Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

First observation of the decay B bar 0s -->; D+-s K-+ and measurement of Br(B bar 0s -->; D+-sK-+)/Br(B bar 0s -->; D+s pi-)

Permalink

https://escholarship.org/uc/item/8sd0d5xh

Author

Shapiro, Marjorie

Publication Date

2007-11-15

eScholarship.org Powered by the California Digital Library
University of California

First observation of the decay $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and measurement of the relative branching fraction $\mathcal{B}(\overline{B}_s^0 \to D_s^{\pm} K^{\mp})/\mathcal{B}(\overline{B}_s^0 \to D_s^{+} \pi^{-})$

by

Johannes Muelmenstaedt

B.S. (Massachusetts Institute of Technology) 2002
M.A. (University of California, Berkeley) 2005

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

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Marjorie Shapiro, Chair
Professor Yury Kolomensky
Professor Stephen Derenzo

Fall 2007

The dissertation of Johannes Muelmenstaedt is approved:

Chair Date

Date

Date

University of California, Berkeley
Fall 2007

First observation of the decay $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and measurement of the relative branching fraction $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm} K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+} \pi^{-}\big)$

© 2007

by Johannes Muelmenstaedt

Abstract

First observation of the decay $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and measurement of the relative branching fraction $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm} K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+}\pi^{-}\big)$

by

Johannes Muelmenstaedt

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Marjorie Shapiro, Chair

We present the first observation of the decay $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and measure the relative branching fraction of $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ to $\overline{B}_s^0 \to D_s^{+}\pi^{-}$. The measurement of the relative branching fraction is performed by applying a fit in invariant mass and specific ionization to 1.2 fb^{-1} of $D_s(\phi\pi)X$ data collected with the CDF II detector in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron collider. We measure $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm} K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+}\pi^{-}\big) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$. The statistical significance of the $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ signal is 7.9σ. To cross-check our analysis method, we also measure $\mathcal{B}\big(\overline{B}^0 \to D^{+}K^{-}\big)/\mathcal{B}\big(\overline{B}^0 \to D^{+}\pi^{-}\big)$ and $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^{-}\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^{-}\big)$ and verify that our results are in agreement with the world average.

Chair Date

For Sylvi, Matthias and Maria

Contents

1 Introduction **1**

2 *CP* violation, the Kobayashi-Maskawa mechanism and the CKM triangle **3**
2.1 *CP* violation . . . 3
2.2 The Standard Model . . . 3
2.3 *CP* violation in the Standard Model: the Kobayashi-Maskawa mechanism 4
2.4 Experimental tests . . . 7
2.5 Towards a measurement of γ from $\overline{B}^0_s \to D^{\pm}_s K^{\mp}$ decays . . . 9

3 The Tevatron collider and the CDF II detector **16**
3.1 The Tevatron collider . . . 16
3.2 Coordinate system . . . 17
3.3 The vertex detector . . . 17
3.4 The Central Outer Tracker . . . 18
3.5 Charged-particle tracking . . . 18
3.6 Trigger system . . . 20

4 Methodology of the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurement **32**
4.1 Statement of the problem . . . 32
4.2 Discriminating variables . . . 33
4.3 Data sample . . . 33
4.4 Reconstruction of B candidates . . . 34
4.5 Candidate selection . . . 35
4.6 Monte Carlo simulation . . . 36
4.7 Mass and dE/dx fit . . . 36
4.8 Mass Templates . . . 38
4.9 Z Templates . . . 39
4.10 Efficiency corrections . . . 40
4.11 Fitter validation . . . 42
4.12 Systematic uncertainties . . . 42

5 $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ control samples **47**
5.1 Selection of candidates . . . 47
5.2 $\overline{B}^0 \to D^+X$ fit configuration . . . 48
5.3 $\overline{B}^0 \to D^+X$ results . . . 51

5.4 $\overline{B}^0 \to D^+X$ validation . . . 53
5.5 $\overline{B}^0 \to D^+X$ residual discrepancies . . . 53
5.6 $\overline{B}^0 \to D^+X$ Monte Carlo–data agreement . . . 57
5.7 $\overline{B}^0 \to D^+X$ systematic uncertainties . . . 58
5.8 $\overline{B}^0 \to D^+X$ comparison with world average . . . 59
5.9 $\overline{B}^0 \to D^{*+}X$ fit configuration . . . 61
5.10 $\overline{B}^0 \to D^{*+}X$ Monte Carlo–data agreement . . . 61
5.11 $\overline{B}^0 \to D^{*+}X$ results . . . 62
5.12 $\overline{B}^0 \to D^{*+}X$ validation . . . 63
5.13 $\overline{B}^0 \to D^{*+}X$ comparison with world average . . . 63
5.14 Summary . . . 63

6 $\overline{B}^0_s \to D^+_s X$ signal sample **105**
6.1 Selection of candidates . . . 105
6.2 Fit configuration . . . 105
6.3 Results . . . 109
6.4 Validation . . . 113
6.5 Systematic uncertainties . . . 114
6.6 Summary . . . 116

7 Summary and conclusions **128**

A dE/dx corrections **135**
A.1 D^{*+} samples . . . 135
A.2 Hit-level and track-level dE/dx corrections . . . 137
A.3 Sample dependence . . . 137
A.4 Reweighting . . . 138

Bibliography **164**

Index **167**

List of Figures

2.1 Vertex factors for the charged weak current 12
2.2 The rescaled unitarity triangle 12
2.3 Current knowledge of the unitarity triangle parameters from the CKM-Fitter global fit 13
2.4 Current knowledge of the unitarity triangle parameters from the UTFit global fit 14
2.5 Diagrams for $B^+ \to \overline{D}^0 K^+$ (left) and $B^+ \to D^0 K^+$ (right) 15
2.6 $\overline{B}_s^0 \to D_s^\pm K^\mp$ decay diagrams 15

3.1 CDF II sectional view 23
3.2 The Fermilab accelerator complex 24
3.3 3D view of SVX II 25
3.4 End view of L00, SVX II, ISL 26
3.5 End view of L00 and the innermost layers of SVX II 27
3.6 A COT cell in superlayer 2 28
3.7 COT sextant 29
3.8 Cross-sectional z–y view of the tracking system 30
3.9 DAQ system 31
3.10 Level 1 and Level 2 block diagram 31

4.1 Cartoon of the mass distributions of $\overline{B}_s^0 \to D_s^+\pi^-$, the radiative tail of $\overline{B}_s^0 \to D_s^+\pi^-$ and $\overline{B}_s^0 \to D_s^\pm K^\mp$ 43
4.2 Z templates for K and π tracks 44
4.3 Composition of $\overline{B}_s^0/B_s^0 \to D_s^+(\phi\pi^+)X$ Monte Carlo sample 44
4.4 D mass cartoon 45
4.5 Signal-subtraction procedure 46

5.1 Mass distribution for $\overline{B}^0 \to D^+\pi^-$ candidates 64
5.2 Mass distribution for $\overline{B}^0 \to D^{*+}\pi^-$ candidates 65
5.3 Composition of $\overline{B}^0 \to D^+\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$ 65
5.4 Composition of $\overline{B}^0 \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$ 66
5.5 Composition of $B^- \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$ 66
5.6 Composition of $B_s \to D_s(\phi\pi)X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$ 67

5.7 Composition of $\overline{B}^0 \to D^{*+}\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^{*+}\pi^-$ . . . 67
5.8 Composition of $B^- \to D^0\pi^-X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^{*+}\pi^-$ . . . 68
5.9 Composition of $\overline{B}^0 \to D^{*+}X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^{*+}\pi^-$ . . . 68
5.10 $\overline{B}^0 \to D^+X$ mass templates . . . 69
5.11 $\overline{B}^0 \to D^{*+}X$ mass templates . . . 70
5.12 Z templates for K and π tracks reweighted for $\overline{B}^0 \to D^+X$ . . . 71
5.13 Z templates for K and π tracks in reweighted for $\overline{B}^0 \to D^{*+}X$ . . . 71
5.14 Fake-D^+ combinatorial template fit (`xbhd0d`) . . . 72
5.15 Fake-D^+ combinatorial template fit (`xbhd0h`) . . . 73
5.16 Fake-D^+ combinatorial template fit (`xbhd0i`) . . . 74
5.17 Fake-D^+ combinatorial template fit (combined) . . . 75
5.18 Real-D^+ combinatorial template fit (`xbhd0d`) . . . 76
5.19 Real-D^+ combinatorial template fit (`xbhd0h`) . . . 76
5.20 Real-D^+ combinatorial template fit (`xbhd0i`) . . . 77
5.21 Real-D^+ combinatorial template fit (combined) . . . 77
5.22 Mass projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0d`) . . . 78
5.23 Mass projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0h`) . . . 79
5.24 Mass projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0i`) . . . 80
5.25 Mass projection for $\overline{B}^0 \to D^+X$ fit (combined) . . . 81
5.26 Mass projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0d`) . . . 82
5.27 Mass projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0h`) . . . 83
5.28 Mass projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0i`) . . . 84
5.29 Mass projection for $\overline{B}^0 \to D^{*+}X$ fit (combined) . . . 85
5.30 Z projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0d`) . . . 86
5.31 Z projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0h`) . . . 87
5.32 Z projection for $\overline{B}^0 \to D^+X$ fit (`xbhd0i`) . . . 88
5.33 Z projection for $\overline{B}^0 \to D^+X$ fit (combined) . . . 89
5.34 Z projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0d`) . . . 90
5.35 Z projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0h`) . . . 91
5.36 Z projection for $\overline{B}^0 \to D^{*+}X$ fit (`xbhd0i`) . . . 92
5.37 Z projection for $\overline{B}^0 \to D^{*+}X$ fit (combined) . . . 93
5.38 Scale factor scan for $\overline{B}^0 \to D^+X$ . . . 94
5.39 Smearing scan for $\overline{B}^0 \to D^{*+}X$ . . . 95
5.40 Scale factor scan for $\overline{B}^0 \to D^{*+}X$ . . . 96
5.41 Smearing scan for $\overline{B}^0 \to D^{*+}X$ . . . 97
5.42 Likelihood scans for $\overline{B}^0 \to D^+X$ . . . 98
5.43 Pull distributions for $\overline{B}^0 \to D^+X$ . . . 99
5.44 Likelihood scans for $\overline{B}^0 \to D^{*+}X$ . . . 100
5.45 Pull distributions for $\overline{B}^0 \to D^{*+}X$ . . . 101
5.46 Z projection for $\overline{B}^0 \to D^+X$ fit (combined), $D\pi$ mass region . . . 102
5.47 Kaon projection of the $\overline{B}^0 \to D^+X$ data sample . . . 103

5.48 Excess of "kaon" tracks in $\overline{B}^0 \to D^+X$ data over the $\overline{B}^0 \to D^+X$ fit . 103
5.49 Fake-D^+ combinatorial normalization fit 104

6.1 Mass distribution for $\overline{B}^0_s \to D^+_s\pi^-$ candidates 117
6.2 Composition of $\overline{B}^0_s \to D^+_s(\phi\pi^+)\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 117
6.3 Composition of $\overline{B}^0_s/B^0_s \to D^+_s(\phi\pi^+)X$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 118
6.4 Composition of $\overline{B}^0/B^0 \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 118
6.5 Composition of $B^-/B^+ \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 119
6.6 Composition of $\overline{B}^0 \to D^{(*)+}_s(\phi\pi^+(\gamma/\pi^0))X$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 119
6.7 Composition of $\Lambda^0_b \to \Lambda^+_c(pK^-\pi^+)X$ Monte Carlo sample reconstructed as $\overline{B}^0_s \to D^+_s\pi^-$ 120
6.8 $\overline{B}^0_s \to D^+_sX$ mass templates 121
6.9 Z templates for K and π tracks reweighted for $\overline{B}^0_s \to D^+_sX$ 122
6.10 Fake-D^+_s combinatorial template fit 123
6.11 Mass projection for $\overline{B}^0_s \to D^+_sX$ fit 124
6.12 Z projection for $\overline{B}^0_s \to D^+_sX$ fit 125
6.13 Likelihood scans for $\overline{B}^0_s \to D^+_sX$ 126
6.14 Pull distributions for $\overline{B}^0_s \to D^+_sX$ 127

7.1 Mass projection for the $\overline{B}^0 \to D^+X$ fit (summary view) 129
7.2 Z projection for the $\overline{B}^0 \to D^+X$ fit (summary view) 130
7.3 Mass projection for the $\overline{B}^0 \to D^{*+}X$ fit (summary view) 131
7.4 Z projection for the $\overline{B}^0 \to D^{*+}X$ fit (summary view) 132
7.5 Mass projection for the $\overline{B}^0_s \to D^+_sX$ fit (summary view) 133
7.6 Z projection for the $\overline{B}^0_s \to D^+_sX$ fit (summary view) 134

A.1 Δm distribution in the `xbhd0d` prompt D^{*+} sample 144
A.2 Δm distribution in the `xbhd0h` prompt D^{*+} sample 145
A.3 Δm distribution in the `xbhd0i` prompt D^{*+} sample 146
A.4 Z versus run number in the prompt D^* sample for the three datasets `xbhd0d`, `xbhd0h` and `xbhd0i` 147
A.5 Pion Z profiles for the `xbhd0d` prompt D^* sample after sideband subtraction in Δm 148
A.6 Pion Z profiles for the `xbhd0h` prompt D^{*+} sample after sideband subtraction in Δm 149
A.7 Pion Z profiles for the `xbhd0i` prompt D^{*+} sample after sideband subtraction in Δm 150
A.8 Kaon Z profiles for the `xbhd0d` prompt D^{*+} sample after sideband subtraction in Δm 151

A.9 Kaon Z profiles for the `xbhd0h` prompt D^{*+} sample after sideband subtraction in Δm 152
A.10 Kaon Z profiles for the `xbhd0i` prompt D^{*+} sample after sideband subtraction in Δm 153
A.11 Pion Z profiles for the `xbhd0d` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 154
A.12 Pion Z profiles for the `xbhd0h` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 155
A.13 Pion Z profiles for the `xbhd0i` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 156
A.14 Kaon Z profiles for the `xbhd0d` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 157
A.15 Kaon Z profiles for the `xbhd0h` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 158
A.16 Kaon Z profiles for the `xbhd0i` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm 159
A.17 Kinematic variable distributions of tracks 160
A.18 Z templates in `xbhd0d` before and after reweighting 161
A.19 Z templates in `xbhd0d` before and after reweighting 162
A.20 Z templates in `xbhd0d` before and after reweighting 163

List of Tables

2.1 Leptons and their electroweak couplings 4
2.2 Quarks and their electroweak couplings 4

3.1 Properties of the SVX II modules 19
3.2 Properties of the L00 and ISL modules 20
3.3 Track quality cuts 21
3.4 `B_CHARM` trigger scenario cuts 22

5.1 Skimming cuts for $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ candidates 48
5.2 Selection cuts for $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ candidates 49
5.3 Fit components for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$fits 52
5.4 Monte Carlo samples used in the $\overline{B}^0 \to D^+X$ fit 53
5.5 Fit results for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ 54
5.6 Correlation coefficients for $\overline{B}^0 \to D^+X$ (combined 1.2 fb^{-1} dataset) . 55
5.7 Pull distributions for $\overline{B}^0 \to D^+X$ 56
5.8 Systematic uncertainties on $\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big)$ 59
5.9 Comparison with PDG for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ fits 60
5.10 Monte Carlo samples used in the $\overline{B}^0 \to D^{*+}X$ fit 61
5.11 Correlation coefficients for $\overline{B}^0 \to D^{*+}X$ (combined 1.2 fb^{-1} dataset) . 62
5.12 Pull distributions for $\overline{B}^0 \to D^{*+}X$ 63

6.1 Skimming cuts for $\overline{B}^0_s \to D^+_s\pi^-$ candidates 106
6.2 Selection cuts for $\overline{B}^0_s \to D^+_s\pi^-$ candidates 107
6.3 Fit components for $\overline{B}^0_s \to D^+_sX$ fit 110
6.4 Monte Carlo samples used in the $\overline{B}^0_s \to D^+_sX$ fit 111
6.5 Fit results for $\overline{B}^0_s \to D^+_sX$ 111
6.6 Correlation coefficients for $\overline{B}^0_s \to D^+_sX$ 112
6.7 Pull distributions for $\overline{B}^0_s \to D^+_sX$ (5102 events per pseudoexperiment) 113
6.8 Pull distributions for $\overline{B}^0_s \to D^+_sX$ (51020 events per pseudoexperiment) 114
6.9 Systematic uncertainties on $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_sK^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s\pi^-\big)$ 116

A.1 Skimming cuts for $\overline{B}^0 \to D^{*+}(D^0\pi^+)\pi^-$ and prompt $D^{*+} \to D^0\pi^+$ candidates 136
A.2 Analysis cuts for $D^0 \to K^-\pi^+$ candidates and D^0 daughter tracks . . 137
A.3 $\langle Z\rangle$ for pions and kaons and kaon–pion separation before and after reweighting in secance and momentum to the $\overline{B}^0 \to D^{*+}X$ sample . . 140

A.4 $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in N_{trks} and momentum to the $\overline{B}^0 \to D^{*+}X$ sample . . . 141
A.5 $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in secance and momentum to the $\overline{B}^0 \to D^{+}X$ sample . . 142
A.6 $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in N_{trks} and momentum to the $\overline{B}^0 \to D^{+}X$ sample 143

Acknowledgments

Advisors sometimes tell their students the parable of the fox and the hare. One day a hare is on his way the library burrow to submit his dissertation, the story goes, when he is accosted by a fox. "I want to eat you", says the fox. "But you can't! So much depends on my submitting this dissertation", pleads the hare. The fox, intelligent creature that he is, is intrigued that he has encountered so educated a lunch. He asks of the hare: "What is your topic?" Comes the reply: "Oh, it's about how hares eat foxes." The fox is sufficiently amused by this bold choice of topic that instead of eating the hare on the spot, he demands to be shown the lab where the hare is conducting his hare-brained research. The hare consents to this demand, and soon the two enter the hare's research burrow. Deafening shrieks ensue, followed by silence. A short while later, the hare emerges, dissertation under his arm, gnawing on a fox shank, and heads for the library burrow again. The moral of this story, we are told, is that it doesn't matter how smart you are; it certainly doesn't matter what your research topic is; the only thing that does matter is that you have a lion as your advisor.

This document is submitted in support of the lion-advisor hypothesis. Marjorie never failed to lift me out by my ears when I found myself in a hole and decided to keep burrowing. Not once in my graduate career did she let her myriad other commitments interfere with her advisor's duties. She also had an amazingly hands-on attitude towards this analysis, often jumping in to help when we students were swamped. Most importantly, she taught me how to do an analysis. It was a pleasure to work for her.

I could not possibly have reached the verge of graduation without the help of Anne Takizawa, who has the uncanny ability to make the university's bureaucracy work for students instead of against them.

The analysis described in this dissertation is closely connected to the analyses of Marjorie's other CDF students, Hung-Chung Fang and Amanda Deisher. Hung-Chung and I share the analysis method, dE/dx corrections and control samples, which he is using for a CP-violation measurement in $B^- \to D^0 K^-$ decays. Amanda and I also share the same signal sample, which she is using for a measurement of the B_s^0 lifetime. This thesis, like many things in high-energy physics, is the result of a collaborative effort. Many thanks to Hung-Chung, Amanda, Marjorie and Alessandro Cerri for their patience with me and their dedication to making the collaboration work.

On a related subject, thanks to Hung-Chung not only for all the effort he put into our analysis and the diligence with which he picked the bugs out of the procedure,

but also for the years of friendship and camaraderie and for all the fun things I learnt from him.

Many of our CDF collaborators played an important role in this analysis. I would like to thank Petar Maksimović, who by now probably knows more about this analysis than I do. I would also like to thank Stefano Giagu and Aseet Mukherjee for helping us understand the dE/dx.

My thesis committee has provided valuable comments on the analysis and on this dissertation, for which I am grateful. Yury also deserves thanks in another matter: it was during his course on particle physics that I first thought I understood something about the Standard Model. Of course, the more I learn, the less I understand, but I remember the beginning of that journey with fondness.

I am also indebted to the members of the CDF and ATLAS groups at LBL for memorable and educational experiences over the years, in particular to Maurice Garcia-Sciveres for all the fun I had working on ATLAS pixels.

Warm thanks go to my many housemates over the years. A happy home is a precondition for everything else, and Tore, Lisa, Amanda, Hera, Rachel, Erin, Henri, Aaron, Jill, Vicky, Thomas, Lieke, Molly, Lindsay, Inge and Caitlin made sure home was always happy.

Finally, I would like to thank two characters whose massive personalities spill out over whatever category I try to put them in. Joe Virzi and Aathavan, mein brauner Bruder, were sometimes coworkers, sometimes twins separated at birth (by several years), sometimes housemates, sometimes running companions, sometimes fellow cattle rustlers. They say that friends help you move, but real friends help you move bodies; with Joe and Aathi, that's not hard to imagine.

Chapter 1

Introduction

One of the most surprising features of nature lies in minute differences between the behavior of particles and antiparticles. *CP* violation, the violation of the discrete symmetry relating particles and antiparticles, is described in the standard model of particle physics (the "Standard Model") by the Kobayashi-Maskawa (KM) mechanism. The Standard Model makes one experimentally falsifiable prediction about *CP* violation, namely that all flavor-changing processes are due to a unitary matrix (the "CKM matrix") that transforms quark mass eigenstates into flavor eigenstates, and that a *CP*-violating phase in the CKM matrix for three quark generations is responsible for all quark-sector *CP* violation.

In recent years, many of the elements of the CKM matrix have been measured, and most aspects of the self-consistency of the KM mechanism have been tested. It is customary to represent the unitarity relations among the columns of the CKM matrix as triangles in the complex plane, and in particular one such triangle whose side lengths are all of the same order of magnitude. If the CKM picture of flavor physics is correct, the different constraints that experimental observables place on the angles and side-lengths of this "unitarity triangle" will all agree. This is indeed the scenario that has played out experimentally, with the exception of one angle, γ, which continues to resist precise determination.

