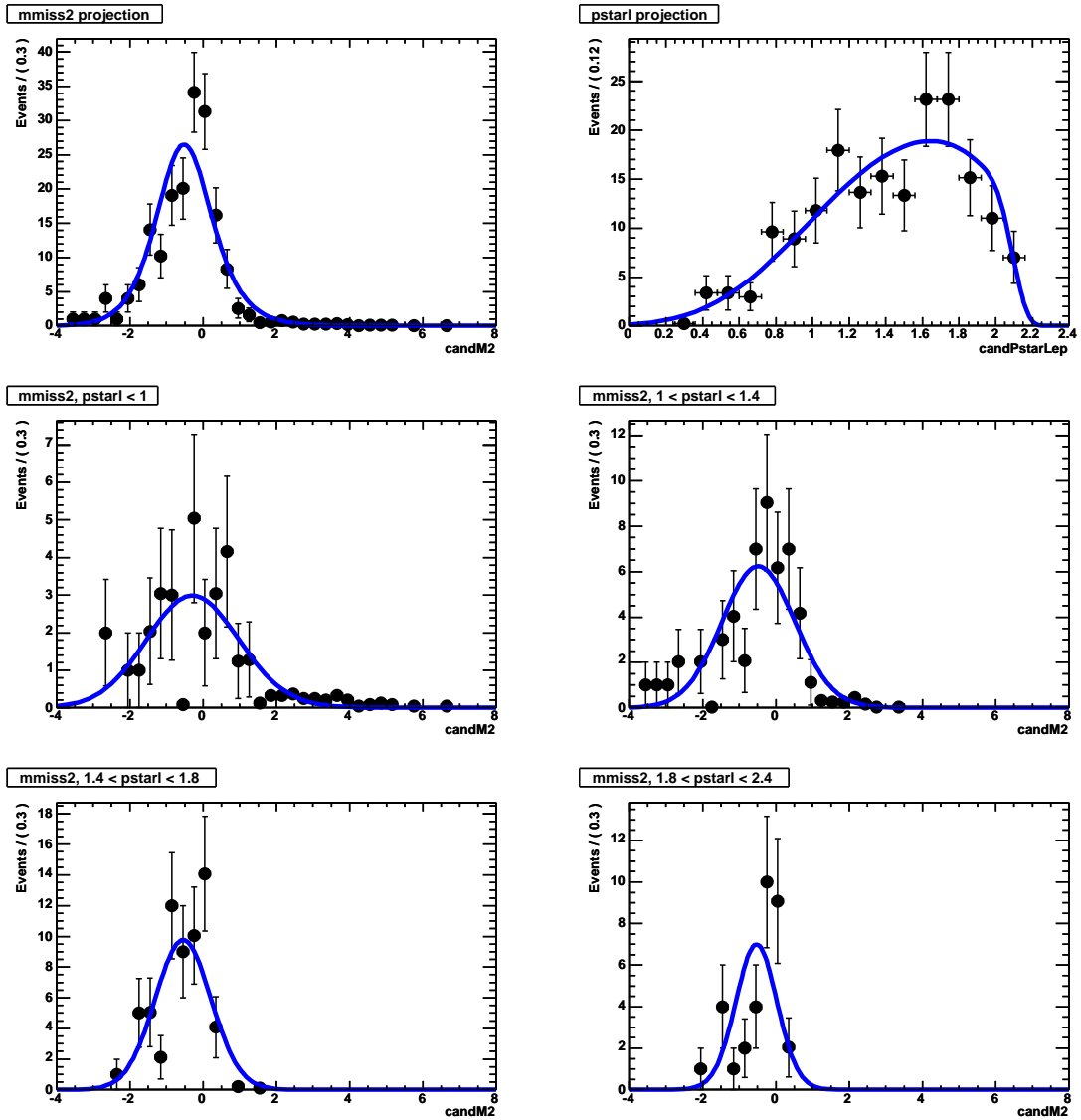


Figure 8.28: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*0} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^0 \pi^0$.

$$D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$$

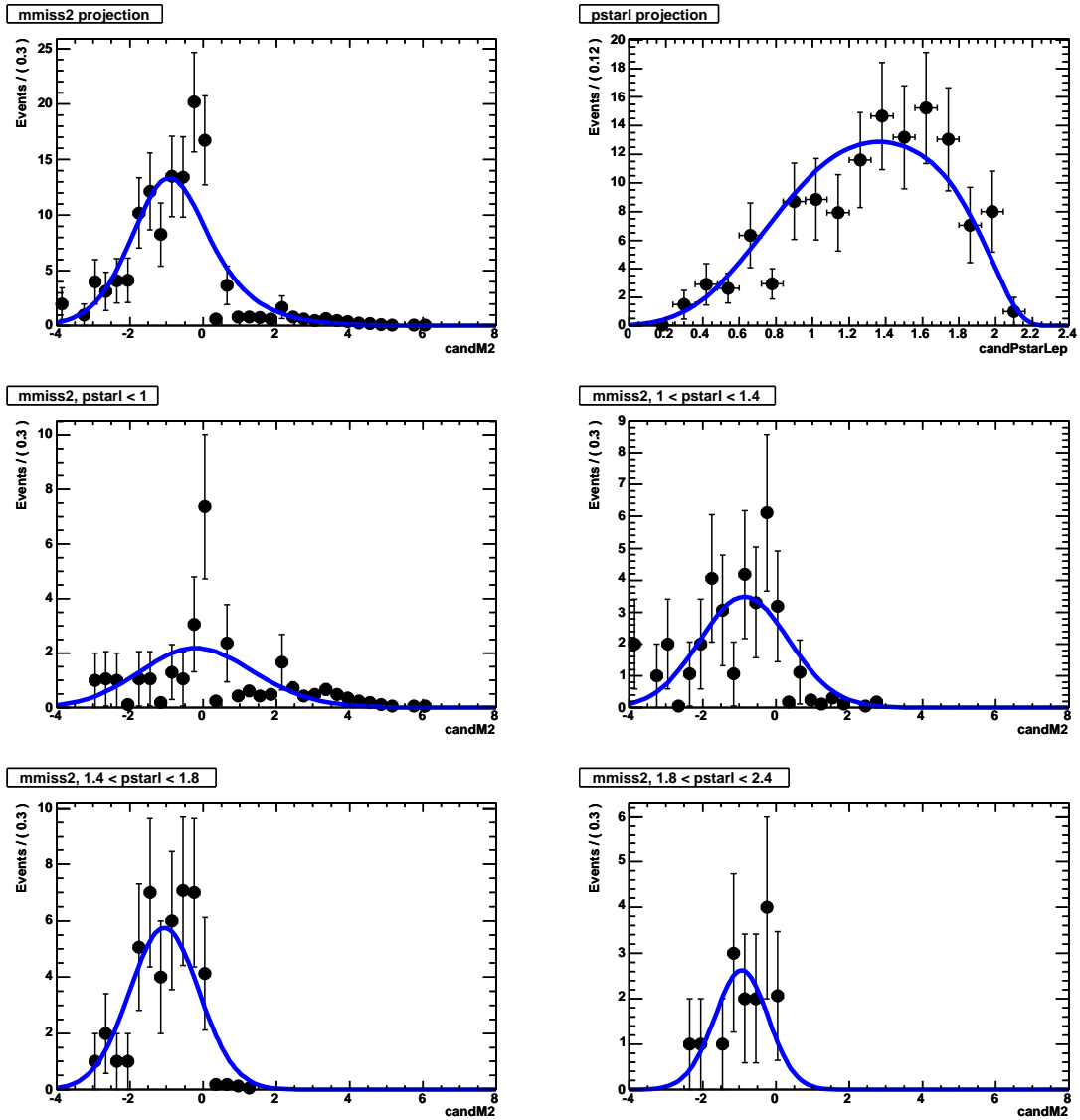


Figure 8.29: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^0 \ell^- \bar{\nu}_\ell \Rightarrow D^0 \pi^0$.

$$D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+} \pi^0$$

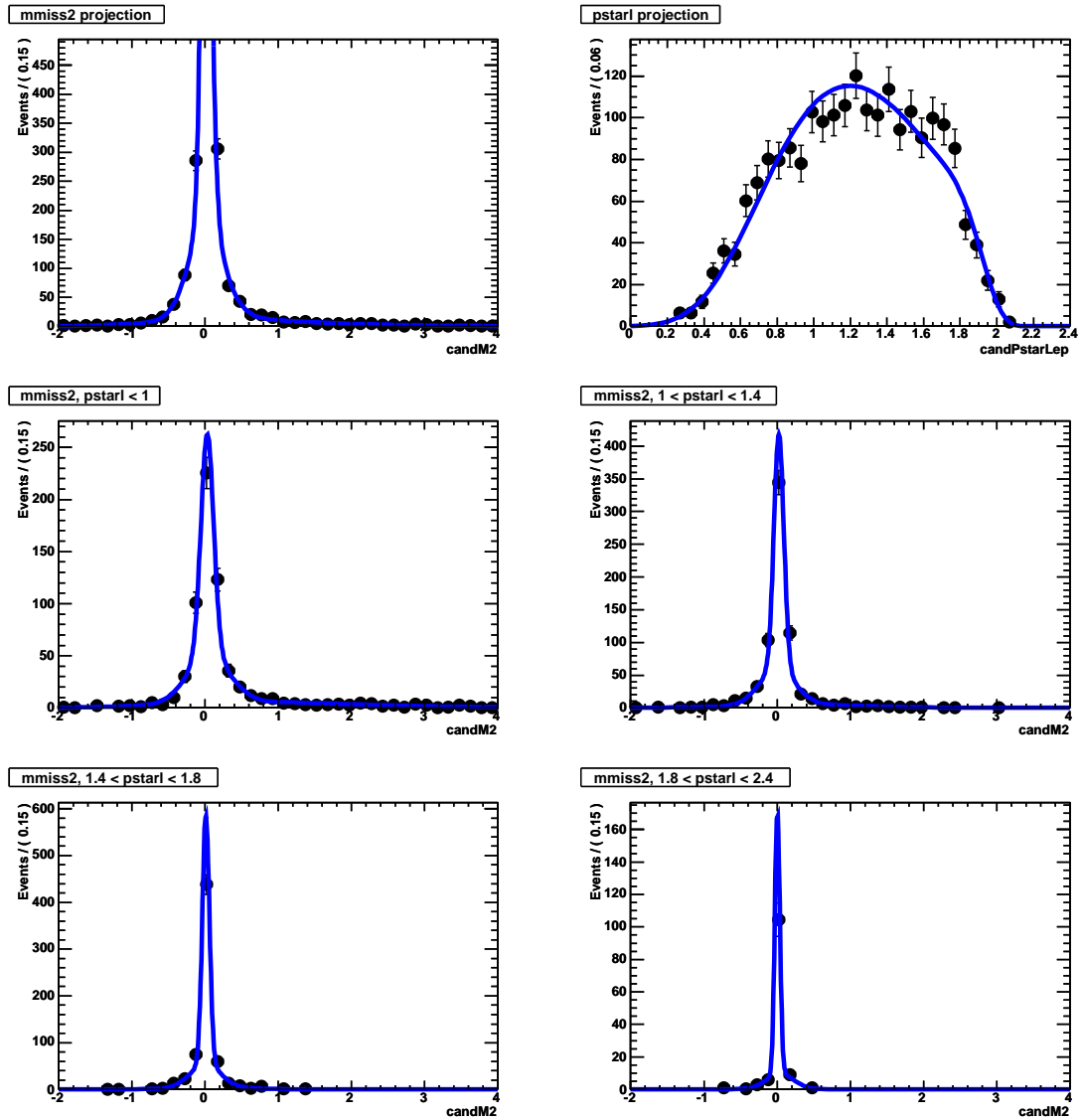


Figure 8.30: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{**} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+} \pi^0$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{*+} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+} \pi^0$$

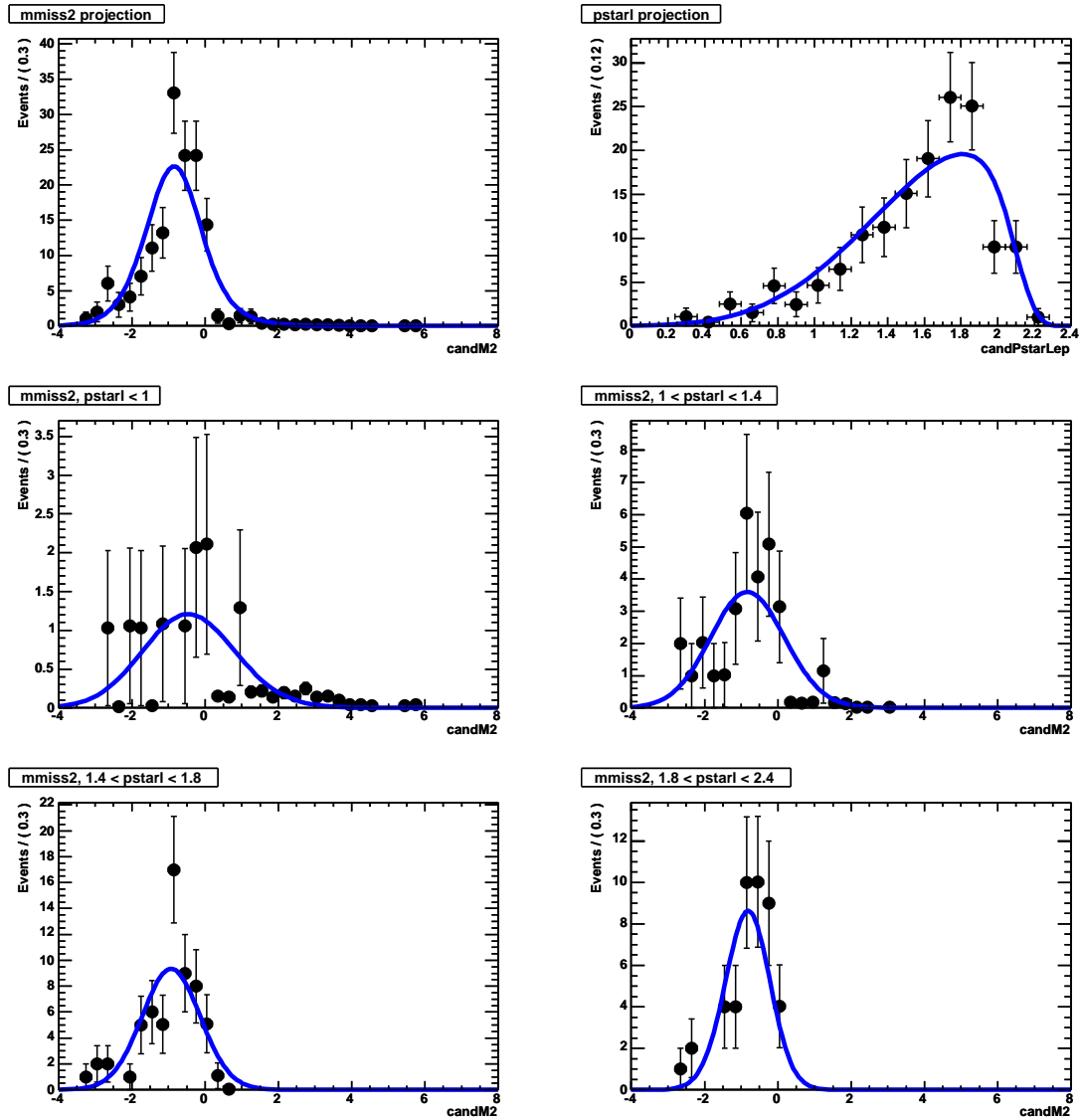


Figure 8.31: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*+} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{*+} \pi^0$.

$$D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+} \pi^0$$

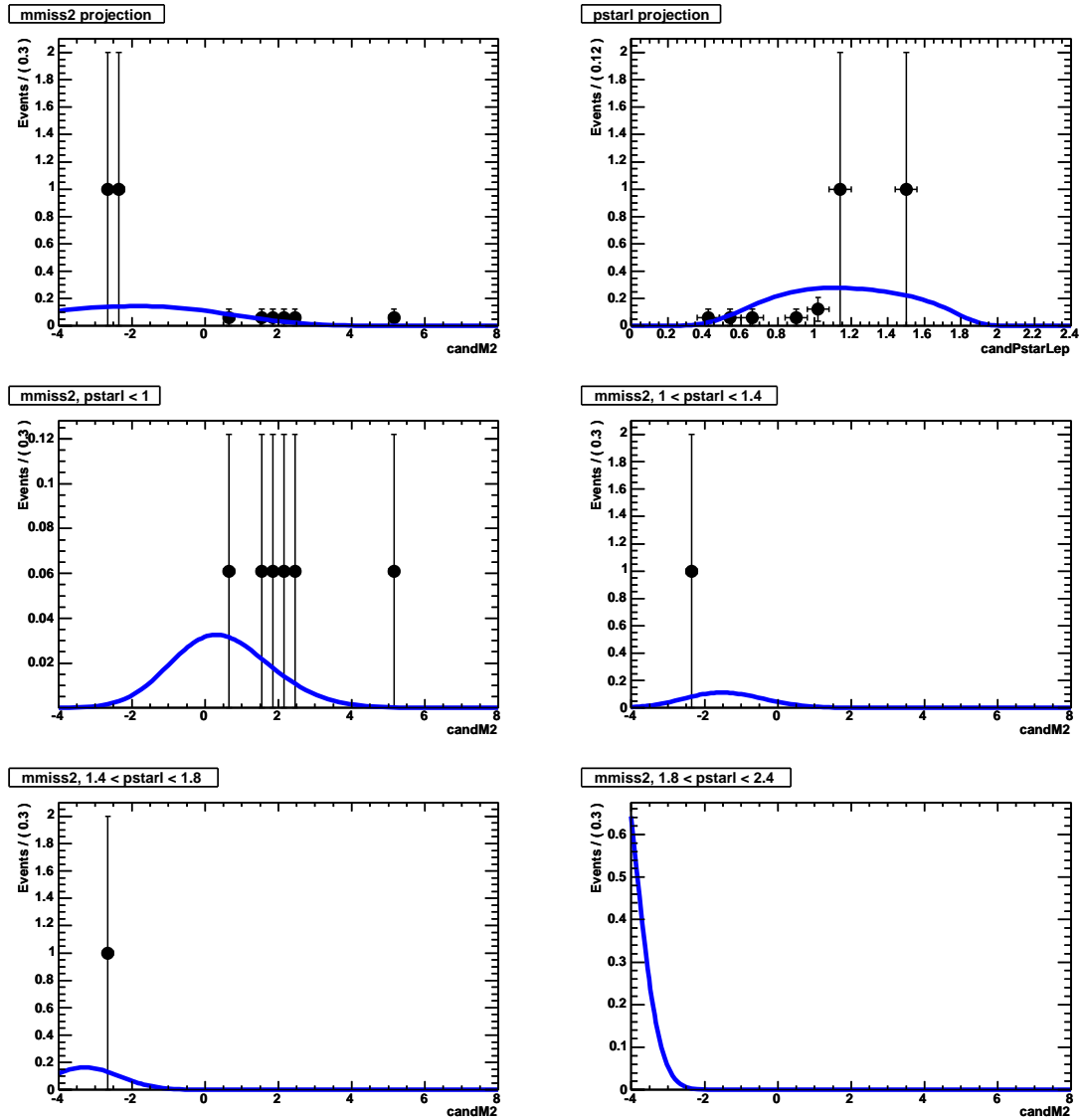


Figure 8.32: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^{*+} \pi^0$.

$$D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$$

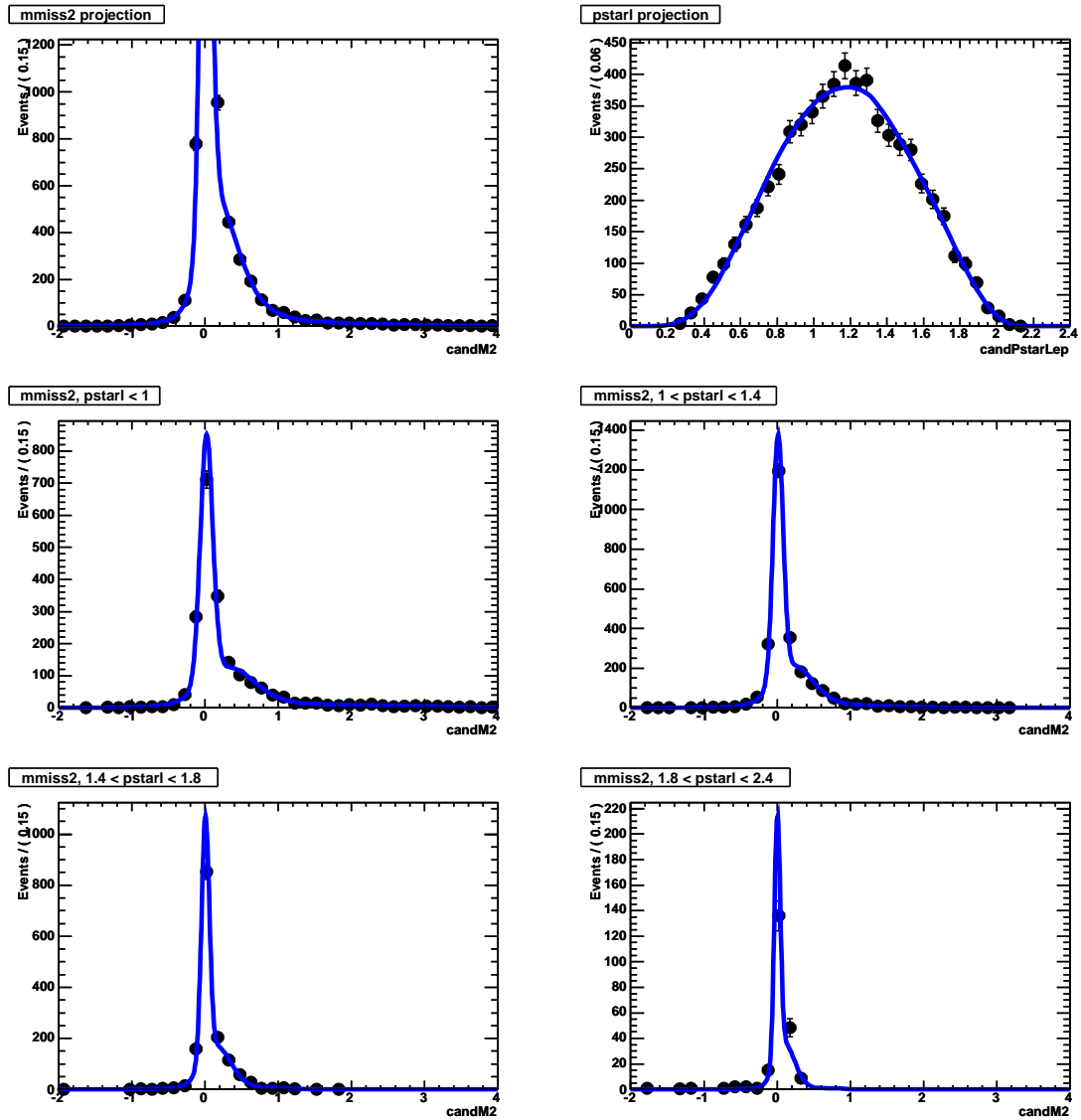


Figure 8.33: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^{**} \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$. The MC sample includes generically-simulated events, as documented in Section 8.5.3, which also explains the fluctuations visible in the points.

$$D^{*+} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{+} \pi^0$$

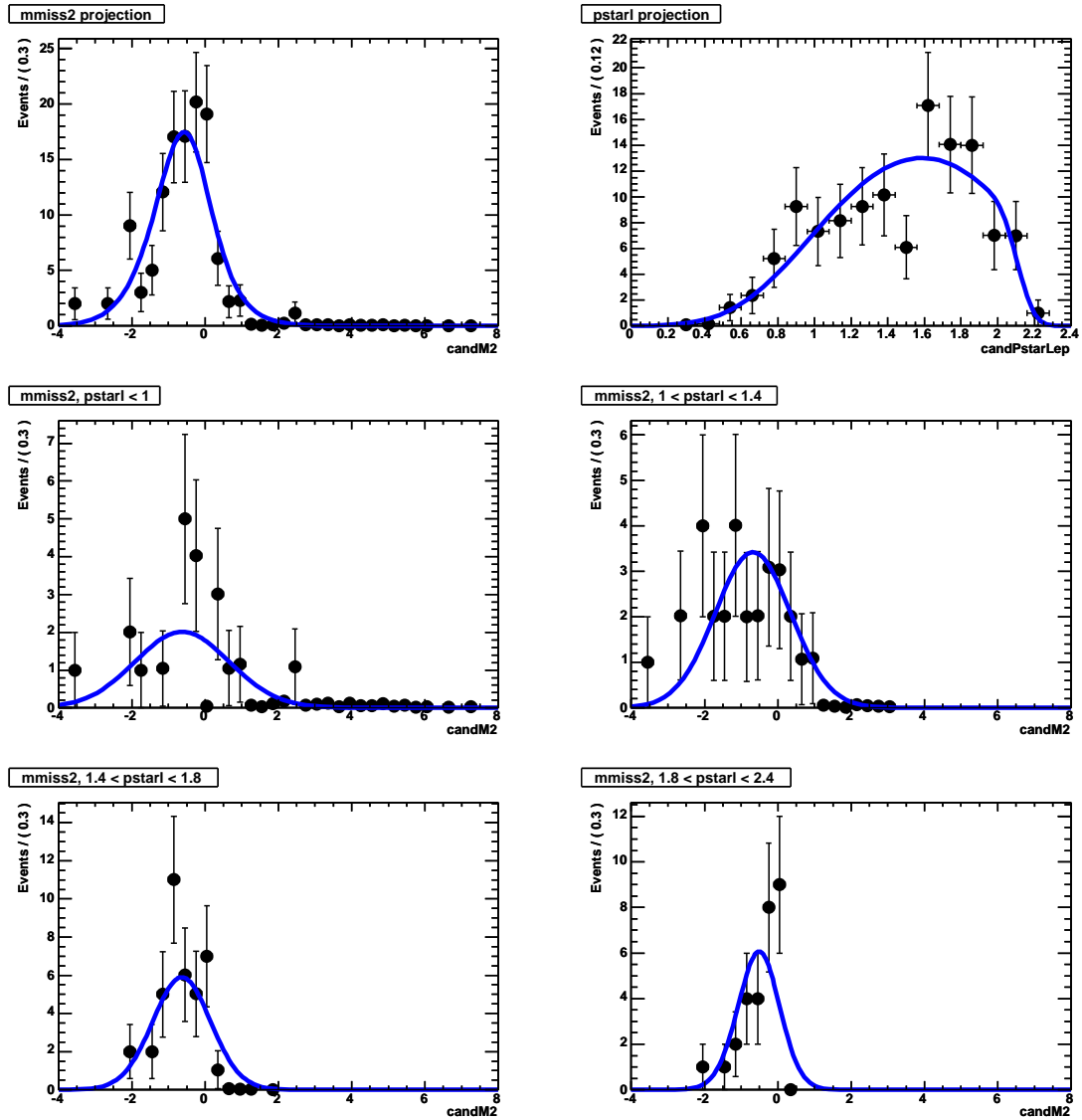


Figure 8.34: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for $D^{*+} \ell^{-} \bar{\nu}_{\ell} \Rightarrow D^{+} \pi^0$.

$$D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$$

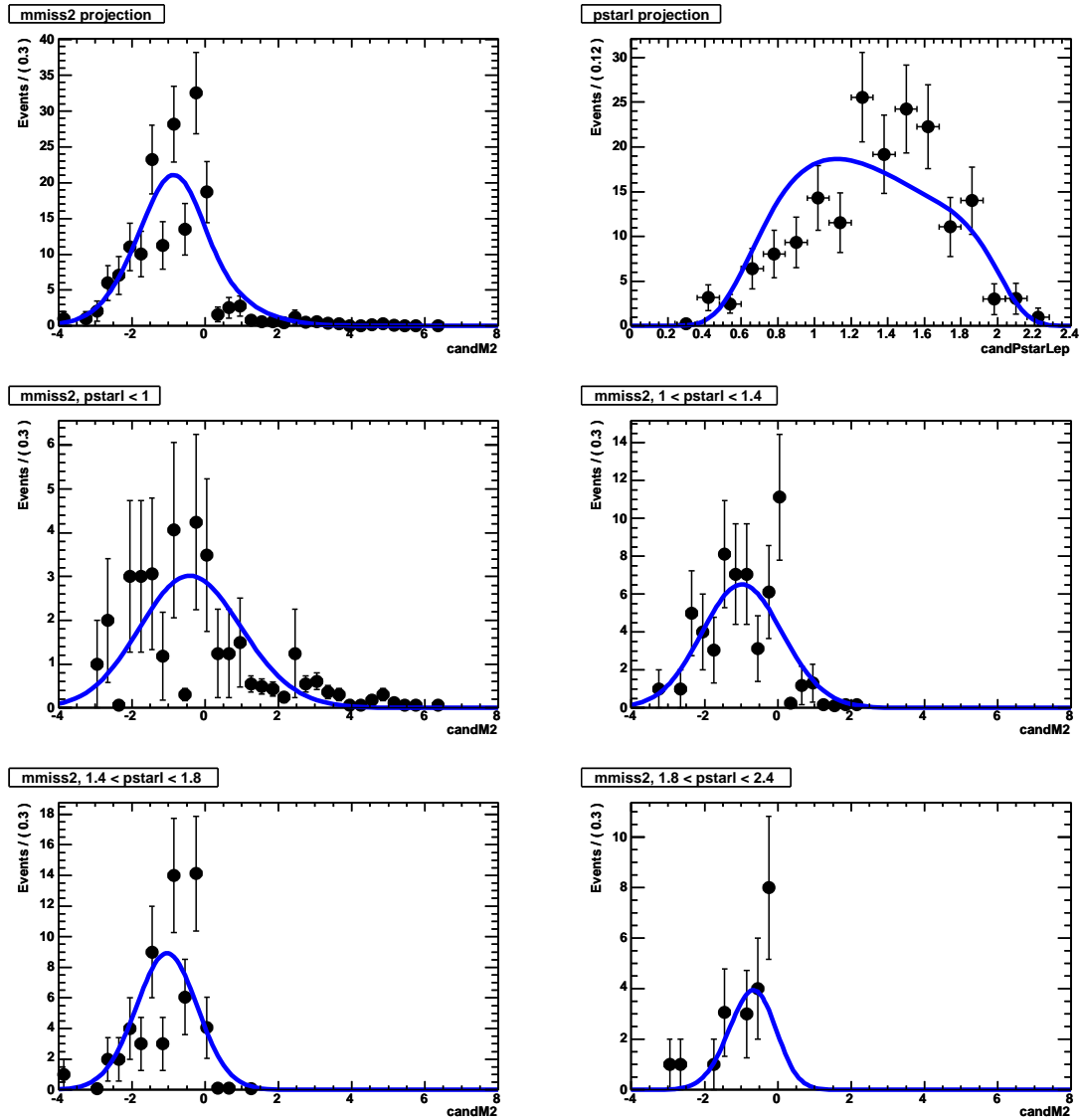


Figure 8.35: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for $D^+ \ell^- \bar{\nu}_\ell \Rightarrow D^+ \pi^0$.

8.7 $p_\ell^* - m_{\text{miss}}^2$ PDFs for Combinatoric Events

The two-dimensional PDF \mathcal{P}_1 used for signal and semileptonic events is general enough to be used for combinatoric background events as well.

Since we have no “signal” MC for combinatoric background, all of the shapes are derived from the generic MC sample, using the true MC decay to categorize events. The composition of combinatoric backgrounds is described in Section 5.2, and will be described in more detail in Section 12.5. Because the number of combinatoric events is rather small, even in the full generic MC sample, we combine events in related channels to extract the combinatoric BG shapes. In total, we fit five samples:

- signal channels: charge crossfeed reconstructed as D
- signal channels: charge crossfeed reconstructed as D^*
- signal channels: all other combinatoric
- D^{**} control sample: charge crossfeed (both D and D^*)
- D^{**} control sample: all other combinatoric.

Note that, as detailed in Tables 8.1 and 8.2, this parameterization means that these PDFs are used multiple times in the complete fit.

One-dimensional projections of these fits are shown in Figures 8.36– 8.40 for each of the 5 PDFs. These figures show the same projections as the signal and semileptonic PDFs above.

Signal D channels, charge crossfeed

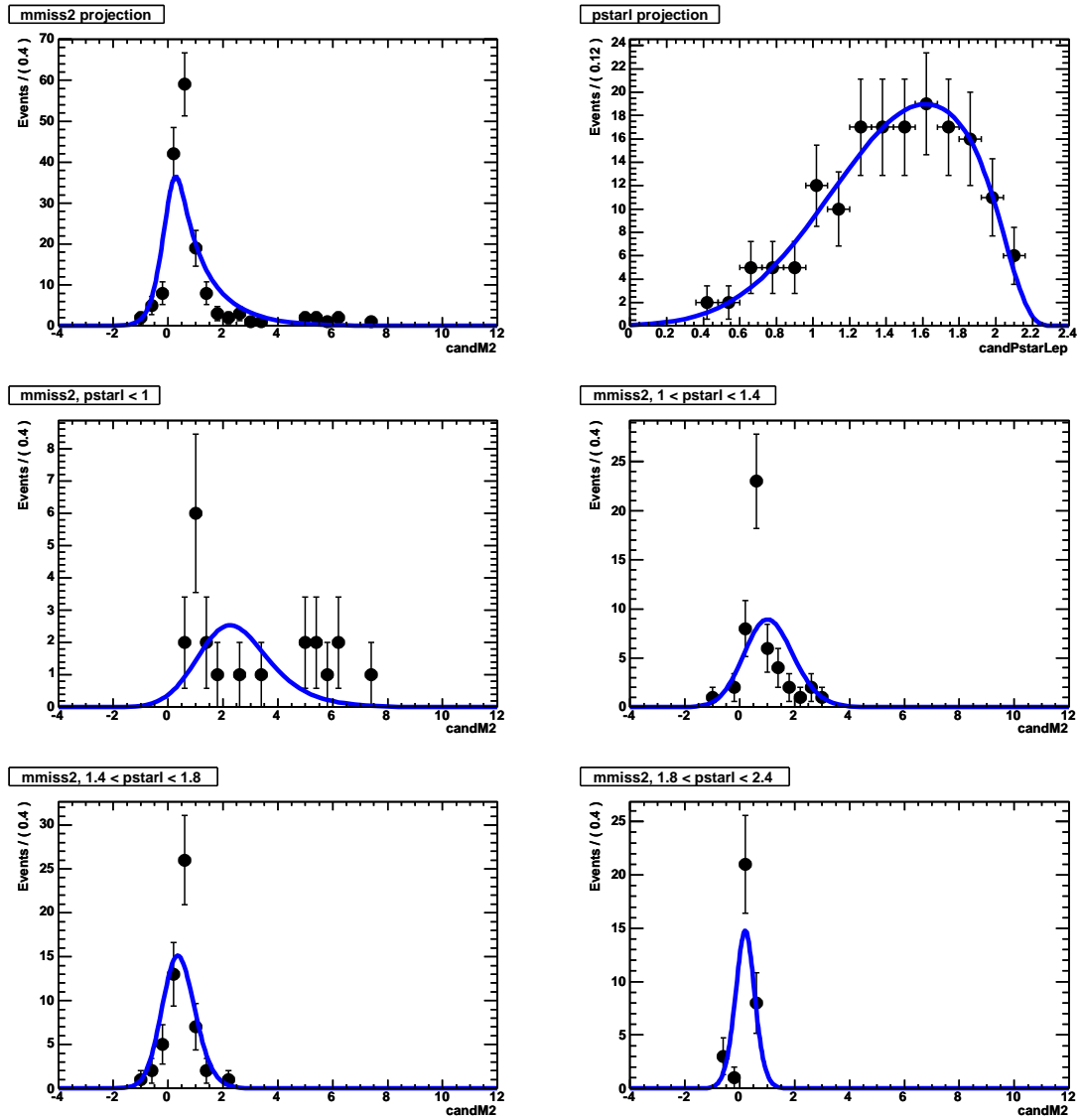


Figure 8.36: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for charge crossfeed in the signal D channels.

Signal D^* channels, charge crossfeed

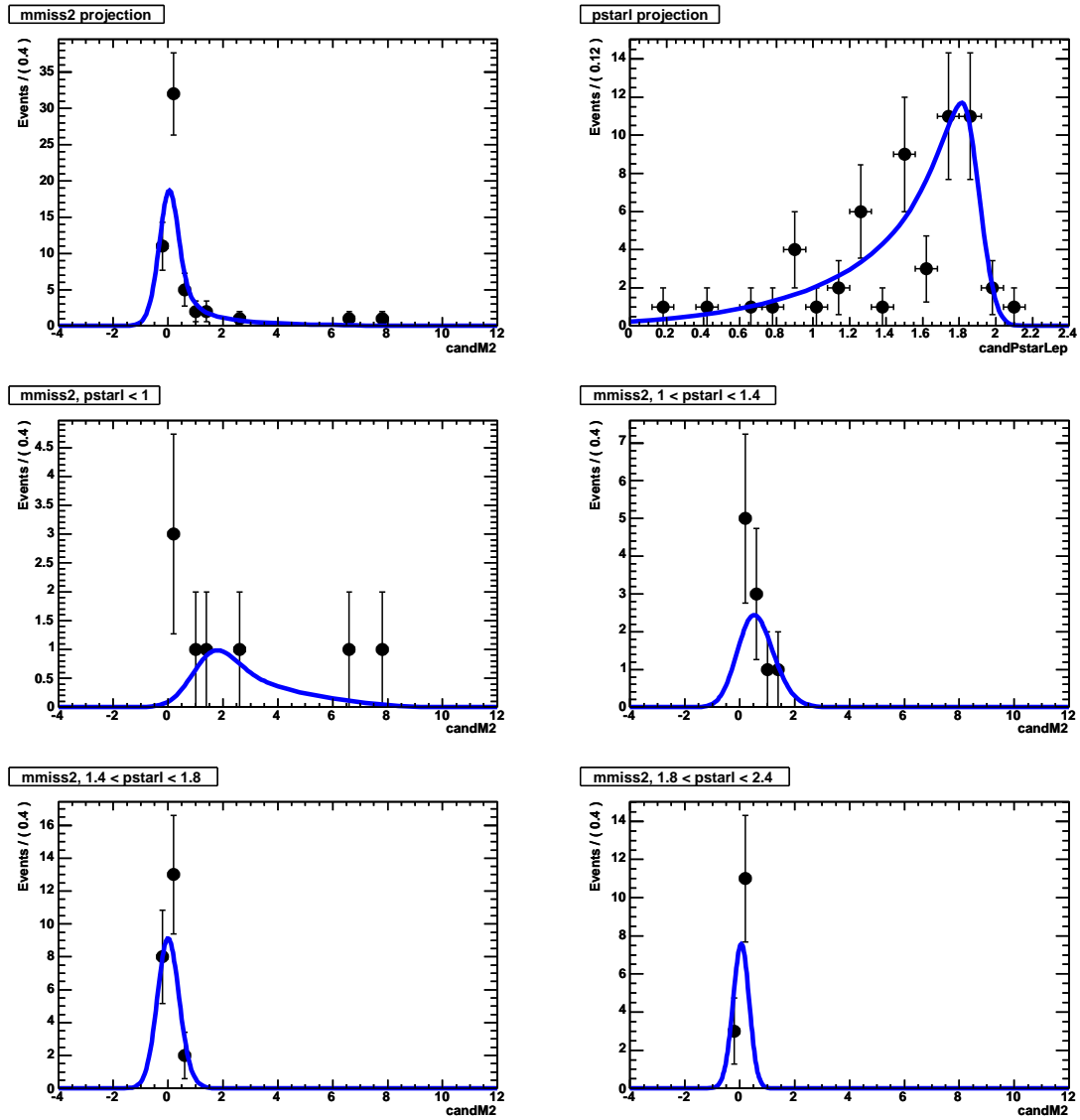


Figure 8.37: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for charge crossfeed in the signal D^* channels.