The reason that the CKM angle γ is difficult to measure is that the *CP*-violating effects are suppressed in the decays of light B mesons (B^+ and B^0). B_s decays, namely $\overline{B}^0_s \to D^\pm_s K^\mp$, offer hope. This hope, however, is tied to significant technical challenges. The only $\overline{B}^0_s \to D^\pm_s K^\mp$ sample large enough to attempt a γ measurement was collected by the CDF II detector at the Fermilab Tevatron. In the present work, we extract the $\overline{B}^0_s \to D^\pm_s K^\mp$ signal from the large backgrounds introduced both by the hadron-collider environment and by decays of B_s mesons and other B hadrons. The result of this work is the first observation of $\overline{B}^0_s \to D^\pm_s K^\mp$ (with a 7.9σ statistical significance) and the first measurement of the relative branching fractions $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_s K^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s \pi^-\big)$. Our result demonstrates that it is possible to extract $\overline{B}^0_s \to D^\pm_s K^\mp$ decays from the large backgrounds; it also shows that the $\overline{B}^0_s \to D^\pm_s K^\mp$ branching fraction, which could have been strongly suppressed, is large enough that a γ measurement at the Tevatron or at the LHC is feasible.

We begin this document with a review of *CP* violation in the Standard Model (Chapter 2), which discusses the present knowledge of the unitarity triangle parameters and how a γ measurement can be performed in the $\overline{B}^0_s \to D^\pm_s K^\mp$ system.

Chapter 3 describes the Fermilab Tevatron collider and the CDF II detector. In Chapter 4 we present the method for measuring the branching fraction of $B \to DK$ decays relative to $B \to D\pi$ decays. We apply the method to control samples in Chapter 5; then we present the first observation of $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and the measurement of $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm} K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+} \pi^{-}\big)$ in Chapter 6.

Chapter 2

CP violation, the Kobayashi-Maskawa mechanism and the CKM triangle

Let us begin by showing where the work presented in this dissertation falls in the landscape of high-energy physics. To do this, we will first review just enough of the Standard Model to show how Standard-Model *CP* violation works. Then we will discuss the present state of experimental tests of the Standard-Model description of *CP* violation, focusing on the least-tested parameter. We will then show how the decay mode to which our work is dedicated can contribute to the determination of this parameter. Finally, we will sketch the experimental program that is required to attain this goal; the remainder of the thesis will then describe in detail that part of the experimental program which we have accomplished.

2.1 *CP* violation

CP is the combined operation of parity inversion (P), which changes $x^\mu = (t, \boldsymbol{x})$ to $Px^\mu = (t, -\boldsymbol{x})$, and charge conjugation (C), which inverts the charge of each particle. It was believed that these operations were symmetries of all fundamental interactions until Lee and Yang [1] pointed out the parity conservation in weak decays had not been verified experimentally. Subsequent experimental probing showed that the weak interaction maximally violates both P [2] and C and exhibits small asymmetries under *CP* [3].

How does one observe *CP* violation experimentally? Generically, a process by which a particle X decays into daughter particles $Y_1, \ldots, Y_N$ has a *CP*-conjugate process in which $\overline{X} = CP\ X$ decays into $\overline{Y}_1, \ldots, \overline{Y}_N$. If the process is *CP*-violating, the rate of $X \to Y_1, \ldots, Y_N$ will differ from the rate of $\overline{X} \to \overline{Y}_1, \ldots, \overline{Y}_N$. A necessary requirement for different rates is that the amplitudes for the process and the *CP*-conjugate process differ in magnitude, not only in phase. We will see in Section 2.3 how such amplitudes arise in the Standard Model.

2.2 The Standard Model

The Standard Model is a renormalizable relativistic quantum field theory that accommodates (and in many cases had predicted) all experimental observations in particle

Lepton	T	T^3	Q	Y
ν_e, ν_μ, ν_τ	$1/2$	$1/2$	0	-1
e_L^-, μ_L^-, τ_L^-	$1/2$	$-1/2$	-1	-1
e_R^-, μ_R^-, τ_R^-	0	0	-1	-2

Table 2.1: Leptons and their electroweak couplings

Quark	T	T^3	Q	Y
u_L, c_L, t_L	$1/2$	$1/2$	$2/3$	$1/3$
d_L, s_L, b_L	$1/2$	$-1/2$	$-1/3$	$1/3$
u_R, c_R, t_R	0	0	$2/3$	$4/3$
d_R, s_R, b_R	0	0	$-1/3$	$-2/3$

Table 2.2: Quarks and their electroweak couplings

physics, with the exception of the neutrino mixing sector. The experimental observations relevant for our work are the following.

Fundamental particles: Matter consists of fermions. Quarks participate in the strong interaction, leptons do not. The quarks and leptons form generations, which are listed in Tables 2.1 and 2.2.

Local gauge invariance: Interactions between the fermions respect certain local gauge symmetries. For the model to be invariant under these local symmetry transformations, it is necessary to introduce vector bosons which mediate the interactions between the fermions. The local gauge symmetries underlying the Standard Model are color $SU(3)$, weak $SU(2)$ and hypercharge $U(1)$.

Chirality of the weak interaction: The left-handed fermions form weak-isospin doublets, while the right-handed fermions form singlets.

2.3 *CP* violation in the Standard Model: the Kobayashi-Maskawa mechanism

Let us now consider the Standard-Model Lagrangian for N generations of quarks. Call the up-type quarks $u_i = \{u, c, t\}$ and the down-type quarks $d_i = \{d, s, b\}$. The left-handed quarks form doublets

$$\begin{pmatrix} u_i \\ d_i \end{pmatrix}_L = \frac{1+\gamma^5}{2} \begin{pmatrix} u_i \\ d_i \end{pmatrix},$$

while the right-handed quarks form singlets $u_{iR} = \frac{1}{2}(1-\gamma^5)u_i$, $d_{iR} = \frac{1}{2}(1-\gamma^5)d_i$. The "kinetic" (in the language of Kobayashi and Maskawa [4]) part of the Lagrangian $\mathcal{L}_{\text{kin}}$ contains the quark kinetic energies and couplings to the gauge fields:

$$\begin{aligned}\mathcal{L}_{\text{kin}} = \left(\overline{u}_i, \overline{d}_i\right)_L \gamma^\mu \left(iD_\mu - \frac{g}{2}\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu - \frac{g'}{2}YB_\mu\right)\begin{pmatrix}u_i\\ d_i\end{pmatrix}_L \\ + \overline{u}_{iR}\gamma^\mu\left(iD_\mu - \frac{g'}{2}YB_\mu\right)u_{iR} + \overline{d}_{iR}\gamma^\mu\left(iD_\mu - \frac{g'}{2}YB_\mu\right)d_{iR}.\end{aligned} \tag{2·1}$$

(Summation over $i = 1, \ldots, N$ is implied.)

Gauge invariance forbids constructing "bare" mass terms for the fermions: the fermion bilinear combinations $\overline{u}_L u_R$, $\overline{u}_R u_L$, $\overline{d}_L d_R$ and $\overline{d}_R d_L$ mix $SU(2)$ doublets and singlets, so that an $SU(2)$ doublet term would be introduced into the Lagrangian. Nothing, however, forbids fermion bilinears that are coupled to another field with the appropriate $SU(2)$ and $U(1)$ quantum numbers. If we introduce a weak-isospin doublet of scalar fields

$$\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \tag{2·2}$$

with hypercharge 1 (called the "Higgs doublet"), then a term like

$$\left(\overline{u}_i, \overline{d}_i\right)_L \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} d_{jR}$$

is an isospin and hypercharge singlet and can be added to the Lagrangian. The same is true for

$$\left(\overline{u}_i, \overline{d}_i\right)_L \begin{pmatrix}-\overline{\phi}^0\\ \phi^-\end{pmatrix} u_{jR}$$

where

$$\tilde{\phi} = \begin{pmatrix}-\overline{\phi}^0\\ \phi^-\end{pmatrix} \tag{2·3}$$

is the charge-conjugate of ϕ. Thus, we can add the following terms to the Lagrangian that couple the Higgs doublet to the quarks:

$$\mathcal{L}_Y = -G_d^{ij}\left(\overline{u}_i, \overline{d}_i\right)_L \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} d_{jR} - G_u^{ij}\left(\overline{u}_i, \overline{d}_i\right)_L \begin{pmatrix}-\overline{\phi}^0\\ \phi^-\end{pmatrix} u_{jR} + \text{hermitian conjugate.} \tag{2·4}$$

In the Standard Model, the Higgs doublet acquires a vacuum expectation value

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \quad \text{and analogously} \quad \tilde{\phi}_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}v\\ 0\end{pmatrix} \tag{2·5}$$

by spontaneous symmetry breaking. The coupling of the quark fields to ϕ_0 leads to

quark mass terms in $(2\cdot 4)$:

$$\mathcal{L}_{\text{mass}} = -G_d^{ij}\left(\overline{u}_i, \overline{d}_i\right)_L \phi_0 d_{jR} - G_u^{ij}\left(\overline{u}_i, \overline{d}_i\right)_L \tilde{\phi}_0 u_{jR} = -m_d^{ij}\overline{d}_{iL}d_{jR} - m_u^{ij}\overline{u}_{iL}u_{jR}, \tag{2·6}$$

with $m_{u,d} = G_{u,d}v/\sqrt{2}$.

To write the Lagrangian in terms of mass eigenstates, the $N \times N$ matrices m_u^{ij} and m_d^{ij} must be diagonalized. Consider the unitary transformations

$$\begin{aligned} u_R &\mapsto R_u u_R & \overline{u}_R &\mapsto \overline{u}_R R_u^\dagger \\ d_R &\mapsto R_d d_R & \overline{d}_R &\mapsto \overline{d}_R R_d^\dagger \\ u_L &\mapsto L_u u_L & \overline{u}_L &\mapsto \overline{u}_L L_u^\dagger \\ d_L &\mapsto L_d d_L & \overline{d}_L &\mapsto \overline{d}_L L_d^\dagger \end{aligned}$$

By an appropriate choice of L_d and R_d, $L_d^\dagger m_d R_d$ can be made diagonal, so that

$$\sum_{i,j} -m_d^{ij}\overline{d}_{iL}d_{jR} \mapsto \sum_i -m_d^i \overline{d}_{iL}d_{iR}$$

and likewise for m_u. $\mathcal{L}_{\text{mass}}$ is thus diagonalized. Because the R_u, R_d, L_u, L_d are unitary, the second line of $(2\cdot 1)$ is still diagonal; in the first line, this is true of all terms except those involving $\tau_1 W_\mu^1$ and $\tau_2 W_\mu^2$, which becomes

$$\begin{aligned} \mathcal{L}_W &= -\frac{g}{\sqrt{2}}\gamma^\mu\Big(\overline{u}_{iL}(L_u^\dagger)_{i\alpha}W_\mu^+(L_d)_{\alpha j}d_{jL} + \overline{d}_{iL}(L_d^\dagger)_{i\alpha}W_\mu^-(L_u)_{\alpha j}u_{jL}\Big) \\ &= -\frac{g}{\sqrt{2}}\gamma^\mu\left(\overline{u}_{iL}W_\mu^+ V_{ij}d_{jL} + \overline{d}_{iL}V_{ij}^\dagger W_\mu^- u_{jL}\right), \end{aligned} \tag{2·7}$$

where

$$W^\pm = \frac{1}{\sqrt{2}}\left(W^1 \mp iW^2\right) \tag{2·8}$$

and

$$V = L_u^\dagger L_d. \tag{2·9}$$

As the product of unitary matrices, V is itself unitary. It transforms the mass eigenstates of the down-type quarks into the eigenstates participating in the charged weak current interaction (it is a conventional choice to transform the down-type quarks and not the up-type quarks). V is therefore called the "weak mixing matrix" or the "CKM matrix" after Cabibbo, who first proposed that weak and mass eigenstates did not coincide [5], and Kobayashi and Maskawa [4], who argued (in essentially the same way as we have just reproduced) that weak mixing occurs naturally in the Standard Model. In the next paragraphs, we will follow their argument to show that in a world with three or more quark generations, the Standard Model accommodates *CP* violation.

The kinetic terms of the Lagrangian $(2\cdot 1)$ are invariant under *CP*. However, the

Yukawa terms $(2 \cdot 4)$ are not: each operator is replaced with its hermitian conjugate, while the coefficients G_d^{ij} and G_u^{ij} are unchanged. This is equivalent to replacing G_d^{ij} with $\left(G_d^{ij}\right)^*$ and G_u^{ij} with $(G_u^{ij})^*$, or V_{ij} with V_{ij}^*. We come to the same conclusion by applying CP to $(2 \cdot 7)$, which replaces every operator with its Hermitian conjugate but leaves V_{ij} and $V_{ij}^\dagger$ unchanged, or equivalently replaces V_{ij} with V_{ij}^*. The Feynman diagrams resulting from $\mathcal{L}_W$, shown in Figure 2.1, confirm this derivation.

Thus, CP violation will occur due to the charged-current interactions if V contains complex elements. In two generations, the "Cabibbo matrix"

$$V = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \tag{2 \cdot 10}$$

has only one physical parameter, the mixing angle θ_C between first- and second-generation quarks. The two-generation matrix is therefore completely real, and no CP violation exists. In three generations, the CKM matrix has three CP-conserving mixing angles and one CP-violating phase. The original parameterization by Kobayashi and Maskawa [4] uses three angles θ_{12}, θ_{13}, θ_{23} and the phase δ:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2 \cdot 11}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$.

Empirically, the CKM matrix exhibits a hierarchical structure in which the diagonal elements are large and the off-diagonal elements become progressively smaller. The expansion proposed by Wolfenstein [6], valid to $\mathcal{O}(\lambda^4)$, makes this hierarchy manifest:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2 \cdot 12}$$

The expansion parameter $\lambda = \sin\theta_C \approx 0.22$ is the Cabibbo suppression factor inherited from the 2×2 Cabibbo matrix.

2.4 Experimental tests

By construction, the CKM matrix is unitary. Out of this property (the only property of flavor mixing predicted by the Standard Model) arise the experimental tests of the KM mechanism. The orthonormality relations amongst the columns and rows of a unitary matrix,

$$\sum V_{ij}V_{kj}^* = \delta_{ik} \tag{2 \cdot 13}$$

and

$$\sum V_{ij}V_{ik}^* = \delta_{jk}, \tag{2 \cdot 14}$$

imply (for example)

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2·15}$$

Relationships of the form (2·15) can be represented graphically as triangles (for a 3×3 matrix) in the complex plane. Of the three such row-wise combinations, the one given in (2·15) lends itself particularly well to this graphical representation, since all three terms are of comparable size $\mathcal{O}(\lambda^3)$.

Various experimental observables are sensitive to the sides and angles of the unitarity triangle (2·15). If the CKM picture of weak mixing is correct, measurements of the same side or angle in different weak decays must give consistent answers. Furthermore, the measured angles and sides must form a closed triangle. To test the latter prediction, it is convenient to rescale all sides of the triangle by $(V_{cd}V_{cb}^*)^{-1}$, as shown in Figure 2.2, so that measurements of the sides and angles can be reported as constraints on the apex of the triangle. The position of the apex in the complex plane is parameterized as

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{2·16}$$

which is related to the Wolfenstein parameters of (2·12) by [7]

$$\rho + i\eta = \frac{\sqrt{1 - A^2\lambda^4}(\overline{\rho} + i\overline{\eta})}{\sqrt{1-\lambda^2}\left[1 - A^2\lambda^4(\overline{\rho} + i\overline{\eta})\right]}. \tag{2·17}$$

The three angles of the unitarity triangle are

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{2·18}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{2·19}$$

and

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{2·20}$$

All measurements sensitive to $\overline{\rho}$ and $\overline{\eta}$ are collected and compiled into a global fit for $\overline{\rho}$ and $\overline{\eta}$ regularly by two collaborations, CKMFitter [7] and UTFit [8]. The current constraints on the unitarity triangle are shown in Figures 2.3 [7] and 2.4 [9]. From these figures, it is evident that γ is determined with worse precision than all other sides and angles: BaBar measures $\gamma = (92 \pm 41(\text{stat}) \pm 11(\text{syst}) \pm 12(\text{Dalitz model}))^\circ$ [10], Belle measures $\gamma = \left(53^{+15}_{-18}(\text{stat}) \pm 3(\text{syst}) \pm 9(\text{Dalitz model})\right)^\circ$ [11]. (These analyses use the Dalitz structure of multi-body D^0 decays; hence the quoted uncertainty due to the Dalitz model.) What can be done about this? We will address that question shortly. First, we will review what limits the current γ measurements.

B^+ decays provide the most accessible diagrams for measuring γ: the relative weak

phase between the two decay diagrams shown in Figure 2.5 is γ. If these diagrams can be made to interfere, then a CP-violating asymmetry will result whose size is related to γ. At first glance, it seems that $B^+ \to \overline{D}^0 K^+$ and $B^+ \to D^0 K^+$ are distinct final states and hence do not interfere; the trick is to let the D meson decay to a final state that is accessible to both $\overline{D}^0$ and D^0. Several approaches have been proposed. CP-eigenstate decay modes such as K^+K^-, $\pi^+\pi^-$ or $K_s^0\phi$ are accessible to $\overline{D}^0$ and D^0 because the CP eigenstates $D^0_{CP\pm}$ are superpositions of the flavor eigenstates $\overline{D}^0$ and D^0; using CP modes was proposed by Gronau, London and Wyler and is known as the GLW method [12, 13, 14]. A second method, proposed by Atwood, Dunietz and Soni [15], considers a final state like $K^-\pi^+$ that is Cabibbo-allowed for the D^0 but doubly Cabibbo-suppressed for the $\overline{D}^0$. Finally, a method proposed by Giri, Grossman, Soffer and Zupan, commonly known as the "Dalitz-plot method" [16], uses multi-body decays like $D \to K_s^0\pi^-\pi^+$ that are Cabibbo-allowed for both flavor eigenstates.

All three methods share an experimental difficulty. $B^+ \to \overline{D}^0 K^+$ is color-allowed, while the internal W emission in $B^+ \to D^0 K^+$ is color-suppressed. Since the size of the interference effect is determined by the ratio between the interfering amplitudes, the expected asymmetry in the $B^+ \to D^0 K$ system is small. This is the reason that the uncertainty on γ is large.

The expectation that far higher statistics will need to be accumulated in the $B^+ \to D^0 K$ system has led to searches for other decays that are sensitive to γ. One such decay is $B^0 \to D^\mp\pi^\pm$, for which experimental results are already available; but here, too, the interference effect suffers from a large ratio between the interfering diagrams. A very promising system, promising because the amplitude ratio could be close to unity, will be described in the next section.

2.5 Towards a measurement of γ from $\overline{B}^0_s \to D_s^\pm K^\mp$ decays

Consider the two diagrams in Figure 2.6. The relative weak phase between these diagrams is once again γ. What sets these diagrams apart from the B^+ and B^0 decays discussed in the previous section is that there is neither a color-suppression factor (both W emissions are external) nor a large ratio of CKM factors (both diagrams are proportional to λ^3). This is not a guarantee that the amplitude ratio is indeed close to unity, as the long-range physics of hadronization could still destroy the balance; it is difficult to predict these non-perturbative QCD effects, so one can only know after having attempted the measurement whether the long-range physics cooperates. But $\overline{B}^0_s \to D_s^\pm K^\mp$ offers hope of a γ measurement.

How can $B_s^0 \to D_s^+ K^-$ and $B_s^0 \to D_s^- K^+$ be harnessed to measure γ? We will describe two promising methods. The first has smaller theoretical uncertainties; the second is experimentally less challenging. The "tagged method" [17] uses interference from B_s mixing: a B_s^0 can reach the final state $D_s^+K^-$ through the decay $B_s^0 \to D_s^+K^-$, or it can first undergo mixing to become a $\overline{B}_s^0$ and then decay through the $\overline{B}_s^0 \to D_s^+K^-$ channel. The interfering diagrams lead to an observable effect related to

γ through the time-dependent evolution of the decay rates $\Gamma(B_s^0 \to D_s^- K^+)$, $\Gamma(B_s^0 \to D_s^+ K^-)$, $\Gamma(\overline{B}_s^0 \to D_s^- K^+)$ and $\Gamma(\overline{B}_s^0 \to D_s^+ K^-)$:

$$\begin{aligned}
\Gamma(B_s^0 \to D_s^- K^+) &= \frac{|A_f|^2 e^{-\Gamma_s t}}{2}\Big\{(1+|\lambda_f|^2)\cosh(\Delta\Gamma_s t/2) + (1-|\lambda_f|^2)\cos(\Delta m_s t) \\
&\quad - 2|\lambda_f|\cos(\delta+\gamma)\sinh(\Delta\Gamma_s t/2) - 2|\lambda_f|\sin(\delta+\gamma)\sin(\Delta m_s t)\Big\}, \\
\Gamma(B_s^0 \to D_s^+ K^-) &= |A_f|^2 e^{-\Gamma_s t}/2\Big\{(1+|\lambda_f|^2)\cosh(\Delta\Gamma_s t/2) - (1-|\lambda_f|^2)\cos(\Delta m_s t) \\
&\quad - 2|\lambda_f|\cos(\delta-\gamma)\sinh(\Delta\Gamma_s t/2) + 2|\lambda_f|\sin(\delta-\gamma)\sin(\Delta m_s t)\Big\}, \\
\Gamma(\overline{B}_s^0 \to D_s^- K^+) &= |A_f|^2 e^{-\Gamma_s t}/2\Big\{(1+|\lambda_f|^2)\cosh(\Delta\Gamma_s t/2) - (1-|\lambda_f|^2)\cos(\Delta m_s t) \\
&\quad - 2|\lambda_f|\cos(\delta+\gamma)\sinh(\Delta\Gamma_s t/2) + 2|\lambda_f|\sin(\delta+\gamma)\sin(\Delta m_s t)\Big\}, \\
\Gamma(\overline{B}_s^0 \to D_s^+ K^-) &= |A_f|^2 e^{-\Gamma_s t}/2\Big\{(1+|\lambda_f|^2)\cosh(\Delta\Gamma_s t/2) + (1-|\lambda_f|^2)\cos(\Delta m_s t) \\
&\quad - 2|\lambda_f|\cos(\delta-\gamma)\sinh(\Delta\Gamma_s t/2) - 2|\lambda_f|\sin(\delta-\gamma)\sin(\Delta m_s t)\Big\}.
\end{aligned} \tag{2·21}$$

These rates depend on the mass and lifetime differences Δm_s and $\Delta\Gamma_s$ between the heavy and light *CP* eigenstates of the B_s,

$$\begin{aligned}
|B_s^L\rangle &= p|B_s^0\rangle + q|\overline{B}_s^0\rangle, \\
|B_s^H\rangle &= p|B_s^0\rangle - q|\overline{B}_s^0\rangle,
\end{aligned} \tag{2·22}$$

on the amplitude $A_f = \langle D_s^- K^+|B_s^0\rangle$ (and its conjugate $\overline{A}_f = \langle D_s^- K^+|\overline{B}_s^0\rangle$), on

$$\lambda_f = \frac{q}{p}\frac{A_f}{\overline{A}_f} \tag{2·23}$$

and on the relative weak and strong phases γ and δ between $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$. This method thus relies on resolving decay times of order $(\Delta m_s)^{-1}$ and on determining the flavor (B_s^0 or $\overline{B}_s^0$) of the B_s meson at production. At CDF, flavor tagging suppresses the sensitivity of the method by a factor of 20. The return on the experimental difficulty of flavor tagging, however, is immense: there are no theoretical inputs in (2·21) that do not either cancel in ratios (in the case of A_f) or can be determined from the data (in the case of δ). A γ measurement using this method, should it be feasible experimentally, will be very clean theoretically.

The "untagged method" [18] does not need to determine the flavor of the B_s meson at production. Rather, it uses the *CP* eigenstates of the B_s. Since these states are superpositions of the B_s^0 and the $\overline{B}_s^0$, decays of the *CP* eigenstates proceed through both diagrams. By tracking the decay rate over a sufficiently long time, the effect of the short-lived and long-lived *CP* eigenstates can be disentangled. A time-dependent measurement of the decay rate of $B_s^0 \to D_s^- K^+$ relative to $B_s^0 \to D_s^- \pi^+$ (using a

normalization mode has the advantage that many systematic uncertainties cancel in the ratio) would depend on γ as follows:

$$\frac{\Gamma(B_s^0 \to D_s^- K^+)}{\Gamma(B_s^0 \to D_s^- \pi^+)} = \frac{A_{D_s^- K^+}}{A_{D_s^- \pi^+}} \left(1 + |\lambda_f|^2\right) \left\{1 + 2\cos(\delta - \gamma) \tanh\left(\frac{\Delta\Gamma\, t}{2}\right)\right\}, \quad (2 \cdot 24)$$

There is neither a need for flavor tagging nor for disentangling fast oscillations. However, the theoretical inputs are not nearly as clean as in the tagged method. Furthermore, the sensitivity of the method is suppressed by the smallness of $\Delta\Gamma$; $\Delta\Gamma$ is expected to be 0.1–0.2 in the Standard Model, which gives a suppression in the untagged method comparable to the suppression due to flavor tagging in the tagged method.

Both methods depend on a large sample of $\overline{B}_s^0 \to D_s^\pm K^\mp$ decays. (By $\overline{B}_s^0 \to D_s^\pm K^\mp$, we indicate that we do not determine the decay-time flavor of the B_s.) Such a sample is not easy to come by. The "B factories" PEP-II and KEKB are mostly run at the Υ(4S) pole, where the only B hadrons produced are B^+ and B^0; the Υ(5S) runs at CESR and KEKB, where $B_s^0\overline{B}_s^0$ pairs are produced, have not yielded a significant-sized sample of $\overline{B}_s^0 \to D_s^+ X$ decays. Thus, until Run II of the Tevatron collider, there was no facility at which $\overline{B}_s^0 \to D_s^\pm K^\mp$ physics could be done. Currently, the Tevatron collider is the world's only B_s^0 factory. Both Tevatron detectors, CDF II and DØ, are accumulating B_s^0 datasets, with a larger sample of hadronic B_s^0 decays available at CDF. The CDF sample used in the present work contains about 1000 $\overline{B}_s^0 \to D_s^+\pi^-$ events after selection cuts in the D_s^+ decay mode with the highest signal-to-background ratio, $D_s^+ \to \phi\pi^+$.

Estimating a ratio of $\mathcal{O}(0.1)$ for the $\overline{B}_s^0 \to D_s^\pm K^\mp$ yield relative to $\overline{B}_s^0 \to D_s^+\pi^-$, we can predict that there are enough $\overline{B}_s^0 \to D_s^\pm K^\mp$ events available at CDF to make the first observation of $\overline{B}_s^0 \to D_s^\pm K^\mp$ and to measure the $\overline{B}_s^0 \to D_s^\pm K^\mp$ branching fraction relative to the $\overline{B}_s^0 \to D_s^+\pi^-$ branching fraction. As we will describe in greater detail in Chapter 4, isolating $\overline{B}_s^0 \to D_s^\pm K^\mp$ decays is a challenging task, both because of the large non-B backgrounds found in the hadron collider environment and because of the intrinsic background from other B decay modes. Our work serves a twofold purpose. First, it is a proof of principle for isolating $\overline{B}_s^0 \to D_s^\pm K^\mp$ decays in a hadron collider environment, providing a stepping stone for further work towards measuring γ. Second, the measurement of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ branching fraction allows an accurate prediction of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ yield available to future efforts in the $\overline{B}_s^0 \to D_s^\pm K^\mp$ sector.

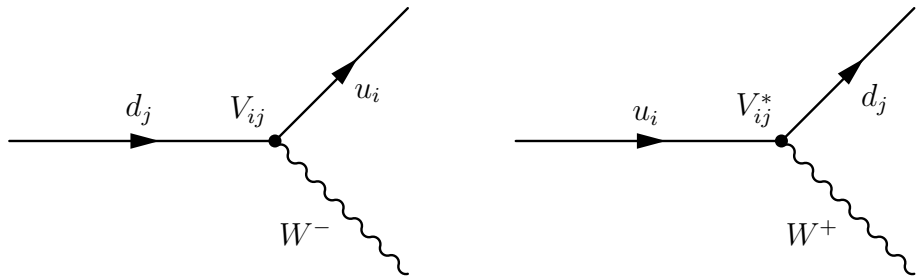

Figure 2.1: Vertex factors for the charged weak current

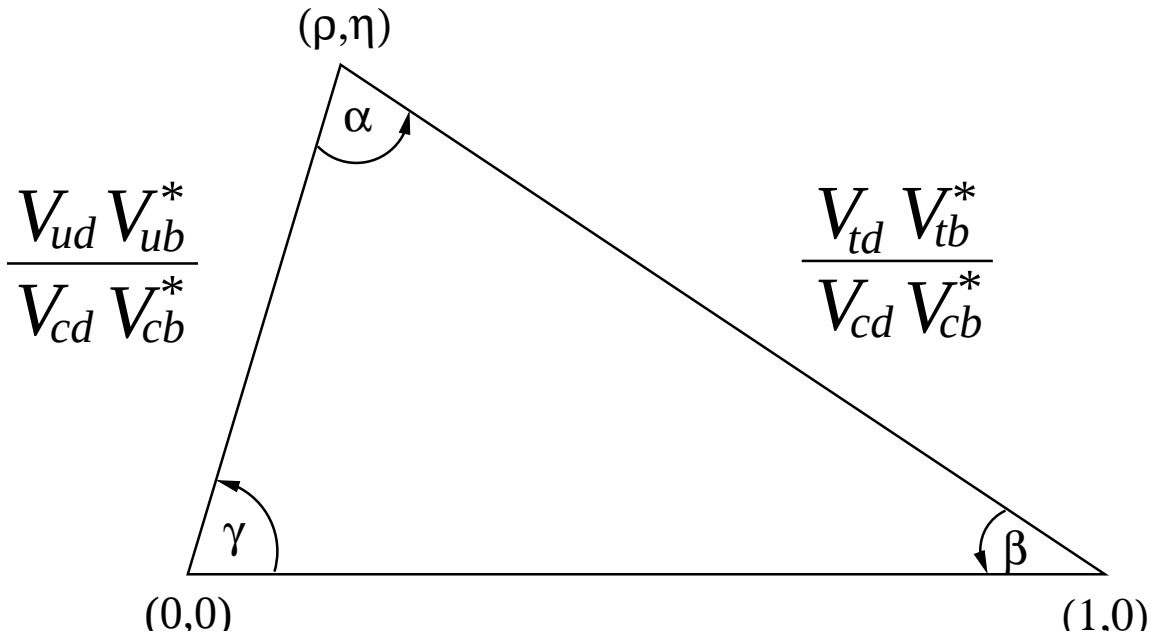

Figure 2.2: The rescaled unitarity triangle

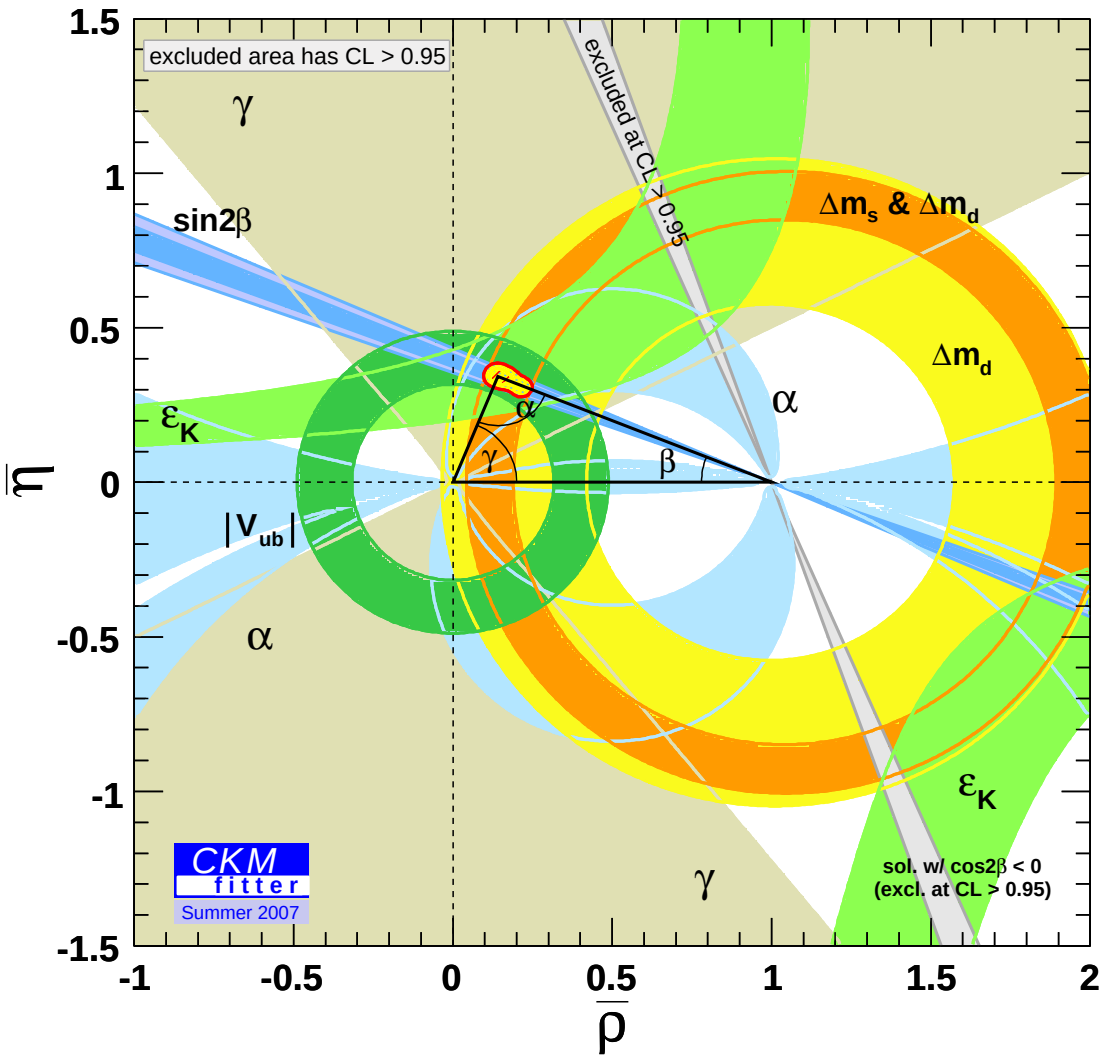

Figure 2.3: Current knowledge of the unitarity triangle parameters from the CKM-Fitter global fit

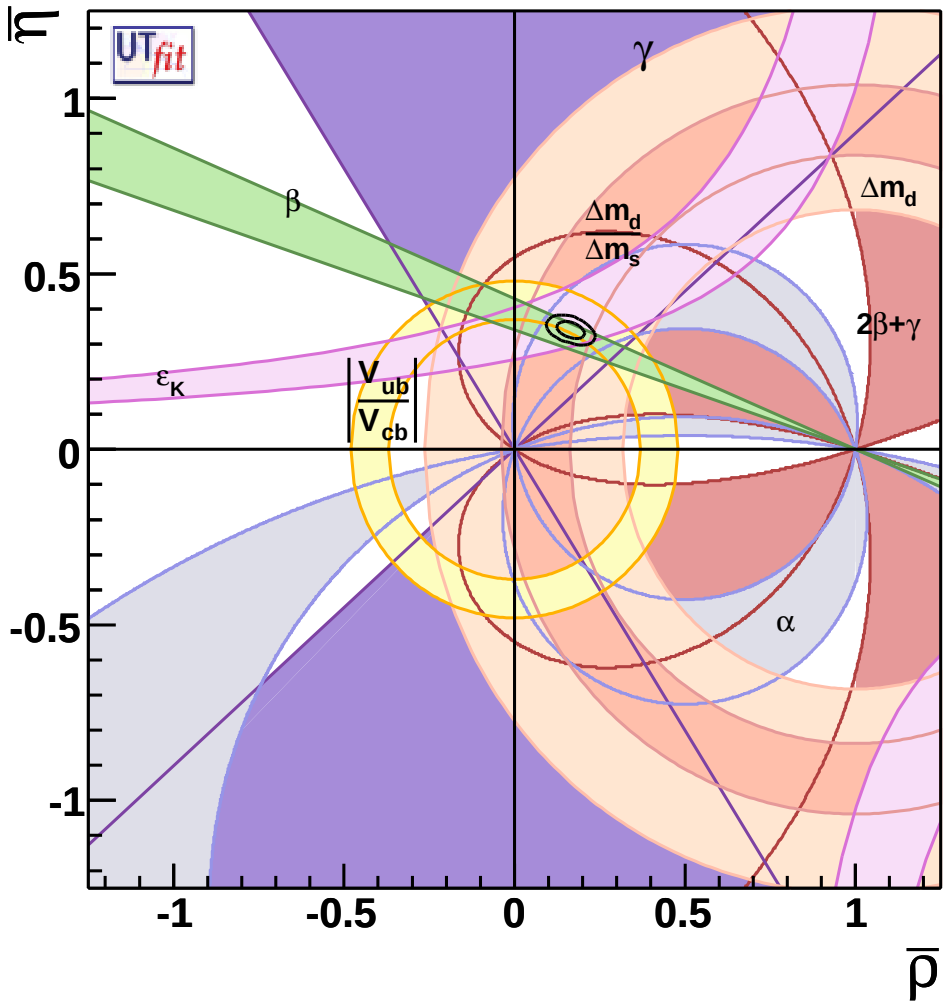

Figure 2.4: Current knowledge of the unitarity triangle parameters from the UTFit global fit

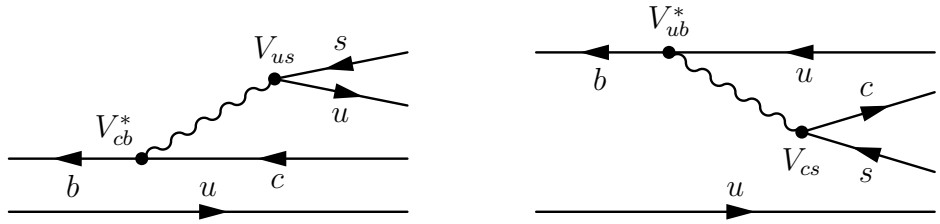

Figure 2.5: Diagrams for $B^+ \to \overline{D}^0 K^+$ (left) and $B^+ \to D^0 K^+$ (right). The relative weak phase between the diagrams is γ.

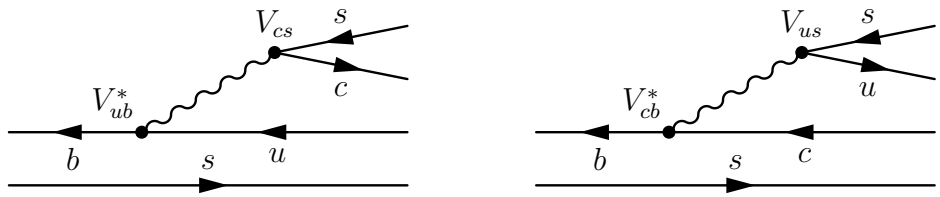

Figure 2.6: $\overline{B}_s^0 \to D_s^\pm K^\mp$ decay diagrams. The relative weak phase between the diagrams is γ, and the amplitude ratio is $\mathcal{O}(1)$.

Chapter 3

The Tevatron collider and the CDF II detector

One of the prerequisites for our work is a large sample of B_s^0 decays. The sample we use is provided by the Tevatron collider and recorded using the CDF II detector. CDF II, as the sectional view in Figure 3.1 shows, is a general-purpose high-energy physics detector; only some of its subsystems (vertex detector, tracking chamber, trigger system) are relevant for our analysis. Those subsystems are described in this chapter.

3.1 The Tevatron collider

When protons and antiprotons collide at the interaction points of the Tevatron, they are experiencing the climax of a journey through a chain of accelerators, shown in Figure 3.2. This journey begins in the Cockroft-Walton pre-accelerator, which ionizes H_2 gas molecules to form H^- ions; these ions are accelerated to 750 keV. The ions are further accelerated to 400 MeV in the Linac. They then pass through a carbon foil, which strips them of their electrons. The bare protons now enter the Booster synchrotron, which accelerates them to 8 GeV. Thence they are sent to the Main Injector.

The Main Injector is the workhorse of the Fermilab experimental particle physics program. Its first task is acceleration of protons to 120 GeV to feed the Antiproton Source and other consumers of protons, mesons and neutrinos. Its second task is the acceleration of protons and antiprotons to 150 GeV for Tevatron injection.

In the Antiproton Source, 120 GeV protons bombard a nickel target; nickel is chosen because it combines a high nuclear interaction cross-section with the ability to withstand and dissipate the heat generated by the intense proton beam. The resulting spray of particles (which contains approximately two $\overline{p}$ for every 10^5 incident protons) is focused by a current-carrying lithium tube. A pulsed magnet selects 8 GeV antiprotons. These antiprotons are "cooled" (i.e., the spread in their kinetic energies is reduced) in the Debuncher. The antiprotons are then "stacked" in Accumulator, an 8 GeV synchrotron, where they await injection into the Tevatron. Another 8 GeV synchrotron, the Recycler, is used to "stash" antiprotons. The stash can be much larger than the stack, allowing more antiprotons to be injected into the Tevatron at once and therefore increasing the instantaneous luminosity; stashing also allows the Accumulator to run at lower antiproton currents, which improves the stacking efficiency. (The name of the Recycler stems from its original purpose, "recycling" the

antiproton beam remaining in the Tevatron at the end of a store.)

A new "store" of protons and antiprotons is delivered to the Tevatron about once daily. The Main Injector first injects a 150 GeV proton beam into the Tevatron, followed by a 150 GeV antiproton beam. Each beam consists of 36 "bunches" of particles; each bunch is a packets of particles trapped in a potential well (called a bucket) created by the RF system. After injection, the countercirculating beams are accelerated to 980 GeV. Some cleaning of the beams then takes place to reduce the halo of particles that escaped the central orbit during injection. Once this is completed, the beams are focused by quadrupole magnets before being brought to collision in two interaction points, B0 (home of the CDF II detector) and D0 (home of the DØ detector).

3.2 Coordinate system

CDF II uses a right-handed coordinate system in which the origin lies at the center of the detector, $\hat{z}$ points in the proton direction and $\hat{y}$ points up. Because of the symmetry of the detector, coordinates transverse to the z axis are usually quoted in the radius from the symmetry axis r and azimuthal angle ϕ; the polar angle θ or the pseudorapidity

$$\eta = -\log\tan(\theta/2) \tag{3·1}$$

is used as the longitudinal coordinate. It is convenient to use variables that respect the fact that the hard scattering process in hadronic collisions is boosted along the beam axis by an unknown amount. Pseudorapidity is chosen because in the relativistic limit it approaches the rapidity

$$y = \frac{1}{2}\log\frac{E+p_z}{E-p_z}, \tag{3·2}$$

which is additive under boosts along the z axis; as a consequence, the difference $\Delta\eta$ between two objects (for example the opening angle between two tracks) is an approximate invariant in hadronic collisions.

3.3 The vertex detector

The detector closest to the $p\overline{p}$ interaction point is the vertex detector. This detector is designed for the precise determination of primary (production) and secondary (decay) vertices. It has to be finely segmented to operate in a high track density, and it has to be radiation-hard to withstand the high radiation doses at the Tevatron.

The vertex detector is composed of L00 (a single layer of axial strip detectors placed close to the beampipe), SVX II (five cylindrical layers of double-sided sensors providing five axial, three 90° stereo and two small-angle stereo measurements) and ISL (two double-sided axial and small-angle stereo layers) [19]. Tables 3.1 and 3.2 list the properties of the strip detectors. Figure 3.3 shows a three-dimensional view

of the SVX II; Figure 3.4 shows an end view of all three silicon detectors, with more detail on the innermost layers in Figure 3.5.

The SVX readout chip is designed for deadtime-free operation, i.e., data taking does not have to be interrupted for data readout. On every bunch crossing, the strips are read out, and the signal is stored in a 46-cell pipeline. When a Level 1 trigger is issued for an event, that event is moved from the pipeline into one of four Level 2 buffers for readout. Readout of the 400000 channels of the SVX II (which is required for the Level 2 displaced track triggers discussed below) is accomplished in under 10 μs. The fast readout is made possible by performing pedestal subtraction and zero suppression in the readout chip.

3.4 The Central Outer Tracker

Surrounding the vertex detector is the Central Outer Tracker (COT), an open-cell drift chamber with 96 layers of axial and small-angle (2°) stereo sense wires [20]. The precise hit position information in r–ϕ at large radii yields a precise p_T measurement. Each cell (an example of which is shown in Figure 3.6) consists of twelve readout wires tilted with respect to the radial direction by the 35° Lorentz angle. The cells are grouped into eight "superlayers", four axial and four stereo, as shown for one sextant in Figure 3.7.

Tracks traversing the COT ionize a mixture of even parts argon and ethane (with 1.7% isopropanol). The deposited charge drifts to the readout wires, where avalanche discharge results in a gain of approximately 10^4. The arrival time of the pulse on the readout and its time over threshold are digitized; the time over threshold is used to calculate the ionization loss (dE/dx) of the track. Because the maximum drift time is 177 ns, the COT can be read out after every bunch crossing, permitting trigger decisions based on track information.