Signal channels, combinatoric

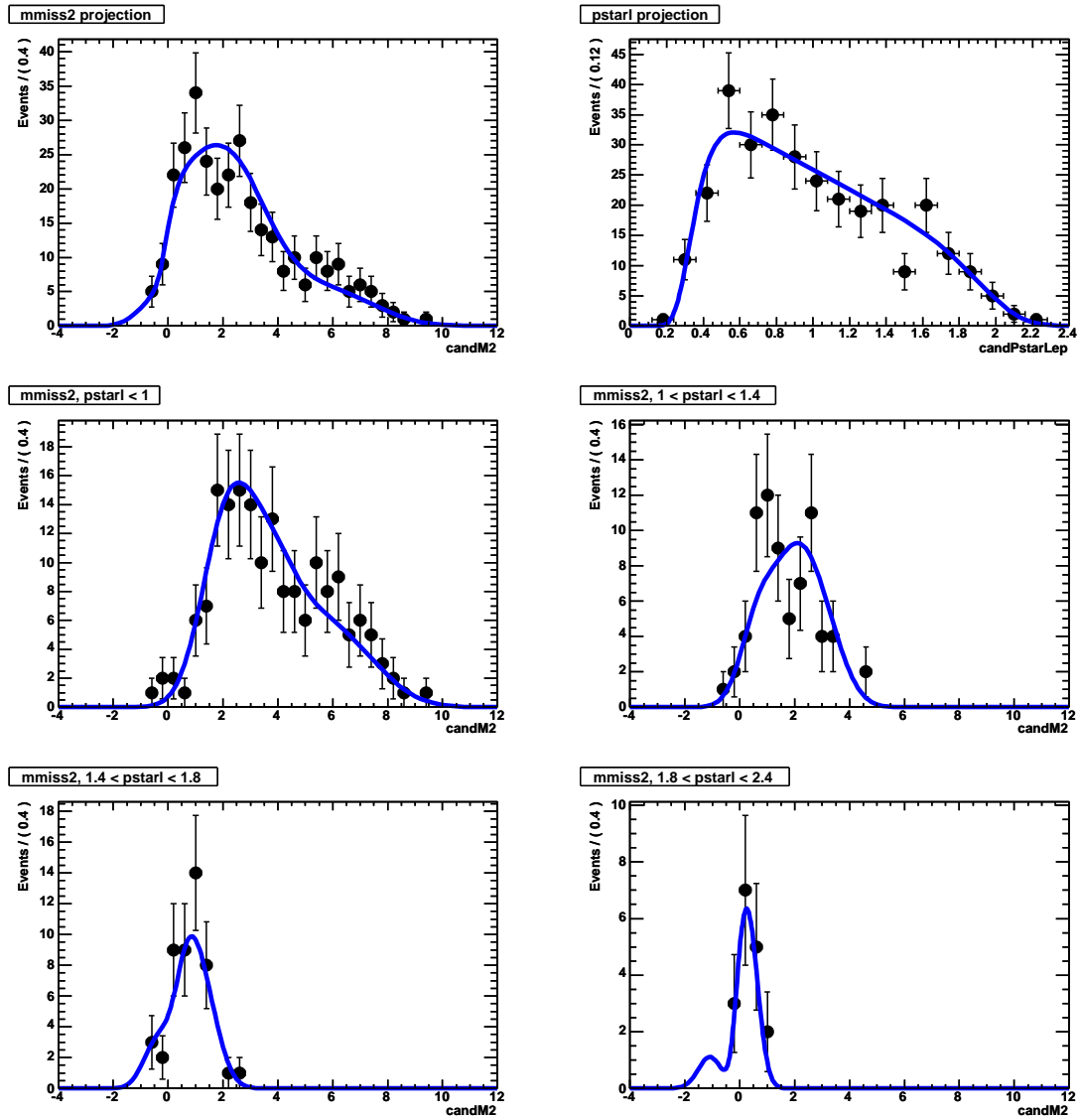


Figure 8.38: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for other combinatoric background in the signal channels.

D^{**} control sample, charge crossfeed

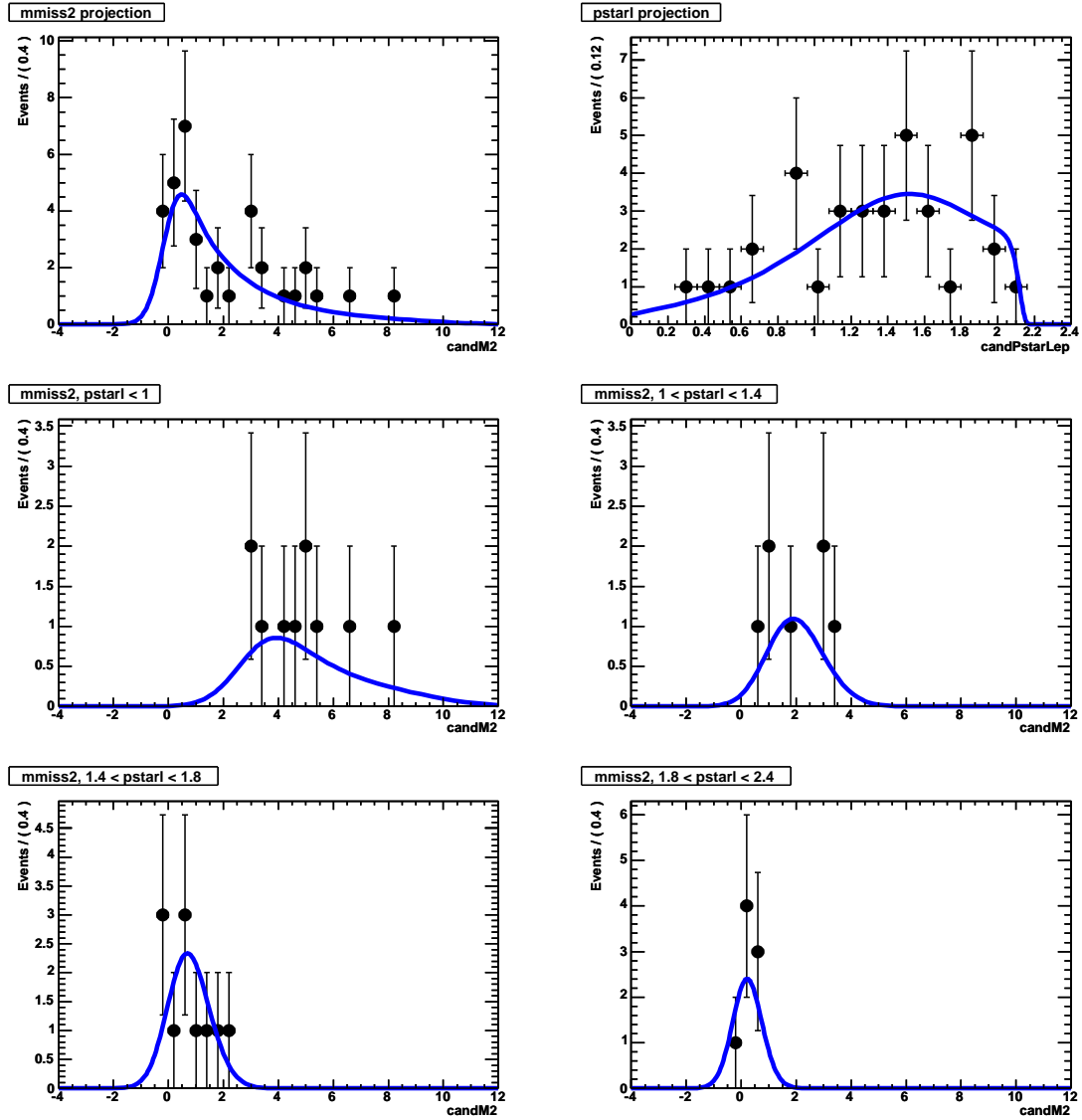


Figure 8.39: Fitted MC distribution of $p_{\ell}^* - m_{\text{miss}}^2$ for charge crossfeed in the D^{**} control sample.

D^{**} control sample, combinatoric

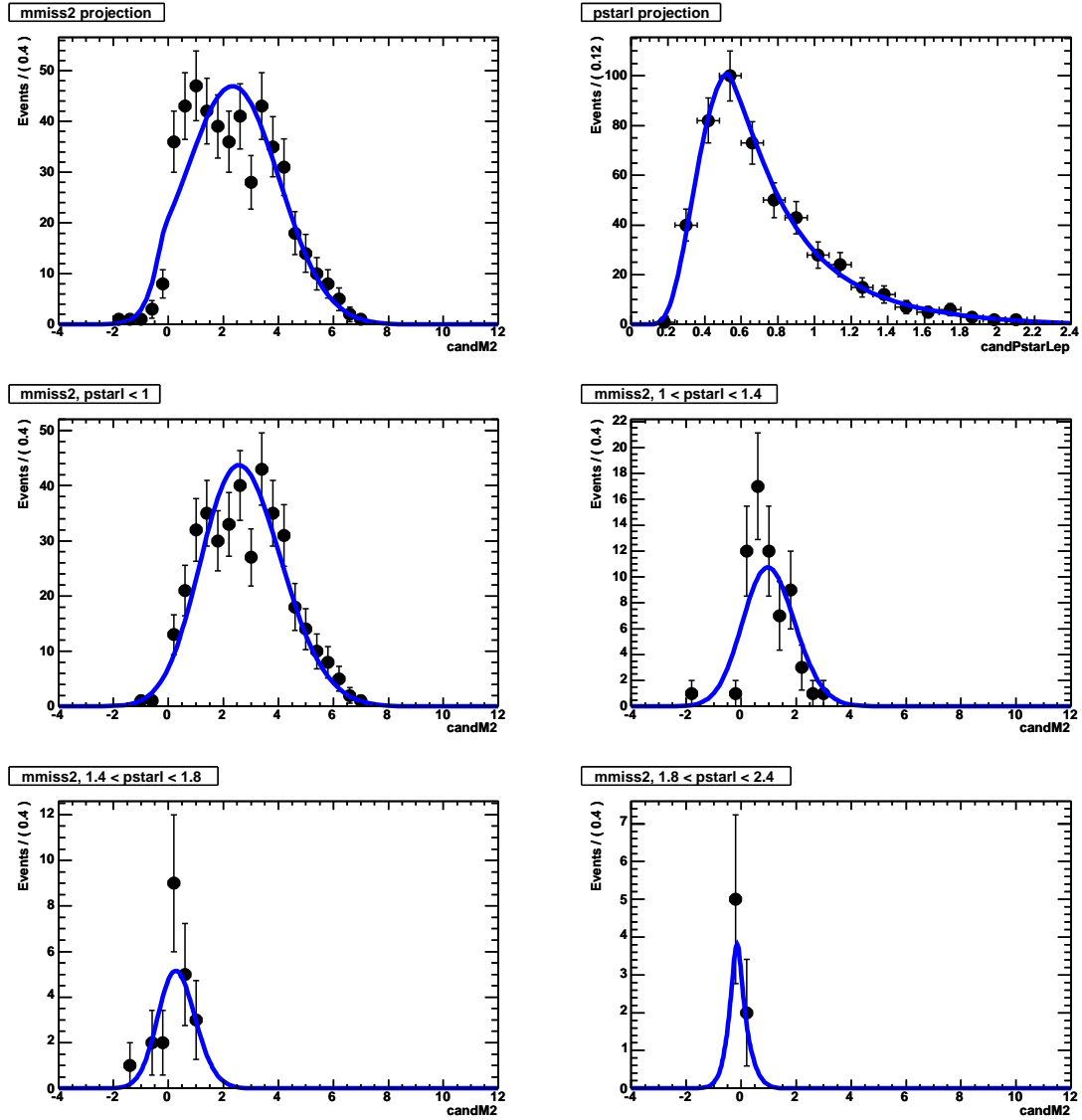


Figure 8.40: Fitted MC distribution of $p_\ell^* - m_{\text{miss}}^2$ for other combinatoric background in the D^{**} control sample.

8.8 Crossfeed Constraints in the Fit

The fit described in the previous sections is able to extract the yield of the various components within each of the signal channels, but we are able to do better by combining information from multiple channels simultaneously. We can use, for instance, the fit results from the D^{*0} channel to constrain the amount of $D^{*0}\ell^-\bar{\nu}_\ell$ feed-down in the D^0 channel because the physical source of the two samples is the same.

We construct 20 such constraints, enumerated in yields; there are 20 such constraints, enumerated in Table 8.5. Each constraint relates the event yield in a crossfeed channel (either feed-up or feed-down) to the yield in the correctly reconstructed channel, and takes the form

$$N_{i'j} = N_{ij} \cdot f_{i'j}^i$$

where N_{ij} is the event yield in the correctly-reconstructed channel, $N_{i'j}$ the yield of the same component j in the crossfeed channel i' , and $f_{i'j}^i$ the expected ratio between the two. The values of these crossfeed constraints in signal and semileptonic MC and can be seen in Table 8.5, although, in practice, we will determine some of these constraints directly from the data, as described in Section 8.9. Note that the feed-down constraints are generally “large” num-

bers, between 0.2 and 1.0, while the feed-up constraints are much smaller, between 0.01 and 0.10. The feed-up constraints correspond to very few events (typically fewer than 10) in the final data samples, and so these constraints do not matter as much as the feed-down constraints.

Each such constraint effectively reduces the number of degrees of freedom in the fit by one.

These constraints introduce correlations between the measurements in the various channels. Because events in the feed-down modes typically have $m_{\text{miss}}^2 > 0$, these channels interfere with the measurement of the signal modes; adding the constraints on the feed-down channels reduces this interference and improves our signal sensitivity.

8.9 D^* Efficiency Correction and Floating Constraints

The constraints described in Section 8.8 are derived from Monte Carlo samples and can be used when fitting the generic MC to produce an unbiased result. Studies in control samples in data, however, indicate that the MC overestimates the D^* reconstruction efficiency, an effect which is not seen in the D channels, suggesting that some of the crossfeed constraints need to be

Table 8.5: Constraints within the signal extraction fit. The errors shown are calculated using \sqrt{N} errors on the direct and crossfeed yields. Note that these errors are correlated since a single event yield (*i.e.*, $D^{*0}\ell^{-}\bar{\nu}_\ell \Rightarrow D^{*0}$) can appear in the denominator of several different constraint terms. This will be treated explicitly when we calculate systematic errors in Section 12.6.

Source	Direct channel	Crossfeed channel	Type	f_{ij}^i
$D^{*0}\tau^{-}\bar{\nu}_\tau$	D^{*0}	D^0	feed-down	0.9139 ± 0.0135
$D^{*0}\ell^{-}\bar{\nu}_\ell$	D^{*0}	D^0	feed-down	1.0326 ± 0.0163
$D^{*+}\tau^{-}\bar{\nu}_\tau$	D^{*+}	D^+	feed-down	0.2150 ± 0.0036
$D^{*+}\ell^{-}\bar{\nu}_\ell$	D^{*+}	D^+	feed-down	0.2304 ± 0.0048
$D^{**}\ell^{-}\bar{\nu}_\ell$	$D^0\pi^0$	D^0	feed-down	0.5844 ± 0.0210
$D^{**}\ell^{-}\bar{\nu}_\ell$	$D^{*0}\pi^0$	D^{*0}	feed-down	0.6015 ± 0.0329
$D^{**}\ell^{-}\bar{\nu}_\ell$	$D^+\pi^0$	D^+	feed-down	0.2705 ± 0.0215
$D^{**}\ell^{-}\bar{\nu}_\ell$	$D^{*+}\pi^0$	D^{*+}	feed-down	0.6049 ± 0.0331
$D^0\tau^{-}\bar{\nu}_\tau$	D^0	D^{*0}	feed-up	0.0818 ± 0.0022
$D^0\ell^{-}\bar{\nu}_\ell$	D^0	D^{*0}	feed-up	0.0616 ± 0.0026
$D^+\tau^{-}\bar{\nu}_\tau$	D^+	D^{*+}	feed-up	0.0273 ± 0.0018
$D^+\ell^{-}\bar{\nu}_\ell$	D^+	D^{*+}	feed-up	0.0182 ± 0.0014
$D^{*0}\ell^{-}\bar{\nu}_\ell$	D^{*0}	$D^0\pi^0$	feed-down and feed-up	0.0273 ± 0.0017
$D^0\ell^{-}\bar{\nu}_\ell$	D^0	$D^0\pi^0$	feed-up	0.0183 ± 0.0012
$D^{*0}\ell^{-}\bar{\nu}_\ell$	D^{*0}	$D^{*0}\pi^0$	feed-up	0.0187 ± 0.0014
$D^0\ell^{-}\bar{\nu}_\ell$	D^0	$D^{*0}\pi^0$	double feed-up	0.0017 ± 0.0004
$D^{*+}\ell^{-}\bar{\nu}_\ell$	D^{*+}	$D^+\pi^0$	feed-down and feed-up	0.0118 ± 0.0009
$D^+\ell^{-}\bar{\nu}_\ell$	D^+	$D^+\pi^0$	feed-up	0.0286 ± 0.0016
$D^{*+}\ell^{-}\bar{\nu}_\ell$	D^{*+}	$D^{*+}\pi^0$	feed-up	0.0147 ± 0.0010
$D^+\ell^{-}\bar{\nu}_\ell$	D^+	$D^{*+}\pi^0$	double feed-up	0.0006 ± 0.0002

modified in order to accurately describe data. In particular, six constraints are affected: two describing $D^*\tau^-\bar{\nu}_\tau$ feed-down into the D channels (one charged and one neutral), two describing $D^*\ell^-\bar{\nu}_\ell$ feed-down into the D channels, and two describing $D^*\ell^-\bar{\nu}_\ell$ feed-down-and-up into the $D\pi^0$ control samples. The feed-up constraints are unchanged since the effect we see appears to be purely related to the efficiency; the feed-down constraints from the D^{**} control samples into the signal channels are unchanged since the π^0 in the control sample has a minimum momentum of 400 MeV/ c and the efficiency loss is believed to only appear at low momentum (see also Section 12.8.5, where we discuss systematic effects related to this assumption).

We determine these six feed-down constraints from the data. The two constraints describing $D^*\ell^-\bar{\nu}_\ell$ feed-down are measured directly by floating them in the fit, while the other four are parameterized as linear functions of the dominant $D^*\ell^-\bar{\nu}_\ell$ constraint. We rely on the Monte Carlo only to relate these constraints to one another, allowing two floating parameters to describe six constraints, as in Equation 8.9 below.

We first make a crude estimate of this effect (which will be improved shortly) by applying a “killing” procedure to our Monte Carlo. We select a relatively-pure sample of $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ events by examining only the two D^* signal channels and selecting events with $|m_{\text{miss}}^2| < 0.7(\text{GeV}/c^2)^2$; neither

of our two previously used control samples (defined by selecting ranges of p_ℓ^* and q^2) can be used since their kinematics are slightly different from the signal region. For each of the four D^* modes reconstructed, we estimate an efficiency correction factor as the ratio of observed events in data to the number expected from MC (including all corrections to be discussed in Section 11). These ratios are found to be:

$$\begin{aligned} D^0\pi^0 & : 372 / 461.413 = 0.806 \\ D^0\gamma & : 335 / 389.373 = 0.860 \\ D^0\pi^+ & : 290 / 405.959 = 0.714 \\ D^+\pi^0 & : 76 / 103.176 = 0.708 . \end{aligned}$$

For D^* candidates reconstructed in signal and semileptonic MC, we randomly reject candidates (“killing”) with a probability given by (1– the efficiency correction above), between 14% and 29%. This killing is applied before the single-candidate selection takes place so that, if the D^* candidate is killed, the corresponding candidate in the D channel still has a chance to get selected. This killing procedure allows us to calculate new values for the six feed-down constraints that must be modified. The killed values of these constraints are shown in Table 8.6, along with the unkilld values for comparison.

This killing procedure, however, is likely to overestimate the actual correction needed to the feed-down constraints. We derived the killing probabilities using the ratio of event yields in data and MC; this ratio contains the D^* efficiency which we wish to correct for, but it also contains unknown efficiency

Table 8.6: Feed-down constraints modified by D^* killing. The unkilld values differ slightly from those listed in Table 8.5 because of the corrections described in Section 11; these corrected values are more appropriate for use with data.

Source	Direct channel	Crossfeed channel	f unkilld	f killd
$D^{*0}\tau^-\bar{\nu}_\tau$	D^{*0}	D^0	0.9387 ± 0.0139	1.2380 ± 0.0189
$D^{*+}\tau^-\bar{\nu}_\tau$	D^{*+}	D^+	0.2206 ± 0.0037	0.3234 ± 0.0055
$D^{*0}\ell^-\bar{\nu}_\ell$	D^{*0}	D^0	1.0404 ± 0.0166	1.3422 ± 0.0220
$D^{*+}\ell^-\bar{\nu}_\ell$	D^{*+}	D^+	0.2325 ± 0.0049	0.3439 ± 0.0073
$D^{*0}\ell^-\bar{\nu}_\ell$	D^{*0}	$D^0\pi^0$	0.0249 ± 0.0017	0.0378 ± 0.0022
$D^{*+}\ell^-\bar{\nu}_\ell$	D^{*+}	$D^+\pi^0$	0.0109 ± 0.0009	0.0186 ± 0.0014

corrections due to the B_{tag} , which will be common to both the D^* and D reconstruction channels and which will cancel when we study the feed-down constraint. We would like to apply the killing only partially, which we will do in a data-driven way by allowing constraints to float in the fit.

The dominant feed-down components in the analysis, $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0$ and $D^{*+}\ell^-\bar{\nu}_\ell \Rightarrow D^+$ have sufficient statistics and a distinctive shape that the two corresponding constraints can be floated in the fit to data. Doing this also gives us a sense of “how much killing” should be applied, and allows us to derive corrections to the remaining four feed-down constraints. We relate the two remaining B^+ feed-down constraints to that for $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0$ by linearly interpolating between the unkilld and killd values, and likewise for the two B^0 constraints and $D^{*+}\ell^-\bar{\nu}_\ell \Rightarrow D^+$. As a specific example, we have (letting

$f^\tau \equiv$ the feed-down constraint for $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0$ and $f^\ell \equiv$ the feed-down constraint for $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0$):

$$\begin{aligned}
 f_{\text{fit}}^\tau &\equiv f_{\text{unkilled}}^\tau + (f_{\text{killed}}^\tau - f_{\text{unkilled}}^\tau) \frac{f_{\text{fit}}^\ell - f_{\text{unkilled}}^\ell}{f_{\text{killed}}^\ell - f_{\text{unkilled}}^\ell} \\
 &= 0.9387 + 0.9917 \times (f_{\text{fit}}^\ell - 1.0404) .
 \end{aligned}
 \tag{8.9}$$

In this way, floating two constraints gives us a good estimate of the other four. This procedure of floating the constraints will be used in the fit to data, as well as for the fits used to estimate systematic uncertainties.

8.10 Fit Results on Generic Monte Carlo Events

To test the fit shapes and procedure, we run a full fit on the generic MC sample, containing about 640 fb^{-1} of simulated events. In this fit, the PDF shapes are all fixed as described above, so that the only free parameters in the fit are the numbers of events in each component in each channel and the two feed-down parameters.

There are 48 event yields in the complete fit. Since some of the combinatoric event yields (2 yields in each of the 4 signal channels, and 1 yield in each D^{**} control sample channel) are fixed (to the true values), there are only 36 floating yields in the fit. The 20 crossfeed constraints reduce the number

of independent event yields to 16, and the two floating feed-down parameters bring the total number of floating parameters in the fit to 18.

The numbers of true and fitted events in each channel modes are given in Table 8.7. In the event-yield table, errors and pulls for feed-up and feed-down yields are shown as “—”; the error and pull shown on the direct measurement reflects all the available information for each component.

We see that, for most components, the fit extracts the number of events quite well. A few channels have pulls greater than 1σ , but the overall distribution of pulls is not unreasonable.

One-dimensional projections of the fit are shown in Figures 8.41–8.46.

In 640 fb^{-1} of generic MC, we have true signal yields of 158, 283, 47, and 114 events in the $D^0\tau^-\bar{\nu}_\tau$, $D^{*0}\tau^-\bar{\nu}_\tau$, $D^+\tau^-\bar{\nu}_\tau$, and $D^{*+}\tau^-\bar{\nu}_\tau$ modes, respectively. In the Run 1–4 data sample, we therefore expect 53, 96, 16, and 38 signal events, respectively, and, scaling the errors from the fit, statistical significances of 2.3σ , 4.9σ , 2.3σ , and 4.9σ . We see that the D^* modes are more sensitive than the D modes, due largely to the fact that $D^* \Rightarrow D$ feeddown contributes significantly to the signal yield.

Table 8.7: Fit results from 640 fb^{-1} of generic MC. Both the true number of events and the result from the fit are given for each mode. In the table, errors and pulls for feed-up and feed-down yields are shown as “—”; the error and pull shown on the direct measurement reflects all the available information for each component.

	Component	N_{true}	N_{fit}	Pull		Component	N_{true}	N_{fit}	Pull
D^{*0}	$D^{*0}\tau^-\bar{\nu}_\tau$	135	144 ± 17	0.53	$D^{*0}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	187	175 ± 14	-0.79
	$D^{*0}\ell^-\bar{\nu}_\ell$	2798	2792 ± 53	-0.09		$D^{*0}\ell^-\bar{\nu}_\ell$	55	$52 \pm \text{—}$	—
	$D^0\tau^-\bar{\nu}_\tau$	16	$10 \pm \text{—}$	—		$D^0\ell^-\bar{\nu}_\ell$	4	$1 \pm \text{—}$	—
	$D^0\ell^-\bar{\nu}_\ell$	72	$68 \pm \text{—}$	—		Charge XF	3	3 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	97	$105 \pm \text{—}$	—		Comb. BG	55	63 ± 9	0.83
	Charge XF	35	35 ± 0	fixed					
	Comb. BG	44	44 ± 0	fixed					
D^0	$D^0\tau^-\bar{\nu}_\tau$	142	133 ± 33	-0.27	$D^0\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	465	481 ± 23	0.69
	$D^{*0}\tau^-\bar{\nu}_\tau$	153	$143 \pm \text{—}$	—		$D^{*0}\ell^-\bar{\nu}_\ell$	76	$80 \pm \text{—}$	—
	$D^0\ell^-\bar{\nu}_\ell$	1143	1116 ± 38	-0.67		$D^0\ell^-\bar{\nu}_\ell$	20	$20 \pm \text{—}$	—
	$D^{*0}\ell^-\bar{\nu}_\ell$	3010	$3037 \pm \text{—}$	—		Charge XF	18	18 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	270	$281 \pm \text{—}$	—		Comb. BG	128	120 ± 14	-0.52
	Charge XF	93	93 ± 0	fixed					
	Comb. BG	94	94 ± 0	fixed					
D^{*+}	$D^{*+}\tau^-\bar{\nu}_\tau$	92	111 ± 13	1.43	$D^{*+}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	75	73 ± 9	-0.17
	$D^{*+}\ell^-\bar{\nu}_\ell$	1695	1680 ± 41	-0.36		$D^{*+}\ell^-\bar{\nu}_\ell$	22	$24 \pm \text{—}$	—
	$D^+\tau^-\bar{\nu}_\tau$	4	$1 \pm \text{—}$	—		$D^+\ell^-\bar{\nu}_\ell$	1	$0 \pm \text{—}$	—
	$D^+\ell^-\bar{\nu}_\ell$	5	$7 \pm \text{—}$	—		Charge XF	0	0 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	50	$44 \pm \text{—}$	—		Comb. BG	17	18 ± 5	0.27
	Charge XF	2	2 ± 0	fixed					
	Comb. BG	27	27 ± 0	fixed					
D^+	$D^+\tau^-\bar{\nu}_\tau$	43	52 ± 13	0.70	$D^+\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	211	202 ± 15	-0.56
	$D^{*+}\tau^-\bar{\nu}_\tau$	22	$22 \pm \text{—}$	—		$D^{*+}\ell^-\bar{\nu}_\ell$	22	$18 \pm \text{—}$	—
	$D^+\ell^-\bar{\nu}_\ell$	427	425 ± 21	-0.06		$D^+\ell^-\bar{\nu}_\ell$	10	$12 \pm \text{—}$	—
	$D^{*+}\ell^-\bar{\nu}_\ell$	380	$355 \pm \text{—}$	—		Charge XF	4	4 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	50	$54 \pm \text{—}$	—		Comb. BG	94	112 ± 12	1.46
	Charge XF	9	9 ± 0	fixed					
	Comb. BG	31	31 ± 0	fixed					

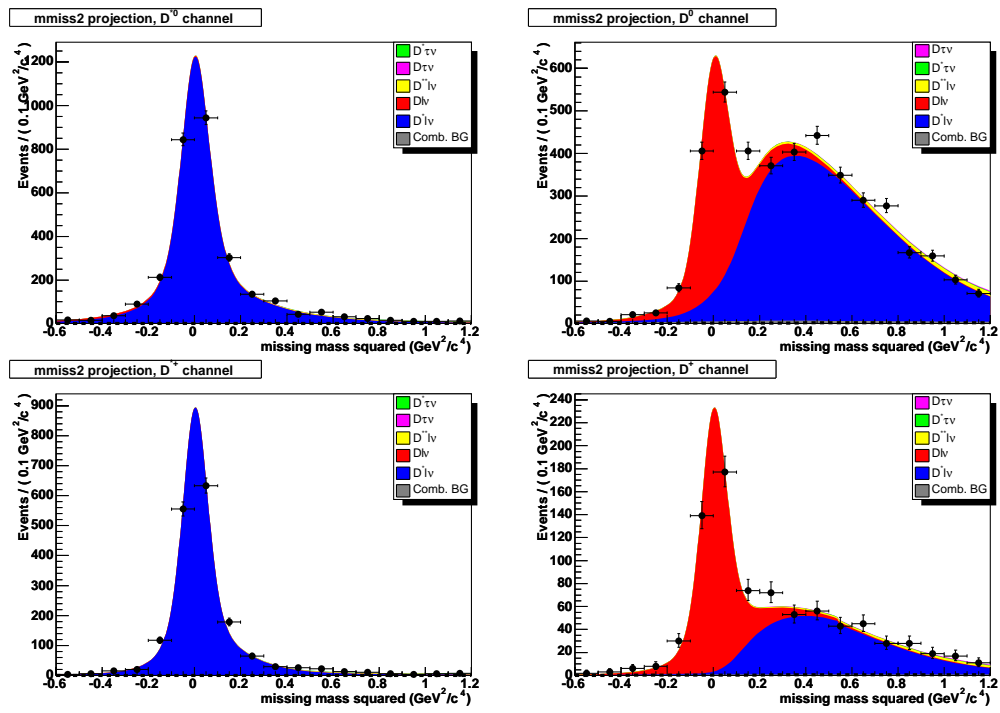


Figure 8.41: m_{miss}^2 projection of the fit to 640 fb⁻¹ of generic MC in the signal channels.

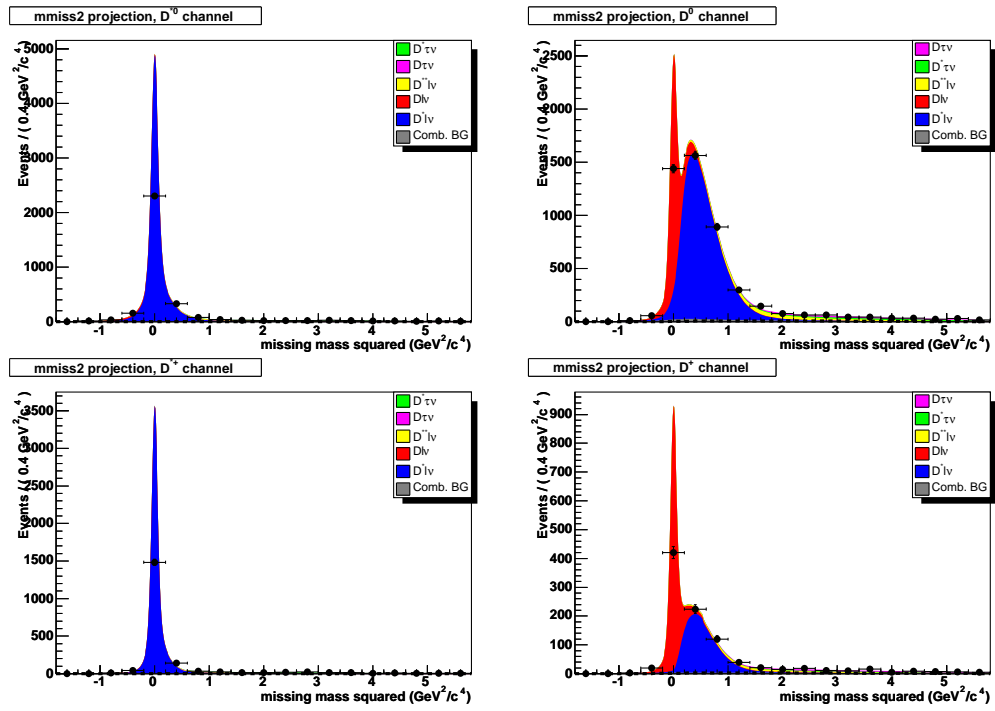


Figure 8.42: Same as Figure 8.41 (vertical zoom).

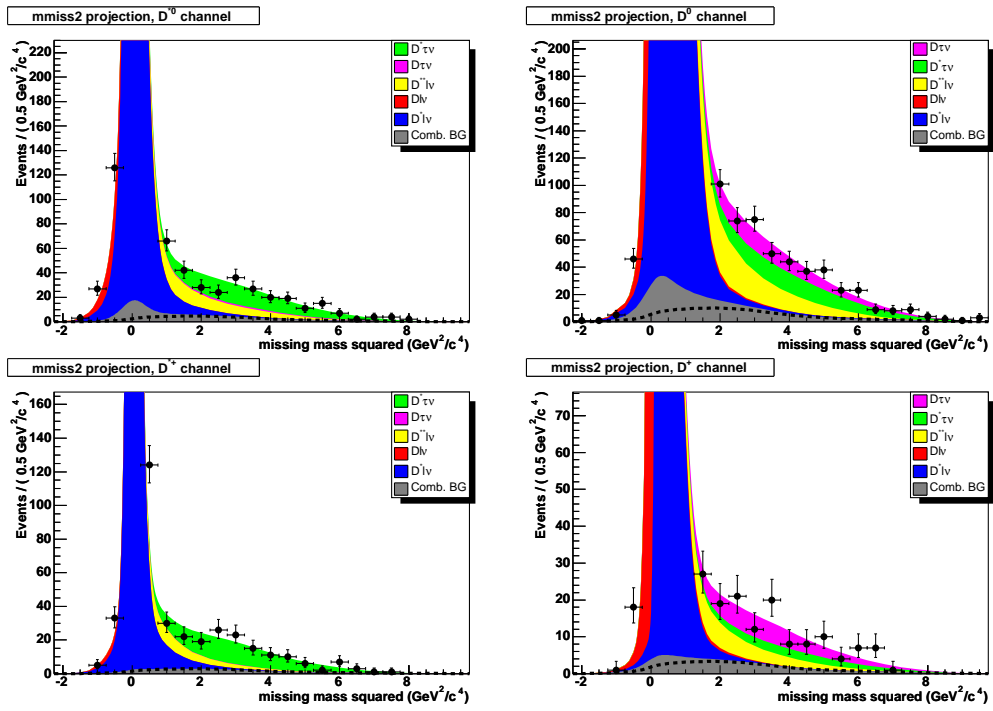


Figure 8.43: Same as Figure 8.41 (vertical zoom, horizontal scale expanded).

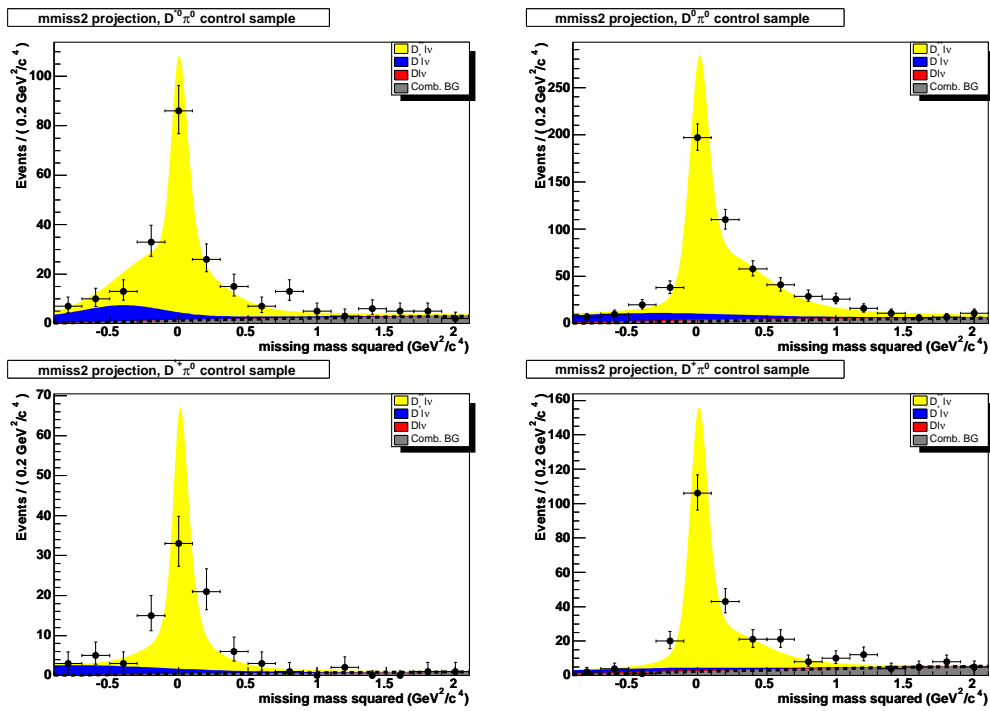


Figure 8.44: m_{miss}^2 projection of the fit to 640 fb^{-1} of generic MC in the D^{**} control sample.