In early 2004, a rapid gain drop in the COT was discovered. Investigation of a COT wire showed that deposits of hydrocarbons had formed. By adding $\mathcal{O}$ (100 ppm) of O_2 to the gas mixture, the wire aging could be reversed; since summer 2004, the COT has been taking data again with stable gain [21].

3.5 Charged-particle tracking

Charged particle tracks in the pseudorapidity range $|\eta| \lesssim 1$ (the tracks we use in this analysis) traverse the vertex detector and all superlayers of the COT. (See Figure 3.8.) Both detectors are immersed in a 1.4 T magnetic field produced by an aluminum-stabilized NbTi superconducting solenoidal magnet.

Most of the tracks we use are reconstructed from the outside in, meaning that tracks are first reconstructed from COT hits, after which silicon hits are attached. In the COT, the tracking is performed as follows [22]. First, "segment finding" combines hits within a superlayer into straight line segments. Then, two axial tracking reconstruction algorithms run on the segments in the axial superlayers: "Segment

Table 3.1: Properties of the SVX II modules

	L0	L1	L2	L3	L4
Sensor	Ham 90	Ham 90	Micron 1.2	Ham 90	Micron 1.2
Stereo angle	90°	90°	−1.2°	90°	+1.2°
Radius	2.55/3.00 cm	4.12/4.57 cm	6.52/7.02 cm	8.22/8.72 cm	10.1/10.6 cm
Number of modules	72	72	72	72	72
Thickness	300 μm	300 μm	300 μm	300 μm	300 μm
Number of ϕ strips	256	384	640	768	896
Number of z strips	512	576	640	512	896
Number of ϕ chips	2	3	5	6	7
Number of z chips	2	3	5	4	7
ϕ pitch	60 μm	62 μm	60 μm	60 μm	65 μm
z pitch	141 μm	125.5 μm	60 μm	141 μm	65 μm

	L00	ISL
Sensor	Ham/SGS	Micron 1.2
Stereo angle		+1.2°
Radius	1.35/1.62 cm	20.0/22.0/28.0 cm
Number of modules	48	296
Thickness	300 μm	300 μm
Number of ϕ strips	128/256	1024
Number of z strips		768
Number of ϕ chips	3/2	4
Number of z chips		3
ϕ pitch	25 μm	55 μm
z pitch		73 μm

Table 3.2: Properties of the L00 and ISL modules

linking" matches segments from different superlayers. "Histogram linking" starts with a segment in one superlayer, extrapolates the segment to the beam position, and histograms the position in curvature space of hits in the other superlayers; if the resulting histogram has a bin with a sufficient number of hits, the corresponding combination of hits is accepted as a track. After linking by either algorithm, a fit is performed on the hits to determine the track parameters. Because a large fraction of tracks missed by one linking algorithm is found by the other, the combination of the algorithms yields a high tracking efficiency: greater than 99% for tracks with $p_T > 1.5$ GeV [23]. Hits from the stereo superlayers are then attached to the two-dimensional track.

In the silicon, strips are first merged into clusters. These clusters are then attached to the COT tracks and the tracks refit in an iterative procedure in decreasing radius. The resulting tracks are known as "outside-in" tracks [24].

Our analysis makes so-called "standard" (or "tight") quality cuts on tracks. Tracks are required to have segments in at least 2 axial and 2 stereo superlayers in the COT with at least 5 hits in each segment, as well as at least 3 axial silicon hits. The only exception to this rule is the daughter track of a D^{*+}, which does not require silicon hits (so-called "soft" cuts). Table 3.3 shows the cuts.

3.6 Trigger system

The trigger system reduces the data rate from the 2.5 MHz bunch-crossing rate (of which 1.7 MHz is actual beam crossing) to 100 Hz, a data volume that can be written to tape. The first two of the trigger system's three levels are customized hardware designed for quick decision based on coarse-grained subset of the entire event information. Level 3 runs full event reconstruction using a modified version of the offline

	Standard	Soft
SVX II radial hits	≥ 3	—
COT axial superlayers	≥ 2	
COT hits per superlayer	≥ 5	
p_T	> 0.35 GeV	
$\|\eta\|$	< 2	

Table 3.3: Track quality cuts. The standard cuts are used for all tracks except the soft daughter tracks of a D^{*+}.

production executable on a farm of commercial PC's. We will discuss the features of the trigger that allow us to collect a sample rich in hadronic decays of bottom mesons. Figure 3.9 shows a schematic of the data acquisition.

At Level 1, tracks with p_T above 1.5 GeV are reconstructed by the Trigger Track Processor (XFT) [25] from axial hits in the COT. Track information is distributed by the XTRP for electron, muon and track triggers. This analysis uses the Level 1 track triggers. Level 1 reduces the input rate of 2.5 MHz to an output rate around 35 kHz; the maximum trigger latency is dictated by the 42-event pipeline. The Level 1 trigger system can subject an event to 64 sets of requirements (called "trigger bits").

At Level 2, the Silicon Vertex Trigger (SVT) [26] associates SVX II r–ϕ position measurements with XFT tracks, providing a precise measurement of the track impact parameter d_0. Decays of heavy flavor particles are identified by requiring two tracks with 0.12 mm $< d_0 <$ 1 mm and making additional cuts on the two-track combination. The latency of the Level 2 trigger is 20 μs (of which 10 μs is due to SXV II readout). Four event buffers are available. The maximum output rate is 500 Hz. 192 trigger bits can be accommodated at Level 2. See Figure 3.10 for a block diagram of Level 1 and Level 2.

The Level 3 trigger [27] performs a full event reconstruction and applies a superset of the Level 1 and 2 selections. Because Level 3 performs full event reconstruction, it has access to more information with better resolution than Level 1 and 2. A significant reduction of background is therefore possible at Level 3. The Level 3 output rate is approximately 100 Hz. At Level 3, the trigger bit concept is replaced by "trigger paths". Trigger paths group the trigger requirements at the three trigger levels. There are typically more than 100 trigger paths in use. Events that fail at a given trigger level but would pass a higher level (called "volunteer events") are discarded.

Our analysis uses data collected by the "two-track trigger", which requires two tracks that do not point back to the beamline. (This requirement is imposed at Level 2, the earliest point at which track impact parameter information from the SVX II is available.) Requiring displaced tracks exploits the long lifetime of bottom hadrons, which is a key discriminating feature against light-quark background. The trigger paths on which hadronic B decays to charm are collected are called the "`B_CHARM`"

	SCEN_LOW	SCEN_A	SCEN_C
opposite charge	not required	required	required
p_T	> 2 GeV	> 2 GeV	> 2.5 GeV
$\sum p_T$	—	> 5.5 GeV	> 6.5 GeV
SVT χ^2		< 25	
L_{xy}		> 200 μm	
$\Delta\phi$		$2^\circ < \Delta\phi < 90^\circ$	
d_0		120 μm $< d_0 <$ 1 mm	

Table 3.4: B_CHARM trigger scenario cuts. The evolution over time towards tighter cuts is not reproduced here.

trigger family. It is easy to saturate the trigger bandwidth with these triggers; at high instantaneous luminosity, the high-rate triggers are fit into the available trigger resources by randomly discarding a variable fraction of the triggered events (the trigger is "prescaled"). There are several methods used to reduce the waste of prescaling: the prescale factor can be dynamically reduced as the instantaneous luminosity decays over the course of a Tevatron store ("dynamic prescaling"), and if a Level 2 event buffer is available, a Level 1 trigger can be accepted regardless of its current prescale factor ("überprescaling"). Nevertheless, at high instantaneous luminosity, it is more efficient to trigger with tighter cuts (which reject more background) than to discard events at random. At low luminosity, where trigger rates are lower, it is more efficient to fill the available bandwidth with looser cuts. Therefore, CDF implements a spectrum of two-track triggers with increasingly stringent cuts (called "scenarios"). With increasing instantaneous luminosity, the looser scenarios are increasingly prescaled away and the tighter scenarios become increasingly important. Table 3.4 lists the cuts in the three B_CHARM trigger scenarios. These cuts are imposed at Level 2 by the SVT and confirmed at Level 3 on the fully reconstructed event. The cut variables are the relative charges of the two tracks; the transverse momentum p_T of each track; the scalar sum of the transverse momenta; L_{xy}, the projection of the B and D decay length along the transverse momentum of the B candidate; the azimuthal opening angle $\Delta\phi$ between the two tracks; and the impact parameter d_0 of each track with respect to the beamline.

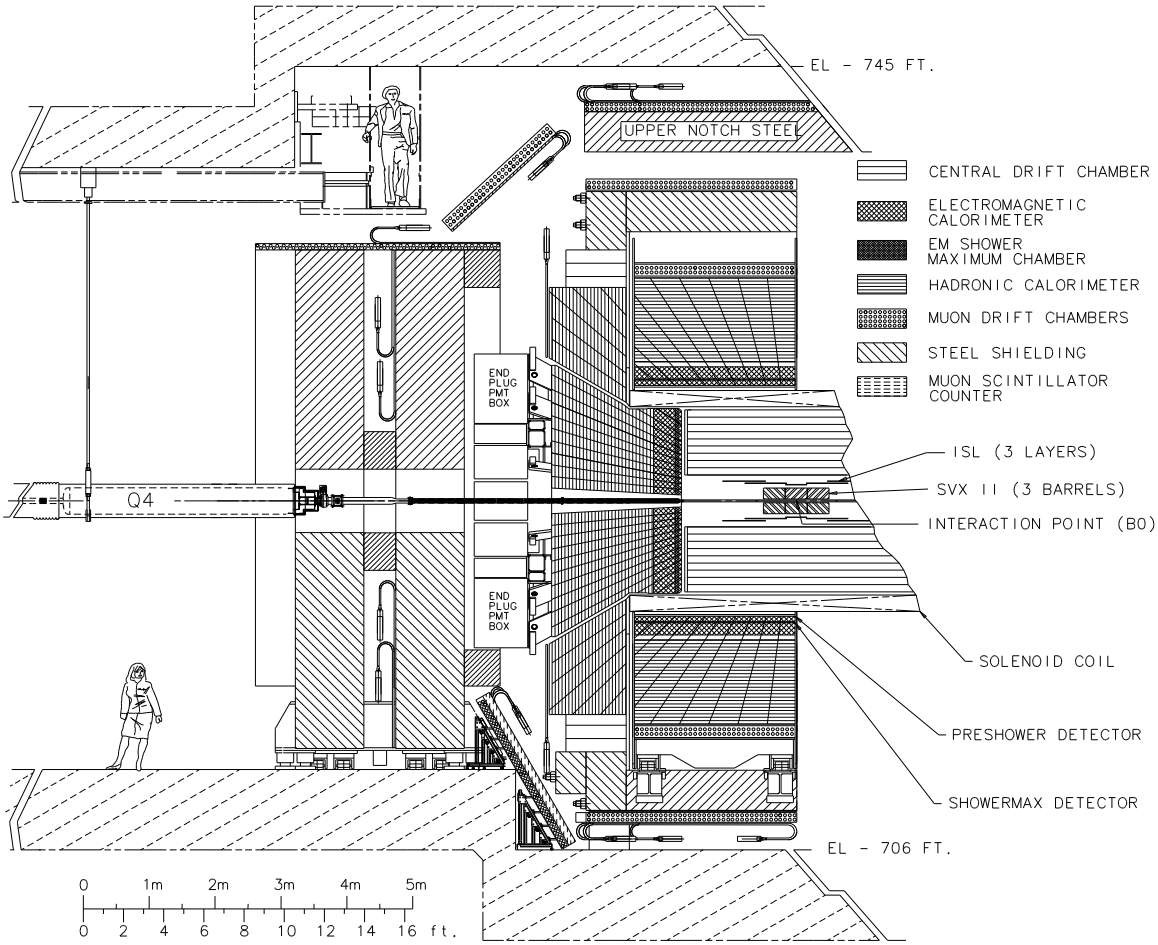

Figure 3.1: CDF II sectional view

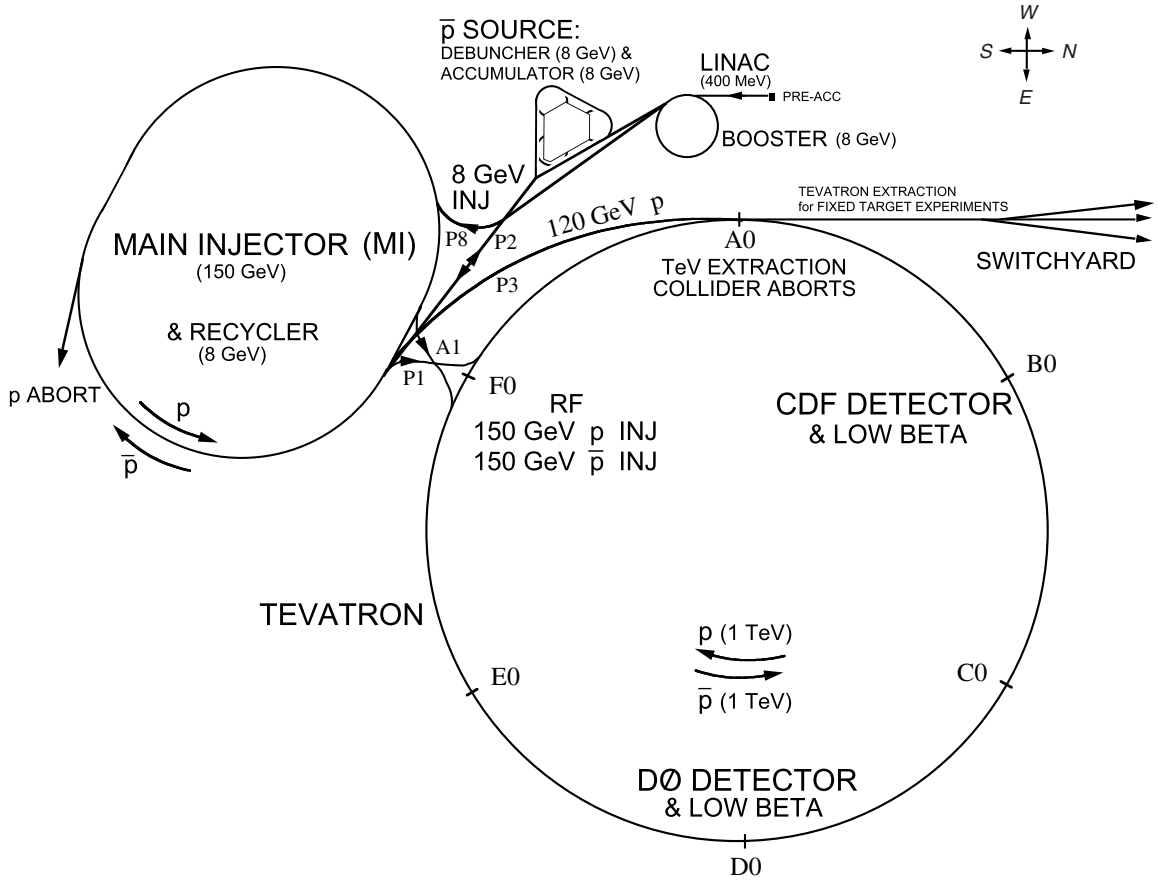

Figure 3.2: The Fermilab accelerator complex

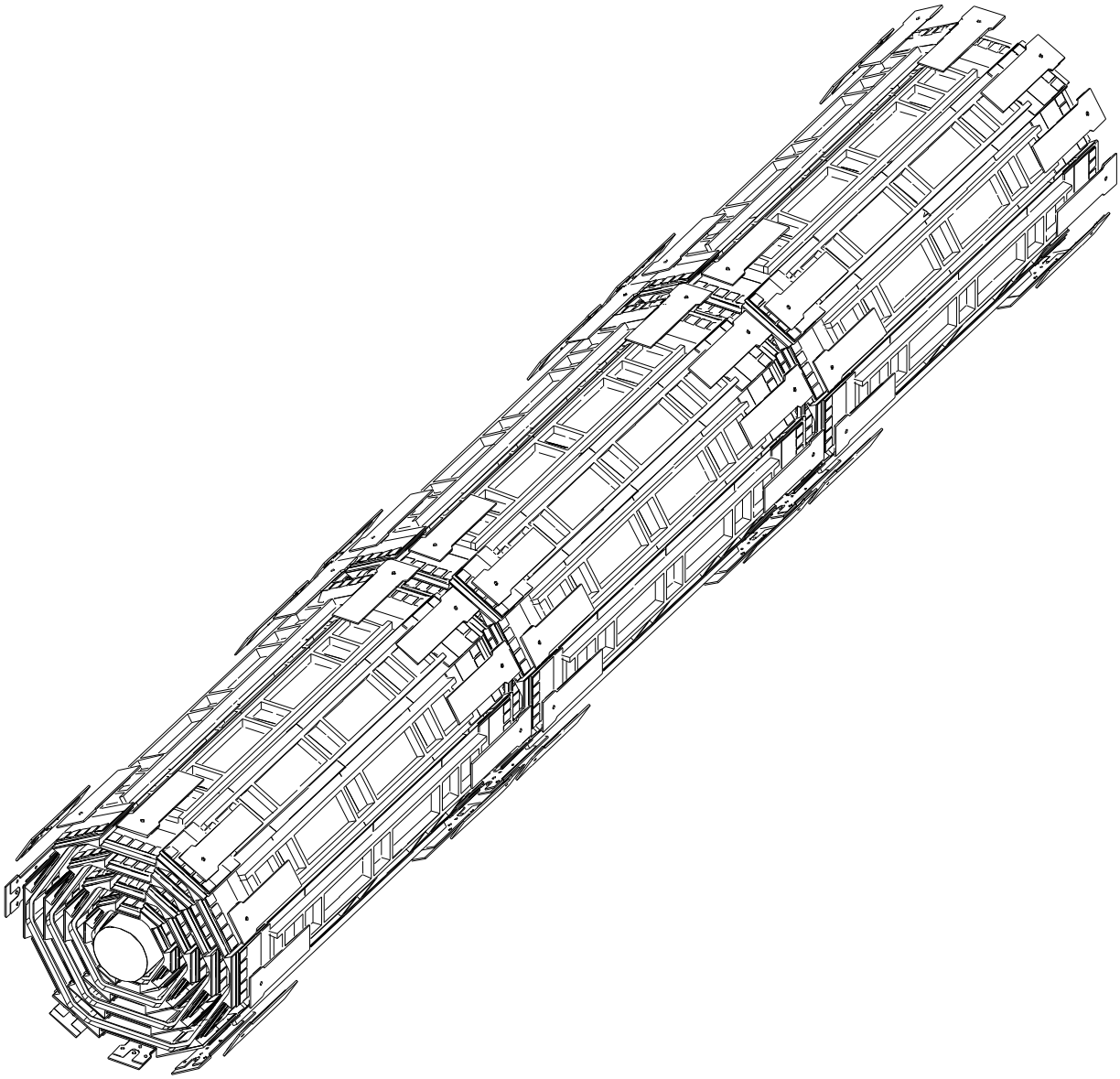

Figure 3.3: 3D view of SVX II

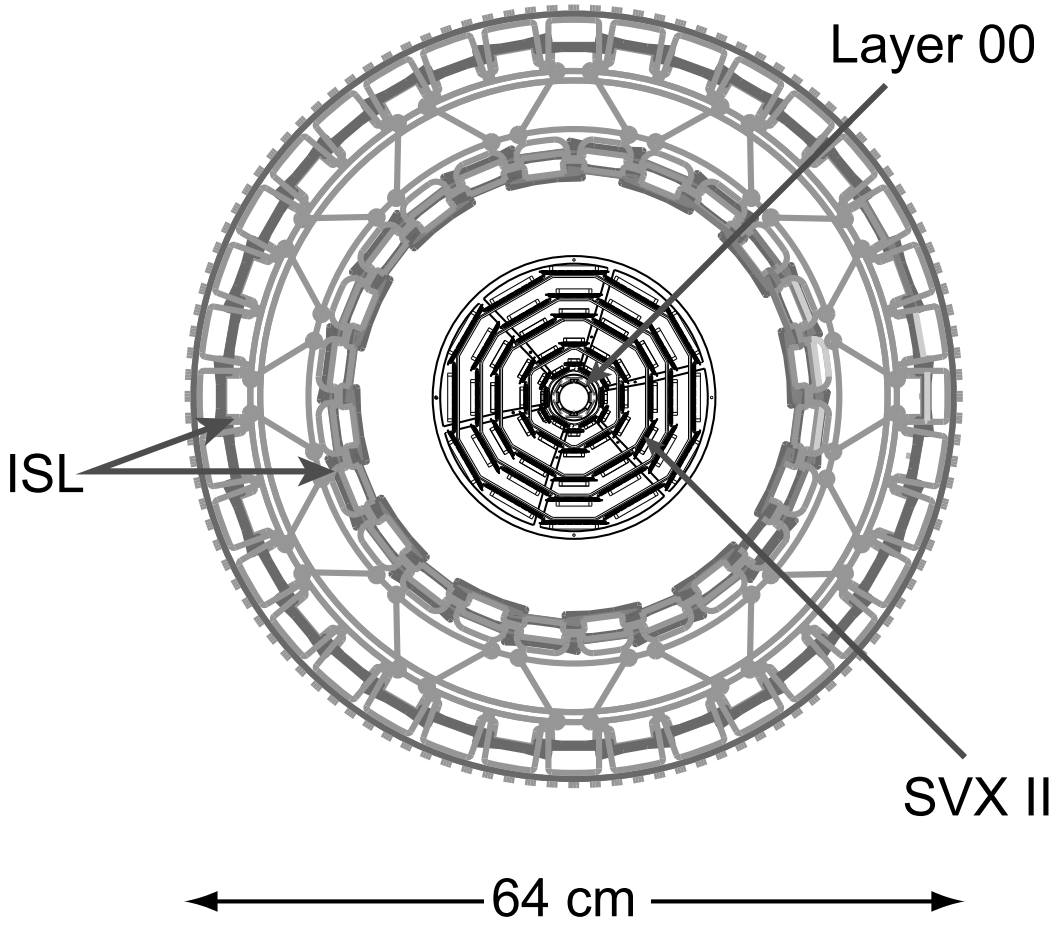

Figure 3.4: End view of L00, SVX II, ISL

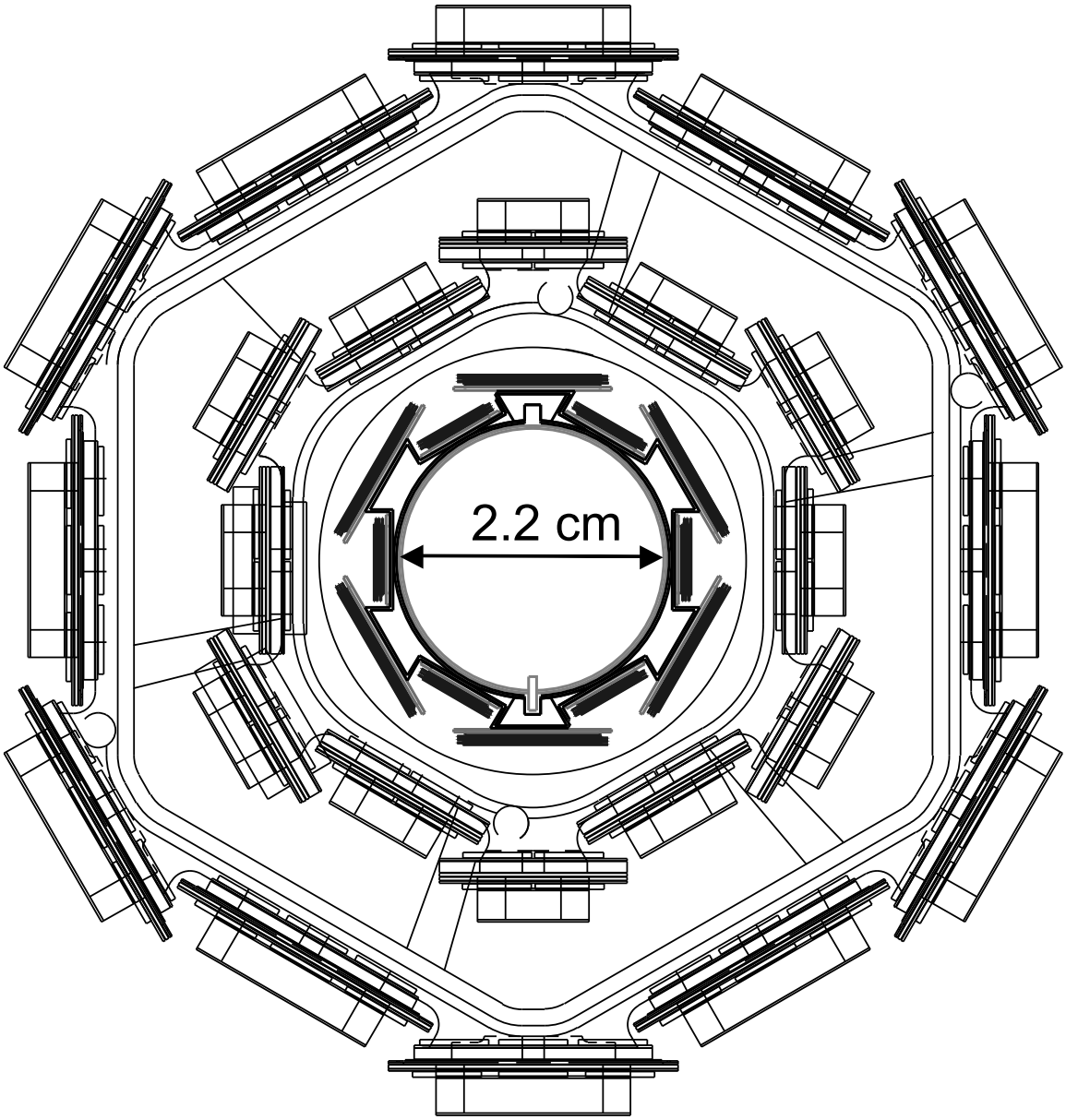

Figure 3.5: End view of L00 and the innermost layers of SVX II

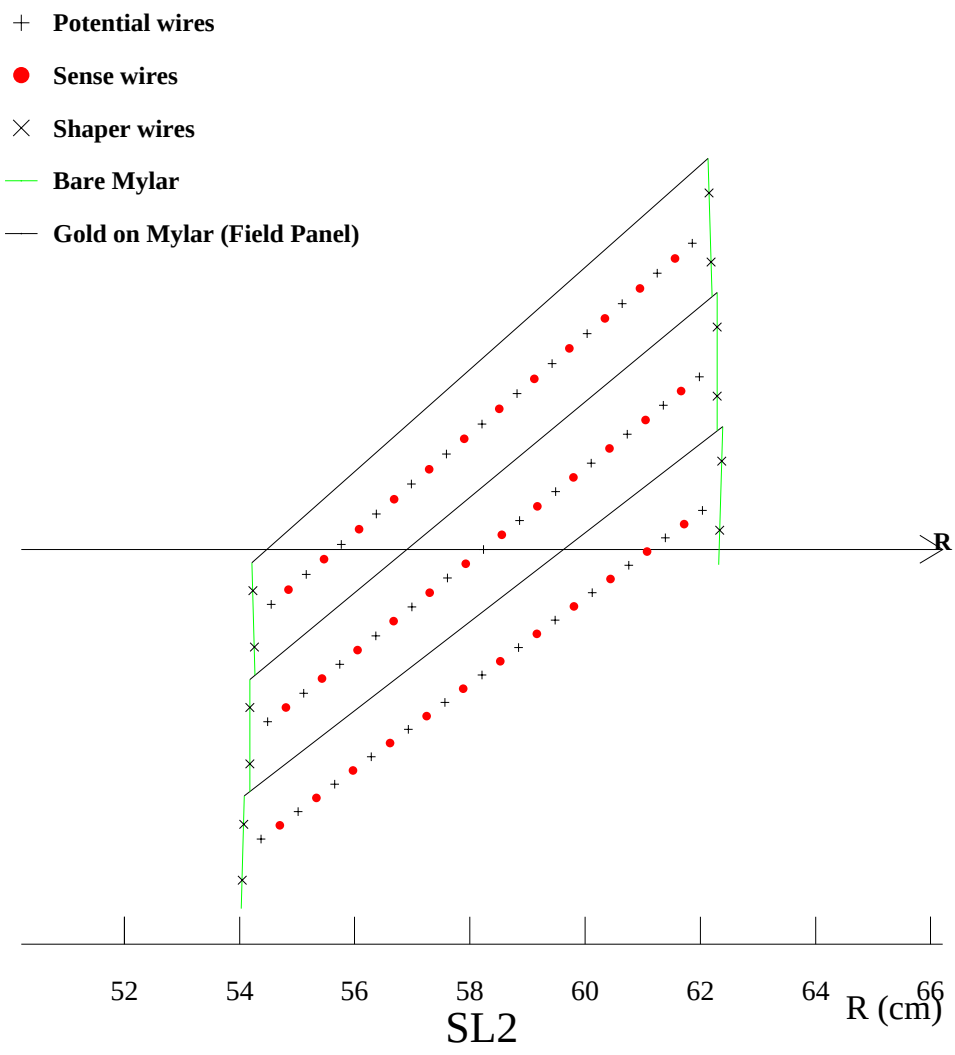

Figure 3.6: A COT cell in superlayer 2

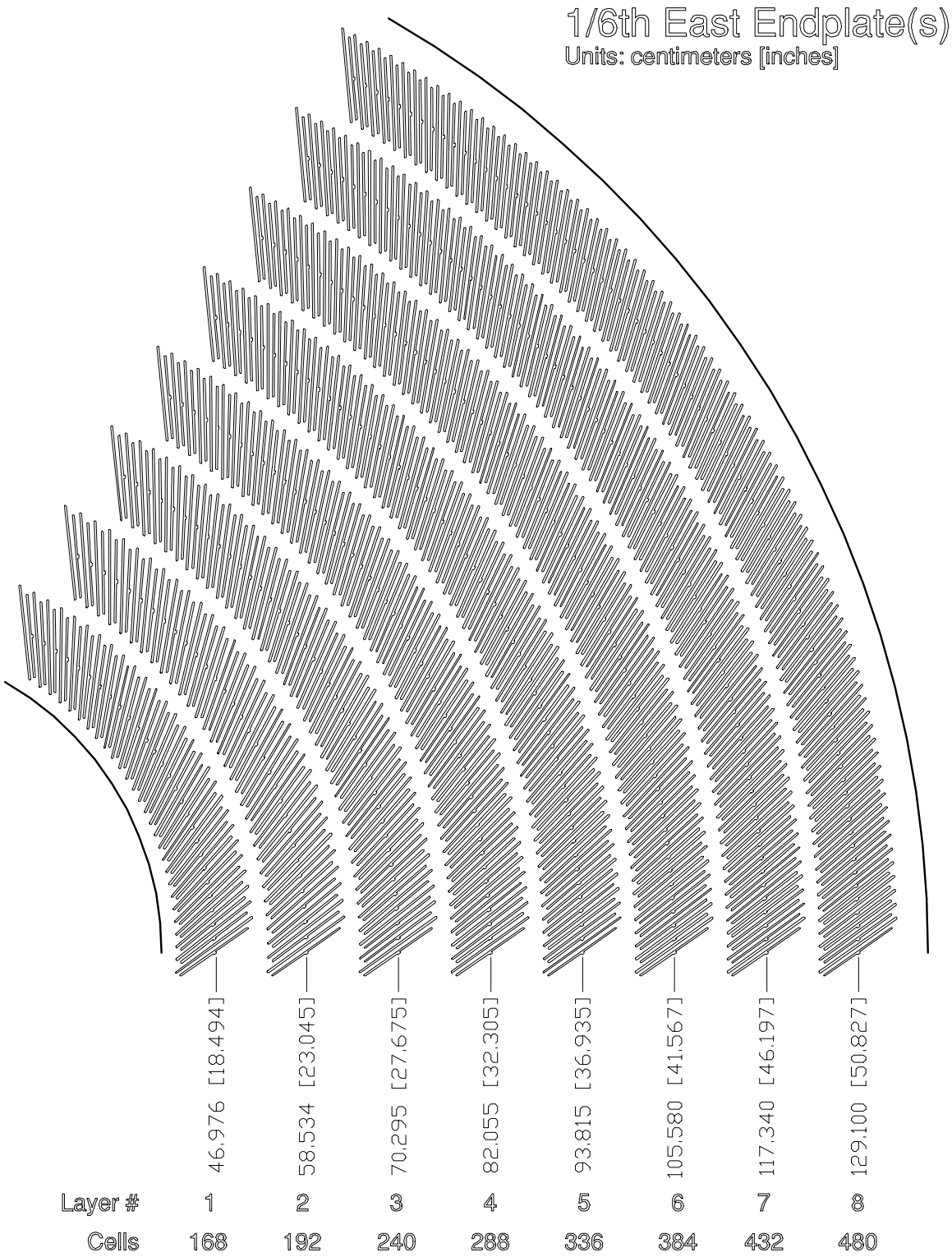

Figure 3.7: COT sextant

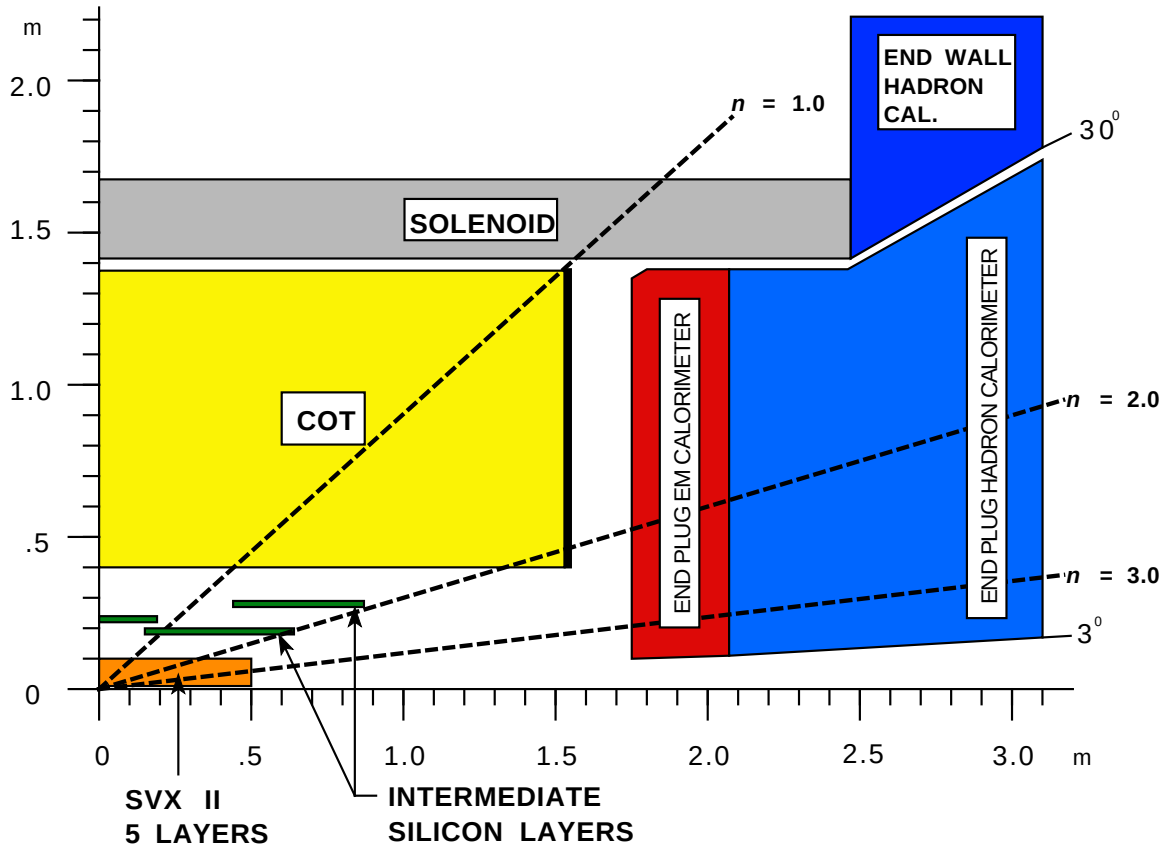

Figure 3.8: Cross-sectional z–y view of the tracking system

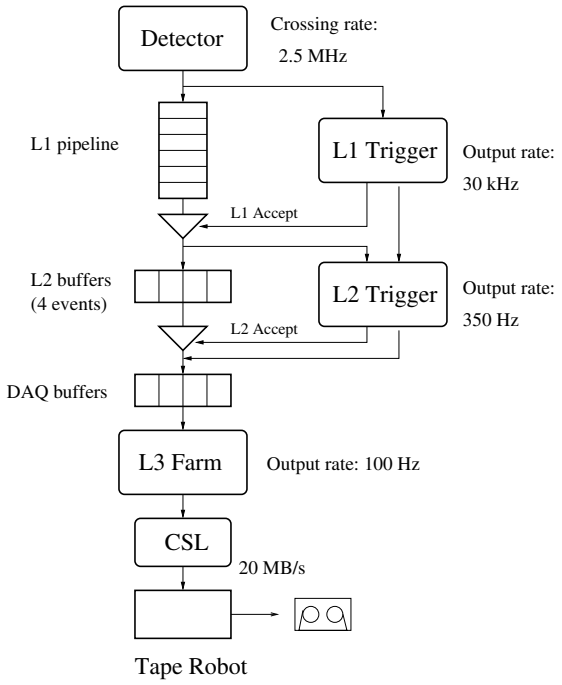

Figure 3.9: DAQ system

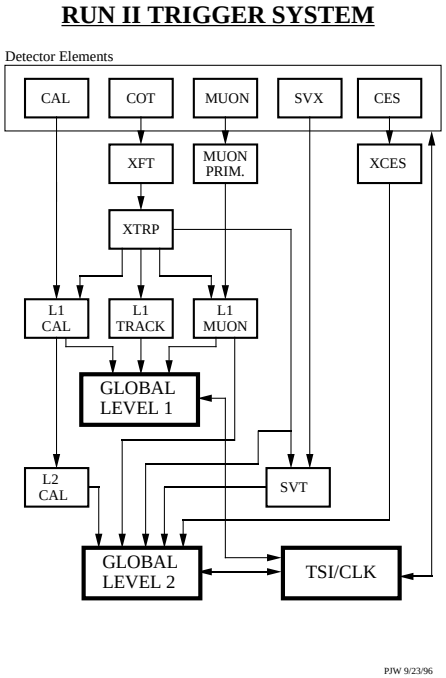

Figure 3.10: Level 1 and Level 2 block diagram

Chapter 4

Methodology of the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurement

The work presented in this thesis consists of three measurements of relative branching fractions of the type $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$: first on the control samples $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$, where the result is known from previous measurements, and then on the signal sample $\overline{B}^0_s \to D^+_s X$. While each sample necessarily receives slightly different treatment, the principle of the analysis is the same for all three samples. This chapter describes the procedure we use for $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurements; sample-specific details are given in the chapter on the control samples (Chapter 5) and the signal sample (Chapter 6).

4.1 Statement of the problem

Several layers of background obscure the $\overline{B}^0_s \to D^\pm_s K^\mp$ signal. The first problem occurs in the production of B^0_s mesons: the $p\overline{p}$ inelastic cross section overwhelms the $b\overline{b}$ cross section by several orders of magnitude. The non-$b\overline{b}$ background is addressed by the Silicon Vertex Trigger described in Section 3.6 and by selection cuts described in this chapter. The next problem occurs in the B^0_s decay: the branching fraction of $\overline{B}^0_s \to D^+_s\pi^-$ is $3 \cdot 10^{-3}$ [28], and the branching fraction of $\overline{B}^0_s \to D^\pm_s K^\mp$ is expected to be an order of magnitude smaller. A combination of selection cuts and modeling of the contributions coming from B hadron decays other than $\overline{B}^0_s \to D^\pm_s K^\mp$ or $\overline{B}^0_s \to D^+_s\pi^-$ is vital to extracting the correct yields for $\overline{B}^0_s \to D^\pm_s K^\mp$ and $\overline{B}^0_s \to D^+_s\pi^-$. The final problem is separating $\overline{B}^0_s \to D^\pm_s K^\mp$ from $\overline{B}^0_s \to D^+_s\pi^-$. There are two discriminating variables available: the mass of the B^0_s candidate reconstructed in the $D_s\pi$ mass hypothesis and the specific ionization (dE/dx) of the track recoiling against the D_s. Figures 4.1 and 4.2 show the difficulty of separating the $\overline{B}^0_s \to D^\pm_s K^\mp$ from the $\overline{B}^0_s \to D^+_s\pi^-$ — in mass, the Cabibbo-suppressed $\overline{B}^0_s \to D^\pm_s K^\mp$ is a small bump on the shoulder of the Cabibbo-allowed $\overline{B}^0_s \to D^+_s\pi^-$, and, to further exacerbate the problem, falls in the same region as the radiative tail of the $\overline{B}^0_s \to D^+_s\pi^-$; in dE/dx, the separation is far too small to allow a candidate-by-candidate cut. The correct procedure therefore is to perform a two-dimensional fit for the $\overline{B}^0_s \to D^\pm_s K^\mp$ and $\overline{B}^0_s \to D^+_s\pi^-$ yields, using mass and dE/dx as fit variables.

4.2 Discriminating variables

We now discuss why invariant mass and dE/dx are well-suited discriminating variables for our purposes. As the first step of B candidate reconstruction, we combine the tracks recorded in an event. No particle identification is performed on the tracks at this point; instead, a mass hypothesis is imposed on the track to match the decay products of the hadron being reconstructed. To reconstruct $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$ candidates (as a concrete example), we first form $D_s^+ \to \phi(K^+K^-)\pi^+$ candidates, which are three-track combinations in the mass hypothesis K^+, K^-, π^+; the invariant mass of the K^+K^- pair is required to be within some distance of the ϕ mass, and the invariant mass of the three-track combination is required to be within some distance of the D_s^+ mass. (Vertex quality cuts, which are crucial to rejecting light-quark combinatorial background, are also applied; see the discussion in Sections 4.4 and 4.5.) A fourth track, under the π^- mass hypothesis, is then added to the D_s^+ candidate to make the $\overline{B}_s^0$ candidate. We will call the invariant mass of the $\overline{B}_s^0$ candidate $m(B_s^0)$ or, when we wish to stress the mass hypothesis used in the reconstruction, $m(D_s\pi)$.

Different contributions to the data sample will have very different $m(D_s\pi)$ distributions; hence, $m(D_s\pi)$ is a powerful discriminating variable in our fit for the yields of the various contributions. $\overline{B}_s^0 \to D_s^+\pi^-$ decays will be distributed around the true B_s^0 mass with a certain detector resolution. $\overline{B}_s^0 \to D_s^\pm K^\mp$ decays, where the kaon track is given the incorrect π mass assignment, have a long tail towards low masses. Decays to more than four particles, such as $\overline{B}_s^0 \to D_s^+\rho^-$, have a mass cutoff below the B_s^0 mass, and sometimes exhibit distinctive features. Finally, combinatorial background has no particular identifying mass characteristics; its distribution is fairly smooth, rising at low $m(D_s\pi)$.

dE/dx is used as a discriminating variable because it allows us to separate kaon tracks from pion tracks and therefore $\overline{B}_s^0 \to D_s^\pm K^\mp$ from $\overline{B}_s^0 \to D_s^+\pi^-$. Since dE/dx depends on $\beta\gamma$, dE/dx for tracks of the same momentum differs between species.

4.3 Data sample

This analysis uses data samples collected between February 2002 and September 2006. This excludes data taken during the early part of Run 2, when the detector configuration was unstable. Also excluded is data from the period in 2004 when the COT was suffering from reduced gain (Section 3.4). The integrated luminosity of the dataset is approximately 1.2 fb^{-1}. All events are collected in the `B_CHARM` family of trigger paths described in detail in Section 3.6; these trigger paths require two displaced tracks at trigger Level 2 with varying levels of stringency in cuts. We require that the runs included in the analysis were marked good by the shift crew and were taken with COT, SVX, SVT, central muon and calorimetry systems operational; this list of good runs is called the B_s mixing good-run list. (Although the muon systems and calorimetry are not required for this analysis, the good-run lists were developed to be usable by all B-physics analyses.)

4.4 Reconstruction of B candidates

We reconstruct three B decay modes with very similar topologies; as a result, the reconstruction method is identical for the three decays, although the selection cuts differ slightly. The B decays are $\overline{B}^0 \to D^+\pi^-$, $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0_s \to D^+_s\pi^-$. (This notation means that the B candidate is reconstructed assuming only one track recoils against the D^+, D^{*+} or D^+_s candidate and that that track is a pion. As noted in Section 4.2, the invariant mass of B candidates that are true $B \to D\pi$ decays will cluster around the B mass, while other contributions to the sample will have different invariant-mass shapes.) The D daughters of the B decay are reconstructed in the following modes: $D^{*+} \to D^0(K^-\pi^+)\pi^+$, $D^+ \to K^-\pi^+\pi^+$ and $D^+_s \to \phi(K^-K^+)\pi^+$.