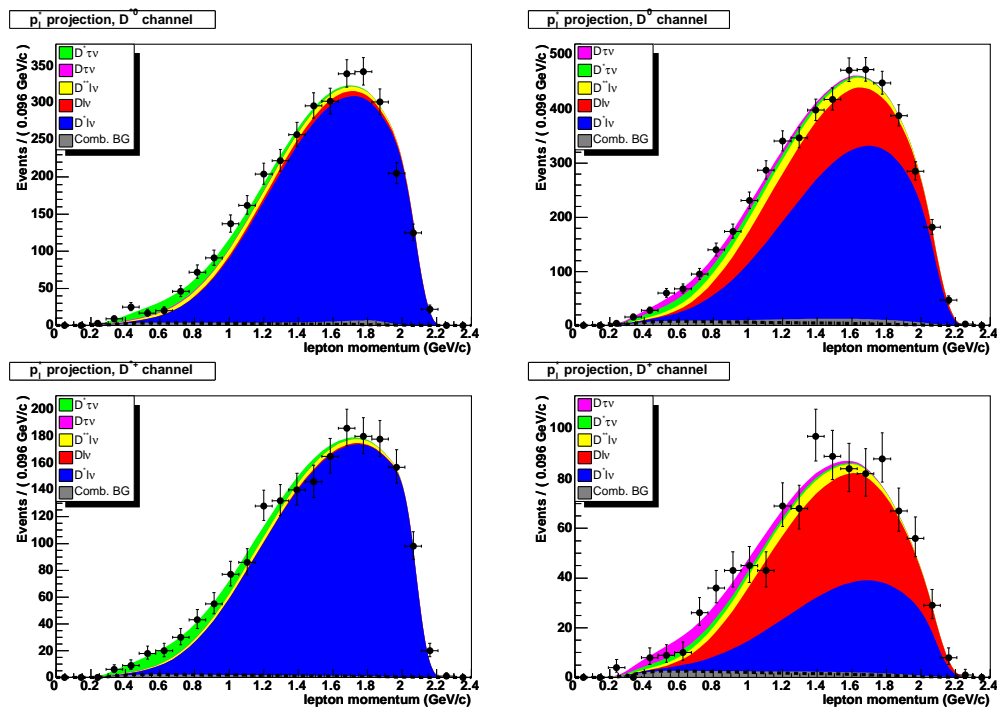


Figure 8.45: p_ℓ^* projection of the fit to 640 fb^{-1} of generic MC in the signal channels.

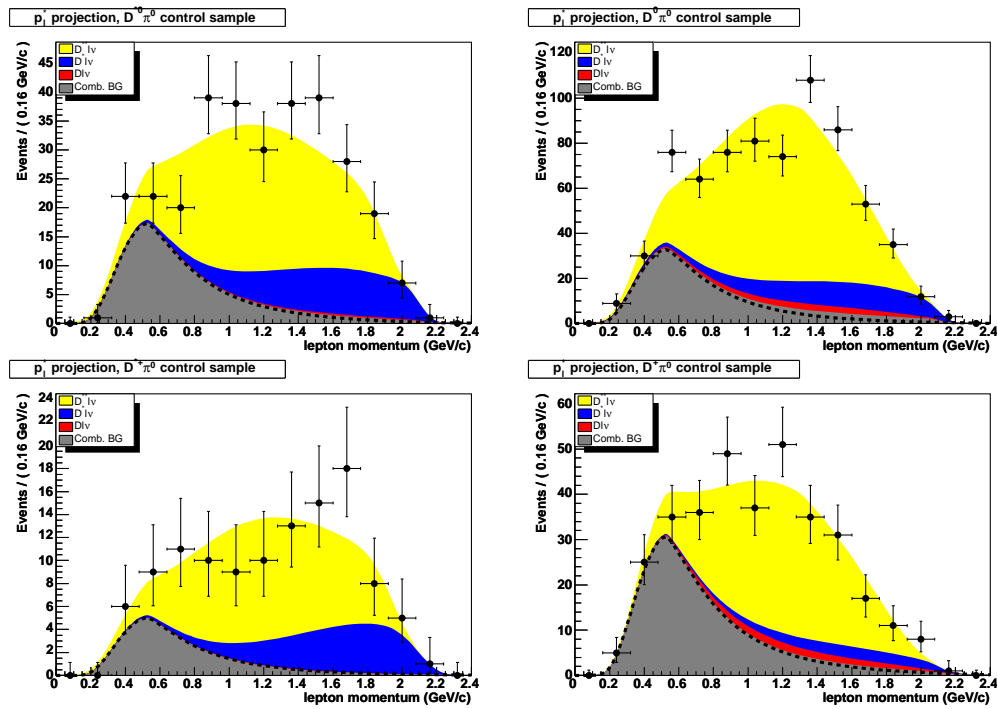


Figure 8.46: p_{ℓ}^* projection of the fit to 640 fb^{-1} of generic MC in the D^{**} control sample.

Chapter 9

Fit Validation

In order to ensure that the fit described in Section 8 is well-behaved, we have performed several validation studies.

9.1 Toy Monte Carlo Studies

We generate a series of 1000 toy MC experiments using the same PDFs as in the fit described above. In all experiments, we generate a number of events equal to the number in the 970 fb^{-1} generic MC sample, with the same expected relative distribution among the various channels and components. We then fit these toy samples using the same fit procedure as before.

Figures 9.1–9.4 show distributions of N_{fitted} and pulls, defined as $(N_{\text{fitted}} - N_{\text{expected}})/\sigma_{\text{fitted}}$ for the four signal channels.

In all four channels, we see that the fit is nearly unbiased — the widths of the pull distributions are all consistent with 1.0 and the mean values are small, with the largest bias being only 0.08. This implies that our fit *procedure*, including the simultaneous fit to multiple channels with crossfeed constraints, does not intrinsically bias the result. The actual fit may still be biased, particularly if the PDFs used do not accurately model real data; this will be considered below.

We have repeated these toy MC studies with a series of smaller experiments, generating and fitting samples equivalent to 200 fb^{-1} of generic MC, approximately equal to the expected number of events in the Run 1–4 data sample. The results of these studies are consistent with those from the larger toys — the means of the four pull distributions are in the range -0.04 – 0.02 , all consistent with zero, and the widths of the pull distributions are in the range 0.92 – 1.03 , all consistent with one.

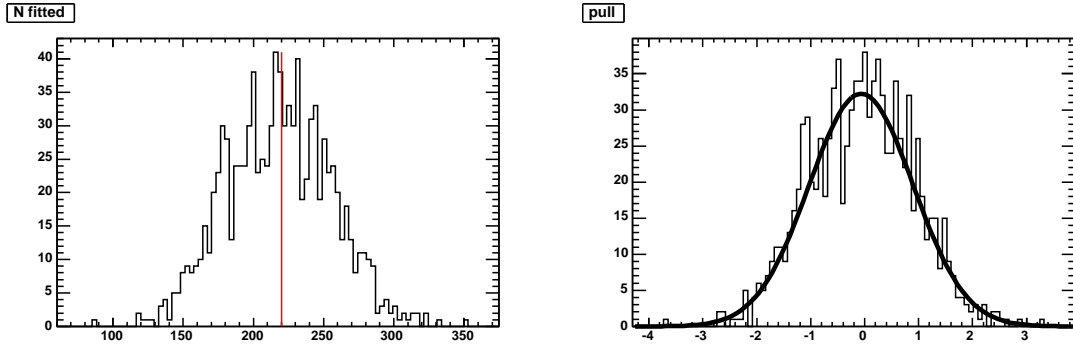


Figure 9.1: Toy MC results for $D^0\tau\nu$. The left plot shows the distribution of N_{fitted} , where the expected value is 220 events. The right plot shows the pull distribution, with a fit to a Gaussian ($\mu = -0.065 \pm 0.033$, $\sigma = 0.965 \pm 0.025$).

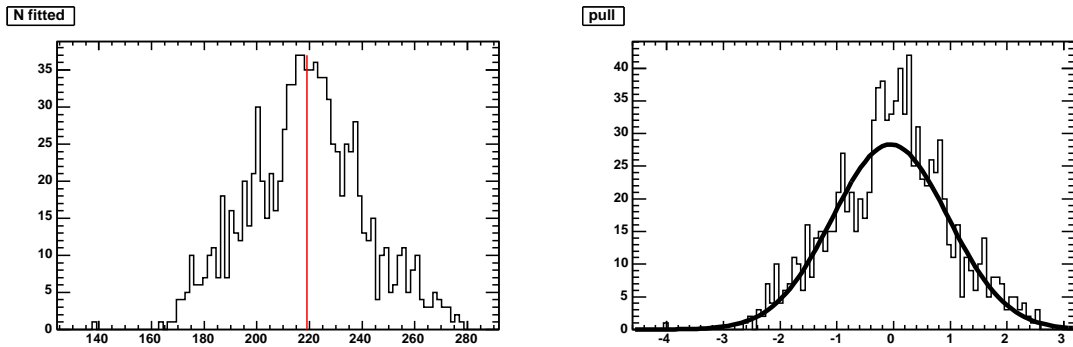


Figure 9.2: Toy MC results for $D^{*0}\tau\nu$. The left plot shows the distribution of N_{fitted} , where the expected value is 219 events. The right plot shows the pull distribution, with a fit to a Gaussian ($\mu = -0.056 \pm 0.035$, $\sigma = 1.019 \pm 0.034$).

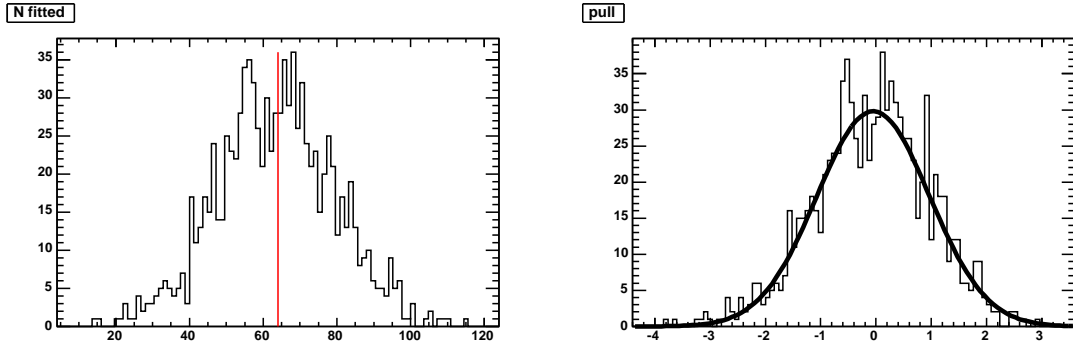


Figure 9.3: Toy MC results for $D^+\tau\nu$. The left plot shows the distribution of N_{fitted} , where the expected value is 64 events. The right plot shows the pull distribution, with a fit to a Gaussian ($\mu = -0.050 \pm 0.034$, $\sigma = 1.018 \pm 0.028$).

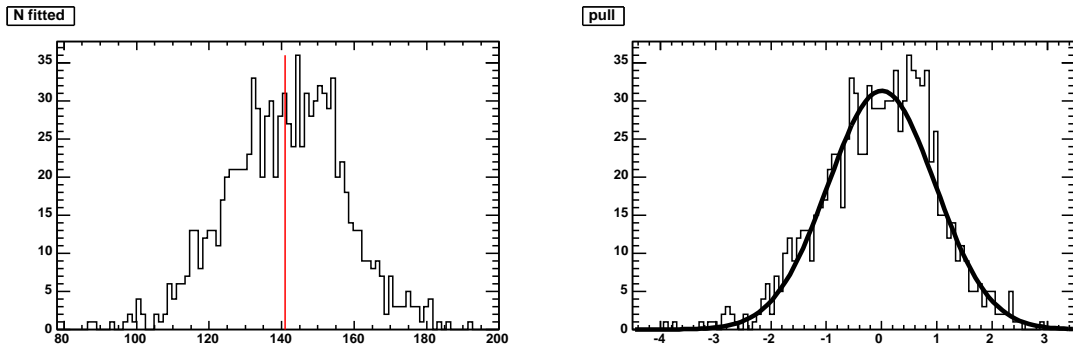


Figure 9.4: Toy MC results for $D^{*+}\tau\nu$. The left plot shows the distribution of N_{fitted} , where the expected value is 141 events. The right plot shows the pull distribution, with a fit to a Gaussian ($\mu = 0.003 \pm 0.033$, $\sigma = 0.970 \pm 0.026$).

9.2 Validation with Statistically Independent Monte Carlo Samples

We have repeatedly used a single 640 fb^{-1} sample of generic Monte Carlo for cut optimization, combinatoric BG shapes and normalizations, and throughout the development of the signal fit.

We have a sample of 330 fb^{-1} of generic SP6 Monte Carlo data which was set aside and never used during cut optimization or fit development. We use this sample as a statistically independent validation of our event selection criteria and our fit procedure. The results of this validation fit are shown in Table 9.1. No significant biases are seen, giving us additional confidence that our fit is inherently unbiased. The results of this validation fit are largely consistent with the fit to the original 640 fb^{-1} sample in terms of the expected numbers of signal and background events, as well as the expected statistical sensitivity of the fit.

Table 9.1: Fit validation from 330 fb^{-1} of generic MC. Both the true number of events and the result from the fit are given for each mode. This fit shows no significant signal biases.

Channel	Component	N_{true}	N_{fit}	Pull	Channel	Component	N_{true}	N_{fit}	Pull
D^{*0}	$D^{*0}\tau^-\bar{\nu}_\tau$	87	82 ± 13	-0.33	$D^{*0}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	97	93 ± 10	-0.34
	$D^{*0}\ell^-\bar{\nu}_\ell$	1540	1530 ± 40	-0.23		$D^{*0}\ell^-\bar{\nu}_\ell$	31	$28 \pm \text{---}$	---
	$D^0\tau^-\bar{\nu}_\tau$	5	$6 \pm \text{---}$	---		$D^0\ell^-\bar{\nu}_\ell$	4	$1 \pm \text{---}$	---
	$D^0\ell^-\bar{\nu}_\ell$	46	$40 \pm \text{---}$	---		Charge XF	2	2 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	49	$56 \pm \text{---}$	---		Comb. BG	59	69 ± 9	1.13
	Charge XF	16	16 ± 0	fixed					
	Comb. BG	22	22 ± 0	fixed					
D^0	$D^0\tau^-\bar{\nu}_\tau$	79	83 ± 24	0.17	$D^0\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	264	262 ± 17	-0.09
	$D^{*0}\tau^-\bar{\nu}_\tau$	75	$71 \pm \text{---}$	---		$D^{*0}\ell^-\bar{\nu}_\ell$	36	$36 \pm \text{---}$	---
	$D^0\ell^-\bar{\nu}_\ell$	652	659 ± 29	0.25		$D^0\ell^-\bar{\nu}_\ell$	10	$12 \pm \text{---}$	---
	$D^{*0}\ell^-\bar{\nu}_\ell$	1493	$1472 \pm \text{---}$	---		Charge XF	6	6 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	132	$153 \pm \text{---}$	---		Comb. BG	67	68 ± 11	0.16
	Charge XF	53	53 ± 0	fixed					
	Comb. BG	57	57 ± 0	fixed					
D^{*+}	$D^{*+}\tau^-\bar{\nu}_\tau$	50	45 ± 9	-0.42	$D^{*+}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	48	49 ± 7	0.23
	$D^{*+}\ell^-\bar{\nu}_\ell$	897	903 ± 30	0.22		$D^{*+}\ell^-\bar{\nu}_\ell$	17	$13 \pm \text{---}$	---
	$D^+\tau^-\bar{\nu}_\tau$	1	$0 \pm \text{---}$	---		$D^+\ell^-\bar{\nu}_\ell$	0	$0 \pm \text{---}$	---
	$D^+\ell^-\bar{\nu}_\ell$	6	$3 \pm \text{---}$	---		Charge XF	0	0 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	32	$30 \pm \text{---}$	---		Comb. BG	15	15 ± 4	0.19
	Charge XF	2	2 ± 0	fixed					
	Comb. BG	17	17 ± 0	fixed					
D^+	$D^+\tau^-\bar{\nu}_\tau$	21	26 ± 9	0.55	$D^+\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	121	115 ± 11	-0.49
	$D^{*+}\tau^-\bar{\nu}_\tau$	11	$8 \pm \text{---}$	---		$D^{*+}\ell^-\bar{\nu}_\ell$	12	$8 \pm \text{---}$	---
	$D^+\ell^-\bar{\nu}_\ell$	187	191 ± 14	0.30		$D^+\ell^-\bar{\nu}_\ell$	7	$5 \pm \text{---}$	---
	$D^{*+}\ell^-\bar{\nu}_\ell$	186	$176 \pm \text{---}$	---		Charge XF	3	3 ± 0	fixed
	$D^{**}\ell^-\bar{\nu}_\ell$	31	$31 \pm \text{---}$	---		Comb. BG	58	70 ± 9	1.28
	Charge XF	6	6 ± 0	fixed					
	Comb. BG	16	16 ± 0	fixed					

Chapter 10

Signal Extraction

In Section 8, we presented a fit which measures the number of signal events in a given data sample. In this Section, we describe how this information is used to extract the branching fractions of $B \rightarrow D^{(*)}\tau^{-}\bar{\nu}_{\tau}$.

10.1 Fitting Branching Ratios

The fit described above parameterizes the signal yields as absolute numbers of events N_{ij} . Because we wish to perform a relative measurement, we

introduce a change of variables in the fit, defining four new quantities:

$$S^0 \equiv N(D^0 \tau^- \bar{\nu}_\tau) / N(D^0 \ell^- \bar{\nu}_\ell)$$

$$S^{*0} \equiv N(D^{*0} \tau^- \bar{\nu}_\tau) / N(D^{*0} \ell^- \bar{\nu}_\ell)$$

$$S^+ \equiv N(D^+ \tau^- \bar{\nu}_\tau) / N(D^+ \ell^- \bar{\nu}_\ell)$$

$$S^{*+} \equiv N(D^{*+} \tau^- \bar{\nu}_\tau) / N(D^{*+} \ell^- \bar{\nu}_\ell) .$$

These quantities replace the absolute signal yields in the fit, but do not actually change the likelihood. Fitting in this way propagates the statistical error on both the signal and normalization (correctly handling the correlation between the two), and makes it easier to perform a single fit to combine the B^+ and B^0 channels, as described below.

10.2 Relative Signal Efficiency and Normalization

We convert the relative signal yields S into branching ratios R as

$$\begin{aligned} R^0 &= S^0 / \varepsilon^0 \cdot \mathcal{B}_\tau \\ R^{*0} &= S^{*0} / \varepsilon^{*0} \cdot \mathcal{B}_\tau \\ R^+ &= S^+ / \varepsilon^+ \cdot \mathcal{B}_\tau \\ R^{*+} &= S^{*+} / \varepsilon^{*+} \cdot \mathcal{B}_\tau \end{aligned}$$

where the ε terms represent the relative signal efficiency in each of the four signal modes and \mathcal{B}_τ is the leptonic τ branching fraction, $\mathcal{B}_\tau \equiv \mathcal{B}(\tau^- \rightarrow$

Table 10.1: Normalization branching fractions of $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell$. The first four are taken from PDG 2006 [10], while the last two are our own averages (using the lifetime ratio, also from PDG 2006, $\tau_B^+/\tau_B^0 = 1.071 \pm 0.009$).

Mode	$\mathcal{B}(\%)$
$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$	2.15 ± 0.22
$B^- \rightarrow D^{*0} \ell^- \bar{\nu}_\ell$	6.5 ± 0.5
$\bar{B}^0 \rightarrow D^+ \ell^- \bar{\nu}_\ell$	2.12 ± 0.20
$\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$	5.35 ± 0.20
$B^- \rightarrow D^+ \ell^- \bar{\nu}_\ell$ (isospin averaged)	2.07 ± 0.14
$B^- \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$ (isospin averaged)	5.46 ± 0.18

$\ell^- \bar{\nu}_\ell \nu_\tau$) = 17.6% [10]. The ratios R are defined (following the convention introduced in Section 1) such that

$$\mathcal{B}(B \rightarrow D^i \tau^- \bar{\nu}_\tau) \equiv R^i \times \mathcal{B}(B \rightarrow D^i \ell^- \bar{\nu}_\ell) . \quad (10.1)$$

These ratios R are physically interesting and can be used to constrain the possible new-physics contribution. They also allow us to present absolute branching fractions by normalizing to the known branching fractions, which are shown in Table 10.1.

The efficiencies are measured in signal and semileptonic MC as

$$\begin{aligned} \varepsilon^0 &= (187 \pm 2.3)\% \\ \varepsilon^{*0} &= (110 \pm 1.2)\% \\ \varepsilon^+ &= (185 \pm 2.7)\% , \\ \varepsilon^{*+} &= (102 \pm 1.1)\% \end{aligned}$$

where the statistical errors here have already been shown in Table 12.1. Note that the relative efficiencies are all greater than one, meaning that a signal event is more likely to be reconstructed than a normalization event.

10.3 Combining Results between Charged and Neutral Channels

To maximize our experimental sensitivity, we also perform a fit in which we combine the B^- and \bar{B}^0 channels, effectively performing an isospin average of the separate measurements. We do this by introducing two additional constraints,

$$S^+ \equiv f_{\text{iso}} S^0$$

$$S^{*+} \equiv f_{\text{iso}}^* S^{*0} ,$$

where the terms $f_{\text{iso}}^{(*)}$ are defined to take into account the differences in efficiency between the charged and neutral channels. We have

$$\begin{aligned} f_{\text{iso}} &\equiv \varepsilon^+ / \varepsilon^0 = (0.99 \pm 0.02) \\ f_{\text{iso}}^* &\equiv \varepsilon^{*+} / \varepsilon^{*0} = (0.92 \pm 0.01) . \end{aligned}$$

Note that constraining the S terms in this way is equivalent to constraining $R^0 \equiv R^+$ and $R^{*0} \equiv R^{*+}$, and so we will also refer to this as the R -constrained fit. This form of the constraint follows from assuming isospin symmetry in both the signal and normalization modes, although it is actually slightly more general: it allows isospin to be violated as long as it is violated in the same way for $B \rightarrow D^{(*)} \ell^- \bar{\nu}_\ell$ and $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ decays.

Table 10.2: Branching fraction extraction in 970 fb^{-1} of generic Monte Carlo. Errors shown are statistical only.

Channel	S (%)	S_{true}	R (%)	R_{true}	\mathcal{B} (%)	$\mathcal{B}_{\text{true}}$
$D^0 \tau^- \bar{\nu}_\tau$	12.2 ± 2.3	12.3	37.0 ± 7.1	33.3	0.78 ± 0.15	0.70
$D^{*0} \tau^- \bar{\nu}_\tau$	5.2 ± 0.5	5.1	27.1 ± 2.7	28.6	1.52 ± 0.15	1.60
$D^+ \tau^- \bar{\nu}_\tau$	12.7 ± 2.7	10.4	39.1 ± 8.2	33.3	0.82 ± 0.17	0.70
$D^{*+} \tau^- \bar{\nu}_\tau$	6.1 ± 0.7	5.5	34.0 ± 3.7	28.6	1.90 ± 0.21	1.60

We fit for S^0 and S^{*0} as above and normalize the result to the isospin averages listed in Table 10.1.

10.4 Validation using the Generic Monte Carlo Sample

We test the procedure outlined above by measuring the signal branching fractions in generic Monte Carlo events, for which the result is known. We use the full 970 fb^{-1} sample and perform two fits, with and without the isospin constraint.

If we fit without the R -constraint, we obtain the results seen in Table 10.2, while, using the R -constraint, we obtain the results in Table 10.3. In both cases, we see that we extract the correct branching fractions within the statistical errors.

Table 10.3: Branching fraction extraction with the R -constraint in 970 fb^{-1} of generic Monte Carlo. Errors shown are statistical only.

Channel	S (%)	R (%)	R_{true}	\mathcal{B} (%)	$\mathcal{B}_{\text{true}}$
$D^0 \tau^- \bar{\nu}_\tau$	12.0 ± 1.7	36.4 ± 5.2	33.3	0.77 ± 0.11	0.70
$D^{*0} \tau^- \bar{\nu}_\tau$	5.7 ± 0.4	29.3 ± 2.1	28.6	1.64 ± 0.12	1.60

Chapter 11

Corrections to the Monte Carlo

We apply a number of correction factors to the *BABAR* Monte Carlo simulation. These corrections are designed to make the simulation more realistic, and more suitable for describing real data. We apply corrections to both the simulation of physical processes (D^{**} modeling, form factors) and detector efficiency (track reconstruction, PID, etc.).

Each correction factor is applied as an event-by-event weight (or, where appropriate, a track-by-track or π^0 -by- π^0 weight) to the MC; the total weight for an event is the product of each individual weight. Reweighted Monte Carlo samples will be used to generate fit shapes and crossfeed constraints when fitting to data; reweighted samples will also be used to calculate the signal efficiencies for the final branching fraction measurements.

We describe below the various correction factors. The same reweighting will also be used in Section 12 to estimate some systematic errors.

Table 11.1: Generated and reweighted D^{**} branching fractions.

Mode	Generated \mathcal{B} (%)	Reweighted \mathcal{B} (%)	
	$B^+ = B^0$	B^0	B^+
D_1	0.56	0.47	0.58
D_2^*	0.37	0.35	0.46
D_0^*	0.20	0.46	0.45
D_1'	0.37	0.85	0.83
$D^*\pi^0$	0.1	0.03	0.029
$D\pi^0$	0.3	0.09	0.088
$D^*\pi^\pm$	0.2	0.06	0.058
$D\pi^\pm$	0.6	0.18	0.175

11.1 D^{**} Modeling

We reweight the branching fractions of the individual D^{**} states as recommended by the AWG [42]. Table 11.1 lists the generated and reweighted branching fractions for the various $B \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ processes. Note that the relative proportion of the four non-resonant states has been kept fixed, while adjusting the total non-resonant rate to saturate the inclusive $b \rightarrow c\ell^-\bar{\nu}_\ell$ branching fraction.

11.2 $B \rightarrow D^{(*)}$ Form Factors

As discussed in Section 4.3, $B \rightarrow D^*\ell^-\bar{\nu}_\ell$ decays are generated using a linear expansion of the HQET form factors and values measured by CLEO,

while $B \rightarrow D\ell^-\bar{\nu}_\ell$ and $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ decays are generated using the ISGW2 model. We reweight these events such that the resulting distribution is consistent with the CLN form factor model, using the recent *BABAR* measurement of the $B \rightarrow D^*$ form factors and the HFAG average for the $B \rightarrow D$ form factor.

Reweightings are performed using the *XslFFReweighting* package [43]. The weights are calculated event-by-event as a function of the four kinematic variables q^2 , θ_ℓ , θ_V , and χ for $B \rightarrow D^*$ decays, or of the two variables q^2 and θ_ℓ for $B \rightarrow D$ decays.

11.3 Detector Efficiency Corrections

11.3.1 Tracking Efficiency

Tracking efficiency corrections have been studied in detail by the Tracking Efficiency Task Force [45]. Following the recommendations of the task force, we apply three sets of efficiency corrections:

- For tracks found in the *GoodTracksLoose* list, we apply the binned, track-by-track correction.
- For tracks found in *GoodTracksVeryLoose* but not in *GoodTracksLoose*, we apply a flat efficiency correction of 0.995 per track (-0.5%).

- For *ChargedTracks* not found in either *GoodTracksLoose* or *GoodTracksVeryLoose*, we apply a flat efficiency correction of 0.9975 per track (-0.25%).

The total tracking efficiency correction for an event is the product of the individual factors over all tracks in the signal side, excluding those from $K_s^0 \rightarrow \pi^+\pi^-$ (a separate K_s^0 correction will be applied below).

11.3.2 PID Efficiency

The PID group has compared the performance of the various selectors in data and Monte Carlo and prepared a set of efficiency correction factors [46]. We apply these PID weights, and the total efficiency correction is the product of the individual correction factors for all tracks in the signal side, excluding those from $K_s^0 \rightarrow \pi^+\pi^-$ and the soft pion from $D^{*+} \rightarrow D^0\pi_s$ (these tracks are not required to pass any PID selection).

11.3.3 π^0 Efficiency

The Neutral AWG has studied the π^0 efficiency and prepared a correction factor which is a linear function of the (laboratory frame) π^0 momentum [47]:

$$w_{\pi^0} = 0.977 + \frac{0.00591}{\text{GeV}/c} p_{\pi^0} . \quad (11.1)$$

We apply these correction factors, and the total correction is the product over all π^0 mesons in the signal side.

We reweight the inefficiency at the same time. Because crossfeed (especially feed-down due to lost π^0 mesons) plays such a large role in this analysis, we apply a correction factor to feed-down events (equal to one over the efficiency correction) using the π^0 momentum taken from the true (generated) event.

11.3.4 K_s^0 Efficiency

The Tracking Efficiency Task Force provides a recipe for correcting the K_s^0 efficiency. This recipe is not provided as an event-by-event weight, but, rather, as an offline correction. As-is, the recipe provides a correction to absolute reconstruction efficiency; we apply the correction to both our signal MC samples and the normalization modes $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ to correct the relative efficiency.

11.4 Corrected Crossfeed constraints

Table 11.2 shows the crossfeed constraints obtained after applying all of the corrections described above. These numbers can be compared with Table 8.5, which shows the Monte Carlo constraints without corrections.

Table 11.2: Constraints within the signal extraction fit after all corrections. These numbers can be compared with Table 8.5, which shows the Monte Carlo constraints without corrections.

Source	Direct channel	Crossfeed channel	Type	f_{ij}^i
$D^{*0} \tau^- \bar{\nu}_\tau$	D^{*0}	D^0	feed-down	0.9446 ± 0.0141
$D^{*0} \ell^- \bar{\nu}_\ell$	D^{*0}	D^0	feed-down	1.0397 ± 0.0165
$D^{*+} \tau^- \bar{\nu}_\tau$	D^{*+}	D^+	feed-down	0.2254 ± 0.0038
$D^{*+} \ell^- \bar{\nu}_\ell$	D^{*+}	D^+	feed-down	0.2328 ± 0.0049
$D^{**} \ell^- \bar{\nu}_\ell$	$D^0 \pi^0$	D^0	feed-down	0.7479 ± 0.0301
$D^{**} \ell^- \bar{\nu}_\ell$	$D^{*0} \pi^0$	D^{*0}	feed-down	0.6558 ± 0.0396
$D^{**} \ell^- \bar{\nu}_\ell$	$D^+ \pi^0$	D^+	feed-down	0.3198 ± 0.0294
$D^{**} \ell^- \bar{\nu}_\ell$	$D^{*+} \pi^0$	D^{*+}	feed-down	0.7136 ± 0.0392
$D^0 \tau^- \bar{\nu}_\tau$	D^0	D^{*0}	feed-up	0.0860 ± 0.0024
$D^0 \ell^- \bar{\nu}_\ell$	D^0	D^{*0}	feed-up	0.0638 ± 0.0027
$D^+ \tau^- \bar{\nu}_\tau$	D^+	D^{*+}	feed-up	0.0292 ± 0.0019
$D^+ \ell^- \bar{\nu}_\ell$	D^+	D^{*+}	feed-up	0.0192 ± 0.0015
$D^{*0} \ell^- \bar{\nu}_\ell$	D^{*0}	$D^0 \pi^0$	feed-down and feed-up	0.0261 ± 0.0017
$D^0 \ell^- \bar{\nu}_\ell$	D^0	$D^0 \pi^0$	feed-up	0.0174 ± 0.0012
$D^{*0} \ell^- \bar{\nu}_\ell$	D^{*0}	$D^{*0} \pi^0$	feed-up	0.0186 ± 0.0014
$D^0 \ell^- \bar{\nu}_\ell$	D^0	$D^{*0} \pi^0$	double feed-up	0.0016 ± 0.0003
$D^{*+} \ell^- \bar{\nu}_\ell$	D^{*+}	$D^+ \pi^0$	feed-down and feed-up	0.0117 ± 0.0009
$D^+ \ell^- \bar{\nu}_\ell$	D^+	$D^+ \pi^0$	feed-up	0.0275 ± 0.0016
$D^{*+} \ell^- \bar{\nu}_\ell$	D^{*+}	$D^{*+} \pi^0$	feed-up	0.0149 ± 0.0010
$D^+ \ell^- \bar{\nu}_\ell$	D^+	$D^{*+} \pi^0$	double feed-up	0.0007 ± 0.0002

Chapter 12

Systematic Uncertainties

12.1 Systematic Error Overview

Because our signal is extracted and normalized relative to $B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_{\ell}$, many sources of systematic error — especially those related to detector efficiency, D branching fractions, and B counting — are expected to cancel, either partially or completely, when we take the ratio. To calculate these systematic errors, we generally follow standard *BABAR* procedures to vary both the signal and normalization efficiencies simultaneously (in a correlated manner), and to study the effect on the relative efficiency.

Systematic sources which are not expected to cancel are those which affect the fit yields. Our strategy here will be to study systematic effects by performing an ensemble of fits. In each “experiment,” we will modify, as appropriate, the fit shapes, crossfeed constraints, and the combinatoric background yields

(all of which are fixed to MC-derived values in the nominal fit), and re-do the complete fit. By doing a large number of such experiments and studying the distribution of signal yields in these ensembles, we estimate the systematic errors due to these sources.

Table 12.1 gives a summary of the expected systematic errors in each channel. Individual contributions are described in the sections below.

12.2 $D^{**}\ell^{-}\bar{\nu}_\ell$ Modeling

The AWG recommendations for D^{**} branching fraction reweighting (see Section 11.1) also specify ranges of variation for systematic error studies. Following these guidelines, we generate an ensemble of D^{**} models, sampling from the distribution in Table 12.2.

To generate each individual D^{**} reweighting model, we follow the procedure outlined here, separately for B^+ and B^0 decays. First, we generate four uncorrelated random numbers, representing the $D\ell^{-}\bar{\nu}_\ell$, $D^*\ell^{-}\bar{\nu}_\ell$, $D_1\ell^{-}\bar{\nu}_\ell$, and $D_2^*\ell^{-}\bar{\nu}_\ell$ branching fractions. Each is sampled from a Gaussian distribution; for D_1 and D_2^* , if a negative branching fraction is generated, we set the branching fraction to zero (the other two distributions have negligible probabilities of generating negative branching fractions). If, after generating these four num-

Table 12.1: Contributions to the total systematic error. The numbers above the line represent uncertainties on the fit yield, and therefore reduce the statistical significance of the results. The numbers below the line represent uncertainties on the normalization, so they affect the numerical results but not the statistical significance. The first four columns summarize errors on the individual branching fractions; the last two columns summarize errors on the R -constrained measurement. The totals here refer to errors on the branching ratios R ; the error on $\mathcal{B}(B \rightarrow D^{(*)}\ell^{-}\bar{\nu}_\ell)$ only applies to the absolute branching fractions, and are not included in the quoted total error.

Source	Fractional error (%)					
	$D^0\tau\nu$	$D^{*0}\tau\nu$	$D^+\tau\nu$	$D^{*+}\tau\nu$	$D\tau\nu$	$D^*\tau\nu$
MC stat. (PDF shape)	11.5	8.4	4.5	1.8	6.9	4.7
MC stat. (constraints)	4.2	1.9	6.1	1.3	3.6	1.4
Comb. BG modeling	7.5	4.1	11.5	2.6	9.1	2.9
D^{**} modeling	5.7	0.5	1.6	0.2	3.0	0.4
$B \rightarrow D^*$ form factors	1.9	0.7	0.8	0.2	1.4	0.4
$B \rightarrow D$ form factors	0.2	0.7	0.6	0.2	0.3	0.4
π^0 crossfeed constraints	0.5	1.1	0.5	0.9	0.5	1.0
D^{**} feed-down	0.4	0.1	0.1	0.3	0.2	0.2
$D^{**}\tau^-\bar{\nu}_\tau$ abundance	0.4	1.3	0.3	0.2	0.3	0.8
m_{miss}^2 tail modeling	1.5	0.5	1.2	0.4	1.6	0.1
MC stat. (efficiency)	1.23	1.09	1.47	1.05	0.96	0.76
Bremsstrahlung/FSR	0.55	0.51	0.26	0.42	0.40	0.47
Tracking ε	0.01	0.02	0.002	0.03	0.01	0.02
e PID ε	0.53	0.54	0.60	0.57	0.61	0.56
μ PID ε	0.53	0.60	0.66	0.58	0.59	0.59
K PID ε	0.15	0.05	0.22	0.03	0.18	0.04
π PID ε	0.07	0.07	0.17	0.04	0.12	0.05
K_s^0 ε	0.07	0.00	0.07	0.08	0.07	0.04
Neutral (π^0 and γ) ε	0.01	0.04	0.02	0.05	0.02	0.04
Daughter \mathcal{B} 's	0.07	0.27	0.04	0.08	0.08	0.30
$\mathcal{B}(\tau^- \rightarrow \ell^-\bar{\nu}_\ell\nu_\tau)$	0.2	0.2	0.2	0.2	0.2	0.2
Total additive	15.6	9.7	14.0	3.6	12.5	5.8
Total multiplicative	1.60	1.49	1.77	1.42	1.38	1.26
Total	15.6	9.9	14.0	3.9	12.5	6.0
$\mathcal{B}(B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell)$	10.2	7.7	9.4	3.7	6.8	3.4

Table 12.2: Ranges of variation for reweighted D^{**} branching fractions. Distributions are taken to be Gaussian. The ranges of variation on nonresonant states are not shown in this table; their distribution is determined by the inclusive rate and the other exclusive modes, as described in the text.

Mode	Generated \mathcal{B} (%)	Reweighted B^0 \mathcal{B} (%)		Reweighted B^+ \mathcal{B} (%)	
	$B^+ = B^0$	μ	σ	μ	σ
D	2.10	2.14	0.14	2.29	0.16
D^*	5.6	5.54	0.25	5.94	0.24
D_1	0.56	0.47	0.08	0.58	0.06
D_2^*	0.37	0.35	0.07	0.46	0.08
D_0^*	0.20	0.46	0.09	0.45	0.09
D_1'	0.37	0.85	0.20	0.83	0.20
$cl\nu$	10.4	10.17	0	10.9	0
$D^*\pi^0$	0.1	0.03	—	0.029	—
$D\pi^0$	0.3	0.09	—	0.088	—
$D^*\pi^\pm$	0.2	0.06	—	0.058	—
$D\pi^\pm$	0.6	0.18	—	0.175	—

bers, the sum of the four exclusive modes exceeds the inclusive rate, the four exclusive modes are scaled down, preserving the relative ratios, such that the total branching fraction is saturated by these four modes. If the first four exclusive modes do not exceed the inclusive rate, we generate two new uncorrelated random numbers, again from Gaussian distributions, corresponding to the broad states $D_0^* \ell^- \bar{\nu}_\ell$ and $D_1' \ell^- \bar{\nu}_\ell$; as before, negative numbers are replaced with zero. If the six exclusive modes now exceed the inclusive rate, the two broad states are scaled down, preserving their ratio, to saturate the total rate (while the first four branching fractions are unchanged). Finally, if the six exclusive modes already generated do not exceed the inclusive rate, we saturate the remainder with the four nonresonant states, maintaining the Monte Carlo ratio of 0.1 : 0.3 : 0.2 : 0.6.

Note that, even though we do not reweight the $D^{(*)} \ell^- \bar{\nu}_\ell$ branching fractions, we need to generate distributions of these numbers to allow for sufficient variation in the broad D^{**} and nonresonant states which are used to saturate the total rate.

Figure 12.1 shows distributions of the various branching fractions generated according to the algorithm above.