Reconstruction of B candidates begins by first reconstructing a D candidate, to which another track is then attached. This reduces the number of track combinations dramatically, since many three-track combinations can be rejected by requiring consistency with a D decay; only for the candidates passing the D selection cuts do four-track combinations need to be considered. D candidates are required to fall within a wide window of the mass of the appropriate D species. In the $D^+_s \to \phi(K^-K^+)\pi^+$ case, a ϕ mass cut on the K^-K^+ pair is imposed; in the $D^{*+} \to D^0(K^-\pi^+)\pi^+$ case, a cut on $\Delta m = m(D^{*+}) - m(D^0)$ is made. (We shall see later that the narrow D^{*+} and ϕ are tremendous background rejectors.) Further cuts are made on the maximum $\Delta\phi$, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ and Δz of the D daughter tracks; the scalar p_T sum of the tracks; the p_T of the D candidate; and on L_{xy}, the projection of the D decay length along the transverse momentum of the D. In addition, one of the D daughter tracks has to pass SVT d_0 and p_T confirmation (see Table 3.4).

B candidates are formed from the D candidates that survive the selection cuts. Further cuts are made on the maximum $\Delta\phi$, ΔR and Δz of the tracks; the scalar p_T sum of the tracks; the p_T of the B candidate; the ΔR between the D and the B daughter track; the impact parameter d_0 of the B candidate; and on L_{xy}, the projection of the B decay length along the transverse momentum of the B. In addition, the B candidate has to pass `SCEN_LOW` SVT confirmation (see Table 3.4). Full tables of skimming cuts can be found in the control-sample chapter (Table 5.1) and signal-sample chapter (Table 6.1).

Three-dimensional vertexing is performed on the tracks that constitute the D candidate using the CTVMFT code [29]; candidates with $\chi^2 > 30$ on the transverse degrees of freedom of the vertex fit are rejected. A further vertex fit is performed on the combination of the D candidate and the B daughter tracks, with the same χ^2_{xy} requirement. The B vertex fit can be performed with or without constraining the D candidate to the appropriate D mass. For the D, the mass resolution is approximately 6 MeV; for the B, the mass-constrained resolution is approximately 17 MeV, while the unconstrained resolution is approximately 20 MeV.

An event-by-event primary vertex is calculated from all reconstructed tracks in the event [30]. This procedure determines the z position of the primary vertex as well as the x–y position. (Due to the small crossing angles of the Tevatron beams, the RMS in z of the beam spot is approximately 10 cm, compared to 30 μm in x–y.)

During candidate reconstruction, the primary vertex is used for a transverse impact parameter cut on the B candidate. In the analysis-level selection cuts, we also cut on the longitudinal impact parameter.

The datasets passing these loose selection cuts are called "skimmed" datasets. We store these datasets for further analysis, dropping many of the raw data banks to reduce the event size and saving the reconstructed events. From the skimmed datasets we generate ntuples. Because the ntuple contains only the information required for the analysis, it is smaller and (since much of the time in processing the data is spent in serializing and de-serializing the file contents) faster than skimmed data. Our ntuple is comprehensive enough that re-ntupling the data was required on only one occasion during the course of the analysis.

4.5 Candidate selection

Two considerations drive the optimization of the candidate selection cuts. The first is statistical: the cuts should maximize $S^2/(S+B)$, where S is the expected $B \to DK$ signal yield and B is the background size. ($B \to D\pi$ is also a signal for us, but since the $B \to DK$ yield is smaller by an order of magnitude, it dominates the statistical uncertainty on $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$.) The second consideration is that combinatorial background contributes to systematic uncertainties; therefore, background-reducing cuts can be beneficial even if they reduce $S^2/(S+B)$.

Each sample uses slightly different cuts; these are listed in the chapters dealing with the specific samples. Here we describe the cut variables, many of which are common to the three samples. The discriminating variables are the transverse (d_0) and longitudinal (z_0) impact parameter of the B candidate with respect to the primary event vertex; the p_T of the B candidate and B daughter track; χ^2_{xy}, the χ^2 of the transverse components of the vertex fit for the B or D vertex; the opening angle ΔR between the D candidate and the track recoiling against the D; L_{xy}, the projection of the B (or D) decay length along the transverse momentum of the B (or D) candidate; and the isolation of the B candidate, defined by

$$\mathrm{iso}(B)\ (\Delta R < 1) = \frac{p_T(B)}{p_T(B) + \sum\limits_{\text{tracks}} p_T(\text{track})} \qquad (4 \cdot 1)$$

where the sum extends over tracks within $\Delta R < 1$ of the B.

A requirement common to all samples is that the B daughter track be an SVT track, since the calibrations on the dE/dx are valid only for trigger tracks. To determine whether an offline track matches one of the SVT tracks, a "match distance" is calculated from the ϕ_0 and curvature of the offline and SVT tracks; this match distance is required to be less than 25. The quality of the SVT track is assured by requiring its χ^2, calculated from the distance of the hits to the track, to be less than 15.

4.6 Monte Carlo simulation

Monte Carlo simulation is used to model signal and background and to determine trigger and reconstruction efficiencies. We generate single B hadrons with `Bgenerator` [31, 32] and simulate their decays with `EvtGen` [33]. A detailed detector and trigger simulation is then performed. The detector simulation is based on `GEANT` 3 [34]; the `TrigSim` package runs an offline implementation of the Level 1 and Level 2 trigger algorithms on the simulated detector response.

It is necessary to correct the Monte Carlo B mass and mass resolution. These corrections take the form of a scale factor on the mass and convolution of the mass shape with a gaussian smearing function. The corrections are determined by comparing the peak mean and width in a clean $\overline{B}^0 \to D^{*+}\pi^-$ data sample with Monte Carlo. For the control samples, the corrections are cross-checked by scanning the maximum likelihood as a function of resolution and mass scale; this procedure is described in greater detail in Sections 5.6 and 5.10.

4.7 Mass and dE/dx fit

As we have explained at the beginning of this chapter, the central challenge in our analysis is to disentangle the various components contributing to the data samples: apart from the $B \to DK$ and $B \to D\pi$ signals, the sample contains various other partially reconstructed B decays, reflections from decays of other bottom hadron species, and combinatorial background. To separate the components and determine the number of candidates of each component type, we perform a maximum-likelihood fit using as the discriminating variables the invariant mass $m(D\pi)$ of the candidate in the $D\pi$ mass hypothesis and the particle-identification variable

$$Z = \log\left(\frac{dE/dx(\text{measured})}{dE/dx(\text{expected for } \pi)}\right) \tag{4·2}$$

for the B daughter track. (Using Z is technically simpler than using dE/dx because it is normally distributed.) The maximization of the likelihood is performed using `MINUIT` [35].

Two different likelihood fit techniques can be used fairly interchangeably. In the first, we fit for the fraction of events in each component; since the fractions must sum to unity, a fit for M components has $M-1$ fit parameters. The likelihood is

$$L(f_1, \ldots, f_{M-1}) = \prod_{i=1}^{N} \sum_{j=1}^{M} f_j\, p_j(m_i)\, q_j(Z_i), \quad \text{where } f_M = 1 - \sum_{j=1}^{M-1} f_j. \tag{4·3}$$

$\prod_i$ is a product over all N candidates and $\sum_j$ is a sum over the M components (decay modes and background) of the fit. The fit parameters f_j are the fractions of candidates in the j-th component. We will describe the mass pdf's $p_j(m)$ and Z pdf's $q_j(Z)$ (also referred to as dE/dx pdf's) in greater detail shortly.

The second fit technique, called "extended likelihood", fits for the candidate yields in each component; for M components, there are M fit parameters. The sum of the yields $\sum_j^M \mu_j$ is constrained to the size of the dataset N through a Poisson constraint:

$$f\Big(\sum_j \mu_j; N\Big) = \frac{(\sum_j \mu_j)^N \exp(-\sum_j \mu_j)}{N!}. \tag{4·4}$$

The extended likelihood function is

$$L(\mu_1, \ldots, \mu_M) = \frac{\exp(-\sum_j \mu_j)}{N!} \prod_{i=1}^{N} \sum_{j=1}^{M} \mu_j \, p_j(m_i) \, q_j(Z_i); \tag{4·5}$$

we have used the $(\sum_j \mu_j)^N$ factor in $(4 \cdot 4)$ to change the parameters f_j to $\mu_j = N f_j$ in the product over candidates.

During the development of the analysis, the ability to test many different fit configurations quickly proved very useful. As a result, the fitting code supports various features in addition to maximizing the likelihoods $(4 \cdot 3)$ and $(4 \cdot 5)$.

- Gaussian constraints on the ratios between components can be imposed to reflect values of known branching fractions.
- A component can be set to a fixed fraction of another component. This feature is useful in the treatment of the $q(Z)$ of the B-physics background.
- Parameters beyond the $\mu_1, \ldots, \mu_M$ can be used; shape parameters for the background mass pdf's are an example.
- Z pdf's can be made mass-dependent.
- Parameters can be fixed.

Each sample has its own set of fit components, which are described more closely in the sample-specific chapters. These components have three possible sources:

- D plus track combinations where the D and the track come from a single B hadron. We will call the pdf's from such samples "single-B templates".
- Cases where the D candidate does not come from a real D. We will call the pdf's for this case "fake-D templates".
- D plus track combinations where a real D is combined with a track coming from fragmentation, the underlying event or the other B. We will call the pdf's from such samples "real-D templates".

4.8 Mass Templates

Mass pdf's for the single-B components are extracted from large Monte Carlo samples of $B \to DX$ decays. These samples are generated with the `Bgenerator` single-B gun. The D is forced to decay in one of the modes ($D_s^+ \to \phi\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $D^{*+} \to D^0(K^-\pi^+)\pi^+$) used in this analysis. Because these samples are inclusive in the B decay modes but exclusive in the D decay mode, we call them "semi-inclusive".

Separate mass templates are extracted according to true B decay mode for the Cabibbo-allowed $B \to D\pi$ and Cabibbo-suppressed $B \to DK$ fully reconstructed decays and for the partially reconstructed modes that overlap in mass with the $B \to DK$ ($B \to D\rho$ and $B \to D^*\pi$). Partially reconstructed modes missing more than one decay product are collected in a catch-all "remainder" template. Depending on the data sample, reflections from other B hadrons also need to be considered.

Figure 4.3 shows a stacked histogram of the mass distributions of some B_s^0 decay modes. It is clear that some of the shapes are not easy to parameterize. Since our Monte Carlo statistics is at least a factor three larger than the data statistics, we represent the mass pdf's for the single-B fit components as histograms and ignore the statistical uncertainty on the templates. We will determine the systematic uncertainty due to the statistical fluctuations in the templates in Section 6.5.

One place where special care has to be taken is in the treatment of the radiative tail in the Cabibbo-allowed decay $B \to D\pi$. Proper treatment of the tail is important for properly determining the number of events in the Cabibbo-allowed mode. This tail also puts additional $D\pi$ events into the region of the Cabibbo-suppressed decay $B \to DK$. The radiative tail on the Cabibbo-allowed mode is modeled in `EvtGen` using the `PHOTOS` algorithm for radiative corrections [36]. This model is expected to describe the shape of the tail well, but its reliability for determining the rate is not known. We leave the size of the radiative tail as a free parameter in our fit. Monte Carlo events in the Cabibbo-allowed mode are separated into a non-radiative subsample and a radiative one, and separate mass templates are created for each subsample; the `PHOTOS` cut-off on the photon energy in `EvtGen` is 10 MeV.

As noted above, the mass resolution in Monte Carlo is slightly better than in the data. We therefore re-smear the mass histograms obtained from the Monte Carlo by convolving them with a Gaussian of width 6.43 MeV (compared with a resolution of 17 MeV for mass-constrained candidates and 20 MeV for unconstrained candidates). Our original skim had a minimum mass cut of 4.65 GeV. After this convolution, we limit our mass fit range to above 4.85 GeV in order to avoid edge effects from the smearing. This mass cut also limits our sensitivity to many-body B decays whose branching fractions are poorly known.

Both combinatorial-background mass pdf's are determined from data. For the fake-D background, B candidates taken from the sidebands of the D are used. For the real-D background, wrong-sign (WS) B candidates (e.g., $\overline{B}^0 \to D^+\pi^+$) are used. In contrast to the single-B templates, the statistics for the combinatorial-background

templates is limited. We therefore describe these templates using a smooth function:

$$p_{\rm bg}(m) = \frac{\alpha(1-\beta)\exp(-\alpha m)}{\exp(-\alpha m_{\rm min}) - \exp(-\alpha m_{\rm max})} + \frac{\beta}{m_{\rm max} - m_{\rm min}}, \qquad (4\cdot 6)$$

where $m_{\rm min}$ and $m_{\rm max}$ are the limits of the fit range.

For the fake-D background, the parameters of $(4\cdot 6)$ are determined by performing an auxiliary fit to the $D\pi$ mass distribution for events in the D sideband. A certain amount of work is involved in determining the fake-D background properties from the D sidebands, owing to the presence of poorly reconstructed real D mesons (coming from real B decays) in the D sideband. Figure 4.4 illustrates the signal "contamination" of the background sample. A procedure called "signal subtraction" removes this contamination: we first determine the $D\pi$ mass shape of the poorly reconstructed D candidates by selecting the D sidebands in the semi-inclusive $B \to DX$ Monte Carlo samples. We then perform an auxiliary fit to the $D\pi$ mass distribution of the D sideband data. This fit includes two fit components, the $D\pi$ mass histogram for the D sideband Monte Carlo and a background shape of the form $(4\cdot 6)$. The auxiliary fit then determines the amount of signal leakage in the sidebands as well as the shape parameters of the fake-D contribution. For illustration, Figure 4.5 shows the result of the auxiliary fit in the sidebands of the D^+.

The real-D combinatorial background is modeled using wrong-sign data. This data sample contains no real-B contributions until at least two pion masses below the B mass, but it does contain real D decays combined with a random track. We plot the invariant mass of the wrong-sign candidates in the $D\pi$ mass assumption and sideband-subtract in the D mass. Then we perform a binned fit of this background-subtracted mass plot using a function of the form $(4\cdot 6)$. We will see in the control-sample fits that the rate of real-D combinatorial background is smaller than the fake-D. In the $\overline{B}^0 \to D^+X$ sample, only about 30% of the rate in the region above 5.4 GeV comes from real D^+.

4.9 Z Templates

We determine the Z templates for pions and kaons from prompt D^{*+} data. D^{*+} decays are a good source for kaon and pion samples: the flavor of the daughter tracks of the D^0 in the decay $D^{*+} \to D^0(K^-\pi^+)\pi^+$ is tagged by the charge of the pion from the D^{*+}; taken together with the large signal-to-background ratio of the $\Delta m = m(D^{*+}) - m(D^0)$ peak, this yields a very clean sample of pions and kaons.

The details of the procedure by which we extract kaon and pion Z templates from the prompt D^{*+} sample are described in Appendix A. In particular, we reweight the daughter tracks of the D^{*+} so that their momentum distribution matches the B daughter tracks'. Furthermore, we correct for the tracking chamber activity in the vicinity of the track; since merging of hits not belonging to the track in question can increase the measured dE/dx, it is necessary to reweight the templates to account for higher activity in the vicinity of the prompt-D^{*+} daughter tracks. Several variables

are useful proxies for the chamber activity; the variable we use is the track "secance", or number of r–ϕ intersections with other tracks between the COT inner and outer radii. We choose this variable because it results in the best agreement between templates derived from the prompt D^{*+} sample and a sample of D^{*+} decays from B mesons. One of the other plausible proxy variables, the total number of tracks in the event, is used to assess the systematic uncertainty due to the local chamber activity.

Because the B data samples contain semileptonic $B \to D\ell\overline{\nu}$ decays, we need to parameterize the dE/dx response for muons and electrons. For muons, we use the same templates as for pions. (Because these semileptonic components are small in the mass region of interest, systematic uncertainties associated with this choice are negligible.) For electrons, which are a small contribution in the mass region of interest, we generate a template from Monte Carlo using the predicted values of the electron dE/dx and its resolution [37].

The Z templates for the combinatorial backgrounds are taken directly from the data. Details vary between samples; a more detailed description can be found in the sample-specific Sections 5.2, 5.9 and 6.2. For the fake-D component, signal subtraction (described in Section 4.8) is performed on the Z templates. To do this, we determine the fraction of signal leakage into the sidebands from the auxiliary mass fit described in the previous section, and then determine what fraction of that leakage is e, μ, π and K from the $B \to DX$ Monte Carlo. The Z pdf of each particle species, with the appropriate normalization, is then subtracted from the D-sideband Z distribution to yield the fake-D Z pdf.

All Z pdf's are represented as histograms. Because the prompt D^{*+} sample is very large (yielding $2.8 \cdot 10^6$ tracks of each species after cuts), the statistical uncertainty on the Z histograms for kaons and pions is negligible. This is not the case for the combinatorial-background Z templates, which are limited by the statistics available in the data samples from which the templates are extracted; systematic uncertainties are assessed on the branching fraction measurements to account for the modeling of the background Z distributions.

4.10 Efficiency corrections

To calculate relative branching fractions from event yields, we proceed as follows. Let μ_1 and μ_2 be the normalizations of the fit components whose ratio of branching ratios $\mathcal{B}_1/\mathcal{B}_2$ we want to measure. The relationship between these quantities is

$$\frac{\mathcal{B}_1}{\mathcal{B}_2} = \frac{\mu_1/(\varepsilon_1\phi_1\eta_1)}{\mu_2/(\varepsilon_2\phi_2\eta_2)}, \tag{4·7}$$

where ε_i is the kinematic efficiency for a channel i, ϕ_i is the fraction of candidates in the fit mass range and η_i is the efficiency for the track species recoiling against the D in channel i to be reconstructed by the XFT track finding in the Level 1 trigger.

Monte Carlo can be used to determine ε_i and ϕ_i: the Monte Carlo yield y_i of a channel is related to the number N of candidates generated and the branching

fraction $\mathcal{B}'_i$ used in the generator by

$$y_i = N\mathcal{B}'_i\varepsilon_i\phi_i. \tag{4·8}$$

Therefore, the relationship between the fit fractions and the ratio of branching ratios is

$$\frac{\mathcal{B}_1}{\mathcal{B}_2} = \frac{\mu_1\mathcal{B}'_1/(y_1\eta_1)}{\mu_2\mathcal{B}'_2/(y_2\eta_2)} \tag{4·9}$$

as long as the two channels are created in the same MC generation.

The relative XFT efficiency η_K/η_π between K and π tracks was determined from a sample of prompt $D^+ \to K^-\pi^+\pi^+$ decays for an earlier analysis [38]. We use the XFT efficiency function from that analysis, which returns a weight $w(p_T|\text{species})$ depending on run-number, p_T and track species (K or π) as well as a $\pm 1\sigma$ uncertainty $\varsigma(p_T|\text{species})$, to calculate the following quantities:

$$\begin{aligned}
\langle w_{DK}\rangle &= \frac{1}{n(DK)}\sum_{i\in K} w(p_T^i|K)\\
\langle w_{DK}^2\rangle &= \frac{1}{n(DK)}\sum_{i\in K} w^2(p_T^i|K)\\
\langle w_{DK}^\pm\rangle &= \frac{1}{n(DK)}\sum_{i\in K} w(p_T^i|K) \pm \varsigma(p_T^i|K)\\
\varsigma_{DK}^\pm &= \langle w_{DK}^\pm\rangle - \langle w_{DK}\rangle\\
\sigma_{DK} &= \sqrt{\langle w_{DK}^2\rangle - \langle w_{DK}\rangle^2}/\sqrt{n(DK)}
\end{aligned} \tag{4·10}$$

and

$$\begin{aligned}
\langle w_{D\pi}\rangle &= \frac{1}{n(D\pi)}\sum_{i\in \pi} w(p_T^i|\pi)\\
\langle w_{D\pi}^2\rangle &= \frac{1}{n(D\pi)}\sum_{i\in \pi} w^2(p_T^i|\pi)\\
\sigma_{D\pi} &= \sqrt{\langle w_{D\pi}^2\rangle - \langle w_{D\pi}\rangle^2}/\sqrt{n(D\pi)}.
\end{aligned} \tag{4·11}$$

The summations run over candidates in the $B \to DX$ Monte-Carlo samples with a track of the indicated species recoiling against the D. The systematic error ς due to the XFT efficiency is calculated only for DK because ς_{DK} and $\varsigma_{D\pi}$ are completely correlated.

From the quantities calculated in (4·10) and (4·11), we can determine the relative XFT efficiency $\langle w_{D\pi}\rangle / \langle w_{D\pi}\rangle$ as well as the statistical uncertainty on the relative efficiency (by propagating σ_{DK} and $\sigma_{D\pi}$) and systematic uncertainty (which is simply $\varsigma_{DK}^\pm$; in practice, the uncertainty is symmetric). The statistical uncertainty can be made arbitrarily small with a sufficiently large sample; the systematic uncertainty is

of the same order of magnitude as the size of the correction itself.

The relative kinematic and XFT efficiencies for $B \to DK$ and $B \to D\pi$ vary slightly between samples; the corrections are smaller than 1% for kinematic efficiency and approximately 4% for the XFT efficiency.

4.11 Fitter validation

We validate the fitter by plotting the pull distributions of the fit parameters. Data is generated by throwing dice against the mass and dE/dx templates with the μ_j determined in the fit as the "true" μ_j, yielding a generated number of candidates that is Poisson-distributed about μ_j. The data is then fit. For each pseudoexperiment, we record the $\tilde{\mu}_j$ returned by the fit and the uncertainty $\sigma(\tilde{\mu}_j)$ returned by the fit. If the fit is unbiased, the average over pseudoexperiments $\langle(\tilde{\mu}_j - \mu_j)/\sigma(\tilde{\mu}_j)\rangle$ will be zero for all fit components; if the error estimate provides coverage, the distribution of $(\tilde{\mu}_j - \mu_j)/\sigma(\tilde{\mu}_j)$ will have unit width. The sample-specific chapters will show that our fits satisfy these conditions.

Before proceeding from the control samples to the signal sample, we will check the validity of our analysis method by comparing our control-sample results with the world average of existing measurements. We will show the agreement with the world averages in the control sample chapter.

4.12 Systematic uncertainties

All systematic uncertainties on $\mathcal{B}\big(B \to DK\big)/\mathcal{B}\big(B \to D\pi\big)$ are determined as follows: An alternative fit model is developed to account for the effect under study. A simple Monte Carlo simulation (using the method described in the previous section) is then performed in which candidates are generated with masses and Z distributed according to the fit model. This data is fit with the alternate fit configuration. The bias on $\mathcal{B}\big(B \to DK\big)/\mathcal{B}\big(B \to D\pi\big)$, averaged over 10000 pseudoexperiments, is taken to be the systematic uncertainty associated with the effect under study.

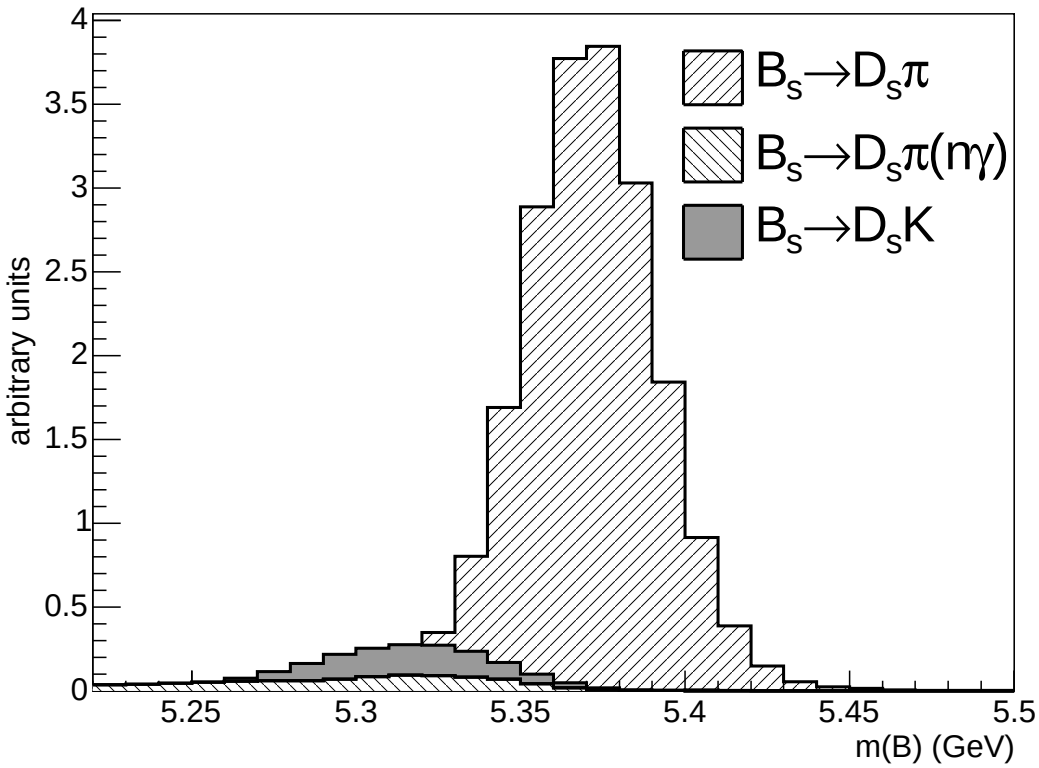

Figure 4.1: Cartoon of the mass distributions of $\overline{B}_s^0 \to D_s^+\pi^-$, the radiative tail of $\overline{B}_s^0 \to D_s^+\pi^-$ (labeled $\overline{B}_s^0 \to D_s^+\pi^-(n\gamma)$) and $\overline{B}_s^0 \to D_s^\pm K^\mp$. (A 10 MeV cutoff on the energy of the emitted photon is used in the modeling of the radiative tail.) It is clear that invariant mass alone cannot discriminate between $\overline{B}_s^0 \to D_s^+\pi^-$ and $\overline{B}_s^0 \to D_s^\pm K^\mp$.

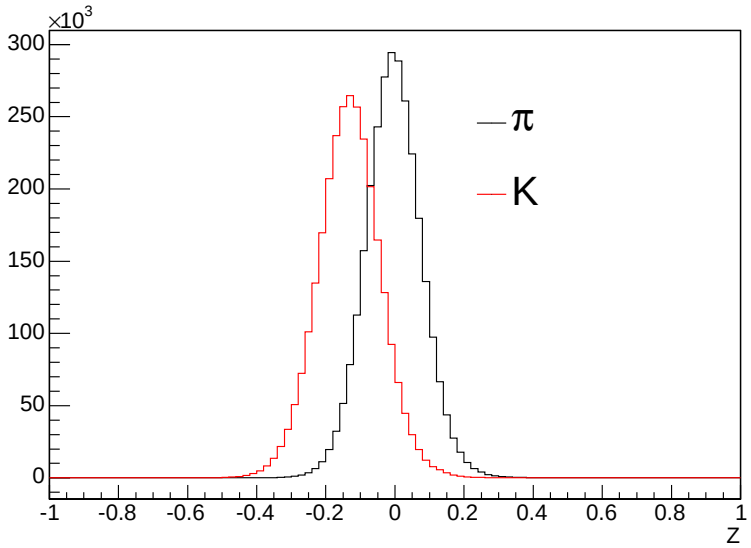

Figure 4.2: Z templates for K and π tracks

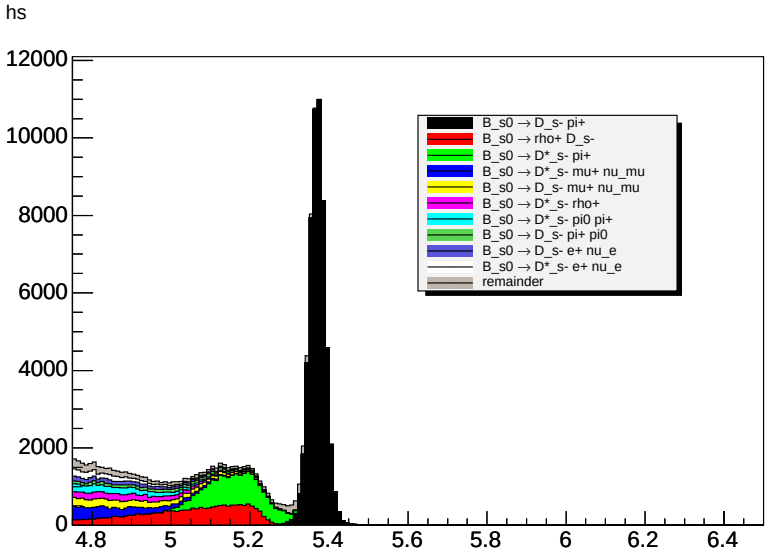

Figure 4.3: Composition of $\overline{B}_s^0/B_s^0 \to D_s^+(\phi\pi^+)X$ Monte Carlo sample. Some decay modes have mass shapes that are not easily parameterized.

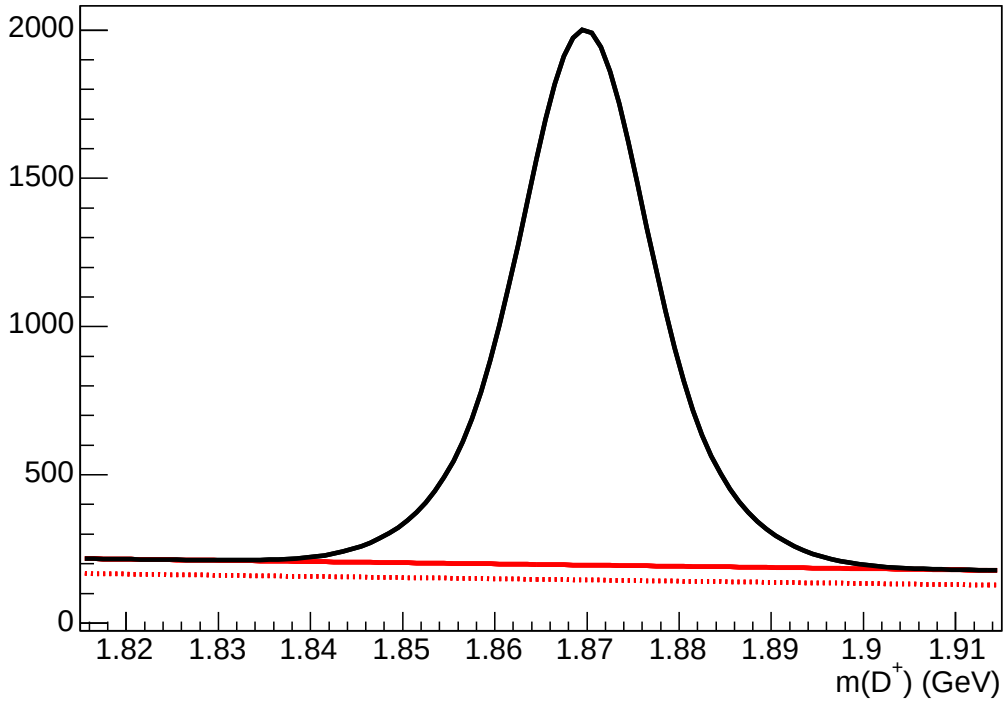

Figure 4.4: D mass cartoon to illustrate signal leakage. The real D signal (illustrated by a double gaussian) sits on top of not only "fake-D" combinatorial background (dashed line) but also poorly measured real D's, some of which come from real B's (solid straight line).

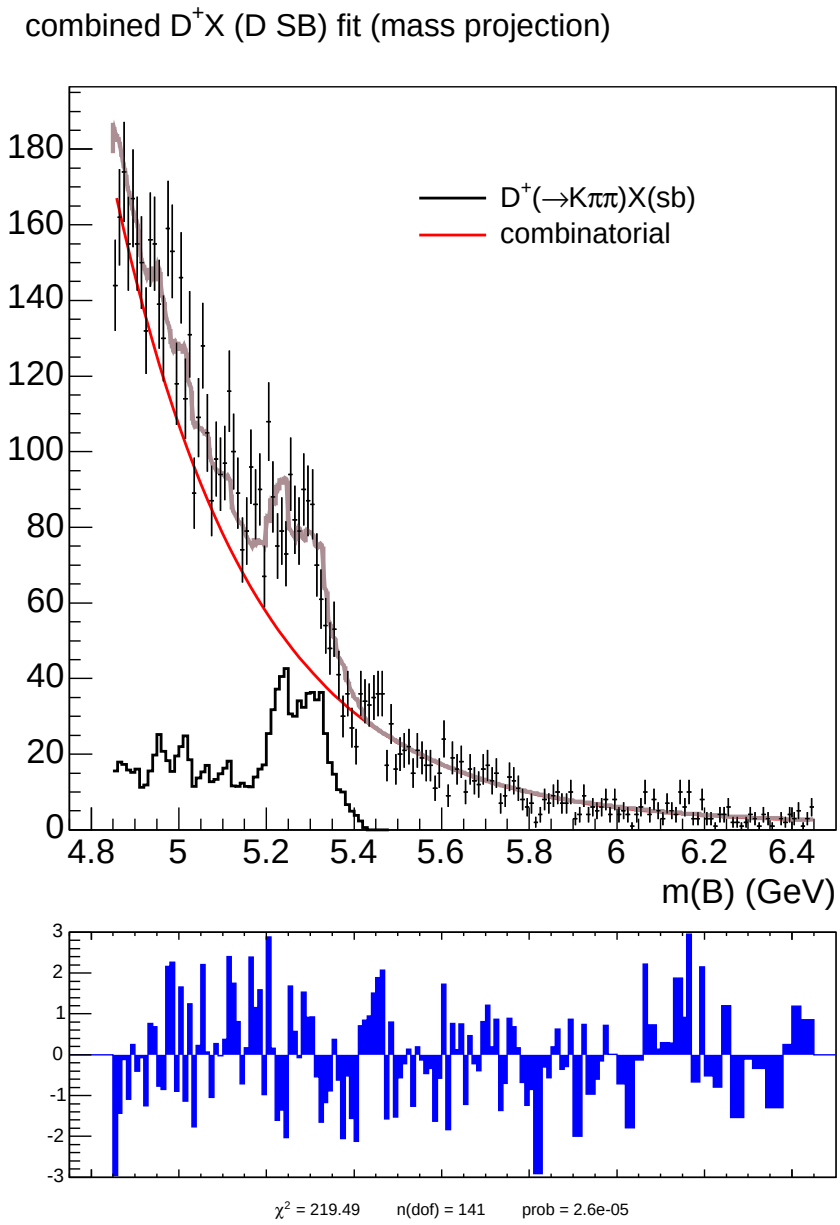

Figure 4.5: Illustration of the signal-subtraction procedure. The B mass distribution of the D candidates from the D sidebands is shown as data points. This distribution is fit with a smooth combinatorial background shape of the form $(4 \cdot 6)$ and a histogram template representing the leakage of real D's from real B's (sketched in Figure 4.4). The signal leakage is thus "subtracted" from the D sidebands.

Chapter 5

$\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ control samples

Having described the analysis method in the previous chapter, we now show the results of its application to the two control samples used in this analysis: $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$. These samples are reconstructed as $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$, with $D^+ \to K^-\pi^+\pi^+$ and $D^{*+} \to D^0(K^-\pi^+)\pi^+$. We begin by listing the fit details that are specific to these samples. Then we will show the fit results for $\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big)$ and $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big)$ and compare them to the world-average values. Because the control samples are large, we can break them up into three different data-taking periods and verify that the fit results in all periods are consistent; these periods, called `xbhd0d`, `xbhd0h` and `xbhd0i`, correspond to approximately 0.4 fb^{-1} each.

5.1 Selection of candidates

Our starting point is the skimmed datasets described in Section 4.4. A list of skimming cuts is found in Table 5.1. We then apply selection cuts to the $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ candidates that are optimized in accordance with the principles of Section 4.5 as follows.

For the $\overline{B}^0 \to D^+X$ selection, we start the optimization on a set of rectangular cuts used in early versions of the B_s mixing analysis [39]. To these cuts, we add the trigger-track requirements from Section 4.5 that allow us to use dE/dx information from the B daughter track. We then perform one-dimensional $S^2/(S+B)$ optimizations of the cut variables described in Section 4.5. The signal yield and background are estimated by selecting the B peak region (5.26 GeV $< m(B) <$ 5.4 GeV) and the D mass peak region (for the signal yield) or D mass sidebands (for the background).

The $\overline{B}^0 \to D^{*+}X$ cut optimization uses $\overline{B}^0 \to D^{*+}X$ Monte Carlo to determine the effects of the cuts on the $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^{*+}K^-$ yield. Combinatorial background is taken from the high sideband of the $\overline{B}^0 \to D^+X$ sample, which is statistically independent from the $\overline{B}^0 \to D^+X$ sample. The signal S is the $\overline{B}^0 \to D^{*+}K^-$ yield, whereas the background B includes $\overline{B}^0 \to D^{*+}\pi^-$, the remaining $\overline{B}^0 \to D^{*+}X$ modes, and the combinatorial background. Some cut variables are not reproduced by the single-B Monte Carlo (iso(B), $z_0(B)$, and $d_0(B)$ significance); we fix these cuts at values that keep most of the signal.

A complete set of analysis cuts for each decay mode is presented in Table 5.2. In addition to the variables described in Section 4.5, vetoes are applied to remove

	$\overline{B}^0 \to D^+\pi^-$	$\overline{B}^0 \to D^{*+}\pi^-$
$m(D^+/D^0)$	$1810 < m/\text{MeV} < 1920$	$1810 < m/\text{MeV} < 1920$
$\Delta\phi(D^+/D^0)$	< 1.5	< 1.5
$\Delta R(D^+/D^0)$	< 2	< 2
$\Delta z(D^+/D^0)$	< 5 cm	< 5 cm
$\sum p_T(D^+/D^0)$	> 2.4 GeV	> 2.4 GeV
$p_T(D^+/D^0)$	> 2.4 GeV	> 2.4 GeV
$L_{xy}(D^+/D^0)$	> 200 µm	> 200 µm
D^+/D^0 SVT confirmation	one track	one track
$\chi^2_{xy}(D^+/D^0)$	< 30	< 30
$m(B)$	$4650 < m/\text{MeV} < 6500$	$4650 < m/\text{MeV} < 6500$
$\Delta\phi(B)$	< 3	< 3
$\Delta R(B)$	< 2	< 2
$\Delta z(B)$	< 5 cm	< 5 cm
$p_T(\pi_B)$	> 1 GeV	> 1 GeV
$\Delta R(D^+/D^0, \pi_B)$	< 1.8	< 1.8
$\sum p_T(B)$	> 5.5 GeV	> 5.5 GeV
$p_T(B)$	> 5 GeV	> 5 GeV
$L_{xy}(B)$	> 300 µm	> 300 µm
$\lvert d_0(B)\rvert$	< 100 µm	< 100 µm
$m(D^{*+}) - m(D^0)$		< 200 MeV
B SVT confirmation	`SCEN_LOW`	`SCEN_LOW`
$\chi^2_{xy}(B)$	< 30	< 30

Table 5.1: Skimming cuts for $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ candidates

reflections of B or D hadrons. To veto the reflection of Λ_c^+ into the D^+ mass peak, we calculate $m(K, \pi \mapsto p, \pi)$, the mass of a D^+ candidate reconstructed as a Λ_c^+ by reassigning one of the pion tracks to a proton hypothesis. The veto applies if reassigning either of the two pions places the candidate in the Λ_c^+ mass window. We veto the $D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$ satellite decay reconstructed as $D^+ \to K^-\pi^+\pi^+$ by calculating $m(K\pi\pi) - m(K\pi)$. The veto applies if either combination of pions yields $\Delta m < 180$ MeV.

The mass distributions for the candidates with these selections are shown in Figures 5.1 and 5.2. The background-rejection power of the narrow Δm peak in $\overline{B}^0 \to D^{*+}X$ decays is striking.

5.2 $\overline{B}^0 \to D^+X$ fit configuration

The $\overline{B}^0 \to D^+X$ fit is an extended fit using the likelihood function $(4 \cdot 5)$. The $\overline{B}^0 \to D^+\pi^-$ candidate reconstruction is performed without a mass constraint on the

	$\overline{B}^0 \to D^+\pi^-$	$\overline{B}^0 \to D^{*+}\pi^-$
$\chi^2_{xy}(B)$	< 10	
$\mathrm{prob}(\chi^2_{xy}(B)/1.7, n(\mathrm{dof}))$		> 0.001
$\chi^2_{xy}(D)$	< 15	
$L_{xy}(B)/\sigma_{L_{xy}}(B)$	> 8	> 8
$L_{xy}(B \leftarrow D)$	> 0	
$L_{xy}(B \leftarrow D)/\sigma_{L_{xy}}(B \leftarrow D)$		> -2
$\lvert d_0(B)\rvert$	< 60 µm	
$\lvert d_0(B)/\sigma_{d_0}(B)\rvert$		< 3
$p_T(B)$	> 5.5 GeV	
$p_T(\pi_B)$	> 2 GeV	> 2 GeV
$\Delta R(D, \pi_B)$	< 1.5	
$\lvert z_0(B)/\sigma_{z_0}(B)\rvert$	< 3	< 3
$\mathrm{iso}(B)$ $(\Delta R < 1)$	> 0.5	> 0.5
$m(D^+)$	$\lvert m/\mathrm{MeV} - 1869.4\rvert < 20$	
$m(D^0)$		$\lvert m/\mathrm{MeV} - 1864.6\rvert < 20$
$m(D^{*+}) - m(D^0)$		$144 < \Delta m/\mathrm{MeV} < 147$
π_B SVT match distance	< 25	< 25
π_B SVT χ^2	< 15	< 15
π_B SVT p_T	> 2 GeV	> 2 GeV
$m(K, \pi \mapsto p, \pi)$	$\lvert m/\mathrm{MeV} - 2285\rvert > 250$	
$m(K\pi\pi) - m(K\pi)$	> 180 MeV	
Z	$\neq -999$	$\neq -999$
Good run list	B_s mixing version 14	B_s mixing version 14

Table 5.2: Selection cuts for $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ candidates. When cutting on χ^2_{xy} probability, we rescale χ^2_{xy} by 1/1.7, which yields a flat probability distribution.

D candidate. (This choice stems from our modeling of the fake-D background; see below.)

Table 5.3 shows the fit components. Apart from the $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^+K^-$ signals, this list includes a number of single-B backgrounds. Most of the partially reconstructed modes are collected in the "remainder" $\overline{B}^0/B^- \to D^+X$ template. $D^+\rho^-$ and $D^{*+}\pi^-$ are separated from the remainder into their own templates because these components leak into the D^+K^- mass region. The branching ratio $\mathcal{B}\big(B^- \to D_2^*\pi^-\big)$ used in our Monte Carlo generation was estimated based on an upper limit. The branching ratio has since been measured [40]. The $D^{(*)}K^{(*)}$ modes are the largest group of modes in which the track recoiling against the D^+ is a kaon; we split this group from the remainder template so we can give it a kaon dE/dx template. Likewise, we split out the $D^+e^-\overline{\nu}_e$ decays and give them an electron dE/dx template. The $D^{(*)}K^{(*)}$ and $D^+e^-\overline{\nu}_e$ components that are split out of the remainder template do not have floating normalizations but are instead fixed to the remainder template with ratios determined from the Monte Carlo. Lastly, a $B_s \to D_s^-(\phi\pi^-)X$ template is also included in the fit. The "parameter" column of Table 5.3 summarizes the connections between the various templates. Also included in the fit are two combinatorial background components, fake-D combinatorial background and real-D combinatorial background.

The provenance of the mass and Z pdf for each fit component is listed in Table 5.3. The pdfs are determined as follows:

Single-B mass templates Table 5.4 lists the single-B Monte Carlo samples from which the mass templates are generated. These Monte Carlo samples are sorted into templates by the true Monte Carlo decay mode according the list in the previous paragraph. The mass distribution of the B decay modes contributing to these samples is shown in Figures 5.3–5.6.

Fake-D mass template The fake-D mass template is obtained from the D^+ sidebands, using the signal-subtraction procedure from Section 4.8. The invariant mass under the $D^+\pi^-$ mass assumption is calculated in the sidebands of the D^+ mass: 1.825 GeV $< m(D^+) <$ 1.84 GeV or 1.9 GeV $< m(D^+) <$ 1.915 GeV. Signal subtraction is performed on this data. To do this, the data is fit with two components, the first of which is a function of the form $(4 \cdot 6)$ and the second of which is a histogram obtained from the $B^0 \to D^-X$ Monte Carlo selecting on the D^+ sideband. Figures 5.14–5.17 show the results of this fit. The shape parameters thus obtained for the background mass pdf are fixed in the $\overline{B}^0 \to D^+X$ fit. Because the fake-D fit gives better agreement with data when performed on candidates that are not mass-constrained, we use unconstrained candidates for all fit components.