Figures 12.2–12.9 show the effect of reweighting on the D^{**} PDFs. In all figures, we show both the m_{miss}^2 and p_ℓ^* projections; unweighted $D^{**} \ell^- \bar{\nu}_\ell$ MC

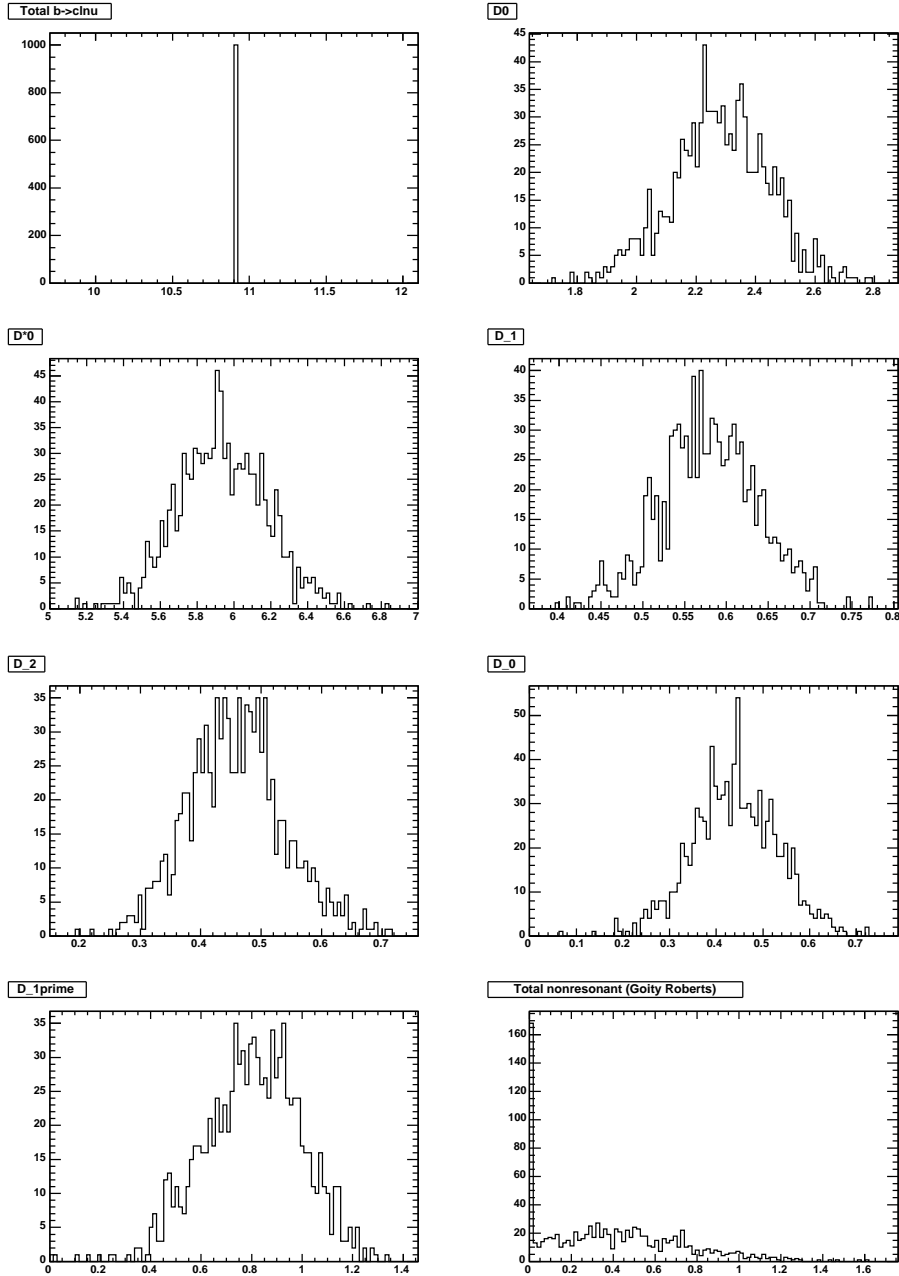


Figure 12.1: Generated branching fraction distributions for D^{**} model reweighting. Plots are shown for the B^+ modes; B^0 is similar.

is shown as points, and the blue bands represent the $1\text{-}\sigma$ and $2\text{-}\sigma$ envelopes of the ensemble of reweighted, refitted PDFs.

After generating the ensemble of D^{**} models, we refit our 970 fb^{-1} sample¹ of generic MC using the refitted PDF shapes and recalculating the D^{**} feed-down constraints. Figure 12.10 shows the distributions of fit results (the relative branching fraction, $R^{(*)}$) for the four signal channels and the two R -averaged measurements in 1000 such experiments; these results are also summarized in Table 12.3. We use the RMS of these fit results, relative to the mean fitted value, as an estimate of the systematic error associated with D^{**} modeling; we see systematic effects of 2% to 6% in the D modes and less than 1% in the D^* modes.

12.3 $B \rightarrow D^*$ Form Factors

To study the systematic dependence on the $B \rightarrow D^*$ form factors, we generate an ensemble of models, reweighting the signal and semileptonic MC accordingly. In each experiment, we simultaneously reweight $\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau$,

¹This sample is actually slightly modified for the D^{**} model studies. The fit to data suggests that the MC overestimates the $D^{**}\ell^-\bar{\nu}_\ell$ background by about a factor of two, so that using the full 970 fb^{-1} sample would likely overestimate the actual systematic uncertainty. In order to obtain a better estimate of the uncertainty, we reject 50% of the $D^{**}\ell^-\bar{\nu}_\ell$ events in the generic MC before performing the ensemble of fits, creating a sample that more closely resembles data.

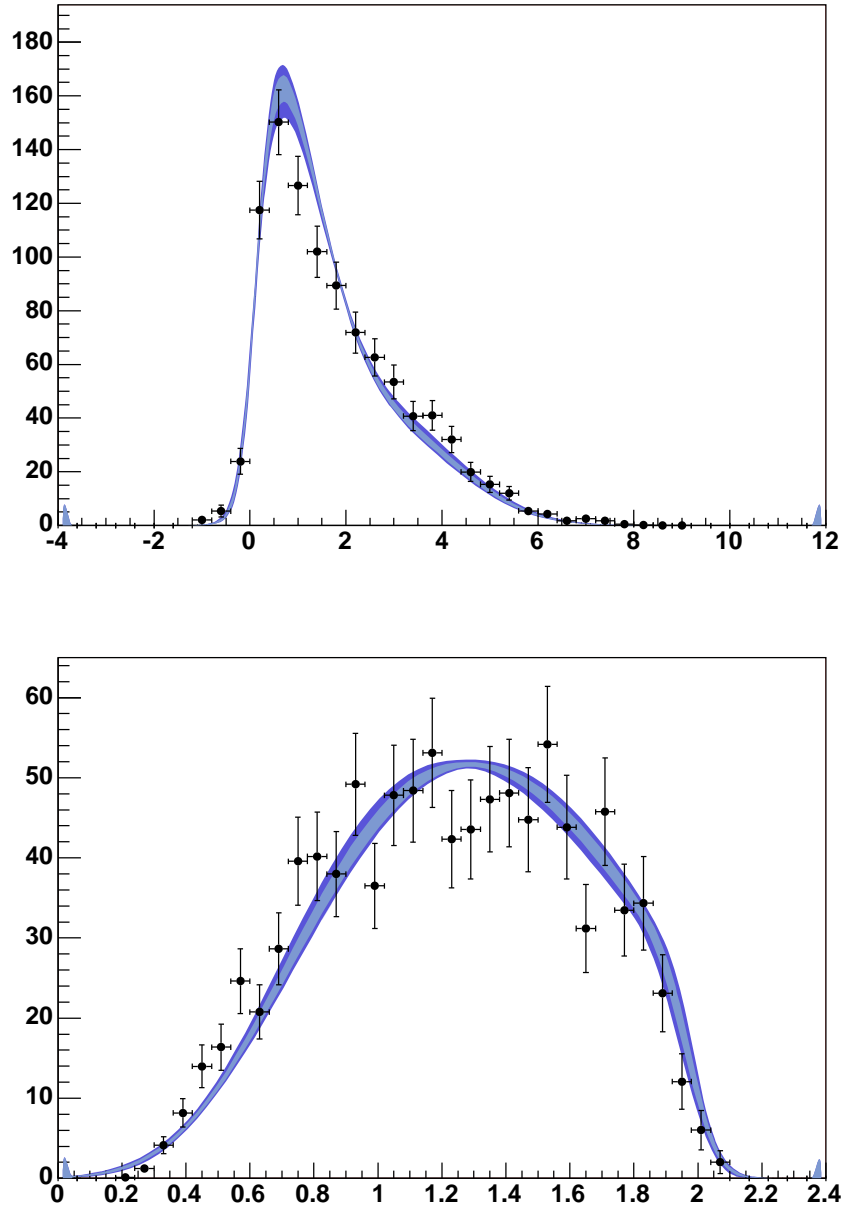


Figure 12.2: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{*0}$.

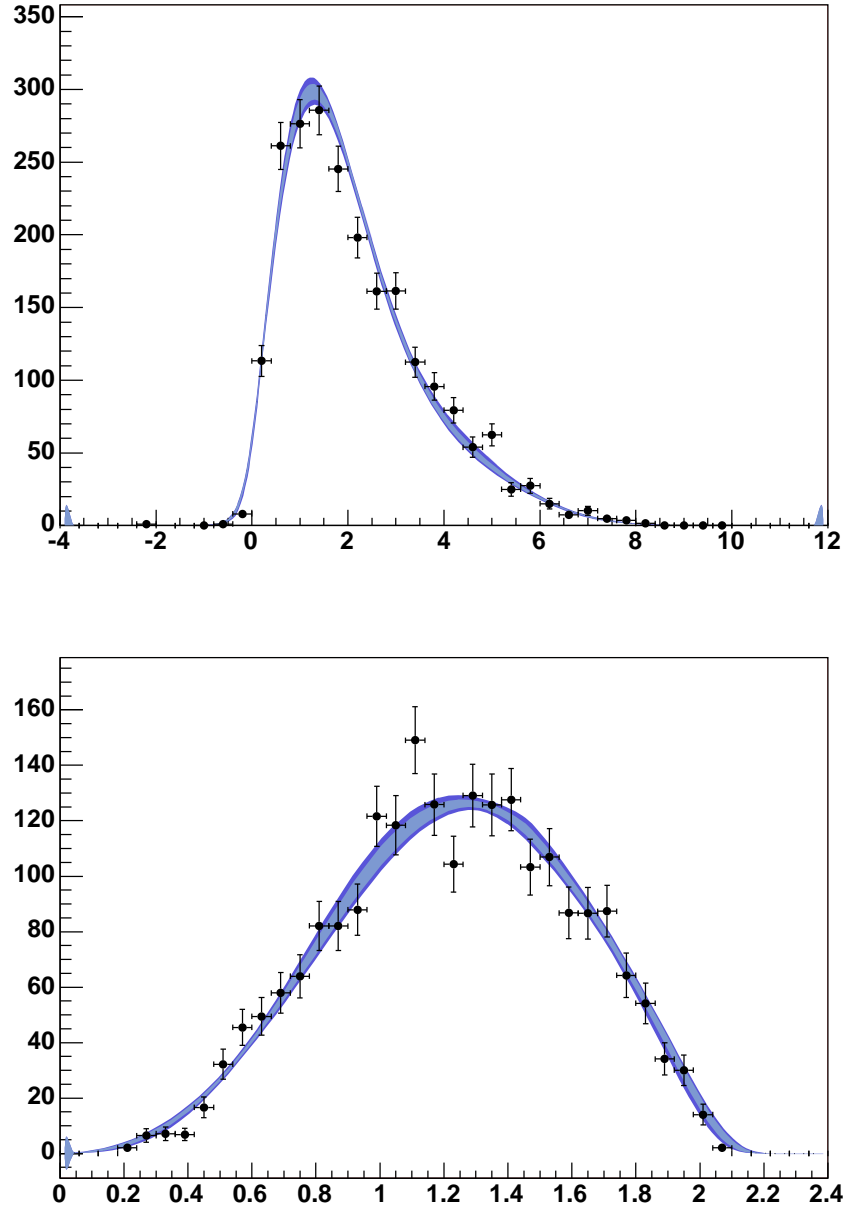


Figure 12.3: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^0$.

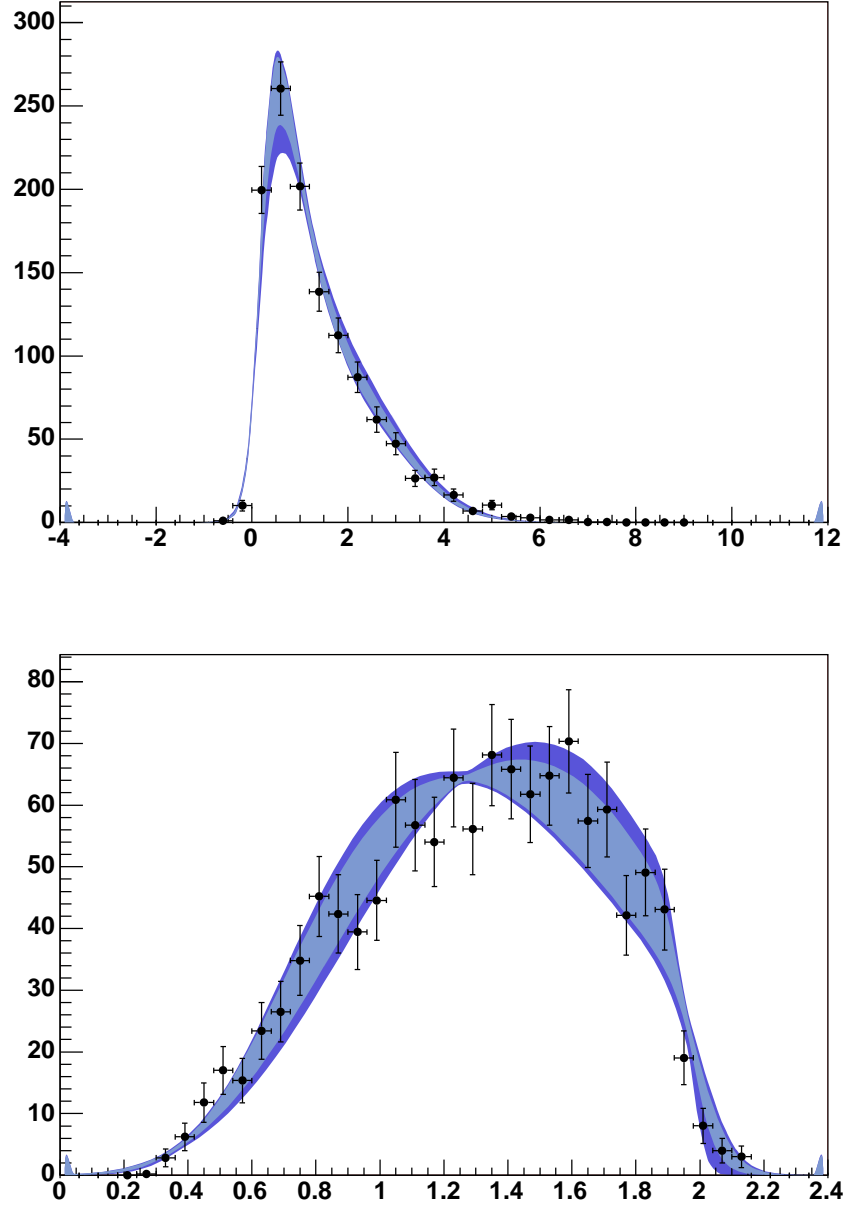


Figure 12.4: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{*+}$.

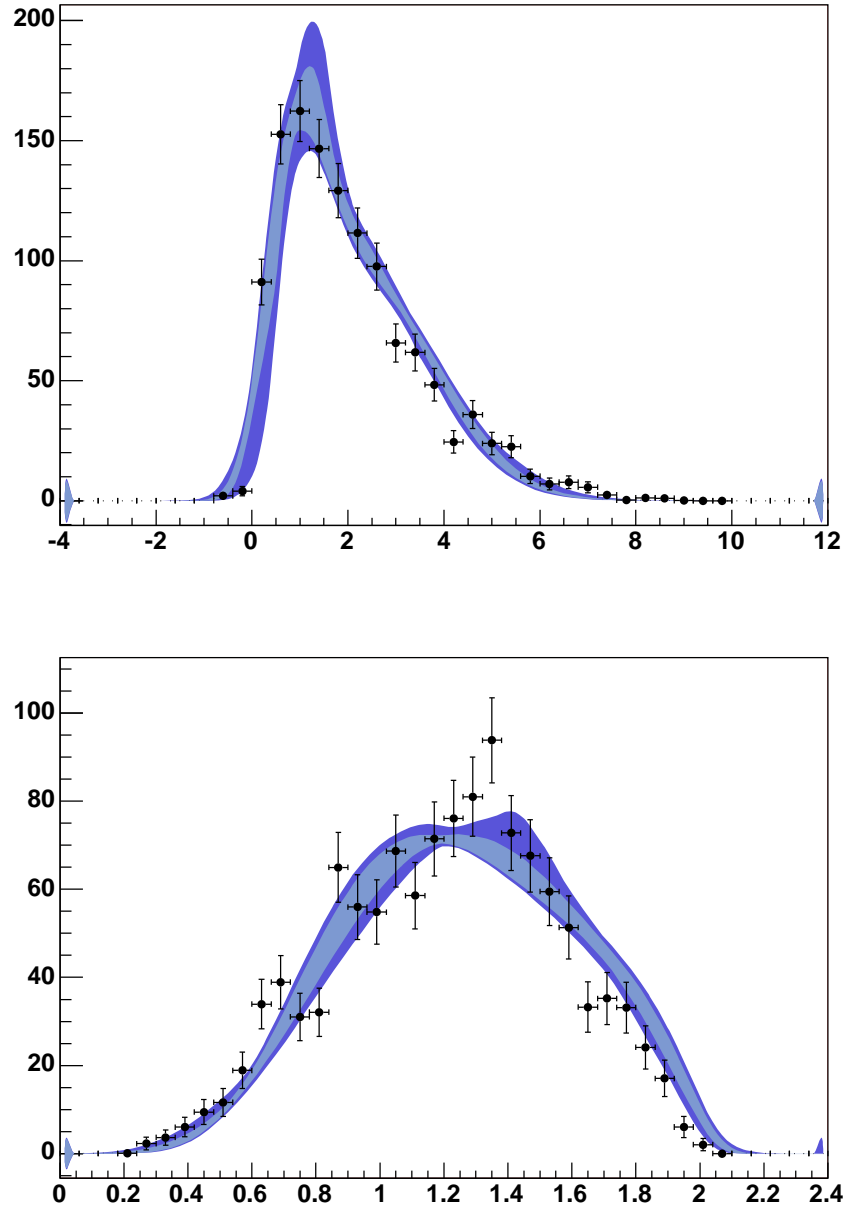


Figure 12.5: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{+}$.

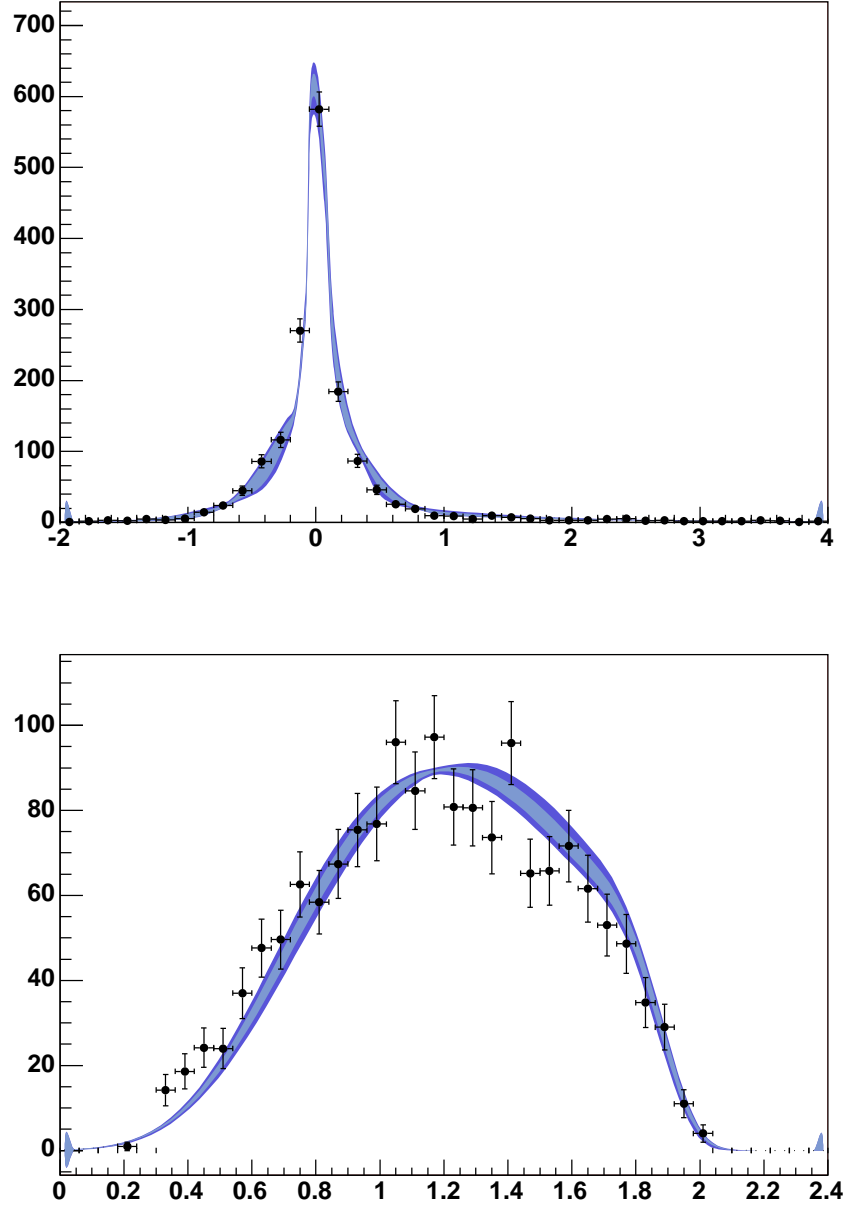


Figure 12.6: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{*0}\pi^0$.

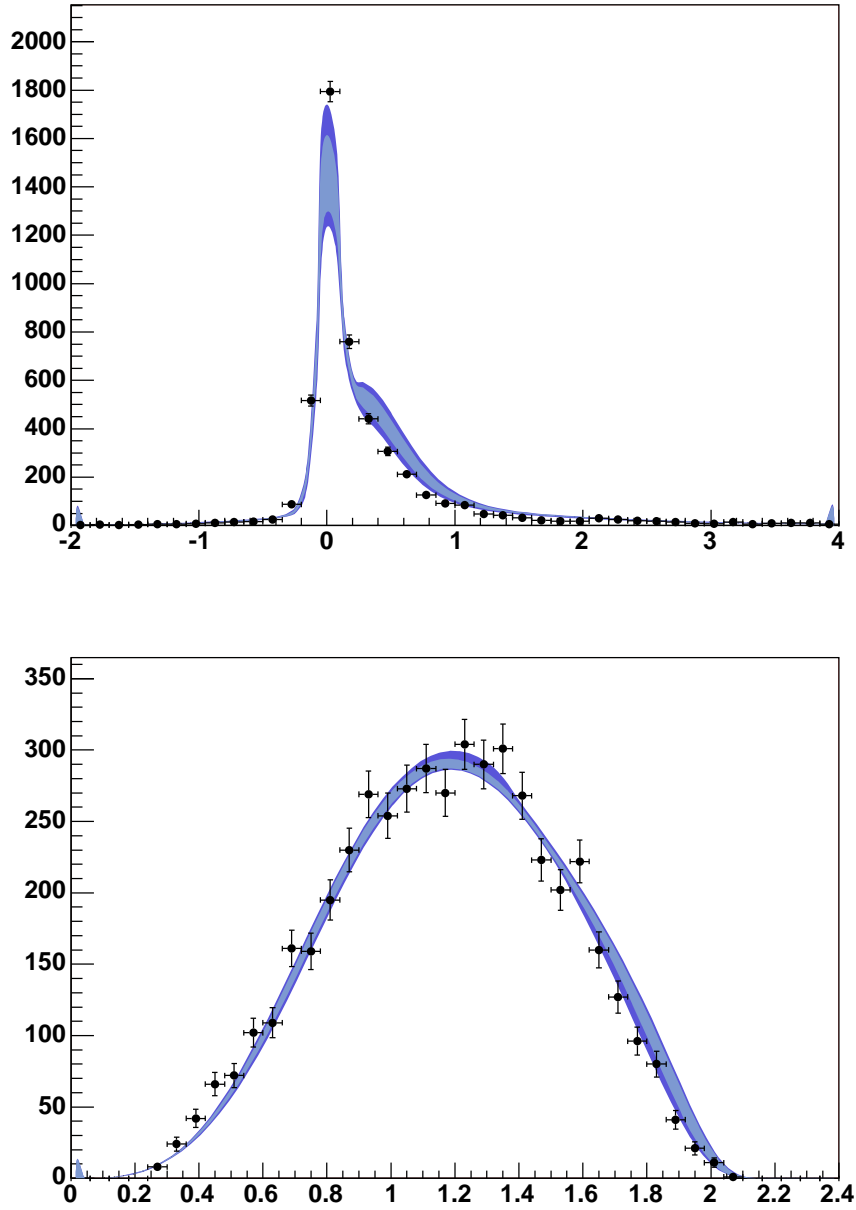


Figure 12.7: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^0\pi^0$.

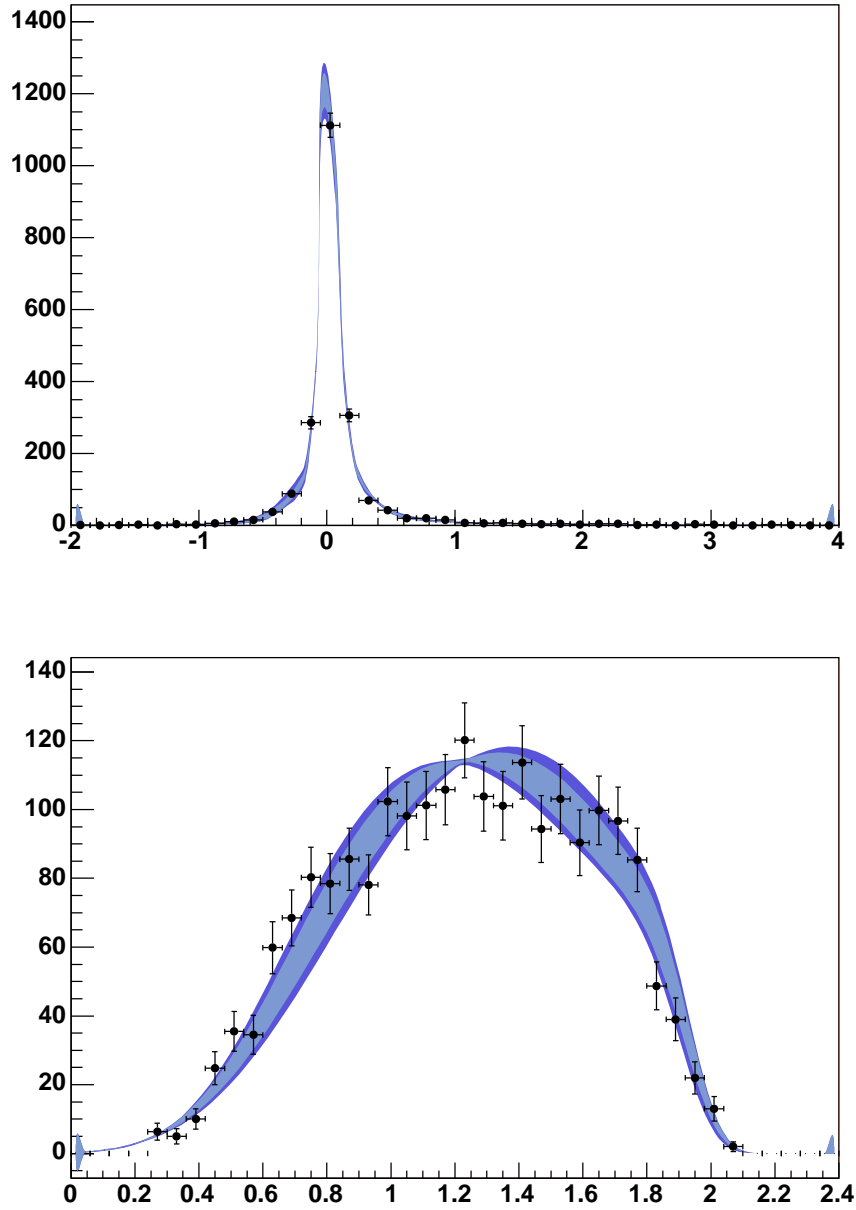


Figure 12.8: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{*+}\pi^0$.

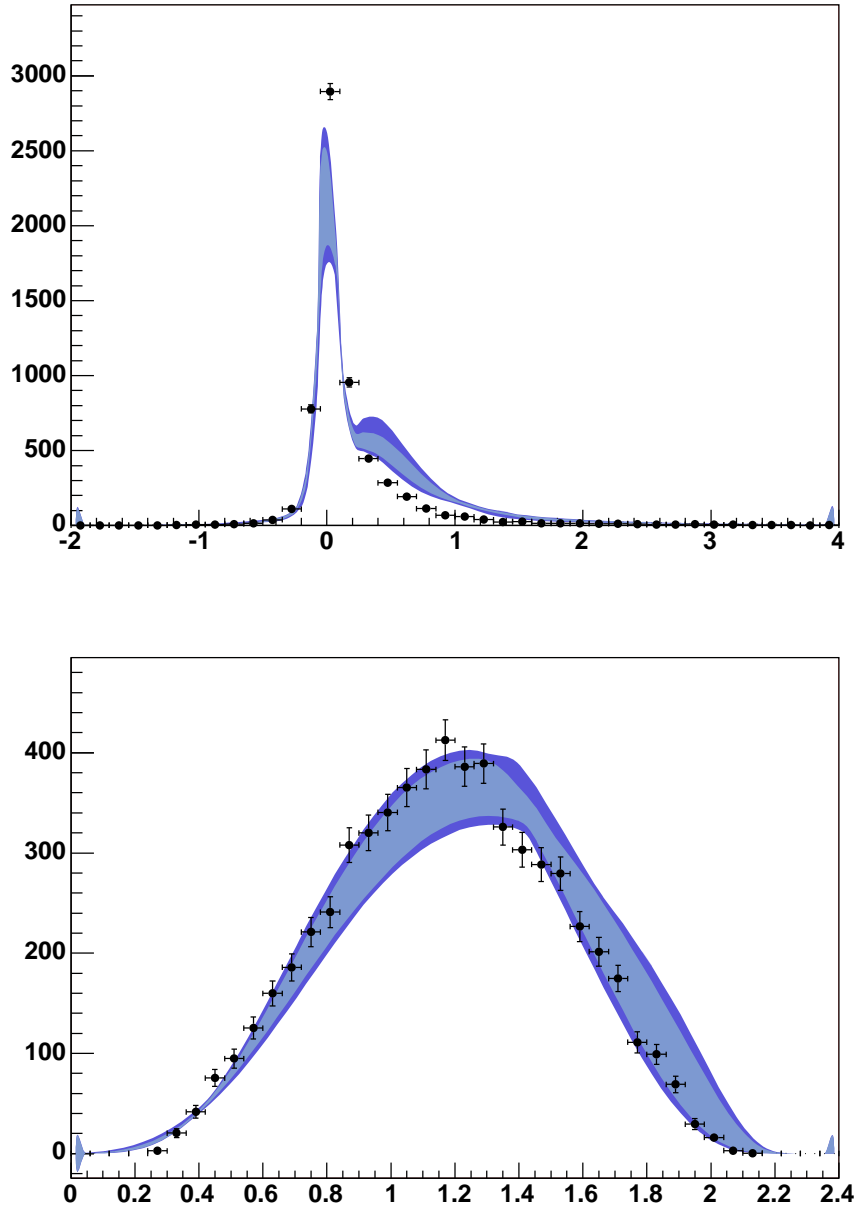


Figure 12.9: D^{**} modeling variation for the channel $D^{**}\ell^{-}\bar{\nu}_{\ell}\Rightarrow D^{+}\pi^{0}$.

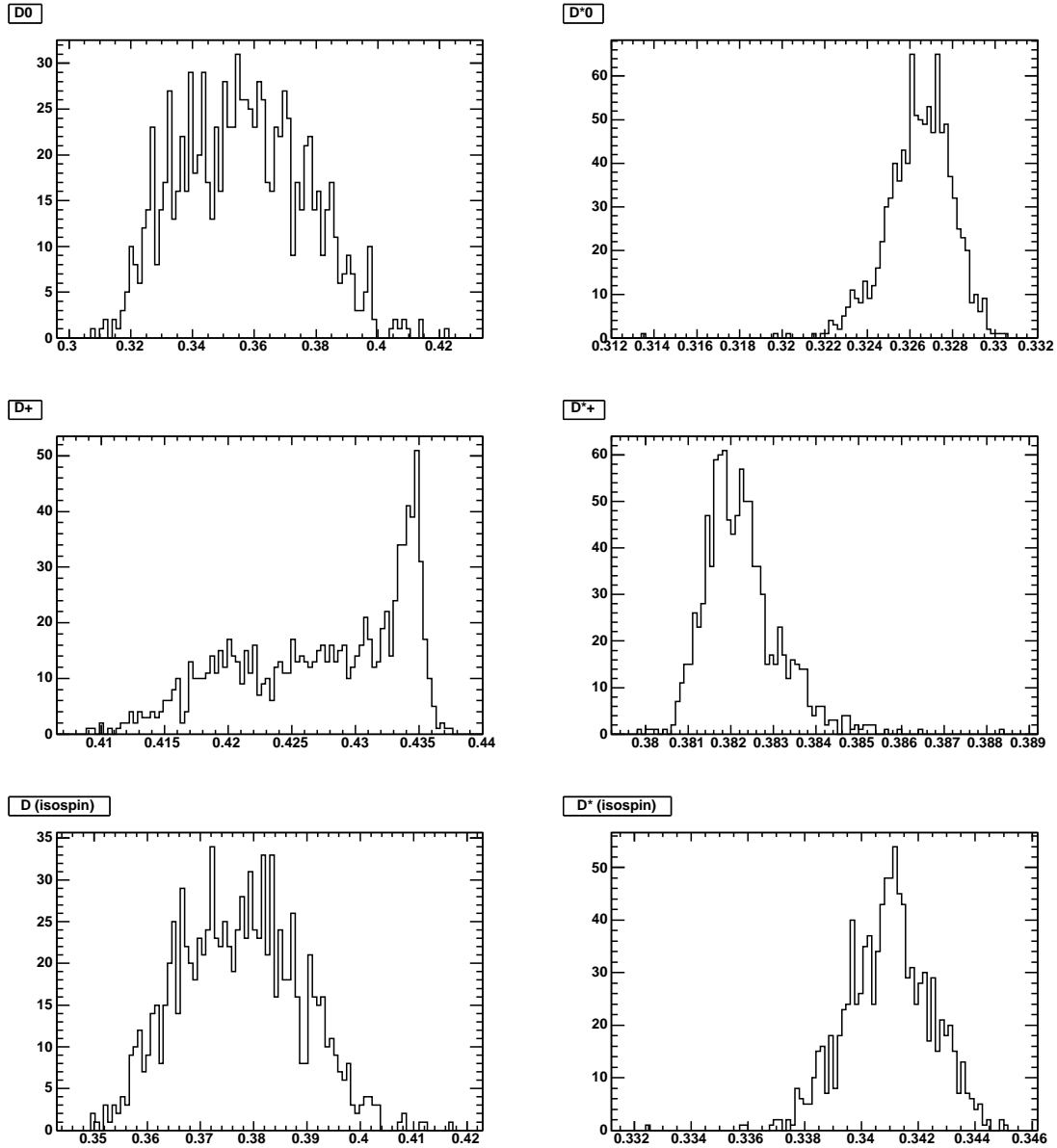


Figure 12.10: Distributions of fit results for D^{**} modeling systematic error. In all figures, the relative branching fraction $R^{(*)}$ is shown.

Table 12.3: Systematic error estimation for D^{**} modeling. For each mode, the mean and RMS of the distribution of the relative branching fraction $R^{(*)}$ (seen in Figure 12.10) is given. We take RMS/mean as an estimate of the systematic error.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	35.6	2.0	5.7
$D^{*0} \tau^- \bar{\nu}_\tau$	32.6	0.2	0.5
$D^+ \tau^- \bar{\nu}_\tau$	42.7	0.7	1.6
$D^{*+} \tau^- \bar{\nu}_\tau$	38.2	0.1	0.2
$D \tau^- \bar{\nu}_\tau$	37.7	1.1	3.0
$D^* \tau^- \bar{\nu}_\tau$	34.1	0.1	0.4

$\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$, $B^+ \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$, and $B^+ \rightarrow D^{*0} \ell^- \bar{\nu}_\ell$ using the same form factor model.

The ensemble of models is sampled from the recent *BABAR* measurement [3] and its errors, taking into account correlations between the three form factors. In particular, we use the CLN form factor expansion, and the form factors:

$$\begin{aligned}
 R_1 &= 1.417 \pm 0.061 \pm 0.044 \\
 R_2 &= 0.836 \pm 0.037 \pm 0.022 \\
 \rho^2 &= 1.179 \pm 0.048 \pm 0.028 ,
 \end{aligned}$$

where the first error is statistical and the second systematic. The errors (both statistical and systematic) are correlated between the three form factors, with

correlation coefficients

$$\rho_{R_1-R_2} = -0.84$$

$$\rho_{R_1-\rho^2} = +0.70$$

$$\rho_{R_2-\rho^2} = -0.83 .$$

We reweight the PDFs for events reconstructed in the direct channels ($D^* \tau^- \bar{\nu}_\tau \Rightarrow D^*$ and $D^* \ell \nu \Rightarrow D^*$, both charged and neutral) and for events reconstructed as feed-down ($D^* \tau^- \bar{\nu}_\tau \Rightarrow D$ and $D^* \ell \nu \Rightarrow D$); feed-up to the D^{**} control sample is relatively suppressed, and reweighting such a small component is not expected to have much of an effect.

We refit the 970 fb^{-1} sample of generic MC using the refitted PDF shapes and recalculating the $D^* \Rightarrow D$ feed-down constraints. The distributions of fit results are summarized in Table 12.4. We see systematic effects of 0.2% to 1.9%.

12.4 $B \rightarrow D$ Form Factors

Similar to what was done for the $B \rightarrow D^*$ form factors, we generate an ensemble of models for the $B \rightarrow D$ form factor slope and reweight the MC, both for signal modes $\bar{B}^0 \rightarrow D^+ \tau^- \bar{\nu}_\tau$ and $B^+ \rightarrow D^0 \tau^- \bar{\nu}_\tau$ and normalization

Table 12.4: Systematic error estimation for $B \rightarrow D^*$ form factors.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	45.4	0.9	1.9
$D^{*0} \tau^- \bar{\nu}_\tau$	25.4	0.2	0.7
$D^+ \tau^- \bar{\nu}_\tau$	42.0	0.3	0.8
$D^{*+} \tau^- \bar{\nu}_\tau$	33.5	0.1	0.2
$D \tau^- \bar{\nu}_\tau$	42.4	0.6	1.4
$D^* \tau^- \bar{\nu}_\tau$	28.2	0.1	0.4

modes $\bar{B}^0 \rightarrow D^+ \ell^- \bar{\nu}_\ell$ and $B^+ \rightarrow D^0 \ell^- \bar{\nu}_\ell$. We again use the CLN expansion of the form factors, with the slope parameter $\rho^2 = 1.17 \pm 0.18$ [44].

We only reweight the PDFs for $D \tau^- \bar{\nu}_\tau \Rightarrow D$ and $D \ell^- \bar{\nu}_\ell \Rightarrow D$ (both charged and neutral); feed-up to the D^* channels and the D^{**} control sample is relatively suppressed, and reweighting the shape of such a small component is not expected to have much of an effect.

We refit the 970 fb^{-1} sample of generic MC using the refitted PDF shapes and recalculating the $D \Rightarrow D^*$ feed-up constraints. The distributions of fit results are summarized in Table 12.5. We see systematic effects of 0.2% to 0.7%.

Table 12.5: Systematic error estimation for $B \rightarrow D$ form factors.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	40.1	0.1	0.2
$D^{*0} \tau^- \bar{\nu}_\tau$	28.8	0.2	0.7
$D^+ \tau^- \bar{\nu}_\tau$	36.9	0.2	0.6
$D^{*+} \tau^- \bar{\nu}_\tau$	36.6	0.1	0.2
$D \tau^- \bar{\nu}_\tau$	37.2	0.1	0.3
$D^* \tau^- \bar{\nu}_\tau$	31.8	0.1	0.4

12.5 Combinatoric Background Modeling

Section 5.2 briefly described the physical sources of combinatoric background seen in this analysis; Table 12.6 summarizes these sources, giving their approximate abundance in Generic MC after all event selection. In order to study systematic effects, we perform an ensemble of fits, reweighting events from the various combinatoric sources.

In total, the two-body B decays $B \rightarrow D_s^{(*)+} D^{(*) (*)}$ and $B \rightarrow D^{(*)} D^{(*)}$ make up approximately 45% of the total combinatoric background yield, while the three-body decays $B \rightarrow D^{(*)} D^{(*)} K$ contribute another 15%. Branching fractions of most of the relevant two-body B decays (and some of the three-body decays as well) have previously been measured. These branching fractions are listed in Table 12.7, along with relevant branching fractions of the D_s^+ meson.

Table 12.6: Sources and approximate abundances of combinatoric background in generic MC. All four signal channels are combined here. The third column shows what fraction of the B decays in each group have previously been observed, and the fourth column is the product of the second and third.

Source	% of total BG	% \mathcal{B} measured	
		(relative)	(absolute)
$B \rightarrow D_s^{(*)+} D^{(*)}$ (+ light hadrons)			
... with $D_s^+ \rightarrow \tau\nu$	30	90	27
... with $D_s^+ \rightarrow \ell\nu(\phi/\eta/\eta')$	10	90	9
$B \rightarrow D^{(*)} D^{(*)}$ (+ light hadrons)	35	65	25
Both $B \rightarrow D^{(*)} \ell\nu$	15	100	15
Fake lepton	5	0	0
Completely misreconstructed	5	0	0
Total			75

To study systematic uncertainties related to combinatoric background modeling, we perform an ensemble of fits. In each fit, we reweight each event in the Monte Carlo according to the following rules:

- If the mode is listed in Table 12.7, the reweighted branching fraction is sampled from the measurement, assuming a Gaussian distribution of the errors. If the decay involves a D_s^+ meson, the weight is the product of both the B and D_s^+ weights. All rows in Table 12.7 are treated as uncorrelated.
- If the event is a charge crossfeed event (a true $B \rightarrow D^{(*)} \ell\nu$ event where the B_{tag} and signal $D^{(*)}$ swap a charged particle), the dominant system-

Table 12.7: Branching fractions of D_s^+ and two- and three-body B decays contributing to combinatoric background. All measurements come from [10], except (\dagger) which come from the 2005 online update of the PDG, and (\ddagger) which come from [48].

Mode		\mathcal{B}	SP6 (same scale)
$D_s^+ \rightarrow \tau\nu$	(\dagger)	$(6.4 \pm 1.5) \times 10^{-2}$	7.0
$D_s^+ \rightarrow \eta\ell\nu$		$(2.5 \pm 0.7) \times 10^{-2}$	2.6
$D_s^+ \rightarrow \eta'\ell\nu$		$(8.9 \pm 3.3) \times 10^{-3}$	8.9
$D_s^+ \rightarrow \phi\ell\nu$		$(2.0 \pm 0.5) \times 10^{-2}$	2.0
$D_s^+ \rightarrow \mu\nu$		$(5.0 \pm 1.9) \times 10^{-3}$	4.6
$B^+ \rightarrow \bar{D}^0 D_s^+$		$(1.3 \pm 0.4) \times 10^{-2}$	1.06
$B^+ \rightarrow \bar{D}^0 D_s^{*+}$		$(9 \pm 4) \times 10^{-3}$	9.1
$B^+ \rightarrow \bar{D}^{*0} D_s^+$		$(1.2 \pm 0.5) \times 10^{-2}$	1.02
$B^+ \rightarrow \bar{D}^{*0} D_s^{*+}$		$(2.7 \pm 1.0) \times 10^{-2}$	2.28
$B^+ \rightarrow D^{*-0} D_s^{(*)+}$		$(2.7 \pm 1.2) \times 10^{-2}$	3.0
$B^+ \rightarrow \bar{D}^0 D^{*+} K^0$		$(5.2 \pm 1.2) \times 10^{-3}$	5.2
$B^+ \rightarrow \bar{D}^{*0} D^{*+} K^0$		$(7.8 \pm 2.6) \times 10^{-3}$	7.8
$B^+ \rightarrow D^0 \bar{D}^0 K^+$	(\dagger)	$(1.37 \pm 0.32) \times 10^{-3}$	1.9
$B^+ \rightarrow D^{*0} \bar{D}^{*0} K^+$		$(5.3 \pm 1.6) \times 10^{-3}$	5.3
$B^+ \rightarrow D^{*0} \bar{D}^0 K^+$		$(4.7 \pm 1.0) \times 10^{-3}$	4.8
$B^+ \rightarrow D^{*-} D^+ K^+$		$(1.5 \pm 0.4) \times 10^{-3}$	0.5
$B^0 \rightarrow D^- D_s^+$		$(8.0 \pm 3.0) \times 10^{-3}$	7.4
$B^0 \rightarrow D^{*-} D_s^+$		$(1.07 \pm 0.29) \times 10^{-2}$	1.03
$B^0 \rightarrow D^- D_s^{*+}$		$(1.0 \pm 0.5) \times 10^{-2}$	0.74
$B^0 \rightarrow D^{*-} D_s^{*+}$		$(1.9 \pm 0.5) \times 10^{-2}$	1.97
$B^0 \rightarrow D^- D^0 K^+$		$(1.7 \pm 0.4) \times 10^{-3}$	1.7
$B^0 \rightarrow D^- D^{*0} K^+$		$(3.1 \pm 0.6) \times 10^{-3}$	3.1
$B^0 \rightarrow D^{*-} D^{*0} K^+$		$(1.18 \pm 0.20) \times 10^{-2}$	1.18
$B^0 \rightarrow D^{*-} D^+ K^0$		$(6.5 \pm 1.6) \times 10^{-3}$	8.1
$B^0 \rightarrow D^{*-} D^{*+} K^0$		$(8.8 \pm 1.9) \times 10^{-3}$	8.8
$B^0 \rightarrow D^{*+} D^{*-}$	(\ddagger)	$(8.1 \pm 1.2) \times 10^{-4}$	8.3
$B^0 \rightarrow D^+ D^{*-}$	(\ddagger)	$(10.4 \pm 2.0) \times 10^{-4}$	6.7
$B^0 \rightarrow D^+ D^-$	(\ddagger)	$(2.8 \pm 0.7) \times 10^{-4}$	2.7

atic uncertainty is not the branching fraction, but rather the efficiency to reconstruct the B_{tag} with the wrong charge. We generate weights for these events using a Gaussian with a mean of 1 and a width of 0.1, *i.e.*, assuming a 10% uncertainty on the MC modeling of this efficiency.

- If the event is a double-semileptonic event, with both B mesons decaying to $D^{(*)}\ell\nu$, again, the dominant uncertainty comes not from the branching fractions, but from the probability to reconstruct a B_{tag} candidate in this event. We again assume a 10% uncertainty on this number.
- For fake leptons (almost exclusively pions masquerading as muons), we again assign a 10% uncertainty, and sample weights accordingly. The typical fake rate is 2–3%, with data-MC discrepancies generally 10% or less (in the momentum ranges of interest).
- For all remaining sources of combinatoric background, including high-multiplicity $B \rightarrow Dhh\dots$ and $B \rightarrow DDhh\dots$ (where D here is any charm meson and h any light meson) and events which are sufficiently misreconstructed as to deny classification, we assume a 50% uncertainty in the relevant rates.

In each experiment, we fit the reweighted Monte Carlo to generate new PDF shapes, and we recalculate the expected yield of combinatoric events in

Table 12.8: Systematic error estimation for combinatoric background modeling.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	37.6	2.8	7.5
$D^{*0} \tau^- \bar{\nu}_\tau$	29.4	1.2	4.1
$D^+ \tau^- \bar{\nu}_\tau$	34.8	4.0	11.5
$D^{*+} \tau^- \bar{\nu}_\tau$	38.6	1.0	2.6
$D \tau^- \bar{\nu}_\tau$	34.6	3.2	9.1
$D^* \tau^- \bar{\nu}_\tau$	32.8	1.0	2.9

each channel. Note that the reweighting does not affect the shape of the charge crossfeed backgrounds, only the normalization.

We refit the 970 fb^{-1} sample of generic MC using the modified shapes and background normalizations for each experiment in the ensemble. The distributions of fit results are summarized in Table 12.8. We see systematic effects of 2.6% to 11.5%.

12.6 Monte Carlo Statistics

To study systematic effects due to limited Monte Carlo statistics, we generate two ensembles of fits, one to study the uncertainty due to the PDF shapes, and another to study the crossfeed constraints.

Table 12.9: Systematic error estimation for MC statistics on PDF shapes.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0\tau^-\bar{\nu}_\tau$	39.4	4.5	11.5
$D^{*0}\tau^-\bar{\nu}_\tau$	29.3	2.5	8.4
$D^+\tau^-\bar{\nu}_\tau$	37.1	1.7	4.5
$D^{*+}\tau^-\bar{\nu}_\tau$	37.2	0.7	1.8
$D\tau^-\bar{\nu}_\tau$	37.0	2.5	6.9
$D^*\tau^-\bar{\nu}_\tau$	32.1	1.5	4.7

12.6.1 PDF Shapes

In the first study, the shape parameters of each of the 37 PDFs (for signal, semileptonic, and combinatoric background) are changed, selecting random points in the parameter space according to the error matrix obtained in the original shape fit. In each experiment, all 37 PDFs are modified simultaneously, treating the statistical fluctuations between the 37 PDFs as completely independent.

With each variation of the PDF shapes, we refit the 970 fb^{-1} sample of generic MC. The distributions of fit results are summarized in Table 12.9. We see systematic effects of 1.8% to 11.5%.

Table 12.10: Systematic error estimation for MC statistics on crossfeed constraints.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	38.2	1.4	3.7
$D^{*0} \tau^- \bar{\nu}_\tau$	29.7	0.5	1.8
$D^+ \tau^- \bar{\nu}_\tau$	36.7	2.2	5.9
$D^{*+} \tau^- \bar{\nu}_\tau$	37.0	0.5	1.3
$D \tau^- \bar{\nu}_\tau$	36.2	1.2	3.2
$D^* \tau^- \bar{\nu}_\tau$	32.3	0.4	1.2

12.6.2 Crossfeed Constraints

In the second study, the crossfeed constraints are varied within their statistical errors. In this study, the constraints are varied in a coherent way: if the fraction of $D^* \ell^- \bar{\nu}_\ell$ events reconstructed in a D^* channel were larger, the fractions reconstructed feeding down into the D channel or feeding up into the D^{**} control samples would be smaller (and, naturally, vice-versa), and three constraints would be affected. To generate these experiments, we allow the number of reconstructed events in each channel to fluctuate independently within their statistical errors, and we then recalculate the crossfeed constraints in the usual manner.

With each set of crossfeed constraints, we refit the 970 fb^{-1} sample of generic MC. The distributions of fit results are summarized in Table 12.10.

We see systematic effects of 1.2% to 5.9%.

Table 12.11: Systematic error estimation for MC statistics on floating crossfeed constraint parameterization.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0 \tau^- \bar{\nu}_\tau$	38.3	0.7	1.8
$D^{*0} \tau^- \bar{\nu}_\tau$	26.9	0.0	0.2
$D^+ \tau^- \bar{\nu}_\tau$	39.9	0.6	1.4
$D^{*+} \tau^- \bar{\nu}_\tau$	34.1	0.0	0.1
$D \tau^- \bar{\nu}_\tau$	37.4	0.5	1.3
$D^* \tau^- \bar{\nu}_\tau$	29.3	0.1	0.3

We perform a similar study in which, instead of varying the fixed crossfeed constraints, we vary the relationship between the floating constraints. We do this by varying all of the “unkilled” terms in Equation 8.9 within their statistical errors. The distributions of fit results are summarized in Table 12.11. We see systematic effects of 0.1% to 1.8%, and these errors are added in quadrature with those listed in Table 12.10 to give the total error on “MC statistics (constraints)” listed in Table 12.1.

12.7 $D^{**} \tau^- \bar{\nu}_\tau$ Abundance

In order to study the effect of the $D^{**} \tau^- \bar{\nu}_\tau$ contribution to the D^{**} components in the fit, we generate an ensemble of models, varying the expected branching fraction of these decays (see also Section 8.5.2). We generate weights

Table 12.12: Systematic error estimation for $D^{**}\tau^-\bar{\nu}_\tau$ systematic error.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0\tau^-\bar{\nu}_\tau$	38.4	0.2	0.4
$D^{*0}\tau^-\bar{\nu}_\tau$	26.9	0.3	1.3
$D^+\tau^-\bar{\nu}_\tau$	40.0	0.1	0.3
$D^{*+}\tau^-\bar{\nu}_\tau$	34.1	0.1	0.2
$D\tau^-\bar{\nu}_\tau$	37.5	0.1	0.3
$D^*\tau^-\bar{\nu}_\tau$	29.3	0.2	0.8

for $D^{**}\tau^-\bar{\nu}_\tau$ by sampling a Gaussian distribution centered at 1.0 with a width of 0.3, so that we effectively vary this contribution by $\pm 30\%$. In each experiment, we vary the $D^{**}\ell^-\bar{\nu}_\ell$ PDFs and feed-down constraints and refit the 970 fb⁻¹ sample of generic MC. The distributions of fit results are summarized in Table 12.12. We see systematic effects of 0.2% to 1.3%.

12.8 Detector Efficiency

We will estimate the detector efficiency systematics by looking at how the relative signal efficiency depends on the various sources: tracking, particle ID, K_S^0 , and π^0 efficiency. These sources largely cancel in the relative efficiency, and make negligible contributions to the total systematic error.

12.8.1 Tracking Efficiency

We divide all tracks in the recoil into three mutually exclusive categories:

- Tracks in *GoodTracksLoose* which satisfy $p_T > 200 \text{ MeV}/c$
- Tracks in *GoodTracksLoose* which satisfy $p_T \leq 200 \text{ MeV}/c$
- All other tracks.

These track categories have different correction factors and systematic recipes, and, because of the different track momentum spectra between signal and normalization modes, there will be different abundances of the three categories.

We calculate the absolute efficiency systematic error for signal and normalization modes separately, following the usual recipe [45]. We then calculate the relative efficiency systematic treating the two absolute systematic errors as 100% correlated.

Figure 12.11 shows the distribution of the number of recoil tracks in the first category, *GoodTracksLoose* with $p_T > 200 \text{ MeV}/c$, comparing $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$ to the normalization mode $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0$. Excellent agreement is seen between the two samples, giving a sense of the expected amount of cancellation in the systematic error.

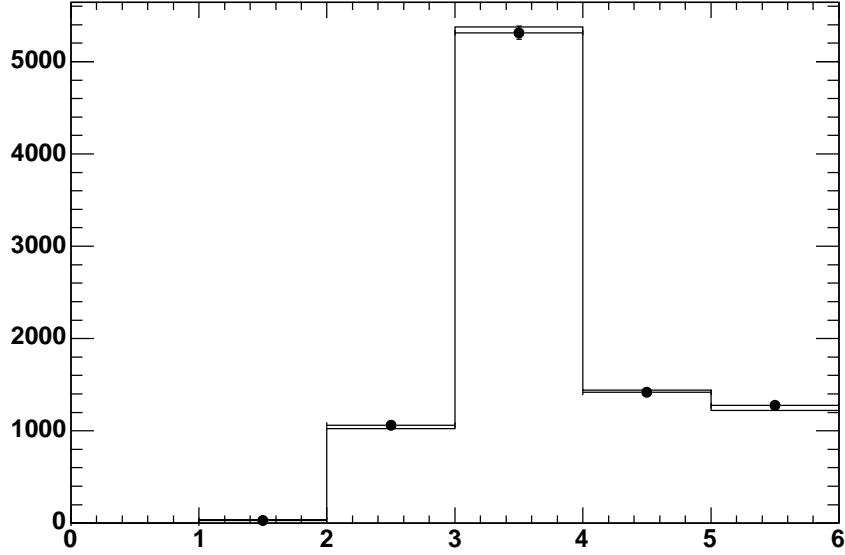


Figure 12.11: Number of *GoodTracksLoose* recoil tracks with $p_T > 200$ MeV/ c , comparing signal and normalization. $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$ is shown as points, while $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0$ is shown as a histogram; the normalization is arbitrary. Plots are taken from signal and semileptonic MC.

We see systematic errors less than 3×10^{-4} on the relative signal efficiency in all four signal channels and in the two R -averaged measurements.

12.8.2 PID Efficiency

We generally follow the same procedure as for tracking. We apply standard PID-table-based correction factors to both the signal and normalization samples and calculate the effect on the relative efficiency, treating the two corrections as correlated.

Figure 12.12 shows the momentum spectrum of recoil kaons and pions, comparing $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}$ to the normalization mode $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}$. Excellent agreement is seen between the two samples, giving a sense of the expected amount of cancellation in the systematic error.

The resulting systematic uncertainties on kaon and pion identification are 2×10^{-3} or less in all channels. For electrons and muons, the momentum spectra between the signal and normalization channels are significantly different, so there is less cancellation; we see systematic uncertainties up to 6×10^{-3} .

12.8.3 K_s^0 Efficiency

The K_s^0 efficiency is handled analogously to tracking and PID. We apply a correction factor which is a function of transverse flight length, transverse momentum, and polar angle in the lab frame. Figure 12.13 shows the transverse flight length of K_s^0 candidates, comparing $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$ to the normalization mode $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0$. Again, excellent agreement is seen between the two samples, giving a sense of the expected amount of cancellation in the systematic error.

As before, we calculate the correction separately for signal and normalization, and treat the resulting corrections as correlated when calculating the

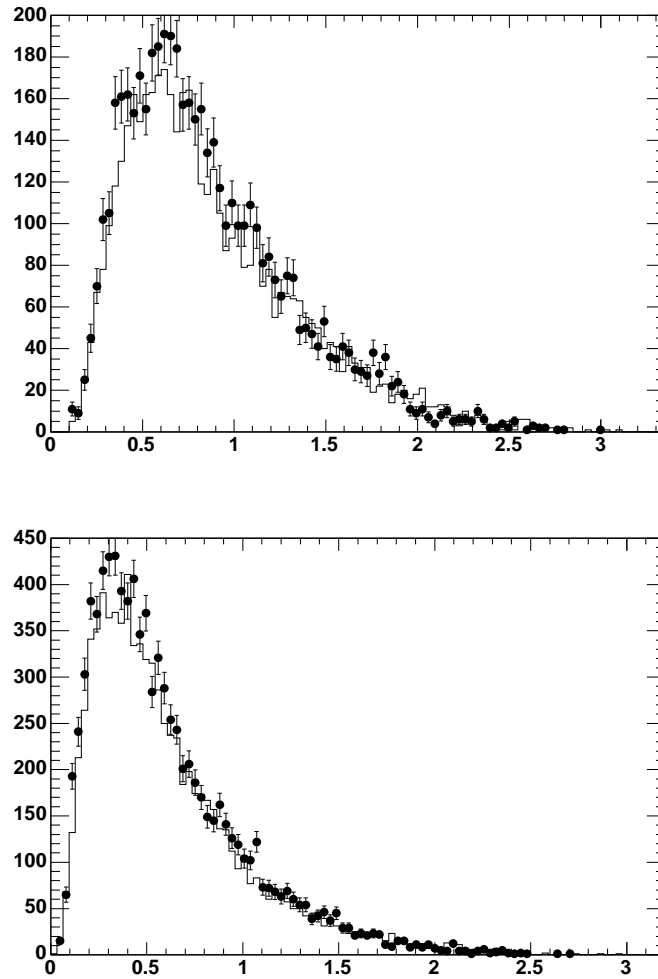


Figure 12.12: Recoil kaon (top) and pion (bottom) momentum spectra, comparing signal and normalization. $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0$ is shown as points, while $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0$ is shown as a histogram; the normalization is arbitrary. Plots are taken from signal and semileptonic MC.

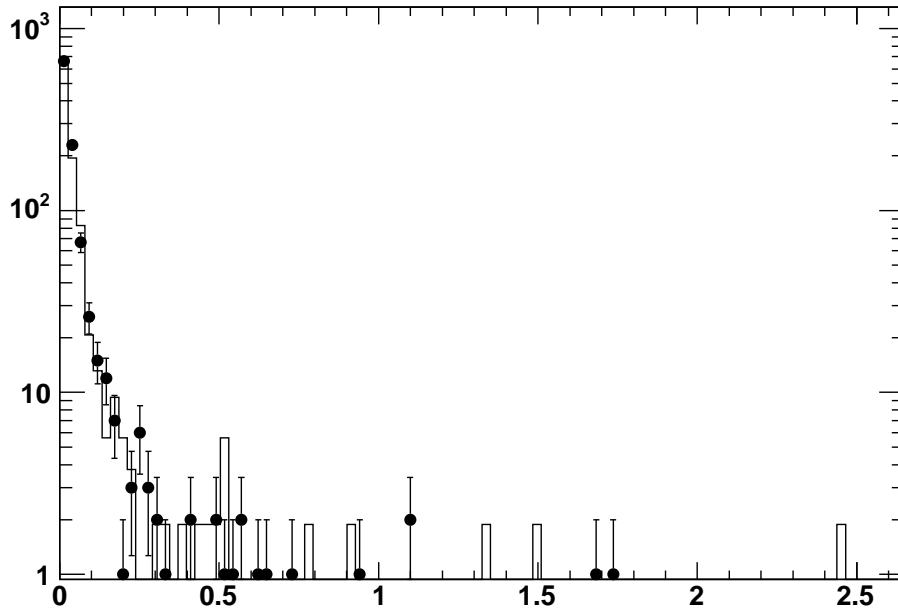


Figure 12.13: Transverse decay length of K_s^0 candidates, comparing signal and normalization. $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0$ is shown as points, while $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0$ is shown as a histogram; the normalization is arbitrary. Plots are taken from signal and semileptonic MC.

effect on the relative efficiency. We see systematic uncertainties less than 1×10^{-3} .

Note that these numbers were calculated using only the Run4 K_s^0 correction tables. In principle, the MC should be split up into different running periods with separate calculations. The Run4 table is the largest block and dominates the calculation. Given the small size of the resulting error, we have not done a complete block-by-block calculation.

12.8.4 Neutral Efficiency

There are two contributions to the systematic uncertainty due to the neutral efficiency. The first is a multiplicative systematic uncertainty due to uncertainties in the relative signal efficiency; this is due to uncertainties in both the π^0 and γ efficiencies. The second uncertainty is due to the fact that the π^0 efficiency largely determines the crossfeed constraints in the fit. If the π^0 efficiency were larger than the MC prediction, our efficiency to reconstruct a D^* meson as a D^* would be higher, while our efficiency to reconstruct a D^* as a D would be smaller; since the feeddown constraint is defined as the ratio of reconstructed D mesons to D^* mesons, this ratio sees the uncertainty in the π^0 magnified.

We estimate the uncertainty due to the efficiency following the same procedure we use for the other systematics. Following the standard recipe, we assign a 3.0% systematic uncertainty per π^0 , and a 1.8% systematic uncertainty per single photon (in the mode $D^{*0} \rightarrow D^0\gamma$). We treat this uncertainty as correlated between signal and normalization and calculate the resulting uncertainty on the relative efficiency; we see effects of 5×10^{-4} or smaller.

To estimate the uncertainty due to the crossfeed constraints, we again perform an ensemble of fits. In each experiment, we construct a new model of the π^0 efficiency, reweight the events accordingly, and recalculate the crossfeed constraints in the fit. The nominal π^0 efficiency correction (see Eq. 11.1) is a linear function of π^0 momentum, with a systematic uncertainty on the resulting efficiency of 3%. We generate the ensemble of constraints by adding a random number, sampled from a Gaussian distribution centered at zero with a width of 0.03, to the offset in the efficiency correction; instead of using 0.977, we use 0.977 ± 0.03 . This is *one* possible reparameterization of the π^0 efficiency which generates a 3% absolute efficiency uncertainty, but is not the only one, and is not necessarily the best one. In each experiment, only one random number is used, so that the π^0 efficiency increases or decreases coherently for all signal and background modes.

Table 12.13: Systematic error estimation for π^0 crossfeed constraints.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0\tau^-\bar{\nu}_\tau$	38.0	0.2	0.5
$D^{*0}\tau^-\bar{\nu}_\tau$	29.5	0.3	1.1
$D^+\tau^-\bar{\nu}_\tau$	36.5	0.2	0.5
$D^{*+}\tau^-\bar{\nu}_\tau$	36.7	0.3	0.9
$D\tau^-\bar{\nu}_\tau$	36.0	0.2	0.5
$D^*\tau^-\bar{\nu}_\tau$	32.1	0.3	1.0

We generate 1000 such experiments and repeat the fit to the 970 fb^{-1} generic MC sample. The distributions of fit results for the four signal channels and two R -averaged measurements are summarized in Table 12.13. We see systematic uncertainties of 0.5% to 1.1%.

12.8.5 Neutral Efficiency and $D^{**}\ell^-\bar{\nu}_\ell$ Feed-down Constraints

We apply an additional systematic error, beyond the standard neutrals systematic described in Section 12.8.4, to $D^{**}\ell^-\bar{\nu}_\ell$ events feeding down into the signal channels. Given that the Monte Carlo does not seem to be reliable at predicting soft pion efficiencies in our high-multiplicity events (see Section 8.9), we use the $D^*\ell^-\bar{\nu}_\ell$ feed-down constraints measured in data to estimate the systematic error on D^{**} feed-down.

The four D^{**} feed-down constraints relate the $B \rightarrow D^{**}\ell^{-}\bar{\nu}_{\ell}$ yield in the four signal channels to the yield in the four D^{**} control sample channels. Note that the four constraints relate pairs of channels that only differ by the “hard” π^0 , the prompt daughter from D^{**} decay (there is a constraint for $D^0\pi^0 \Rightarrow D^0$ and one for $D^{*0}\pi^0 \Rightarrow D^{*0}$, but not one for $D^{*0}\pi^0 \Rightarrow D^0$). Because of this, we do not need to worry about the efficiency for the soft pion from D^* decay, which appears in the same way in both the direct and feed-down channels, so that this efficiency cancels.

For the direct measurement of the D^{**} yield, the D^{**} control samples, we make an explicit requirement that the π^0 momentum be greater than 400 MeV/ c , and, since we expect the efficiency to be well modeled at this momentum, we have no reason to correct these yields. For the feed-down component, where the π^0 is not reconstructed, *most* of the lost pions are still high-momentum and require no correction, but 16% of the lost pions have momentum less than 300 MeV/ c and it may be more appropriate to treat these more like soft pions than hard ones. The neutrals group claims that their efficiency corrections should be valid down to momenta of 300 MeV/ c .

The values of the $D^*\ell^{-}\bar{\nu}_{\ell}$ feed-down constraints measured in data suggest data–MC discrepancies in the soft pion efficiency between 3% and 13%, with

statistical errors of 3%. We therefore choose a number at the high end of this range, 10%, to use as our variation in the D^{**} feed-down systematic error.

We perform an ensemble of fits to estimate the systematic uncertainty. In each experiment, we generate a random number sampled from a Gaussian distribution with mean 1.0 and width 0.1; this will represent our $\pm 10\%$ variation. All $D^{**}\ell^{-}\bar{\nu}_\ell$ events reconstructed in the signal channels are reweighted by this number if a true π^0 with momentum less than 300 MeV/ c was lost. Events in which the lost π^0 has higher momentum are not reweighted, nor are events in which the π^0 was reconstructed — the MC is expected to perform reasonably well at these momenta. We calculate new PDFs and feed-down constraints for the D^{**} contributions and repeat the fit to the 970 fb $^{-1}$ generic MC sample. The distributions of fit results for the four signal channels and the two R -averaged measurements are summarized in Table 12.14. We see systematic effects of 0.1% to 0.4%. These systematics are small since only a small fraction of $D^{**}\ell^{-}\bar{\nu}_\ell$ events produce a low-momentum π^0 .

12.9 m_{miss}^2 Tail Modeling

This analysis depends heavily on our understanding of the m_{miss}^2 distributions of background events. We are searching for a small signal on top of a

Table 12.14: Systematic error estimation for $D^{**}\ell^{-}\bar{\nu}_\ell$ feed-down.

Mode	Mean (%)	RMS (%)	RMS/Mean (%)
$D^0\tau^-\bar{\nu}_\tau$	38.8	0.1	0.4
$D^{*0}\tau^-\bar{\nu}_\tau$	27.3	0.0	0.1
$D^+\tau^-\bar{\nu}_\tau$	40.4	0.0	0.1
$D^{*+}\tau^-\bar{\nu}_\tau$	34.3	0.1	0.3
$D\tau^-\bar{\nu}_\tau$	38.0	0.1	0.2
$D^*\tau^-\bar{\nu}_\tau$	29.6	0.1	0.2

large background; if we underestimate the background at large m_{miss}^2 , we may artificially overestimate the signal.

We use the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ and q^2 control samples introduced in Section 7 to study the m_{miss}^2 tails. Table 12.15 shows the number of events in generic Monte Carlo and in data in the four $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample channels for various cuts on m_{miss}^2 ; Table 12.16 shows the same information for the q^2 control samples, excluding events common to both control samples. These samples are too small to allow a detailed statistical comparison between data and Monte Carlo. To address this problem, we add the four channels together in each of the two control samples, as seen in Table 12.17; this loses some physical meaning, but can still be used as a measurement of the overall performance of the Monte Carlo.

Table 12.15: Events in the m_{miss}^2 tail of the $D^{(*)}\ell^{-}\bar{\nu}_\ell$ control sample. The total number of events in the control samples is shown for both data and Monte Carlo. The number of events passing a cut on m_{miss}^2 (shown in $(\text{GeV}/c^2)^2$) is also shown for each channel, both as an absolute number of events and as a percentage of the total.

	MC						
	Total	> 0.5		> 1.0		> 1.5	
D^0	5079	898	(17.7%)	109	(2.15%)	23	(0.45%)
D^{*0}	3714	111	(2.99%)	18	(0.48%)	2	(0.05%)
D^+	1050	107	(10.2%)	21	(2.0%)	6	(0.57%)
D^{*+}	2338	45	(1.92%)	11	(0.47%)	3	(1.29%)
	Data						
	Total	> 0.5		> 1.0		> 1.5	
D^0	1003	193	(19.2%)	22	(2.2%)	7	(0.7%)
D^{*0}	621	14	(2.25%)	4	(0.64%)	2	(0.32%)
D^+	211	25	(11.8%)	5	(2.37%)	4	(1.9%)
D^{*+}	348	8	(2.30%)	3	(0.86%)	0	(0%)

We see that, on average, the Monte Carlo underestimates the number of events satisfying $m_{\text{miss}}^2 > 1 (\text{GeV}/c^2)^2$ by approximately 10%.

Given that the control samples are dominated by true $D^{(*)}\ell^{-}\bar{\nu}_\ell$ events, we investigate the effect of increasing the m_{miss}^2 tails of $D^{(*)}\ell^{-}\bar{\nu}_\ell$ events on the signal fit. Since the control samples suggest a 10% underestimate of the m_{miss}^2 tails, we reweight $D^*\ell^{-}\bar{\nu}_\ell \Rightarrow D^*$ and $D\ell^{-}\bar{\nu}_\ell \Rightarrow D$ Monte Carlo events with a weight of 1.1 if the event satisfies $m_{\text{miss}}^2 > 1 (\text{GeV}/c^2)^2$. We do not need to reweight the $D^* \Rightarrow D$ feed-down since, in this case, the m_{miss}^2 distribution is not dominated by resolution, but by the physical m_{miss}^2 spectrum of the lost

Table 12.16: Events in the m_{miss}^2 tail of the q^2 control sample, excluding events common to the $D^{(*)}\ell\nu$ control sample. The total number of events in the control samples is shown for both data and Monte Carlo. The number of events passing a cut on m_{miss}^2 (shown in $(\text{GeV}/c^2)^2$) is also shown for each channel, both as an absolute number of events and as a percentage of the total.

	MC						
	Total	> 0.5		> 1.0		> 1.5	
D^0	3250	861	(26.5%)	311	(9.56%)	114	(3.51%)
D^{*0}	1817	181	(9.96%)	99	(5.45%)	58	(3.19%)
D^+	916	129	(14.1%)	54	(5.90%)	25	(2.73%)
D^{*+}	1377	98	(7.12%)	61	(4.43%)	38	(2.76%)
	Data						
	Total	> 0.5		> 1.0		> 1.5	
D^0	536	150	(28.0%)	64	(11.9%)	23	(4.29%)
D^{*0}	263	22	(8.37%)	11	(4.18%)	5	(1.90%)
D^+	148	22	(14.9%)	13	(8.78%)	11	(7.43%)
D^{*+}	181	8	(4.42%)	5	(2.76%)	4	(2.21%)

Table 12.17: Events in the m_{miss}^2 tail of the two control samples with all channels summed together. The total number of events in the control samples is shown for both data and Monte Carlo. The number of events passing a cut on m_{miss}^2 (shown in $(\text{GeV}/c^2)^2$) is also shown for each channel, both as an absolute number of events and as a percentage of the total.

	MC						
	Total	> 0.5		> 1.0		> 1.5	
$D^{(*)}\ell^{-}\bar{\nu}_\ell$	12181	1161	(9.53%)	159	(1.31%)	34	(0.28%)
q^2	7360	1269	(17.2%)	525	(7.13%)	235	(3.19%)
Both together	19541	2430	(12.4%)	684	(3.50%)	269	(1.38%)
	Data						
	Total	> 0.5		> 1.0		> 1.5	
$D^{(*)}\ell^{-}\bar{\nu}_\ell$	2183	240	(11.0%)	34	(1.56%)	13	(0.60%)
q^2	1128	202	(17.9%)	90	(7.80%)	43	(3.81%)
Both together	3311	442	(13.3%)	124	(3.75%)	56	(1.69%)

$\pi_s/\gamma + \bar{\nu}_\ell$ system. We then fit these reweighted samples and derive new PDF shapes. We fit the 970 fb^{-1} generic MC sample with these modified shapes and take the difference between this fit and the nominal fit as a systematic uncertainty. We see effects between 0.1% and 1.6%.

12.10 Bremsstrahlung and Final-State Radiation

Previous studies of bremsstrahlung and final-state radiation [49] have shown a 2.1% systematic uncertainty on semileptonic B decays due to these effects (2.1% is the RMS of the variations listed). In our sample, approximately 11% of both signal and normalization decays are generated with final-state radiation, with the difference between signal and normalization taken as an estimate of the correlation. We estimate systematic errors by taking a 2.1% uncertainty on both the signal and normalization efficiencies, with estimated correlations around 90% to 95%, resulting in uncertainties of 0.2% to 0.6%.

12.11 Hadronic Daughter Branching Fractions

Because of the relative measurement strategy, the uncertainties on the branching fractions of the reconstructed D^* , D , K_S^0 , and π^0 modes cancels almost completely in the ratio. The residual contribution from these uncertainties arises from the fact that the signal and normalization modes may have slightly different efficiencies for the various channels, leading to different abundances of certain modes between the signal and normalization. We estimate the systematic error due to all of these branching fractions by propagating each one through to the relative efficiency, taking into account the correlation between the signal and normalization modes. We see systematic effects of 3×10^{-3} or less.

12.12 Tau Branching Fraction

The branching fraction of $\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\tau$ is known to 0.2% [10]. This uncertainty propagates directly to the relative signal efficiency.

Chapter 13

Unblinded Data

13.1 Numerical Results and Fit Projections

Table 13.1 summarizes the results of the fit to data, for the fits with and without the R -constraint. For the R -constrained fit, the two branching fractions have a correlation of -0.51 , so that the sum of the two is $(2.48 \pm 0.28)\%$. A more detailed breakdown of the fit yields in the unconstrained fit is given in Table 13.2, and for the constrained fit in Table 13.3.

Projections of the unconstrained fit are shown in Figures 13.1–13.8, while projections of the R -constrained fit are shown in Figures 13.9–13.16.

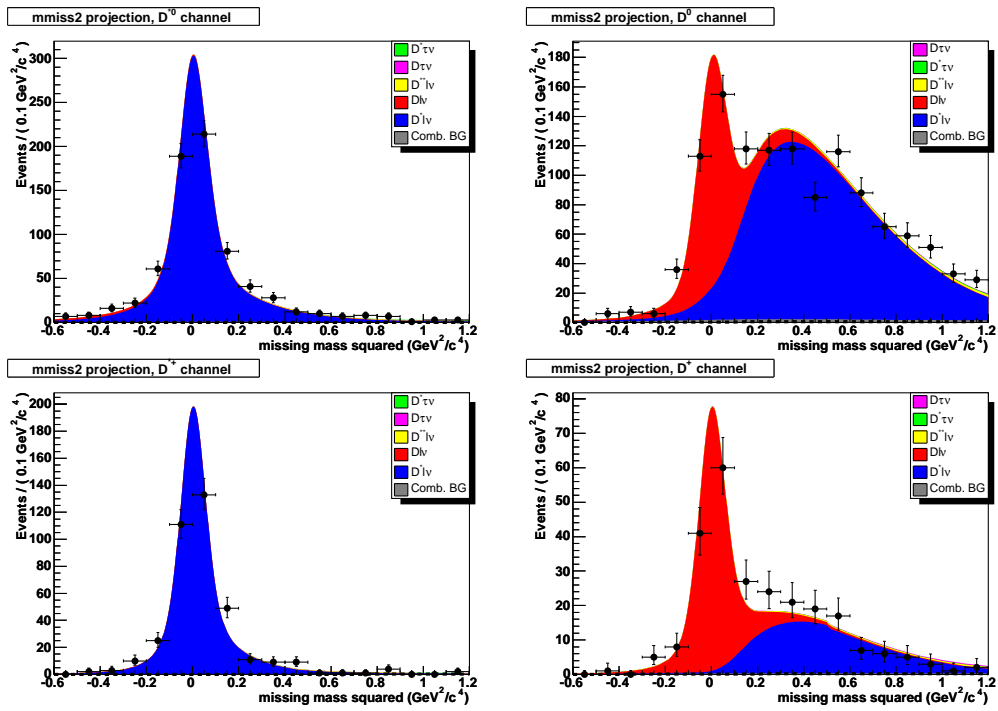


Figure 13.1: m_{miss}^2 projection of the unconstrained fit to data, showing the normalization region in the signal channels.

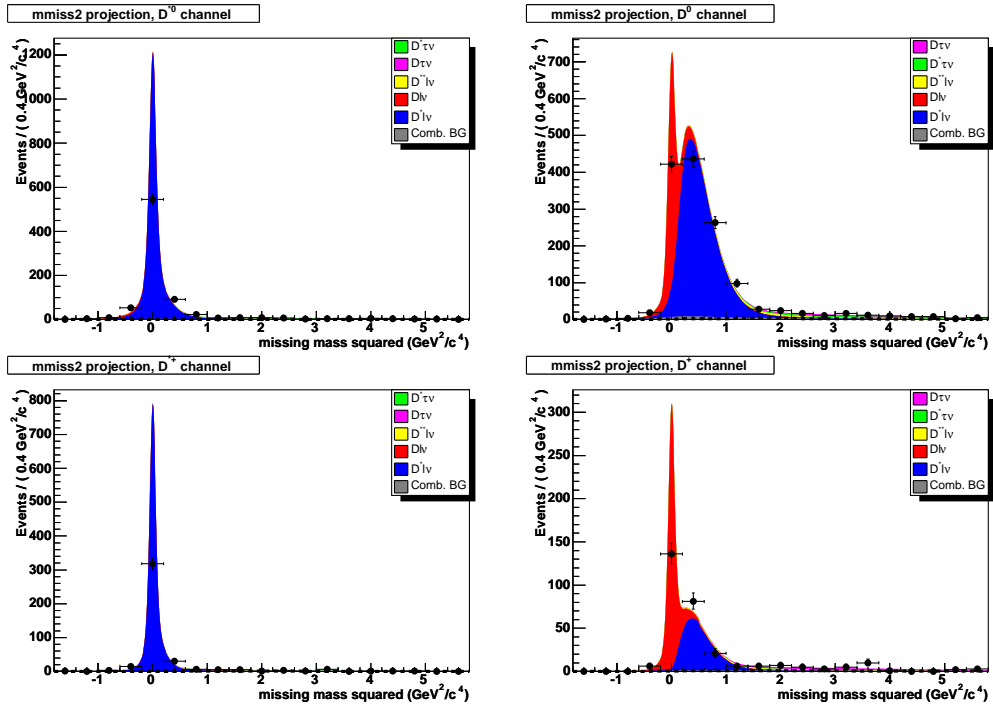


Figure 13.2: m_{miss}^2 projection of the unconstrained fit to data, showing a wider range than Figure 13.1.

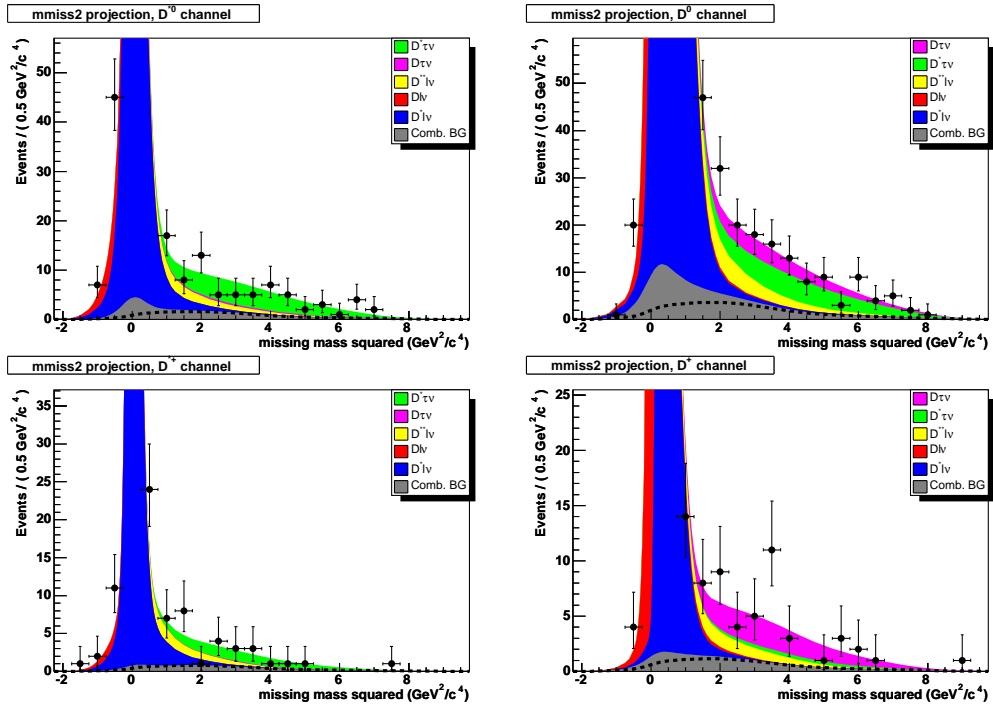


Figure 13.3: m_{miss}^2 projection of the unconstrained fit to data, showing the signal region in more detail.

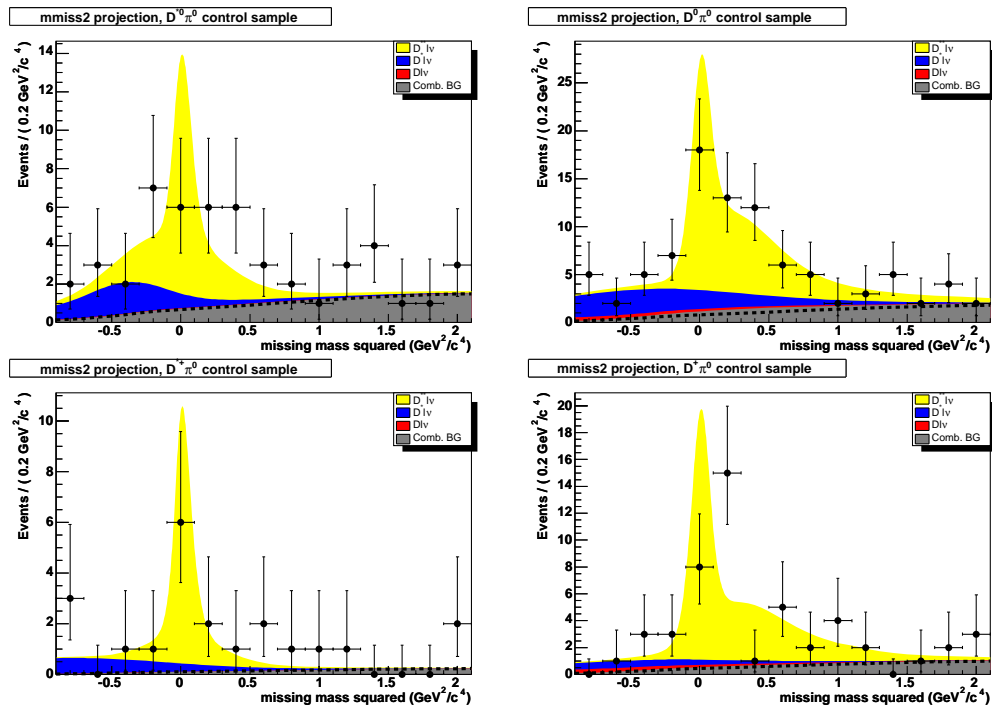


Figure 13.4: m_{miss}^2 projection of the unconstrained fit to data, showing the D^{*0} control samples.

Table 13.1: Results of the branching fraction fits. The event yields listed for both signal and normalization modes include both correctly reconstructed events and crossfeed. The first four rows show the fit without the R -constraint, while the last two show the fit with the constraint imposed (results are quoted for \bar{B}^0). The errors shown here are statistical only.

Channel	Signal yield	Norm yield	R (%)	\mathcal{B} (%)
$D^0 \tau^- \bar{\nu}_\tau$	35.6 ± 19.4	347.9 ± 23.1	31.4 ± 17.0	0.67 ± 0.37
$D^{*0} \tau^- \bar{\nu}_\tau$	92.2 ± 19.6	1629.9 ± 63.6	34.6 ± 7.3	2.25 ± 0.48
$D^+ \tau^- \bar{\nu}_\tau$	23.3 ± 7.8	150.2 ± 13.3	48.9 ± 16.5	1.04 ± 0.35
$D^{*+} \tau^- \bar{\nu}_\tau$	15.5 ± 7.2	482.3 ± 25.5	20.7 ± 9.5	1.11 ± 0.51
$D \tau^- \bar{\nu}_\tau$	66.9 ± 18.9	497.8 ± 26.4	41.6 ± 11.7	0.86 ± 0.24
$D^* \tau^- \bar{\nu}_\tau$	101.4 ± 19.1	2111.5 ± 68.1	29.7 ± 5.6	1.62 ± 0.31

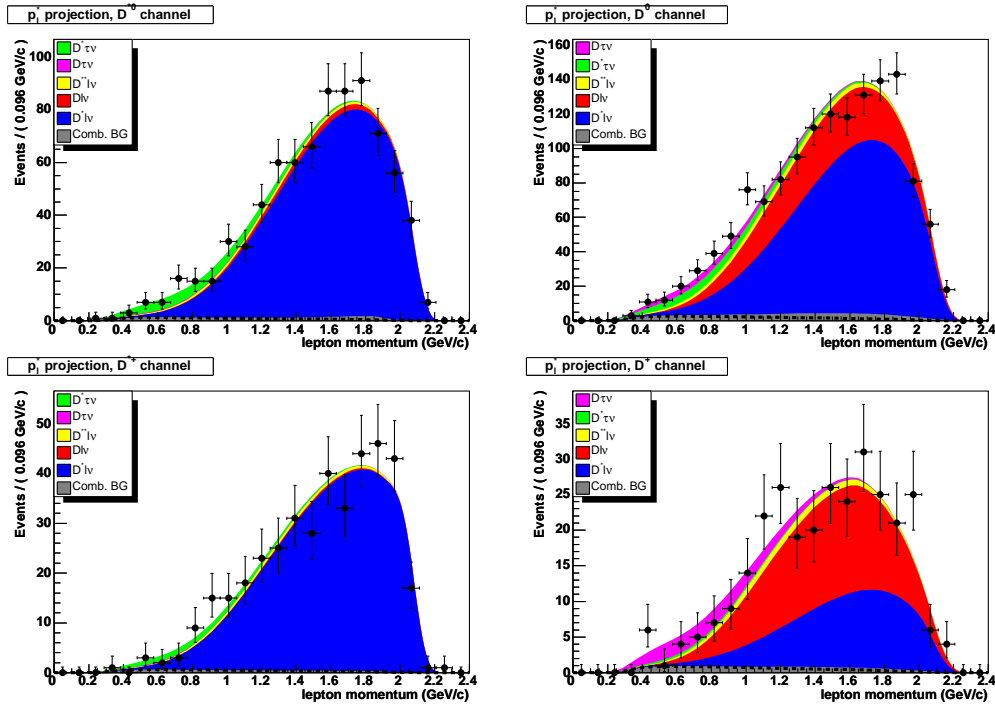


Figure 13.5: p_ℓ^* projection of the unconstrained fit to data, showing the signal channels.

Table 13.2: Event yields in the unconstrained fit.

Component		N_{fit}	Component		N_{fit}
D^{*0}	$D^{*0}\tau^-\bar{\nu}_\tau$	41 ± 8	$D^{*0}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	22 ± 6
	$D^{*0}\ell^-\bar{\nu}_\ell$	688 ± 26		$D^{*0}\ell^-\bar{\nu}_\ell$	$12 \pm \text{—}$
	$D^0\tau^-\bar{\nu}_\tau$	$2 \pm \text{—}$		$D^0\ell^-\bar{\nu}_\ell$	$0 \pm \text{—}$
	$D^0\ell^-\bar{\nu}_\ell$	$20 \pm \text{—}$		Charge XF	0 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$14 \pm \text{—}$		Comb. BG	31 ± 6
	Charge XF	8 ± 0			
	Comb. BG	15 ± 0			
D^0	$D^0\tau^-\bar{\nu}_\tau$	32 ± 17	$D^0\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	56 ± 9
	$D^{*0}\tau^-\bar{\nu}_\tau$	$50 \pm \text{—}$		$D^{*0}\ell^-\bar{\nu}_\ell$	$25 \pm \text{—}$
	$D^0\ell^-\bar{\nu}_\ell$	321 ± 21		$D^0\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^{*0}\ell^-\bar{\nu}_\ell$	$902 \pm \text{—}$		Charge XF	5 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$42 \pm \text{—}$		Comb. BG	37 ± 7
	Charge XF	32 ± 0			
	Comb. BG	33 ± 0			
D^{*+}	$D^{*+}\tau^-\bar{\nu}_\tau$	12 ± 5	$D^{*+}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	11 ± 4
	$D^{*+}\ell^-\bar{\nu}_\ell$	371 ± 19		$D^{*+}\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^+\tau^-\bar{\nu}_\tau$	$0 \pm \text{—}$		$D^+\ell^-\bar{\nu}_\ell$	$0 \pm \text{—}$
	$D^+\ell^-\bar{\nu}_\ell$	$2 \pm \text{—}$		Charge XF	0 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$8 \pm \text{—}$		Comb. BG	4 ± 3
	Charge XF	0 ± 0			
	Comb. BG	7 ± 0			
D^+	$D^+\tau^-\bar{\nu}_\tau$	22 ± 7	$D^+\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	36 ± 7
	$D^{*+}\tau^-\bar{\nu}_\tau$	$3 \pm \text{—}$		$D^{*+}\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^+\ell^-\bar{\nu}_\ell$	143 ± 12		$D^+\ell^-\bar{\nu}_\ell$	$3 \pm \text{—}$
	$D^{*+}\ell^-\bar{\nu}_\ell$	$99 \pm \text{—}$		Charge XF	1 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$11 \pm \text{—}$		Comb. BG	20 ± 6
	Charge XF	3 ± 0			
	Comb. BG	10 ± 0			

Table 13.3: Event yields in the R -constrained fit.

Component		N_{fit}	Component		N_{fit}
D^{*0}	$D^{*0}\tau^-\bar{\nu}_\tau$	35 ± 6	$D^{*0}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	23 ± 6
	$D^{*0}\ell^-\bar{\nu}_\ell$	691 ± 26		$D^{*0}\ell^-\bar{\nu}_\ell$	$12 \pm \text{—}$
	$D^0\tau^-\bar{\nu}_\tau$	$3 \pm \text{—}$		$D^0\ell^-\bar{\nu}_\ell$	$0 \pm \text{—}$
	$D^0\ell^-\bar{\nu}_\ell$	$20 \pm \text{—}$		Charge XF	0 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$15 \pm \text{—}$		Comb. BG	30 ± 6
	Charge XF	8 ± 0			
	Comb. BG	15 ± 0			
D^0	$D^0\tau^-\bar{\nu}_\tau$	43 ± 12	$D^0\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	56 ± 9
	$D^{*0}\tau^-\bar{\nu}_\tau$	$43 \pm \text{—}$		$D^{*0}\ell^-\bar{\nu}_\ell$	$25 \pm \text{—}$
	$D^0\ell^-\bar{\nu}_\ell$	320 ± 21		$D^0\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^{*0}\ell^-\bar{\nu}_\ell$	$901 \pm \text{—}$		Charge XF	5 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$42 \pm \text{—}$		Comb. BG	37 ± 7
	Charge XF	32 ± 0			
	Comb. BG	33 ± 0			
D^{*+}	$D^{*+}\tau^-\bar{\nu}_\tau$	17 ± 3	$D^{*+}\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	10 ± 3
	$D^{*+}\ell^-\bar{\nu}_\ell$	368 ± 19		$D^{*+}\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^+\tau^-\bar{\nu}_\tau$	$0 \pm \text{—}$		$D^+\ell^-\bar{\nu}_\ell$	$0 \pm \text{—}$
	$D^+\ell^-\bar{\nu}_\ell$	$2 \pm \text{—}$		Charge XF	0 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$7 \pm \text{—}$		Comb. BG	5 ± 3
	Charge XF	0 ± 0			
	Comb. BG	7 ± 0			
D^+	$D^+\tau^-\bar{\nu}_\tau$	19 ± 5	$D^+\pi^0$	$D^{**}\ell^-\bar{\nu}_\ell$	36 ± 7
	$D^{*+}\tau^-\bar{\nu}_\tau$	$4 \pm \text{—}$		$D^{*+}\ell^-\bar{\nu}_\ell$	$5 \pm \text{—}$
	$D^+\ell^-\bar{\nu}_\ell$	144 ± 12		$D^+\ell^-\bar{\nu}_\ell$	$3 \pm \text{—}$
	$D^{*+}\ell^-\bar{\nu}_\ell$	$100 \pm \text{—}$		Charge XF	1 ± 0
	$D^{**}\ell^-\bar{\nu}_\ell$	$11 \pm \text{—}$		Comb. BG	20 ± 6
	Charge XF	3 ± 0			
	Comb. BG	10 ± 0			

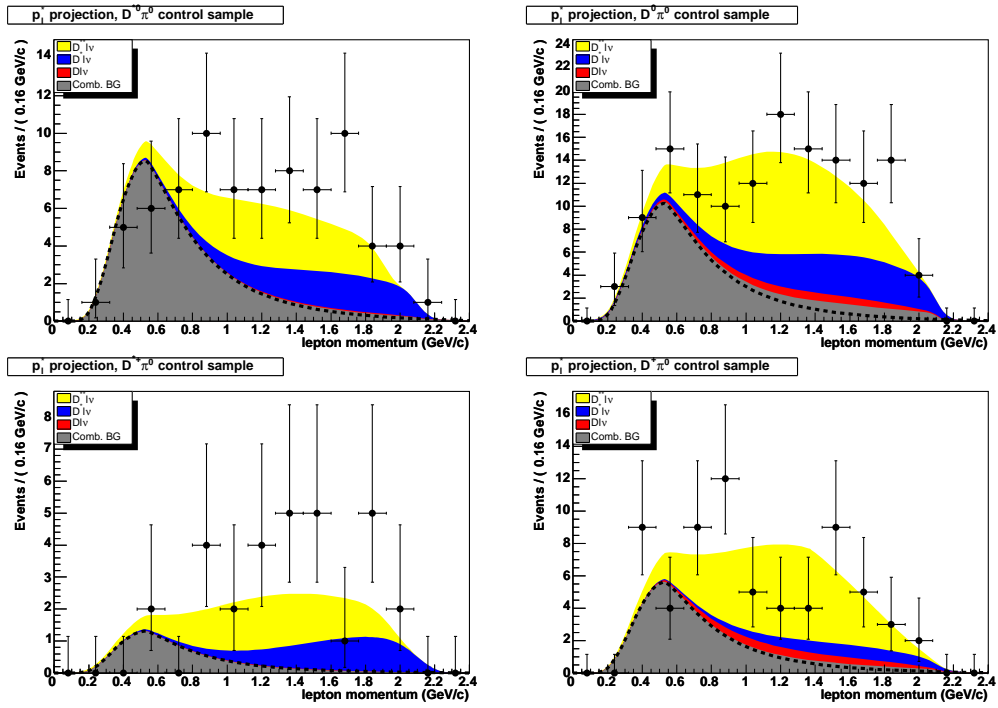


Figure 13.6: p_{ℓ}^* projection of the unconstrained fit to data, showing the D^{**} control samples.

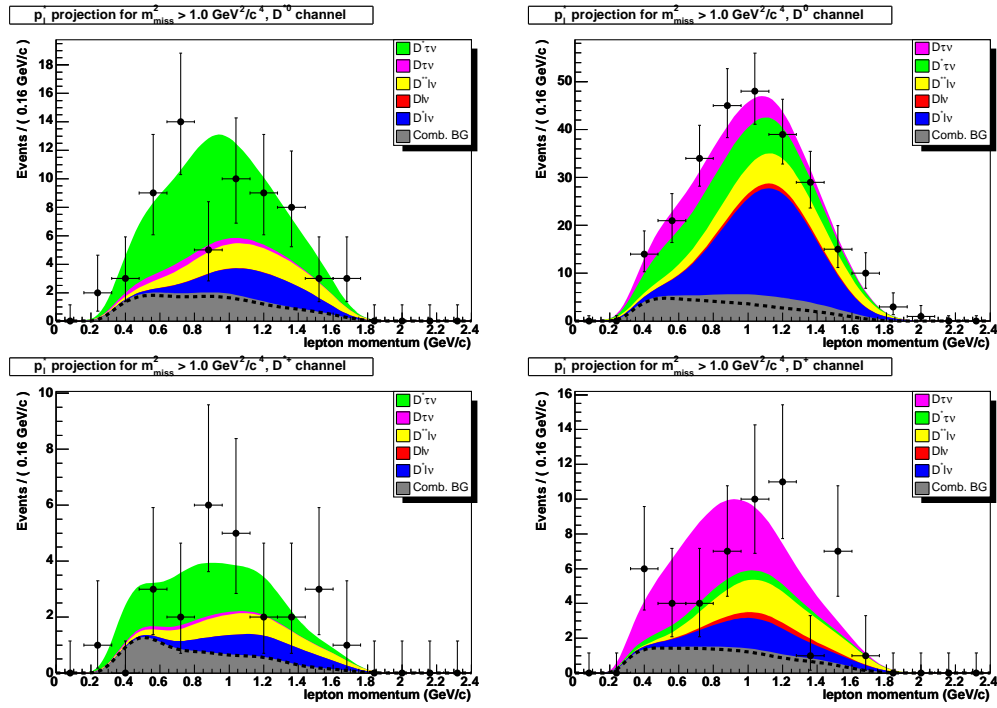


Figure 13.7: p_ℓ^* projection of the unconstrained fit to data, showing the signal channels in the signal-enhanced region $m_{\text{miss}}^2 > 1 \text{ (GeV}/c^2)^2$.

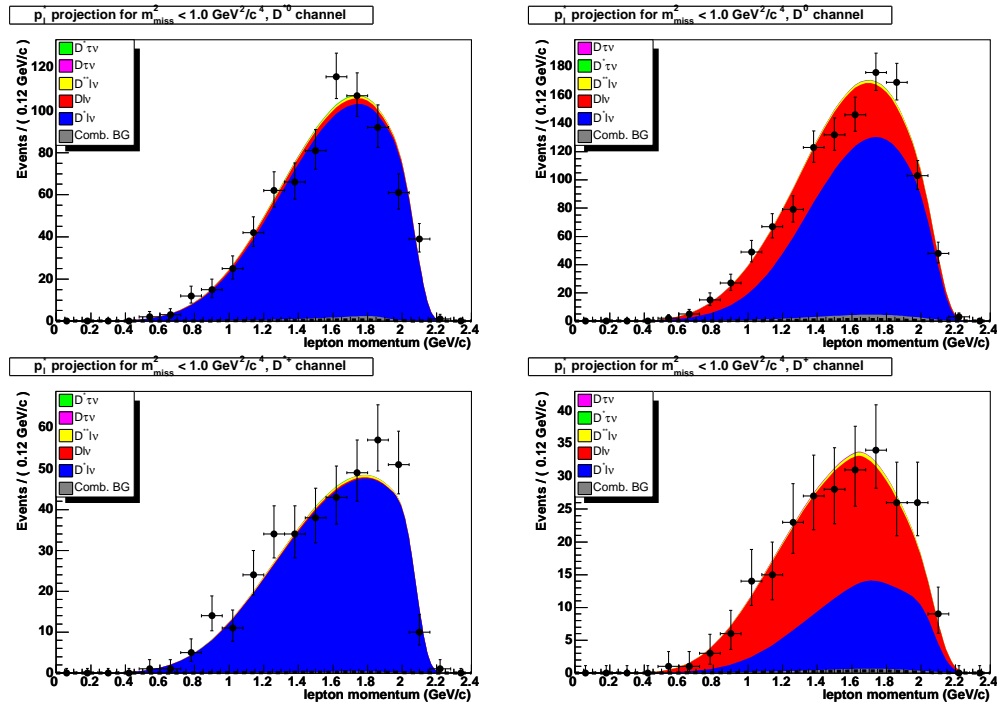


Figure 13.8: p_ℓ^* projection of the unconstrained fit to data, showing the signal channels in the normalization region $m_{\text{miss}}^2 < 1 \text{ (GeV}/c^2)^2$.

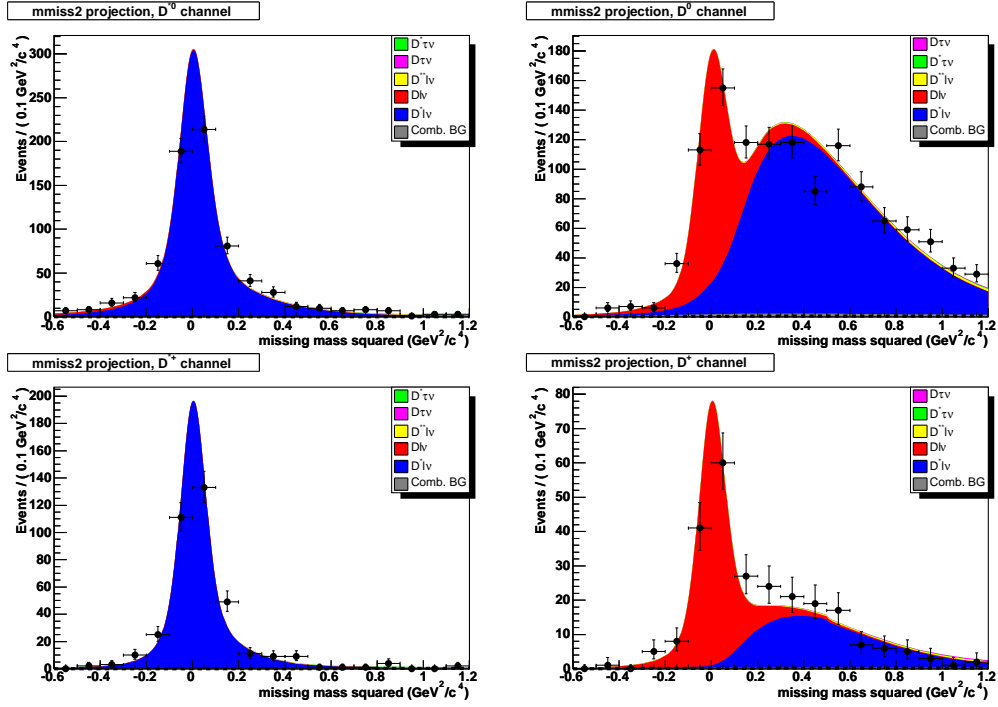


Figure 13.9: m_{miss}^2 projection of the R -constrained fit to data, showing the normalization region in the signal channels.

13.2 Significance of Signals

We calculate the significance of the signals by performing fits in which the signal branching fractions are fixed to zero. We calculate the difference in log-likelihood, $\Delta(\ln \mathcal{L})$, between the nominal fit and the no-signal case, and we take the statistical significance to be $\sqrt{2\Delta(\ln \mathcal{L})}$. We include the systematic errors in this by multiplying the significance by the ratio $\sigma_{\text{stat}}/\sigma_{\text{stat+additivesyst}}$; multiplying the significance like this is equivalent to scaling the log-likelihood

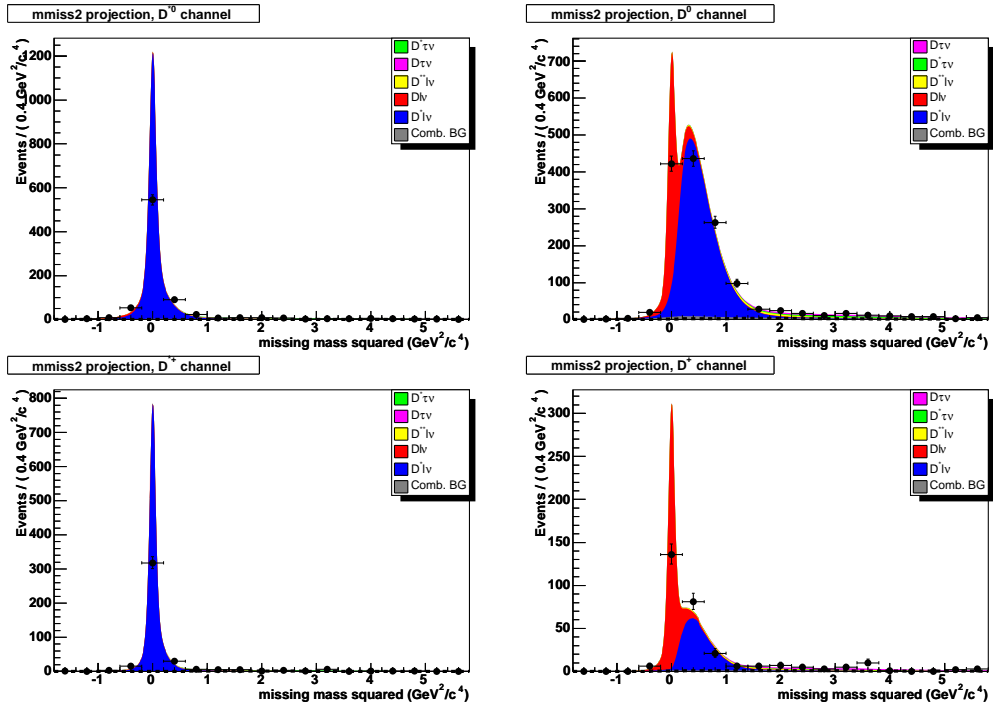


Figure 13.10: m_{miss}^2 projection of the R -constrained fit to data, showing a wider range than Figure 13.9.

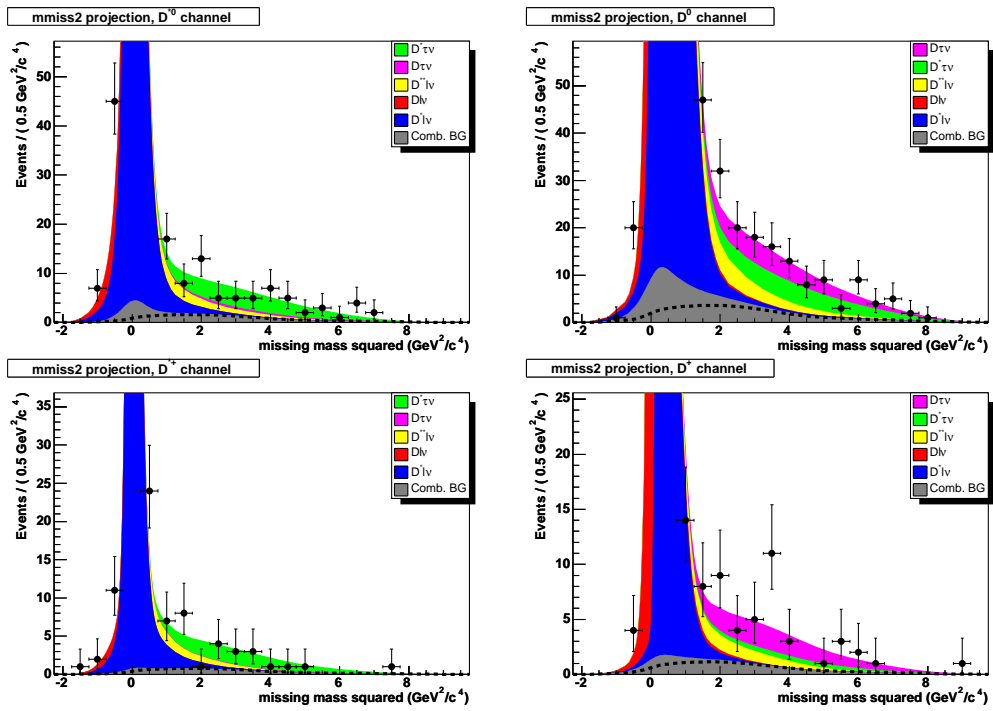


Figure 13.11: m_{miss}^2 projection of the R -constrained fit to data, showing the signal region in more detail.

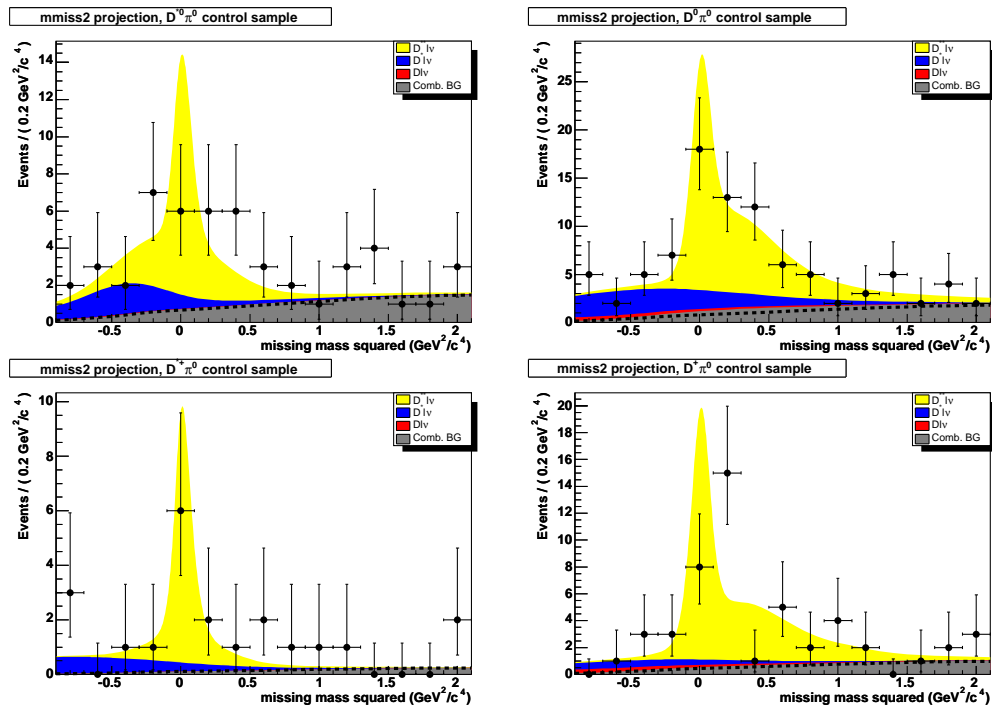


Figure 13.12: m_{miss}^2 projection of the R -constrained fit to data, showing the $D^{**}\pi^0$ control samples.

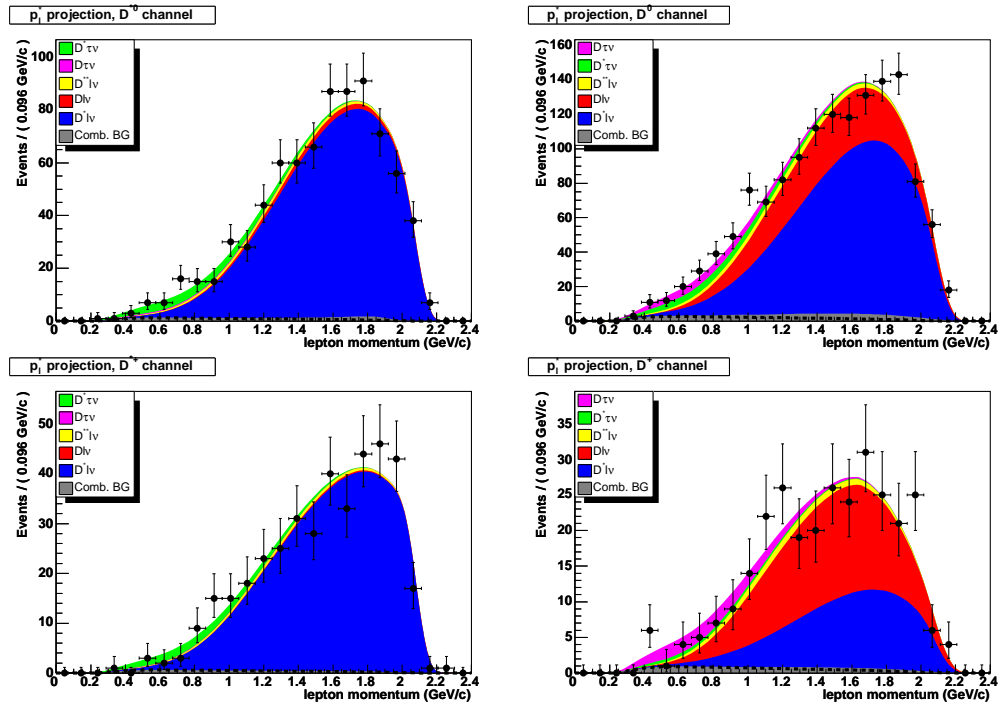


Figure 13.13: p_ℓ^* projection of the R -constrained fit to data, showing the signal channels.

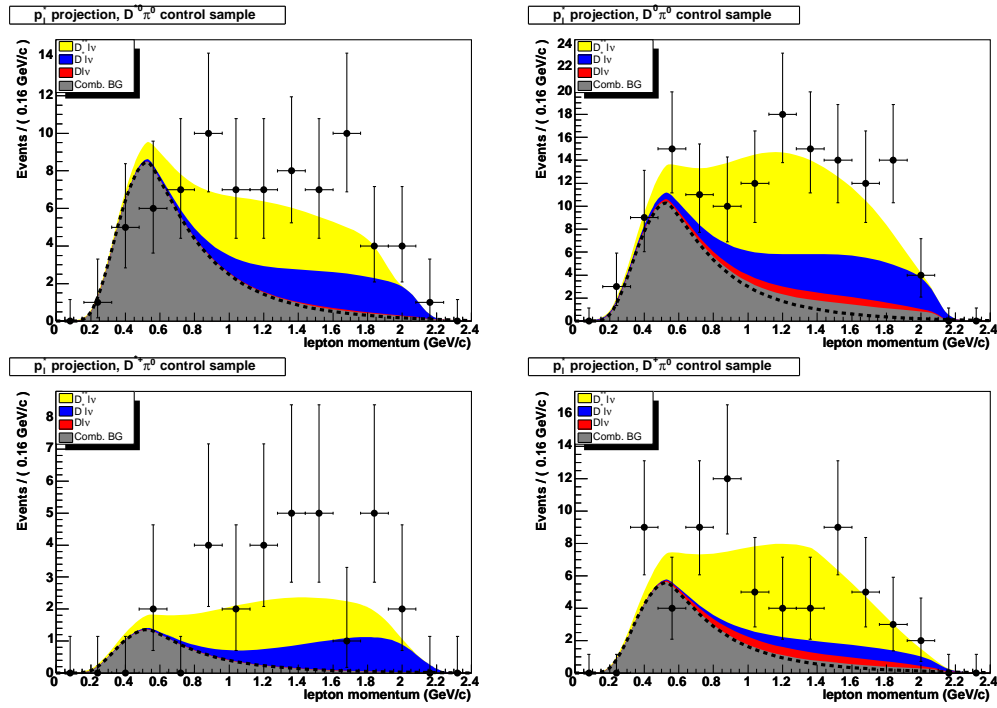


Figure 13.14: p_ℓ^* projection of the R -constrained fit to data, showing the D^{**} control samples.

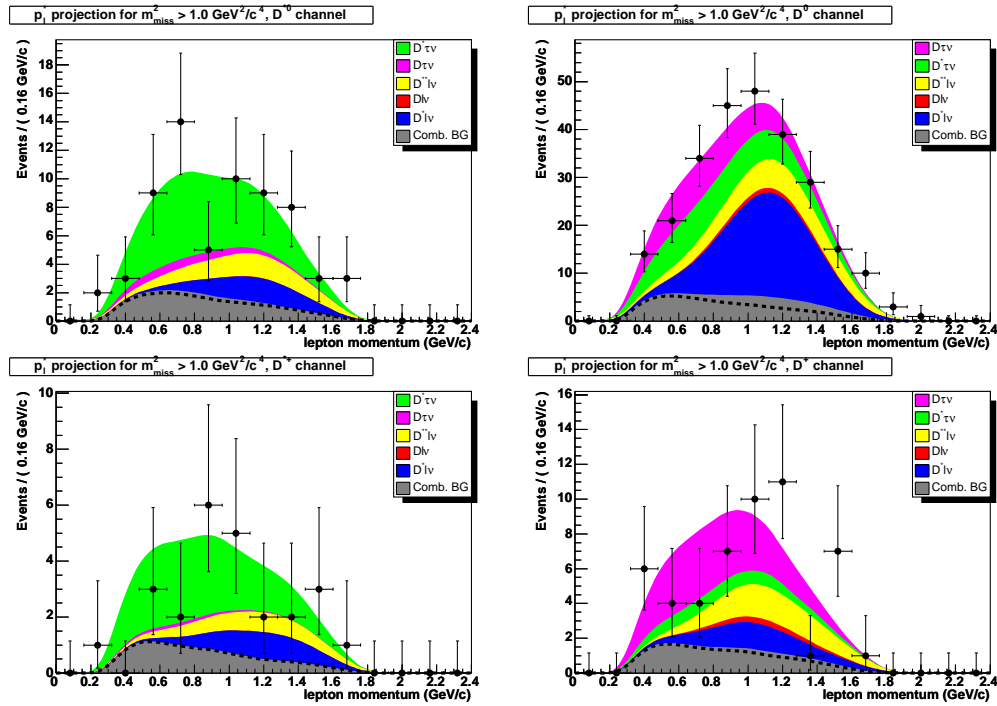


Figure 13.15: p_l^* projection of the R -constrained fit to data, showing the signal channels in the signal-enhanced region $m_{\text{miss}}^2 > 1 \text{ (GeV}/c^2)^2$.

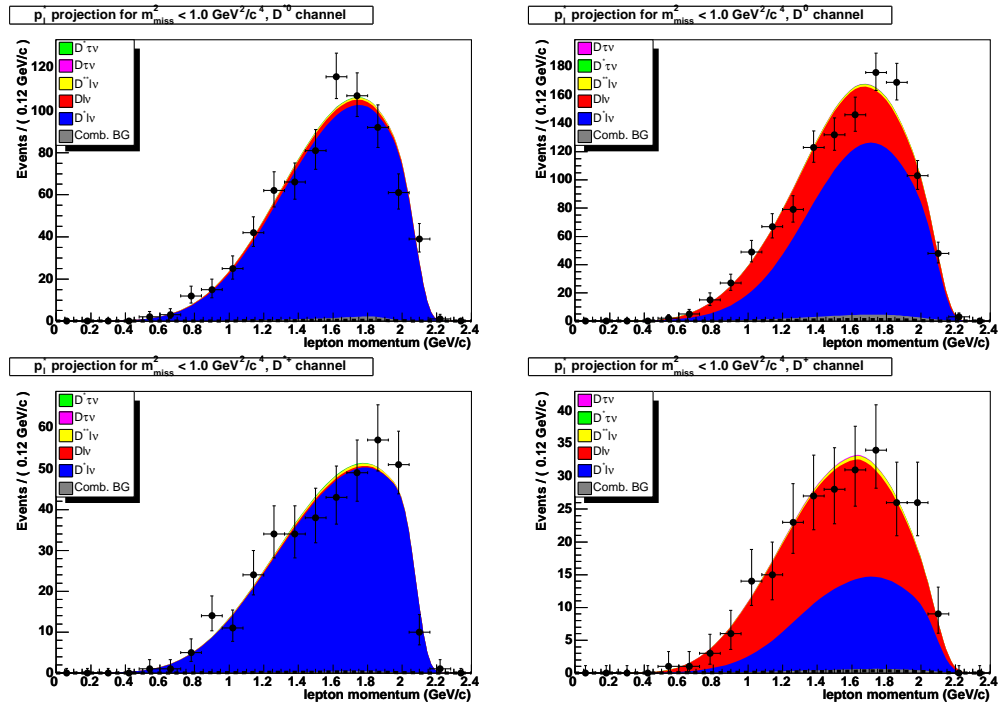


Figure 13.16: p_l^* projection of the R -constrained fit to data, showing the signal channels in the normalization region $m_{\text{miss}}^2 < 1 \text{ (GeV}/c^2)^2$.

Table 13.4: Significances of the signals.

Channel	σ (stat only)	σ (total)
$D^0 \tau^- \bar{\nu}_\tau$	1.8	1.8
$D^{*0} \tau^- \bar{\nu}_\tau$	5.8	5.3
$D^+ \tau^- \bar{\nu}_\tau$	3.6	3.3
$D^{*+} \tau^- \bar{\nu}_\tau$	2.7	2.7
$D \tau^- \bar{\nu}_\tau$	4.0	3.6
$D^* \tau^- \bar{\nu}_\tau$	6.5	6.2

curve. The significances are presented in Table 13.4. We observe significant signals for both $D \tau^- \bar{\nu}_\tau$ and $D^* \tau^- \bar{\nu}_\tau$.

13.3 Stability Checks

We have performed several checks of the composition of the event sample to ensure that our results are stable. We have divided the sample roughly in half in three different ways, according to lepton flavor (e vs μ), lepton charge (ℓ^+ vs ℓ^-), and running period (Runs 1–3 vs Run 4), and, in all three studies we compare the composition between data and Monte Carlo. We perform these checks both in the full data sample and in the signal-enhanced region $m_{\text{miss}}^2 > 1(\text{GeV}/c^2)^2$. These checks are summarized in Table 13.5, and we see that, in all cases, the data abundances are consistent with the Monte Carlo. The ℓ^+ abundances agree with the theoretical expectation of a 50/50 split,

Table 13.5: Stability checks in the final data sample. The first set of numbers corresponds to the fraction of muons in the sample, the second set to the fraction of positively-charged leptons, and the third to the fraction of events which were recorded during Run 4. In all cases, the data abundances agree with the MC prediction. The ℓ^+ abundances agree with the theoretical expectation of a 50/50 split, and the Run 4 abundance is consistent with the ratio of luminosities — Run 4 makes up 47.5% of the total recorded luminosity.

Sample		f_{data} (%)	f_{MC} (%)
μ	full sample	40.0 ± 0.9	40.9 ± 0.4
	high m_{miss}^2	30.7 ± 2.3	31.9 ± 1.0
ℓ^+	full sample	50.2 ± 1.0	49.2 ± 0.4
	high m_{miss}^2	49.3 ± 2.5	48.9 ± 1.0
Run 4	full sample	44.6 ± 1.0	47.6 ± 0.5
	high m_{miss}^2	46.8 ± 2.5	48.5 ± 1.2

and the Run 4 abundance is consistent with the ratio of luminosities — Run 4 makes up 47.5% of the total recorded luminosity.

13.4 Goodness of Fit

We have estimated the goodness of fit using an ensemble of toy MC experiments. We generate 1000 data-equivalent samples, using the nominal PDFs for the fit to data and event yields based on the R -constrained fit to data. We fit each of these samples both with and without the R constraints and study the distribution of $-\ln \mathcal{L}$ in these fits.

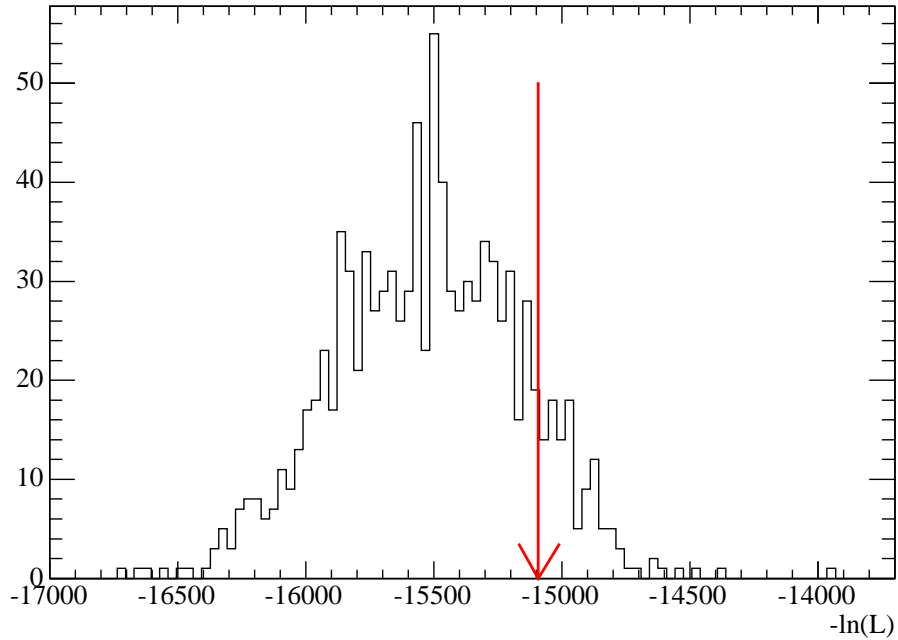


Figure 13.17: Distribution of $-\ln \mathcal{L}$ for the unconstrained fit, used to estimate goodness of fit.

Figure 13.17 shows the distribution of $-\ln \mathcal{L}$ for the ensemble of unconstrained fits, while Figure 13.18 shows the distribution for the R -constrained fits. In both figures, the value of $-\ln \mathcal{L}$ obtained in the fit to data is indicated with an arrow, and, in both cases, this value is found within the central part of the distribution, indicating a reasonably good fit. In the unconstrained fit, 11.7% of the toy experiments have a value of $-\ln \mathcal{L}$ greater than the value observed in data, indicating an acceptable goodness of fit. The corresponding probability for the R -constrained fit is 11.8%, also acceptably large.

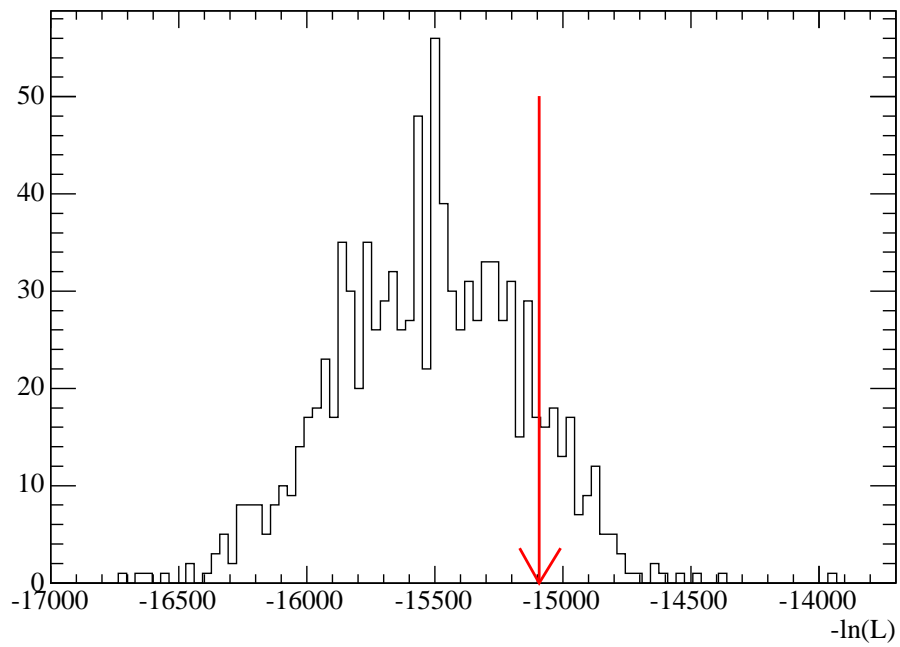


Figure 13.18: Distribution of $-\ln \mathcal{L}$ for the R -constrained fit, used to estimate goodness of fit.

Chapter 14

Results and Conclusions

We have presented measurements of the branching ratios for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ relative to the corresponding light-lepton modes $B \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$:

$$\begin{aligned} R^0 &= (31.4 \pm 17.0 \pm 4.9)\% \\ R^{*0} &= (34.6 \pm 7.3 \pm 3.4)\% \\ R^+ &= (48.9 \pm 16.5 \pm 6.9)\% \\ R^{*+} &= (20.7 \pm 9.5 \pm 0.8)\% , \end{aligned}$$

and, combining B^- and \bar{B}^0 channels,

$$\begin{aligned} R &= (41.6 \pm 11.7 \pm 5.2)\% \\ R^* &= (29.7 \pm 5.6 \pm 1.8)\% , \end{aligned}$$

where the first error is statistical and the second is systematic. The significances of the signals are 1.8σ , 5.3σ , 3.3σ , and 2.7σ , respectively, for the individual branching ratios, and 3.6σ and 6.2σ for the combined measurement.

From these branching ratios and the known branching fractions of the normalization modes, we obtain the absolute branching fractions

$$\begin{aligned}
 \mathcal{B}(B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau) &= (0.67 \pm 0.37 \pm 0.11 \pm 0.07)\% \\
 \mathcal{B}(B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau) &= (2.25 \pm 0.48 \pm 0.22 \pm 0.17)\% \\
 \mathcal{B}(\bar{B}^0 \rightarrow D^+ \tau^- \bar{\nu}_\tau) &= (1.04 \pm 0.35 \pm 0.15 \pm 0.10)\% \\
 \mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau) &= (1.11 \pm 0.51 \pm 0.04 \pm 0.04)\% ,
 \end{aligned}$$

and, for the combined measurement,

$$\begin{aligned}
 \mathcal{B}(B \rightarrow D \tau^- \bar{\nu}_\tau) &= (0.86 \pm 0.24 \pm 0.11 \pm 0.06)\% \\
 \mathcal{B}(B \rightarrow D^* \tau^- \bar{\nu}_\tau) &= (1.62 \pm 0.31 \pm 0.10 \pm 0.05)\% ,
 \end{aligned}$$

where the third error is due to uncertainties on the normalization mode. The branching fractions for the combined measurement are quoted for the \bar{B}^0 lifetime.

This represents the first evidence for $B \rightarrow D \tau^- \bar{\nu}_\tau$, which is important due to the sensitivity to potential new physics contributions. The measurements of $B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$ and $\bar{B}^0 \rightarrow D^+ \tau^- \bar{\nu}_\tau$ represent the first observation and first evidence for these modes, respectively. The measurement of $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ is consistent with a preliminary result from the Belle Collaboration [50].

The branching ratios R and R^* are about 1σ higher than predictions based on the Standard Model, but, given the uncertainties, there is still room for a sizeable non-SM contribution.

Bibliography

- [1] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **73** 092001 (2006).
- [2] CLEO Collaboration, J.E. Duboscq *et al.*, Phys. Rev. Lett. **76** 3898 (1996).
- [3] BAD Note #1395 (to be published soon); to supercede the publication: B. Aubert *et al.*, the *BABAR* Collaboration; Phys. Rev. D **74** 092004 (2006).
- [4] J. Richman, In: R. Gupta *et al.* (eds.), *Les Houches, Session LXVIII*, 28 July-5 September 1997.
- [5] M. Neubert, Phys. Rep. **245** 259 (1994).
- [6] C.-H. Chen and C.-Q. Geng, JHEP **0610**, 053 (2006).
- [7] A.F. Falk *et al.*, Phys. Lett. B **326** 145 (1994).
- [8] M. Tanaka, Z. Phys. C **67** 321 (1995).
- [9] J.G. Körner and G.A. Schuler, Z. Phys. C **46** 93 (1990).
- [10] Particle Data Group, W.-M. Yau *et al.*, J. Phys. G: Nucl. Part. Phys. **33**, 1 (2006).
- [11] ALEPH collab., Eur. Phys. Jour. C **19** 213 (2001).
- [12] L3 collab., Z. Phys. C **71** 379 (1996).
- [13] L3 collab., Phys. Lett. B **332** 201 (1994).
- [14] OPAL collab., Phys. Lett. B **520** 1 (2001).
- [15] T. Miura and M. Tanaka, Talk at Workshop on Higher Luminosity B Factory, KEK, Japan, 23-24 August 2001; preprint hep-ph/0109244.

Bibliography

- [16] Y. Grossman, H. Haber and Y. Nir; preprint hep-ph/9507213.
- [17] A.G. Akeroyd *et al*; preprint hep-ex/0406071.
- [18] BAD Note #540.
- [19] M. Wirbel, B. Stech, and M. Bauer, Z. Phys. C **29** 637 (1985).
- [20] N. Isgur, D. Scora, B. Grinstein, and M.B. Wise, Phys. Rev. D **39** 799 (1989).
- [21] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. B **530** 153 (1998).
- [22] J.G. Körner and G.A. Schuler, Z. Phys. C **38** 511 (1988) [Erratum: Z. Phys. C **41** 690 (1989)].
- [23] R. Garisto, Phys. Rev. D **51** 1107 (1995).
- [24] M. Neubert and V. Rieckert, Nucl. Phys. B **382** 97 (1992).
- [25] N. Isgur and M.B. Wise, Phys. Lett. B **232** 113 (1989); Phys. Lett. B **237** 527 (1990).
- [26] H. Itoh, S. Komine, and Y. Okada, Prog. Theor. Phys. **14** 179 (2005).
- [27] D. Lange *et al*; Nucl. Instr. Methods Phys. Res., Sect. A **462** 152 (2001).
- [28] S. Agostinelli *et al*, the GEANT Collaboration; Nucl. Instr. Methods Phys. Res., Sect. A **506** 250 (2003).
- [29] E. Richter-Was; Phys. Lett. B **303** 163 (1993).
- [30] D. Scora and N. Isgur, Phys. Rev. D **52** 2783 (1995).
- [31] J.L. Goity and W. Roberts, Phys. Rev. D **51** 3459 (1995).
- [32] BAD Note #271.
- [33] The Crystal Ball function is given by

$$C(m_{\text{ES}}) \propto \begin{cases} \left(\frac{n}{\alpha}\right)^n \frac{\exp(-\alpha^2/2)}{\left(\frac{(\mu-m_{\text{ES}})/\sigma + n/\alpha - \alpha}{\sigma}\right)^n} & m_{\text{ES}} < \mu - \alpha\sigma \\ \exp\left[\frac{-1}{2} \left(\frac{m_{\text{ES}}-\mu}{\sigma}\right)^2\right] & m_{\text{ES}} > \mu - \alpha\sigma \end{cases},$$

where the parameters μ and σ are the mean and resolution of a Gaussian and the parameter n describes a power-law function. The Gaussian and the power-law functions are joined continuously at the point $\mu - \alpha\sigma$. The four parameters μ , σ , n , and α are free in the fit; D. Antreasyan, Crystal Ball Note 321 (1983).

- [34] The ARGUS function is $A(m_{\text{ES}}) \propto m_{\text{ES}} \sqrt{1-x^2} \exp[-\zeta(1-x^2)]$, where $x \equiv 2m_{\text{ES}}/\sqrt{s}$ and ζ is a free parameter in the fit; ARGUS Collaboration., H. Albrecht *et al*, Z. Phys. C **48** 543 (1990).
- [35] <http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/EHBDoc/skims/SemiExclskimdefinition.html>
- [36] BAD Note #347.
- [37] BAD Note #12, Appendix B.
- [38] W. Hulsburgen; preprint physics/0503191.
- [39] BAD Note #474.
- [40] BAD Note #514.
- [41] BAD Note #633.
- [42] <http://www.slac.stanford.edu/BFROOT/www/doc/BbrMeetingOrganizer/vol02/pro/mtg000833/itm0002993/Reweight.pdf>. In particular, we use the second approach described in this talk, which replaces a previous study found here: <http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/InclusiveSL/common/main.html>
- [43] D. Côté *et al.*; Eur. Phys. Jour. C **38** 105 (2004).
- [44] HFAG winter 2006 average, <http://www.slac.stanford.edu/xorg/hfag/semi/winter06/winter06.shtml>
- [45] <http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2004.html>
- [46] <http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/PidOnMc/pidonmc.html>

Bibliography

- [47] http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals/validation/recipe14_improved.html
- [48] BAD Note #1136, publication pending.
- [49] BAD Note #664.
- [50] Belle Collab., A. Matyja *et al.*, arXiv:0706:4429, submitted to Physical Review Letters.

Appendices

Appendix A

Definitions of Standard Lists

A.1 Raw Lists

The two most basic lists of reconstructed objects are *ChargedTracks*, which consists of all candidates found in the tracking subsystems, and *CalorClusterNeutral*, which consists of all clusters in the EMC not matched to a *ChargedTracks* object. Candidates from these two lists are used throughout the analysis.

A third raw list, *CalorNeutral*, consists of single-bump *CalorClusterNeutral* candidates. This list will be used for all subsequent photon lists; the *CalorClusterNeutral* list is only used in merged π^0 reconstruction.

A.2 Refined Track Lists

The *GoodTracksLoose* list (GTL) is a subset of *ChargedTracks* which is used for reconstructing the signal-side ℓ candidate. The definition of GTL is based on the following criteria:

- The point of closest approach to the beam spot (POCA) must lie within 1.5 cm of the center of the beam spot, measured in the x - y plane.
- The POCA must lie within 10 cm of the center of the beam spot, measured along the z -axis.
- The track must have a momentum less than 10 GeV/ c in the lab frame.
- The track must have a transverse momentum greater than 100 MeV/ c in the lab frame.
- The track must have at least 12 DCH hits.

A.3 Refined Neutral Lists

The *GoodPhotonLoose* list is a subset of *CalorNeutral* which is used in the Bremsstrahlung recovery algorithm. The γ candidates in this list must satisfy the following criteria:

- $E_{\text{cluster}} > 30 \text{ MeV}$.
- $\text{LAT} < 0.8$.

A.4 Electron ID

For electron identification, we use the *PidLHElectrons* selector, both during the Vcb reskim and the final analysis. This selector runs over the *Charged-Tracks* list and requires e candidates to satisfy:

- $p > 300 \text{ MeV}/c$.
- $0.5 < E/p < 1.5$.
- $n_{\text{crystal}} \geq 4$.
- $500 < dE/dx < 1000$ as measured in the DCH.
- $-10 < Z_{42} < 10$.
- $-0.05 < \delta\phi < 0.15$, where $\delta\phi \equiv q_{\text{trk}} \cdot (\phi_{\text{trk}}^{\text{EMC}} - \phi_{\text{cluster}}^{\text{EMC}})$.
- The fractional likelihood $\mathcal{L}_{\text{frac}} > 0.95$.

The fractional likelihood is defined as

Appendix A. Definitions of Standard Lists

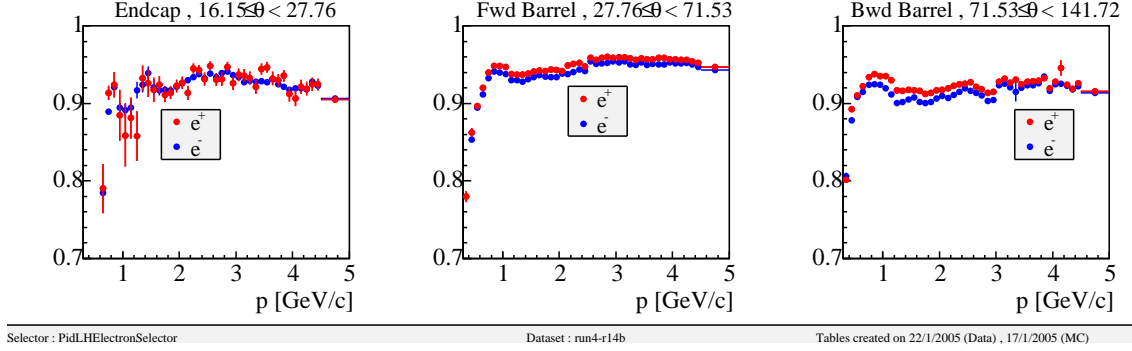


Figure A.1: Efficiency of the *PidLHElectrons* selector as a function of electron momentum and polar angle.

$$\mathcal{L}_{\text{frac}} \equiv \frac{f_e \mathcal{L}_e}{f_e \mathcal{L}_e + f_\pi \mathcal{L}_\pi + f_K \mathcal{L}_K + f_p \mathcal{L}_p}$$

where the f_i terms are the expected abundances of electrons, pions, kaons, and protons, and the \mathcal{L}_i terms are the likelihood for this track to match the i^{th} particle hypothesis. The likelihoods are calculated using E/p , LAT, $\delta\phi$, dE/dx , and θ_c .

Figure A.1 shows the efficiency of the *PidLHElectrons* selector as a function of electron momentum and polar angle.

A.5 Muon ID

A.5.1 muMicroTight

The primary muon list used in this analysis is *muMicroTight* which is a subset of the *ChargedTracks* list. Muon candidates are divided into five bins of polar angle in the lab frame in order to capture the geometric variation in IFR performance; these bins will be labeled, in order of increasing polar angle, forward, mid-forward, barrel, mid-backward, and backward.

Each μ candidate is required to pass several cuts, the values of which are, in general, different in each of the five θ bins. For the *muMicroTight* selector, only one of these cuts actually differs across the polar-angle bins. The actual cuts used are:

- $50 < E_{\text{EMC}} < 400$ MeV for muons inside the EMC acceptance.
- $N_{\text{IFRlayers}} > 1$.
- The difference between the measured and expected interaction lengths of material traversed by the track, λ_{meas} and λ_{exp} , must satisfy both

$$- (\lambda_{\text{exp}} - \lambda_{\text{meas}}) < 1 \text{ \textbf{and}}$$

$$- (\lambda_{\text{exp}} - \lambda_{\text{meas}}) < \frac{p-500 \text{ MeV}/c}{700 \text{ MeV}/c}.$$

- The measured interaction length traversed must also satisfy either
 - $\lambda_{\text{meas}} > 2.2$ **or**
 - $\lambda_{\text{meas}} > \frac{p+1 \text{ GeV}/c}{0.45 \text{ GeV}/c}$.
- The $\chi^2/\text{degree-of-freedom}$ of the match between the track and the IFR cluster must be less than 5.
- The $\chi^2/\text{degree-of-freedom}$ of the IFR cluster position must be less than 3.
- The average number of IFR strip hits per layer must be less than 8.
- The IFR continuity is defined as the fraction of IFR layers between the first and last layers hit, inclusive, which show a hit on this track. The IFR continuity must be greater than 0.3 for the forward and mid-forward bins; no cut on continuity is applied in the other three bins.
- The μ candidate must also fail the criteria for the kaon selector *KMicroTight*. This selector identifies kaons with cuts using dE/dx in the SVT and DCH and θ_c ; the algorithm is similar to the kaon likelihood selectors described below, but the likelihood cuts are optimized differently.

Figure A.2 shows the efficiency of the *muMicroTight* selector as a function of electron momentum and polar angle.

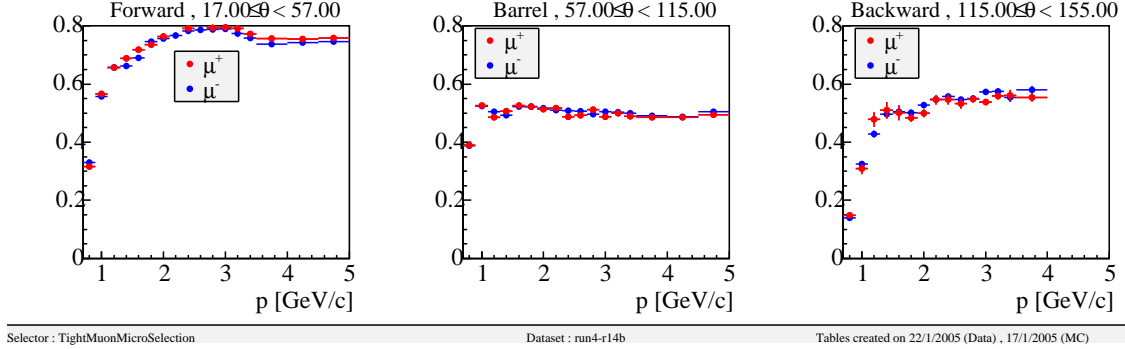


Figure A.2: Efficiency of the muMicroTight selector as a function of electron momentum and polar angle.

A.5.2 Neural Net Muon ID

Two additional muon selectors were used in the Vcb reskim and were considered for use in the analysis: *muNNTight* and *muNNTightFakeRate*.

These selectors use the same variables as the selector described above but, rather than cutting on the variables to select candidates, the variables are fed into a neural network and the output of the network is used to make the decision. Two configurations of the neural net were used, having different optimizations for signal and background efficiency: *muNNTight* is designed to be 70% efficient for typical muons in *BABAR*, while *muNNTightFakeRate* is designed to have a pion fake rate less than 2% for typical pions. The neural net-based muon selectors are described in [39].

A.6 Kaon ID

We use two selectors for kaon identification: *KLHVeryLoose* during the Vcb reskim and *KLHTight* during the final analysis. As with muons, we use a loose cut during the reskim phase to remain flexible with our choice of selectors as long as possible. Both selectors run over *ChargedTracks*.

A likelihood is constructed for each track in three particle hypotheses: pion, kaon, and proton. The likelihood uses dE/dx from both the SVT and DCH and the Cherenkov angle θ_c . Two likelihood ratios are then constructed as

$$\mathcal{L}_{K\text{-vs-}\pi} \equiv \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \quad \text{and} \quad \mathcal{L}_{K\text{-vs-}p} \equiv \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_p}.$$

Kaon candidates are required to satisfy:

- $\mathcal{L}_{K\text{-vs-}\pi} > 0.5$ for the *VeryLoose* selection and 0.9 for the *Tight* selection.
- $\mathcal{L}_{K\text{-vs-}p} > 0.018$ for the *VeryLoose* selection and 0.2 for the *Tight* selection.
- Kaon candidates with a momentum $p > 400$ MeV/ c must also fail the *PidLHElectrons* selector.

Figure A.3 shows the efficiency of the *KLHTight* selector as a function of electron momentum and polar angle.

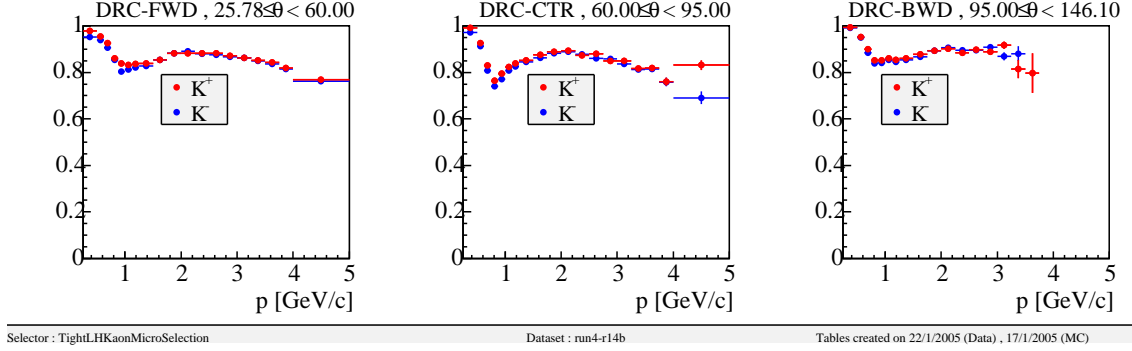


Figure A.3: Efficiency of the KLHTight selector as a function of electron momentum and polar angle.

A.7 K_S^0 Reconstruction

K_S^0 candidates on the signal-side are reconstructed using the *KsDefault* list.

This list consists of $K_S^0 \rightarrow \pi^+\pi^-$ candidates, with both pions coming from the *ChargedTracks* list, which satisfy the following requirements:

- A cut is applied to the raw mass $300 < m_{\pi\pi} < 700 \text{ MeV}/c^2$.
- The K_S^0 candidate is vertexed using GeoKin, and a cut is applied to the fitted mass $472.67 < m_{\pi\pi}^{\text{fit}} < 522.67 \text{ MeV}/c^2$. Note that this cut is further refined in the analysis; see Sec. 6.7.2.
- The flight length of the K_S^0 candidate (measured from the nominal beam spot location) must satisfy $0 < f < 999 \text{ cm}$.

A.8 π^0 Reconstruction

We use two lists of π^0 candidates, *Pi0AllDefault* and *pi0VeryLoose*, as described in Sec. 6.7.1.

A.8.1 pi0VeryLoose

The simpler of the two π^0 lists is *pi0VeryLoose*, which consists of $\pi^0 \rightarrow \gamma\gamma$ candidates, where both γ candidates are taken from the *GoodPhotonLoose* list. The two γ candidates are combined using 4-vector addition, and any π^0 candidate satisfying $90 < m_{\gamma\gamma} < 165 \text{ MeV}/c^2$ is accepted.

A.8.2 Pi0AllDefault

The *Pi0AllDefault* list is formed as the union of two separate lists, *Pi0AllLoose* and *MergedPi0Loose*. The first list, *Pi0AllLoose*, consists of $\pi^0 \rightarrow \gamma\gamma$ candidates, where both γ candidates are taken from the *GoodPhotonLoose* list. The π^0 candidates must satisfy the following requirements:

- $115 < m_{\gamma\gamma} < 150 \text{ MeV}/c^2$. Note that this cut is further refined in the analysis.
- $E_{\gamma\gamma} > 200 \text{ MeV}$.

Appendix A. Definitions of Standard Lists

- The π^0 candidate is kinematically fitted using GeoKin with a mass constraint and assuming the primary vertex as the production point.

The *MergedPi0Loose* list is documented in [40] and consists of *CalorClusterNeutral* clusters satisfying the following requirements:

- $E_{\text{cluster}} > 1 \text{ GeV}$.
- The candidate cluster must have consistency with the π^0 hypothesis greater than 0.01. The consistency is defined using four quantities: E_{cluster} , the second moment of the cluster, the position of the cluster centroid, and the number of bumps in the cluster.

Appendix B

$D^{(*)}$ Candidate Selection

We show here the selection criteria applied to $D^{(*)}$ candidates. In each D mode, we fit the invariant mass distribution from signal Monte Carlo to the sum of two Gaussian components. The mass window for all modes is then set to $\pm 4\sigma$ about the fitted mean, where σ here is the width of the core Gaussian.

Not every mode reconstructed during the reskim is used in every signal channel. A mode is only used if the ratio of the Gaussian peak height to the background level (as measured in generic Monte Carlo) is greater than one.

In Section B.3, we show the invariant mass distributions from signal Monte Carlo for the various D and D^* modes.

B.1 D Candidates

The selection for D^0 candidates is given in Table B.1; the selection for D^+ candidates is given in Table B.2. In both modes, we select candidates based on the invariant mass of the D candidate. Note that we have no signal Monte Carlo for the $D^+ \rightarrow K^+K^-\pi^+$ mode; the mass resolution in this mode is nearly identical to that in $D^+ \rightarrow K^-\pi^+\pi^+$, so we use the values derived from $K\pi\pi$ for $KK\pi$.

B.2 D^* Candidates

The selection for D^{*0} candidates is given in Table B.3; the selection for D^{*+} candidates is given in Table B.4. In both modes, we select candidates based on the invariant mass of the D candidate and the mass difference δm between the D^* and the D candidates.

Table B.1: D^0 candidate selection criteria.

Decay mode	m_D (GeV/ c^2)
$K\pi$	1.839 – 1.891
$K\pi\pi^0$	1.814 – 1.912
$K\pi\pi\pi$	1.849 – 1.881
$K_S^0\pi\pi$	1.841 – 1.889
$K_S^0\pi\pi\pi^0$	not used

Table B.2: D^+ candidate selection criteria.

Decay mode	m_D (GeV/ c^2)
$K\pi\pi$	1.848 – 1.890
$K\pi\pi\pi^0$	not used
$K_S^0\pi$	1.850 – 1.890
$K_S^0\pi\pi^0$	1.831 – 1.907
$K_S^0\pi\pi\pi$	not used
$K^+K^-\pi^+$	1.848 – 1.890

Table B.3: D^{*0} candidate selection criteria.

D^0 decay mode	m_D (GeV/ c^2)	D^{*0} decay mode	δm (MeV/ c^2)
$K\pi$	1.839 – 1.891	$D^0\pi^0$	134 – 148.4
$K\pi\pi^0$	1.814 – 1.912	$D^0\gamma$	125 – 152.4
$K\pi\pi\pi$	1.849 – 1.881		
$K_S^0\pi\pi$	not used		
$K_S^0\pi\pi\pi^0$	not used		

Table B.4: D^{*+} candidate selection criteria.

D decay mode	m_D (GeV/ c^2)	D^{*+} decay mode	δm (MeV/ c^2)
$K\pi$	1.839 – 1.891	$D^0\pi^+$	141.5 – 149.5
$K\pi\pi^0$	1.814 – 1.912		
$K\pi\pi\pi$	1.849 – 1.881		
$K_S^0\pi\pi$	1.841 – 1.889		
$K_S^0\pi\pi\pi^0$	not used		
$K\pi\pi$	1.848 – 1.890	$D^+\pi^0$	136.7 – 144.7
$K\pi\pi\pi^0$	not used		
$K_S^0\pi$	1.850 – 1.890		
$K_S^0\pi\pi^0$	not used		
$K_S^0\pi\pi\pi$	not used		
$K^+K^-\pi$	1.848 – 1.890		

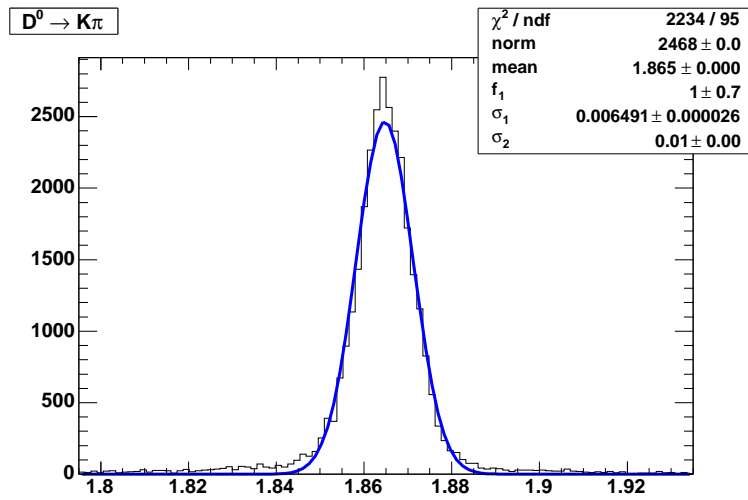


Figure B.1: Invariant mass of $D^0 \rightarrow K^- \pi^+$ candidates from signal Monte Carlo.

B.3 $D^{(*)}$ Mass Distributions

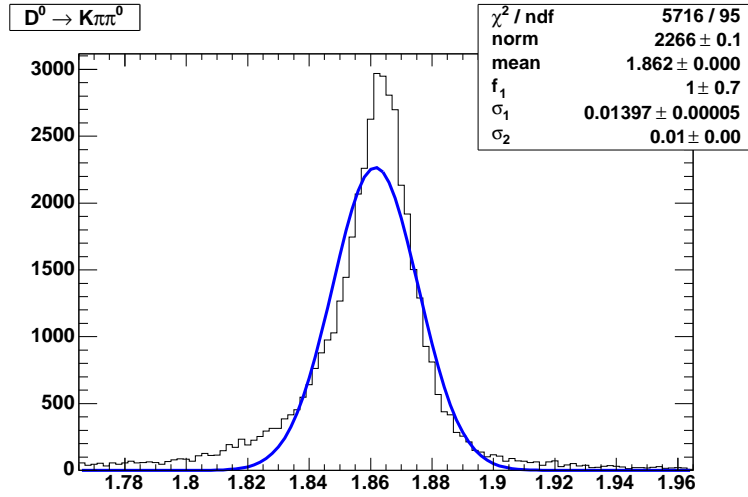


Figure B.2: Invariant mass of $D^0 \rightarrow K^- \pi^+ \pi^0$ candidates from signal Monte Carlo.

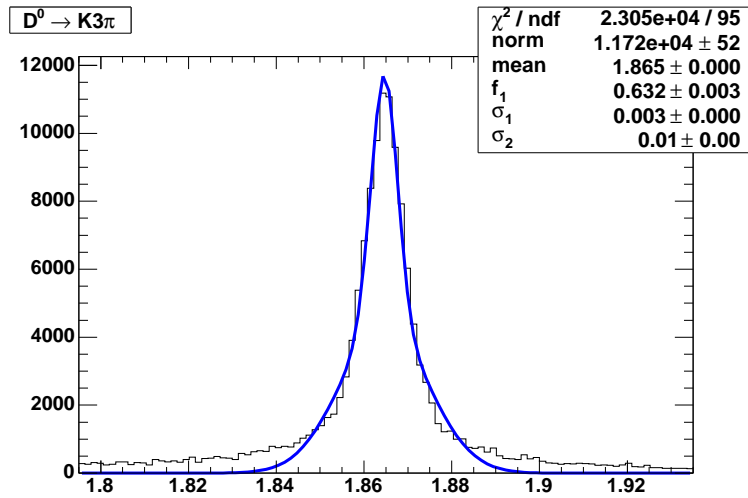


Figure B.3: Invariant mass of $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ candidates from signal Monte Carlo.

Appendix B. $D^{(*)}$ Candidate Selection

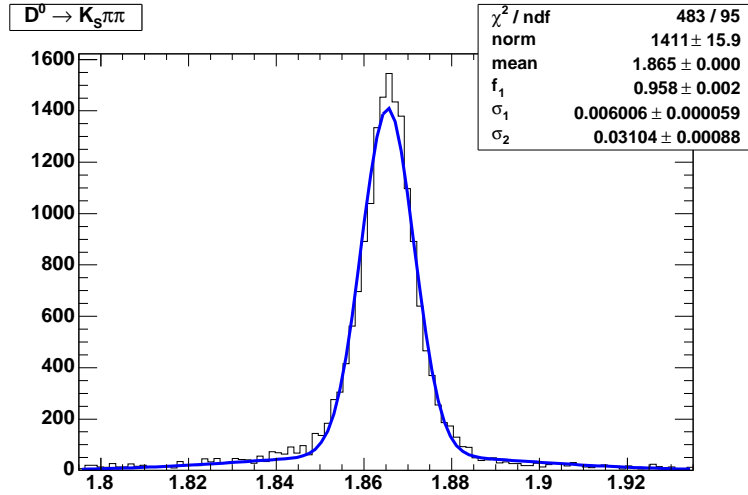


Figure B.4: Invariant mass of $D^0 \rightarrow K_S^0 \pi^+ \pi^-$ candidates from signal Monte Carlo.

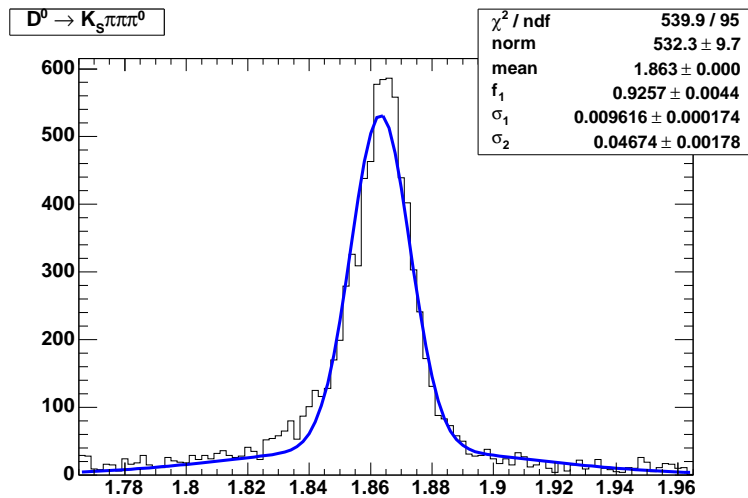


Figure B.5: Invariant mass of $D^0 \rightarrow K_S^0 \pi^+ \pi^- \pi^0$ candidates from signal Monte Carlo.

Appendix B. $D^{(*)}$ Candidate Selection

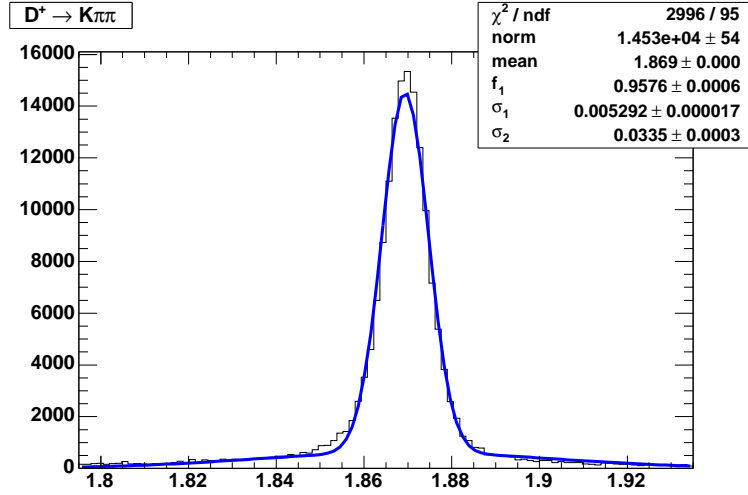


Figure B.6: Invariant mass of $D^+ \rightarrow K^-\pi^+\pi^+$ candidates from signal Monte Carlo.

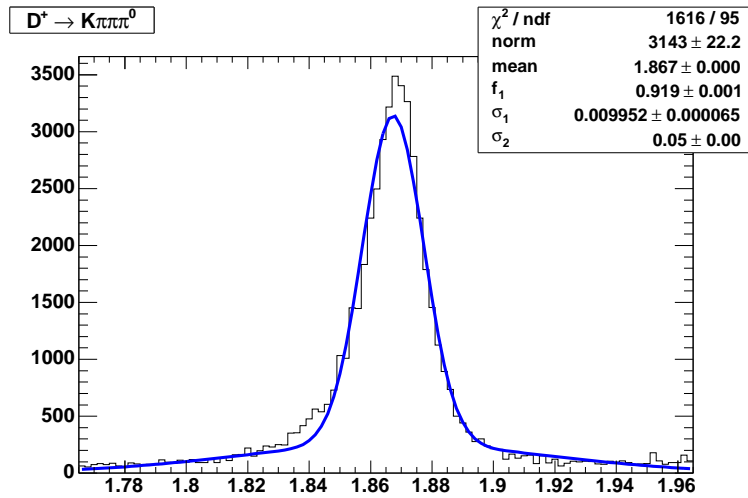


Figure B.7: Invariant mass of $D^+ \rightarrow K^-\pi^+\pi^+\pi^0$ candidates from signal Monte Carlo.

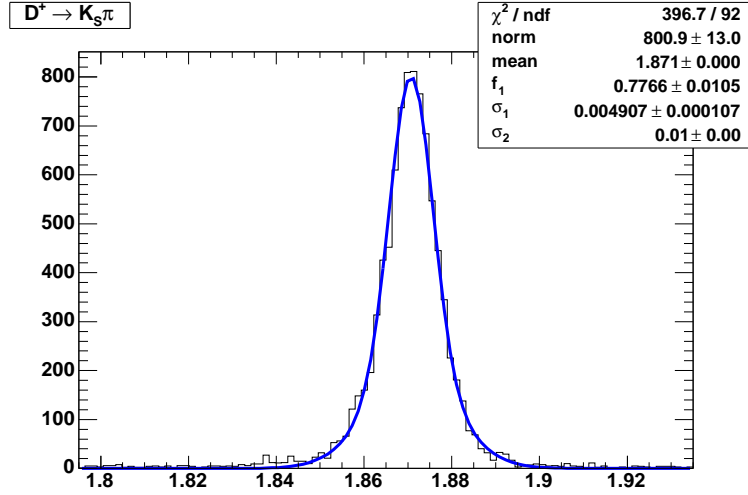


Figure B.8: Invariant mass of $D^+ \rightarrow K_S^0 \pi^+$ candidates from signal Monte Carlo.

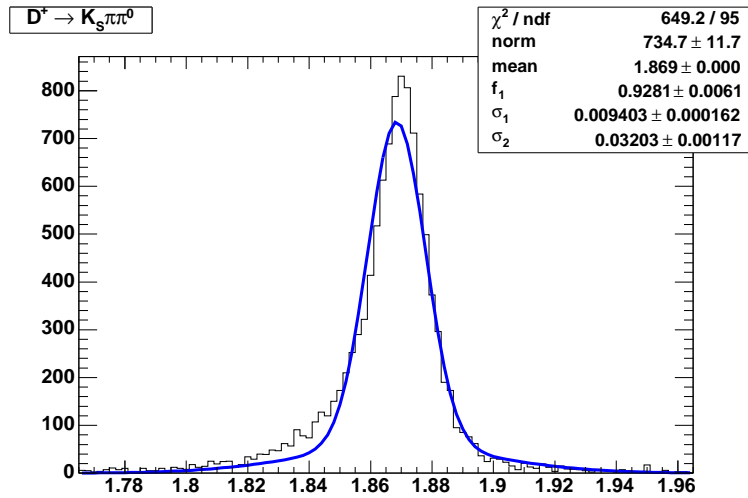


Figure B.9: Invariant mass of $D^+ \rightarrow K_S^0 \pi^+ \pi^0$ candidates from signal Monte Carlo.

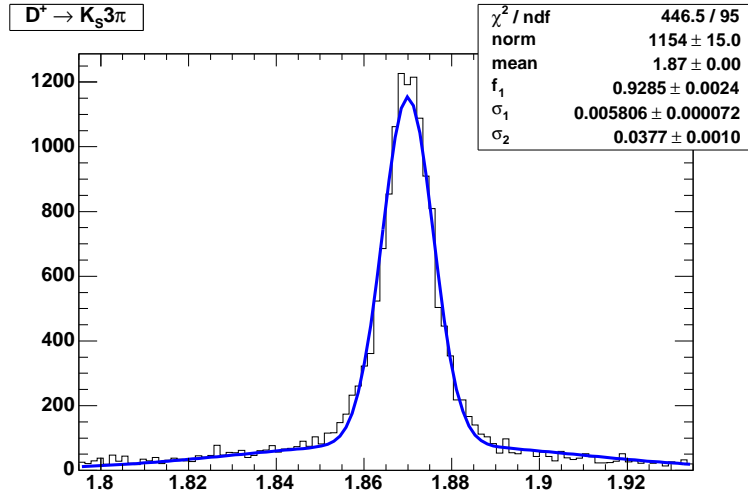


Figure B.10: Invariant mass of $D^+ \rightarrow K_S^0 \pi^+ \pi^- \pi^+$ candidates from signal Monte Carlo.

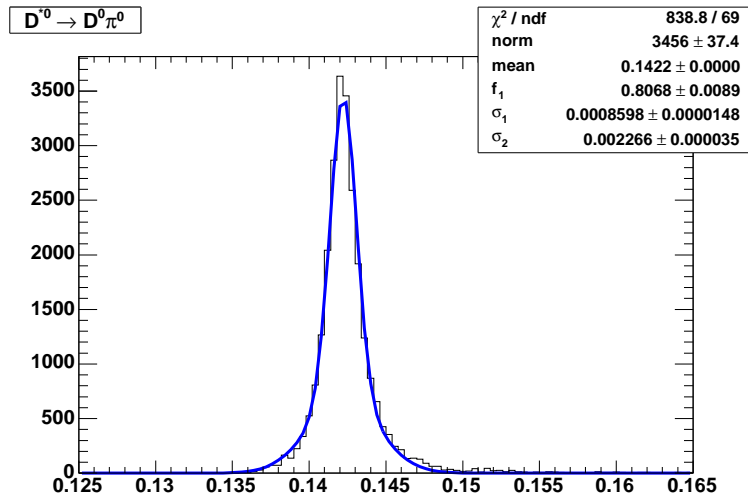


Figure B.11: D^*-D mass difference of $D^{*0} \rightarrow D^0 \pi^0$ candidates from signal Monte Carlo.

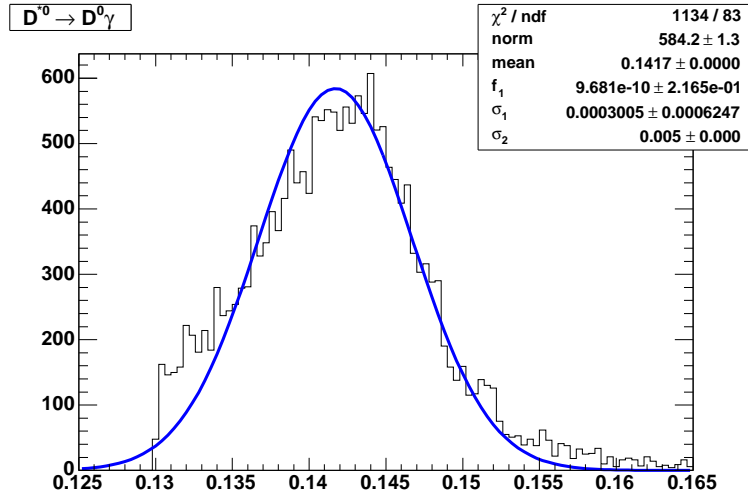


Figure B.12: $D^{*0} \rightarrow D^0 \gamma$ mass difference of $D^{*0} \rightarrow D^0 \gamma$ candidates from signal Monte Carlo.

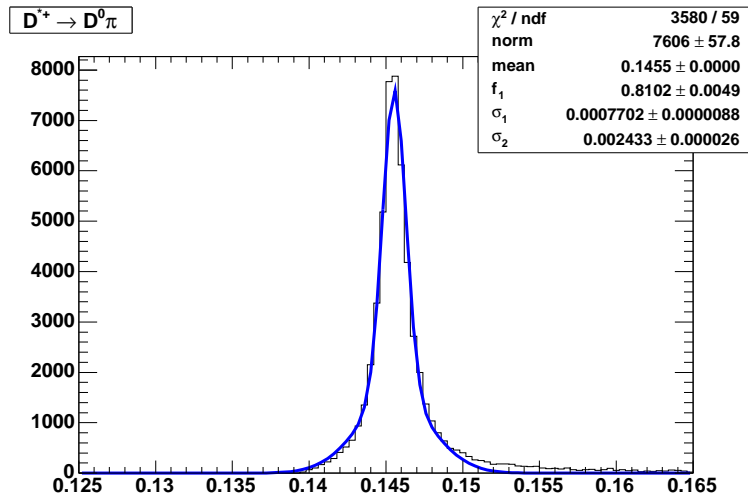


Figure B.13: $D^{*+} \rightarrow D^0 \pi^+$ mass difference of $D^{*+} \rightarrow D^0 \pi^+$ candidates from signal Monte Carlo.

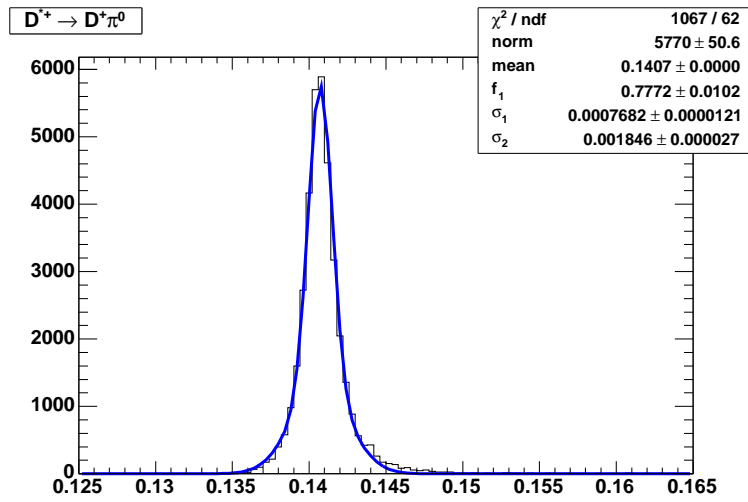


Figure B.14: $D^{*+} - D$ mass difference of $D^{*+} \rightarrow D^+ \pi^0$ candidates from signal Monte Carlo.

Appendix C

Good Track and Neutral Selection

For the definitions of good tracks and neutrals, we closely follow the procedure outlined in [41]. We have changed the sizes of the veto windows for loopers and ghost tracks (see below) to better separate real tracks and fakes, but we have not changed the *spirit* of the original cuts. The only procedural change we have introduced is in the set of tracks used for splitoff rejection in the definition of the good photon list.

C.1 Track Selection

A good track, as it is defined for this analysis, is a candidate from the *ChargedTracks* list which satisfies all of the following criteria:

- $0.41 < \theta_{\text{lab}} < 2.54$

- $p_{\text{lab}} < 10 \text{ GeV}/c$
- $p_{t,\text{lab}} > 0.06 \text{ GeV}/c$
- POCA calculated
- $|\text{DOCA}| < 1.5 \text{ cm}$
- $|z_0| < 5 \text{ cm}$
- $p_{t,\text{lab}} < 0.2 \text{ GeV}/c$ for SVT-only tracks
- track is not a looper (see below)
- track is not a ghost (see below)

C.1.1 Looped Tracks

Charged particles with $p_{t,\text{lab}} \lesssim 360 \text{ MeV}/c$ can loop within the tracking systems of *BABAR*; the tracking algorithms, however, usually reconstruct a looped track as a number of smaller segments, each one describing one half-turn of the helix. Because each turn is each reconstructed as a separate candidate, care must be taken to remove all but one version of the reconstructed track from the list of good tracks.

When multiple turns of a loop are reconstructed, the tracking software produces multiple track candidates with the same (within resolution) p and

Appendix C. Good Track and Neutral Selection

ϕ_{lab} ; half of the turns will be reconstructed as positively-charged tracks with a polar coordinate θ_0 , and half as negatively-charged tracks at $\theta_0 \pm \pi$. Because of this, we will categorize pairs of looper candidates as either “same-charge” or “opposite-charge,” and treat them separately.

We identify looped-track sets by looking for pairs of track candidates satisfying the following criteria:

- $p_{t,\text{lab}}^i < 0.25 \text{ GeV}/c$
- $|\cos \theta_{\text{lab}}^i| < 0.2$
- $|\Delta p_{t,\text{lab}}| \equiv |p_{t,\text{lab}}^2 - p_{t,\text{lab}}^1| < 0.12 \text{ GeV}/c$
- $|\Delta \phi_{\text{lab}}| < 0.1$ or $|\pi - |\Delta \phi_{\text{lab}}|| < 0.1$
- $|\theta_{\text{lab}}^1 - \theta_{\text{lab}}^2| < 0.1$ for same-charge pairs
- $|\pi - \theta_{\text{lab}}^1 - \theta_{\text{lab}}^2| < 0.1$ for opposite-charge pairs.

While $360 \text{ MeV}/c$ represents the theoretical maximum $p_{t,\text{lab}}$ for a track to loop within the radius of the DCH, the first two cuts here represent the practical limits of finding a second loop.

Figures C.1 and C.2 show distributions of $\Delta\theta$ vs. $\Delta\phi$ for looper candidates, in both SP5 MC and Run 1–3 data. Excellent agreement can be seen between

the two samples; SP6 MC and Run 4 data (not pictured) show the same behavior.

When a set of loopers is found, only the single turn with the smallest $|z_0|$, and therefore most likely to have come from the e^+e^- collision, is kept in the good tracks list.

C.1.2 Ghost Tracks

Like loopers, ghost tracks are multiple reconstruction candidates which originate from a single physical particle. In the case of ghost tracks, the tracking algorithms simply split the DCH hits associated with the particle into two separate tracks, each more-or-less following the correct trajectory. Because the two tracks follow the same trajectory, we identify ghost tracks as:

- $p_{t,\text{lab}}^i < 0.35 \text{ GeV}/c$
- $|\Delta p_{t,\text{lab}}| < 0.15 \text{ GeV}/c$
- $|\Delta \phi_{\text{lab}}| < 0.1$
- $|\Delta \theta_{\text{lab}}| < 0.1$
- $N_{\text{DCH}}^1 < 45 - N_{\text{DCH}}^2$ (track 1 is taken to have the greater number of hits)

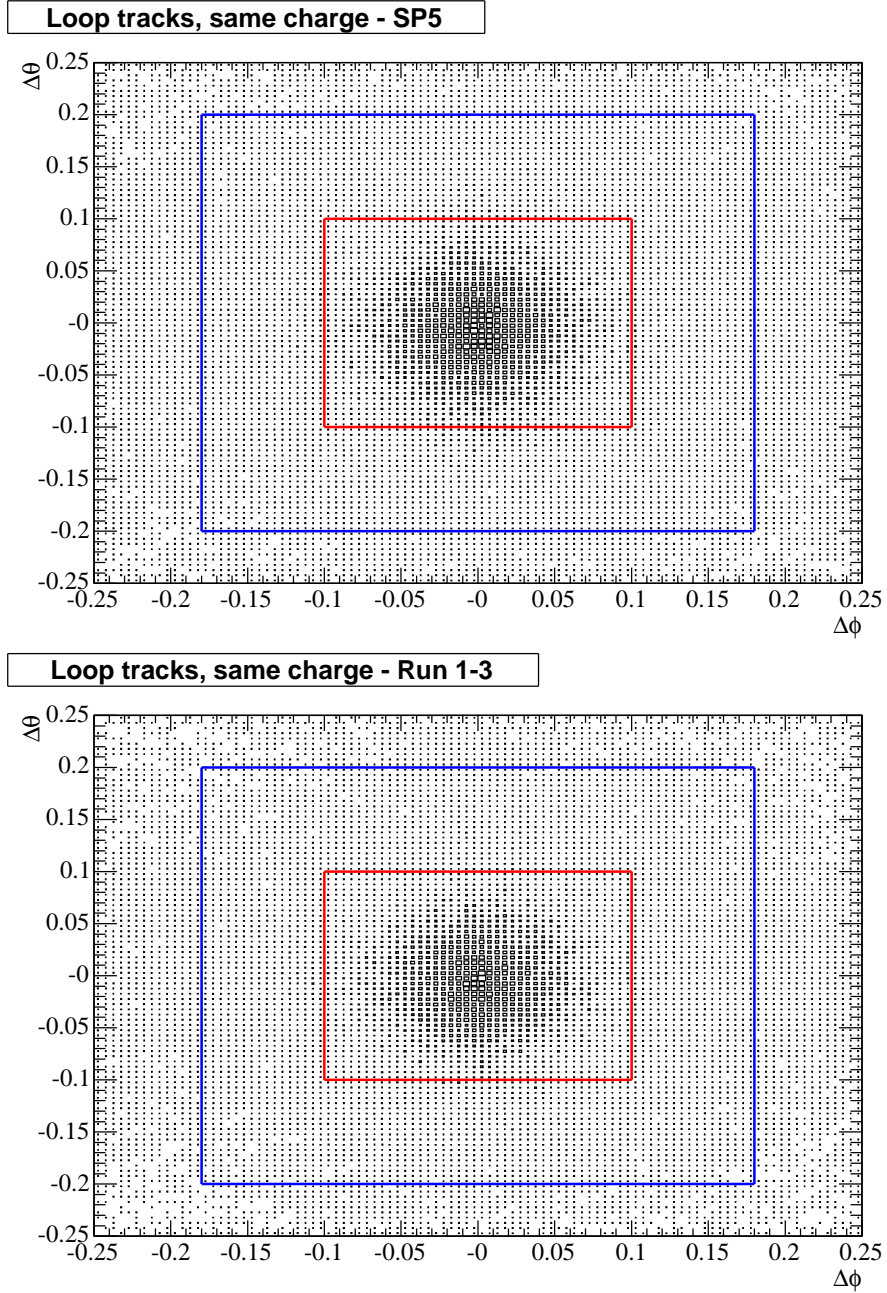


Figure C.1: $\Delta\theta$ vs. $\Delta\phi$ for same-charge loopers. The red (blue) box indicates the cut used in this analysis (BAD 633).

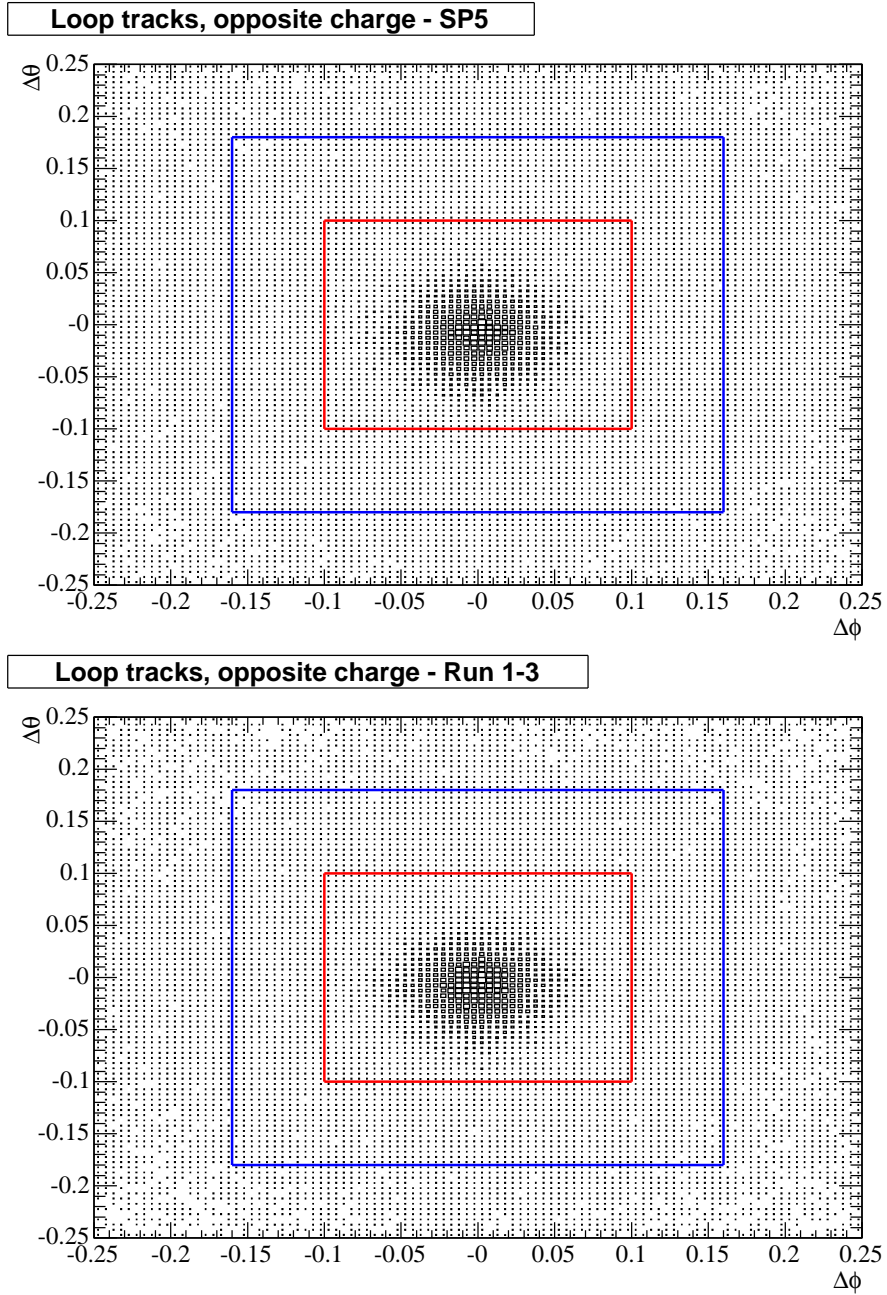


Figure C.2: $\Delta\theta$ vs. $\Delta\phi$ for opposite-charge loopers. The red (blue) box indicates the cut used in this analysis (BAD 633). (N.B. $\Delta\theta$ has been shifted by π to center the excess at zero)

Figure C.3 shows the distribution of $\Delta\theta$ vs. $\Delta\phi$ for ghost candidates, in both SP5 MC and Run 1–3 data. Excellent agreement can be seen between the two samples; SP6 MC and Run 4 data (not pictured) show the same behavior.

When a pair of ghost tracks is found, track 1, the one with more DCH hits is kept in the good tracks list.

C.2 Neutral Selection

A good photon is a candidate from the *CalorNeutral* list which satisfies all of the following:

- $N_{\text{crys}} > 2$
- $E_{\text{clus}} > 50$ MeV
- $\text{LAT} < 0.6$
- $0.32 < \theta_{\text{clus}} < 2.44$
- not identified as an unmatched track (see below)

C.2.1 Unmatched Track Rejection

Frequently, a shower (electromagnetic or hadronic) in the EMC is not correctly associated with the charged track that caused it, which makes the shower

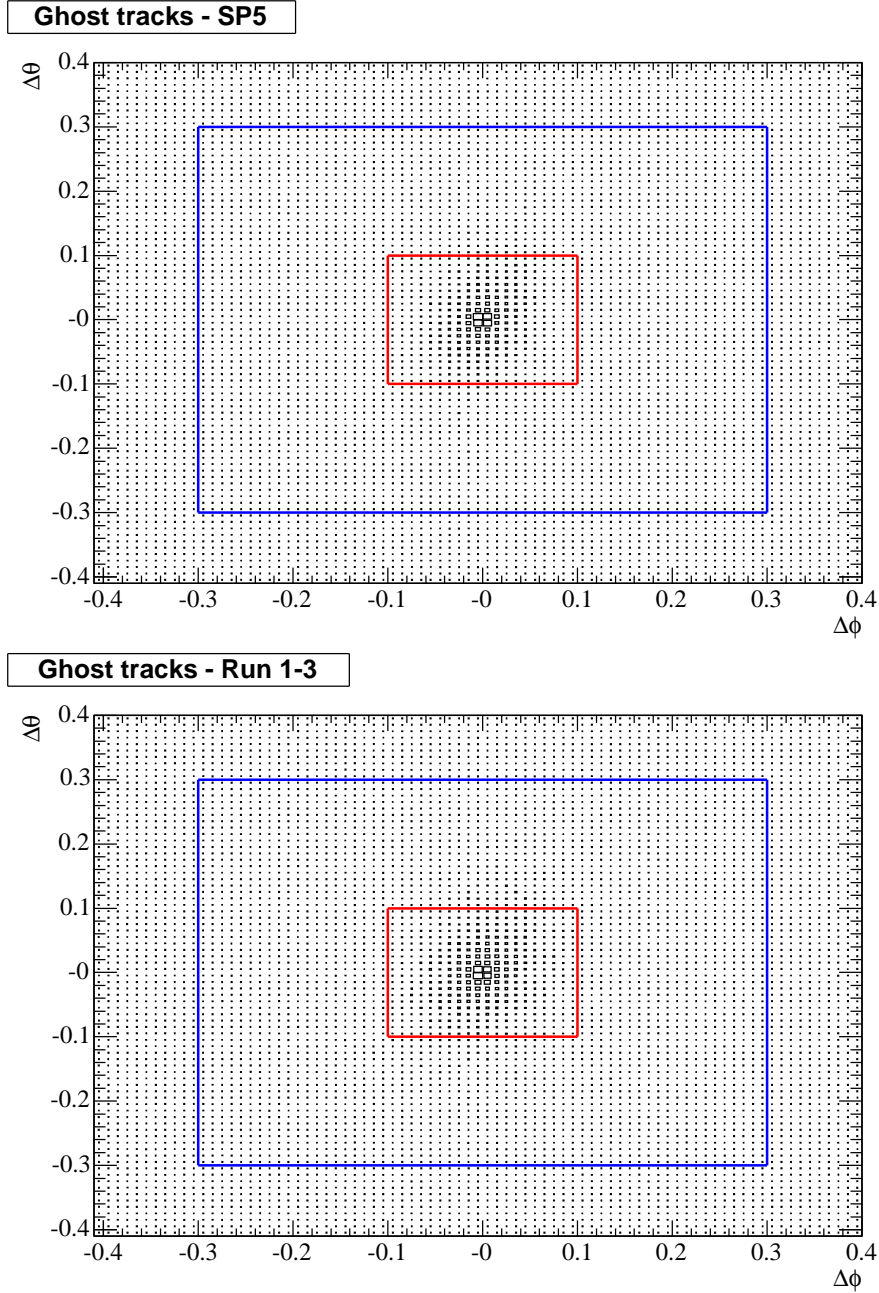


Figure C.3: $\Delta\theta$ vs. $\Delta\phi$ for ghost tracks. The red (blue) box indicates the cut used in this analysis (BAD 633).

appear to be a photon from the interaction region. In order to reject these “photons” from our good photon list, we try to identify these failures of the track-matching algorithm.

For every photon candidate in the event, we identify the nearest charged track (*ChargedTracks*, not restricting ourselves to our refined good tracks list) by calculating the expected intersection between the track and the EMC face.

If the nearest track satisfies:

- $\Delta\alpha \equiv \cos^{-1}[\cos\theta_{\text{clus}}\cos\theta_{\text{trk}} + \sin\theta_{\text{clus}}\sin\theta_{\text{trk}}\cos(\phi_{\text{clus}} - \phi_{\text{trk}})] < 0.8$
- The track is not already matched to an EMC cluster

then the photon candidate is considered to be an unmatched cluster and is not included in the good photon list.

The second requirement above, that the track not already be matched, is the only significant difference between this analysis and that of [41]; the motivation for this can be seen in Figure C.4, which shows a comparison of this algorithm to that in [41]. The default algorithm selects many photons within this veto region simply by virtue of combinatorics — because the algorithm chooses the *closest* track and then cuts on how close it is, this sample is artificially biased towards the veto region (particularly for high-multiplicity events such as we are working with).

Appendix C. Good Track and Neutral Selection

Futher, because we need to discriminate between the various D , D^* , and D^{**} states, we rely on efficient photon identification, and we have found that the default algorithm rejects too many real transition photons.

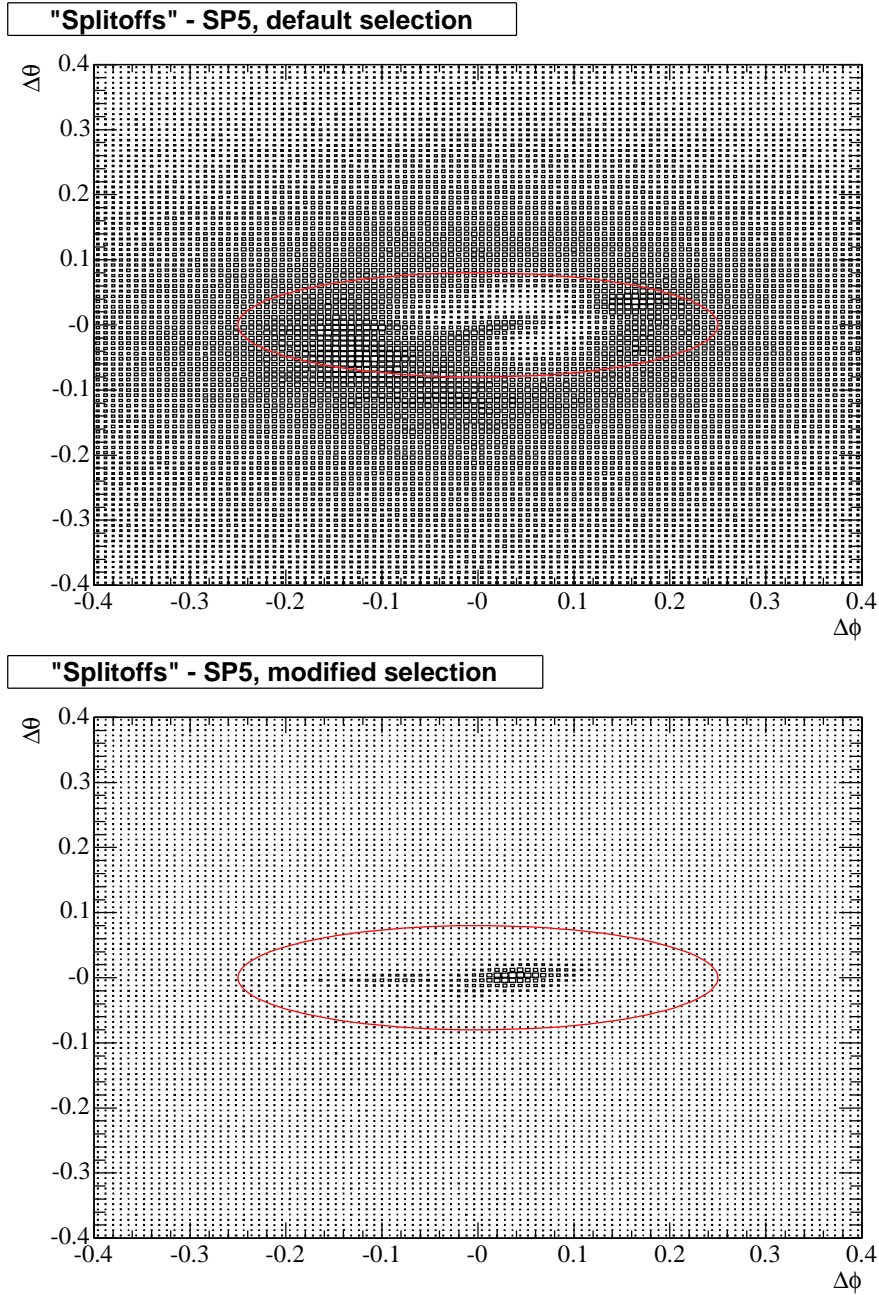


Figure C.4: $\Delta\theta$ vs. $\Delta\phi$ for unmatched tracks. The red ellipse shows the 3-dimensional $\Delta\alpha$ cut projected onto the 2-dimensional EMC face.