Real-D mass template The real-D mass template is obtained from a binned fit of the $D^+\pi$ mass of wrong-sign candidates selected with the same cuts as the right-sign candidates but sideband-subtracted in the D^+ mass. The fit function is of the form $(4 \cdot 6)$. These fits are shown in Figures 5.18–5.21. The shape

parameters obtained from this auxiliary fit are fixed in the $\overline{B}^0 \to D^+X$ fit; the normalization is floating.

Single-B Z templates The Z templates marked "prompt D^*" are described in Section 4.9. Two-dimensional reweighting in secance (s) and momentum (p) is applied. The DeX template uses the parametric dE/dx simulation described in Section 4.9. The D_sX template is composed from e, μ (using π as proxy), π and K templates according to species fractions determined from the $B_s \to D_sX$ Monte Carlo sample. Because the species composition of the D_sX contribution varies by B mass, the D_sX Z template is divided into four mass regions: the region below the D^+K^- (4.85 GeV $< m(B) <$ 5.17 GeV); the D^+K^- region (5.17 GeV $< m(B) <$ 5.26 GeV); the $D^+\pi^-$ region (5.26 GeV $< m(B) <$ 5.4 GeV); and the region above the $D^+\pi^-$ (5.4 GeV $< m(B) <$ 6.45 GeV).

Fake-D Z template The Z template for the fake-D component is taken directly from the D sideband data. Signal subtraction is performed on the fake-D template to account for the leakage of poorly measured real D into the D sidebands. We divide the fake-D Z template into two B-mass regions (4.85 GeV $< m(B) <$ 5.4 GeV and 5.4 GeV $< m(B) <$ 6.45 GeV) to account for its B-mass dependence.

Real-D Z template The Z template for the real-D component is taken from the wrong-sign D data, selecting on the D^+ peak. Because of limited statistics, and because the Z distribution in the D^+ sideband exhibits no large difference from the Z distribution in the D^+ peak, sideband subtraction is not performed on the real-D Z template. We divide the real-D Z template into the same mass regions as the fake-D Z template.

5.3 $\overline{B}^0 \to D^+X$ results

Results of the extended likelihood fit are given in Table 5.5. Results are shown for each of the three data-taking periods and for the combined dataset. The correlation coefficients for the combined dataset are listed in Table 5.6. Figure 5.42 shows the likelihood profiles in all fit parameters. Mass projections are shown in Figures 5.22–5.25. Z projections are shown in Figures 5.30–5.33. We divide the Z projections into four mass regions: the region below the D^+K^- (4.85 GeV $< m(B) <$ 5.17 GeV); the D^+K^- region (5.17 GeV $< m(B) <$ 5.26 GeV); the $D^+\pi^-$ region (5.26 GeV $< m(B) <$ 5.4 GeV); and the region above the $D^+\pi^-$ (5.4 GeV $< m(B) <$ 6.45 GeV). This is done so that the modeling of the combinatorial background Z distribution and the modeling of the pion Z distribution can be tested in the mass regions where these components dominate; we isolate the D^+K^- region to verify that the Z fit is good in the mass region that dominates the final analysis result.

To calculate $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$ from the event yields returned by the fit, we apply the corrections discussed in Section 4.10. We determine a kinematic

Decay modes	Mass pdf	dE/dx pdf	fit parameters
		$\overline{B}^0 \to D^+\pi^-$ fit	
$\overline{B}^0 \to D^+\pi^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$\overline{B}^0 \to D^+\pi^-(n\gamma)$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$\overline{B}^0 \to D^+K^-$	MC	prompt $D^* \to K$, $s:p$ reweight	norm floating
$\overline{B}^0 \to D^+\rho^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$\overline{B}^0 \to D^{*+}\pi^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$D_2^*\pi^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm fixed to $\overline{B}^0/B^- \to D^+X$
$D^{(*)}K^{(*)}$	MC	prompt $D^* \to K$, $s:p$ reweight	norm fixed to $\overline{B}^0/B^- \to D^+X$
$D^{(*)}eX$	MC	parametric e simulation	norm fixed to $\overline{B}^0/B^- \to D^+X$
$\overline{B}^0/B^- \to D^+X$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$\overline{B}_s^0 \to D_s^+X$	MC	species fractions from MC	norm floating
D^+(sb) (fake-D bg)	D SB	D SB, signal-subtracted	norm floating, shape fixed
D^+X^+ (real-D bg)	WS (SBS)	WS (D signal region)	norm floating, shape fixed
		$\overline{B}^0 \to D^{*+}\pi^-$ fit	
$D^{*+}\pi^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$D^{*+}K^-$	MC	prompt $D^* \to K$, $s:p$ reweight	norm floating
$D^{*+}\rho^-$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$D^{*+}\pi^-(n\gamma)$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
remainder ($D^{*+}X$)	MC	parametric simulation	norm floating
combinatorial bg	D^+ SB	D^+ SB	norm $= 1 - \sum_i f_i$, shape fixed

Table 5.3: Fit components for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ fits. See the text for a description of the "mass pdf" and "dE/dx pdf" categories.

Channels		Sample size
$\overline{B}^0$	$\to D^+\pi^-(n\gamma)$	50k
$\overline{B}^0/B^0$	$\to D^+X$	200k
B^-/B^+	$\to D^+X$	10k
B_s	$\to D_s^-(\phi\pi^-)X$	5k

Table 5.4: Monte Carlo samples used in the $\overline{B}^0 \to D^+X$ fit

efficiency correction $(\varepsilon_K\phi_K)/(\varepsilon_\pi\phi_\pi) = 0.997$ and an XFT correction $\eta_\pi/\eta_K = 1.0499\pm 0.0281$. After corrections, we find

$$\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big) = 0.092 \pm 0.005(\text{stat}).$$

The denominator contains both the $\overline{B}^0 \to D^+\pi^-$ main peak and the radiative tail, whose size we measure to be 5.5%±0.9% relative to the main peak; the statistical error on $\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big)$ takes into account the correlations between the candidate yields. We will estimate systematic uncertainties in Section 5.7 and compare $\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big)$ to the world average in Section 5.8.

5.4 $\overline{B}^0 \to D^+X$ validation

We validate the fitter by plotting the pull distributions of the fit parameters. Data is generated by throwing dice against the mass and dE/dx templates with the μ_j determined in the fit as the "true" μ_j, yielding a generated number of candidates that is Poisson-distributed about μ_j. The data is then fit. Figure 5.43 shows that the fit provides unbiased results and that the error estimate is correct. Table 5.7 lists the pull means and widths for 10000 pseudoexperiments of 46498 events (the size of our dataset) each.

5.5 $\overline{B}^0 \to D^+X$ residual discrepancies

One large and unexpected discrepancy between the data and the fit model remains. Figure 5.46, the Z residual plot in the mass region dominated by $D^+\pi^-$ (5.26 GeV $< m(B) <$ 5.4 GeV), shows an excess between $Z = -0.2$ and $Z = -0.1$. To investigate this discrepancy, we first perform a "kaon projection" of the mass plot, where each candidate is given a weight

$$w(Z) = \frac{Z - \langle Z(\pi)\rangle}{\langle Z(K)\rangle - \langle Z(\pi)\rangle} \tag{5·1}$$

according to the Z of the track recoiling against the D^+. Z is measured; $\langle Z(\pi)\rangle$ and $\langle Z(K)\rangle$ are the expected Z in the pion hypothesis for pions and kaons (averaged over

Decay modes	par no	Fit result			
		`xbhd0d`	`xbhd0h`	`xbhd0i`	comb
		$\overline{B}^0 \to D^+\pi^-$ fit			
$\overline{B}^0 \to D^+\pi^-$	2	4130.0 ± 81.2	5085.5 ± 89.6	4128.2 ± 80.8	13336.0 ± 145.4
$\overline{B}^0 \to D^+\pi^-(n\gamma)$	3	256.5 ± 67.0	231.5 ± 72.7	287.7 ± 68.3	734.5 ± 119.8
$\overline{B}^0 \to D^+K^-$	4	334.5 ± 31.0	526.4 ± 36.5	361.6 ± 31.8	1230.0 ± 57.6
$\overline{B}^0 \to D^+\rho^-$	5	2314.4 ± 290.2	4083.6 ± 321.6	2680.7 ± 286.9	8799.9 ± 519.2
$\overline{B}^0 \to D^{*+}\pi^-$	6	1406.3 ± 137.9	1302.4 ± 153.8	1331.0 ± 138.0	4118.4 ± 248.5
$D_2^*\pi^-$	7	fixed to $\overline{B}^0/B^- \to D^+X$			
$D^{(*)}K^{(*)}$	8	fixed to $\overline{B}^0/B^- \to D^+X$			
$D^{(*)}eX$	9	fixed to $\overline{B}^0/B^- \to D^+X$			
$\overline{B}^0/B^- \to D^+X$	10	2471.1 ± 165.6	2434.4 ± 183.1	2300.9 ± 165.3	7262.7 ± 296.6
$\overline{B}^0_s \to D^+_sX$	11	715.9 ± 179.7	808.4 ± 194.5	644.8 ± 176.8	2156.9 ± 318.6
D^+(sb) (fake-D bg)	12	2376.1 ± 168.5	2855.9 ± 176.1	2105.2 ± 236.6	7361.2 ± 321.4
D^+X^+ (real-D bg)	13	333.3 ± 70.4	497.1 ± 79.4	494.9 ± 150.8	1498.5 ± 180.1
		$\overline{B}^0 \to D^{*+}\pi^-$ fit			
$D^{*+}\pi^-$	2	0.342 ± 0.009	0.367 ± 0.008	0.367 ± 0.009	0.359 ± 0.005
$D^{*+}K^-$	3	0.035 ± 0.005	0.028 ± 0.004	0.032 ± 0.004	0.032 ± 0.003
$D^{*+}\rho^-$	4	0.299 ± 0.021	0.287 ± 0.018	0.283 ± 0.020	0.287 ± 0.011
$D^{*+}\pi^-(n\gamma)$	5	0.030 ± 0.009	0.024 ± 0.008	0.025 ± 0.009	0.026 ± 0.005
remainder ($D^{*+}X$)	6	0.250 ± 0.020	0.255 ± 0.017	0.260 ± 0.019	0.257 ± 0.011
combinatorial bg		norm $= 1 - \sum_i f_i$			

Table 5.5: Fit parameters determined for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ by the combined mass and Z fit. The parameters are candidate yields in the $\overline{B}^0 \to D^+X$ case and fractions in the $\overline{B}^0 \to D^{*+}X$ case.

Table 5.6: Correlation coefficients for $\overline{B}^0 \to D^+ X$ (combined 1.2 fb^{-1} dataset)

Parameter		Correlation coefficients								
	global	2	3	4	5	6	10	11	12	13
2	0.52875	1.000	0.008	−0.014	0.157	0.015	0.025	−0.491	0.045	−0.024
3	0.59392	0.008	1.000	−0.341	−0.012	0.013	0.095	−0.254	−0.180	0.101
4	0.41565	−0.014	−0.341	1.000	0.051	−0.003	0.005	−0.016	−0.071	0.039
5	0.95049	0.157	−0.012	0.051	1.000	−0.823	−0.768	−0.253	−0.054	0.022
6	0.88356	0.015	0.013	−0.003	−0.823	1.000	0.659	−0.046	0.039	−0.009
10	0.87800	0.025	0.095	0.005	−0.768	0.659	1.000	−0.037	−0.152	0.046
11	0.83216	−0.491	−0.254	−0.016	−0.253	−0.046	−0.037	1.000	−0.244	0.129
12	0.87826	0.045	−0.180	−0.071	−0.054	0.039	−0.152	−0.244	1.000	−0.767
13	0.80764	−0.024	0.101	0.039	0.022	−0.009	0.046	0.129	−0.767	1.000

Channels	Pull mean (%)	Pull width (%)
$B^0 \to D^-\pi^+$	-2.194 ± 1.006	99.631 ± 0.729
$B^0 \to D^-\pi^+(n\gamma)$	-0.726 ± 1.001	99.500 ± 0.716
$B^0 \to D^-K^+$	-1.450 ± 1.004	99.887 ± 0.713
$B^0 \to D^-\rho^+$	-1.406 ± 1.027	102.151 ± 0.730
$B^0 \to D^{*-}\pi^+$	1.101 ± 1.026	101.983 ± 0.734
$B^0/B^+ \to D^-X$	0.825 ± 1.024	101.908 ± 0.724
$B_s \to D_sX$	0.753 ± 1.012	100.605 ± 0.721
D^-(sb)	-0.847 ± 1.022	101.772 ± 0.725
D^-X^-	-1.340 ± 1.025	101.809 ± 0.734

Table 5.7: Pull distributions for $\overline{B}^0 \to D^+X$

the momenta in the data sample). The weight $w(Z)$ is 0 on average for pions and 1 for kaons. Figure 5.47 shows the kaon projection of the dataset overlaid with the mass pdf of the $\overline{B}^0 \to D^+K^-$ contribution (scaled to the yield determined by the fit). We find an excess of "kaons" (tracks whose Z is smaller than the expected pion Z) in the data centered on the $\overline{B}^0 \to D^+\pi^-$ mass, shown in Figure 5.48. We have considered and ruled out a number of potential sources for this excess. The following are effects we have considered which could add a kaon or proton background in the region in question:

- $\Lambda_b^0 \to D^-p$ decays have too low a yield.
- $\Lambda_b^0 \to \Lambda_c^+(p\pi^+\pi^-)K^-$, which is not covered by the $\Lambda_c^+ \to pK^-\pi^+$ veto, does not contribute appreciably in the D^+ mass window.
- $\Lambda_b^0 \to (pK^-\pi^+)K^-$ (non-resonant) is too small a contribution.
- Neither the $D_s^+K^-$ nor the $\Lambda_c^+K^-$ reflection has the right mass.
- Finally, we have looked for $\overline{B}^0 \to D^+\pi^-$ candidates in Monte Carlo where the K_D^- and the π_B^- are swapped and still pass the D^+ mass cut; this effect occurs with negligible frequency ($< 10/50000$ MC events).

We also note that both protons and kaons have a wider Z distribution than the bump we observe.

We have also considered effects which could deform the pion Z template with respect to the pions in the $\overline{B}^0 \to D^+X$ sample. We have checked for a secance dependence on mass, which we did not find. A comparison of the track parameters of tracks selected from the appropriate mass and Z region shows no difference with tracks from the appropriate mass region with no Z requirement.

A useful clue in this mystery is the absence of a discrepancy in the $D^{*+}\pi^-$ fit. Is the difference between D^+ and D^{*+} the absence of combinatorial background or the

absence of statistics? To answer this question, add the bump observed in $D^+\pi^-$ to the pion Z template used in the $D^{*+}\pi^-$ fit. The χ^2 of the Z projection of the fit in the 5.26 GeV $< m(B) <$ 5.4 GeV region increases from 12.6 before the template modification to 24.0 after. The corresponding probabilities are 99% and 29%. Since $\Delta\chi^2$ is large but the χ^2 probability is still adequate, this study is inconclusive.

Since the present work was performed, a parallel analysis has shown the same effect in $B^- \to D^0(K^-\pi^+)X$, a sample with even higher statistics than $\overline{B}^0 \to D^+X$ [41]. It is unlikely that the three-track $B^- \to D^0(K^-\pi^+)X$ contains the same unaccounted-for background mode as $\overline{B}^0 \to D^+X$. A far more likely scenario is that the environment of the B daughter track is different from the environment of the D^{*+} daughter tracks in even more ways than we have accounted for in Section 4.9. The position of the excess in the kaon projection of the $\overline{B}^0 \to D^+X$ data, which coincides with the $\overline{B}^0 \to D^+\pi^-$ mass, supports this hypothesis. In the next section, we therefore treat the bump as a discrepancy between pions from $\overline{B}^0 \to D^+\pi^-$ and the D^{*+} daughter pions for the purpose of estimating the systematic uncertainty due to the bump. We do not change the central value of relative branching fraction.

5.6 $\overline{B}^0 \to D^+X$ Monte Carlo–data agreement

The reconstructed D and B masses in the Monte Carlo samples differ slightly from those in data. By comparing the $\overline{B}^0 \to D^{*+}\pi^-$ peak position in Monte Carlo and data (with tight selection cuts to reduce the effect of background as much as possible), we find an overall relative scale factor of 1.00025 ± 0.00012. We apply this scale factor to all Monte Carlo mass distributions. In addition, the mass resolution in the Monte Carlo is slightly better than in the data. Looking at the clean $\overline{B}^0 \to D^{*+}\pi^-$ sample, we find that Monte Carlo mass distributions agree with data if the Monte Carlo data is convoluted with a gaussian resolution function of width 6.43 ± 2.09 MeV. (The resolution in data is 17 MeV for mass-constrained B candidates, 20 MeV for non-constrained candidates.) In Section 5.7 we assign a systematic uncertainty on the measured branching ratios by varying the smearing and scaling within their errors.

Thanks to the sharp peak of the fully reconstructed $\overline{B}^0 \to D^+\pi^-$ and the kinematic cutoff ledge of the $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^+\rho^-$, the mass fit itself should be able to determine both the scale factor and the smearing well. In practice, however, it is not possible to let the scale factor and smearing width float; because the mass pdf's are represented as histograms, the likelihood as a function of scale factor has a discrete step whenever an event crosses a histogram bin boundary, thwarting `MINUIT`'s gradient-based minimum search. Nevertheless, we can use the fitter to cross-check the externally supplied scale factor and smearing by scanning the likelihood; at each scan point, the scale factor and smearing are fixed parameters, so the fit is unproblematic.

Figure 5.38 shows the scale factor scan, Figure 5.39 the smearing scan. The scan measures a scale factor of 1.00030 ± 0.00007 and a smearing width of 7.24 ± 0.67 MeV. Both values are comfortably within the errors on the default scale factor and smearing.

5.7 $\overline{B}^0 \to D^+X$ systematic uncertainties

In this section we describe the various studies performed to determine the systematic uncertainties on $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$. Even though the $\overline{B}^0 \to D^+X$ sample is only a control sample, we would like to reassure ourselves before we proceed to the signal sample that no large systematic uncertainties are lurking in the shadows.

All systematic uncertainties are determined as follows. An alternative fit model is developed to account for the effect under study. Toy Monte Carlo data is then generated by sampling the mass and Z templates in the default fit configuration. This data is fit with the alternate fit configuration. The bias on $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$, averaged over 10000 pseudoexperiments, is used as the systematic uncertainty associated with the effect under study. Table 5.8 summarizes the systematic uncertainties.

Combinatorial background mass shape Our knowledge of the combinatorial background mass shape is imperfect. We assess the systematic uncertainty associated with the background shape by comparing different background models. The default model uses two parameterized background components whose shapes are determined by fits to the two background samples (fake-D + track from the right-sign D sidebands and real-D + track from the wrong-sign D peak), letting the normalizations of both components float. The alternate model fixes the normalization of the fake-D background to the value determined in a D^+ mass fit (shown in Figure 5.49) and lets the shape parameters of the wrong-sign function float instead.

Monte Carlo smearing and scaling As explained in Section 5.6, it is necessary to scale the Monte Carlo templates by a factor of 1.00025 ± 0.00012 and smear them with gaussian resolution function of width 6.43 ± 2.09 MeV to achieve agreement with the data. We assess the systematic uncertainty associated with the errors on the scale factor by varying the scale factor up and down within its error. We use the greater of the changes in fit value resulting from an upward or downward variation as the systematic uncertainty due to the scale factor. The same method is employed to determine the systematic uncertainty due to the smearing.

D_sX and fake-D species composition Two other sources of systematic uncertainty are a result of the species composition of the D_sX Z template and the fake-D Z template. We assess the systematic uncertainty due to the K content of the D_sX Z template by reducing the K content by 50% with respect to the content predicted by Monte Carlo. This covers the expected range of possible deviations of the $B_s \to D_sK$ branching fraction from the value in the generator table. To assess the systematic uncertainty due to the composition of the fake-D Z template, we use a fake-D Z template on which no signal subtraction has been performed.

Effect	$D^+K^-/D^+\pi^-$ uncertainty (%)
Residual discrepancy in π Z template	0.45
Z reweighting in $N_{\text{trks}} : p$	0.31
Signal subtraction in fake-D Z template	0.30
Combinatorial bg model	0.17
MC mass smearing	0.16
Composition of D_sX Z template	0.16
MC mass scale factor	0.15
Sum in quadrature	0.70

Table 5.8: Systematic uncertainties on $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$

Choice of dE/dx reweighting variable As described in Section 4.9, the dE/dx templates need to be corrected for the chamber activity in the vicinity of the B daughter track. Appendix A describes in detail how we choose a proxy variable (secance) for the chamber activity and perform a secance-based reweighting. Secance is only one of several variables, all described in the appendix, that can be used for the reweighting. A variable which behaves almost as well is the total number of tracks N_{trks} in the event; this is also the most "global" proxy conceivable for the chamber activity. We assess a systematic uncertainty by fitting with an alternate model in which dE/dx reweighting in N_{trks} is used.

Z residuals in the $D^+\pi^-$-dominated region Even after reweighting to account for the local chamber activity, residual discrepancies between the template D^{*+}-daughter and physics B-daughter Z distributions persist, as we discuss in Section 5.5 and show in Figure 5.46. We quantify the effect of the discrepancy between data and fit Z distributions in the $D^+\pi^-$-dominated region as follows. We add the difference between data and fit in the region 5.26 GeV $< m(B) <$ 5.4 GeV to the Z templates for $D^+\pi^-$ and $D^+\pi^-(n\gamma)$ and use these modified templates as an alternate fit model.

5.8 $\overline{B}^0 \to D^+X$ comparison with world average

Table 5.9 shows our measurement of relative branching fractions for several modes after all corrections described in Section 4.10 have been applied. The agreement between our measurement and the world average is satisfactory not only in $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$ but also in the leading partially reconstructed modes ($D^+\rho^-$ and $D^{*+}\pi^-$), where we are relying purely on the mass shape.

Sample	$D^+K^-/D^+\pi^-$ (%)		$D^+\rho^-/D^+\pi^-$		$D^{*+}\pi^-/D^+\pi^-$	
PDG	6.80 ± 1.66		2.21 ± 0.68		0.81 ± 0.22	
`xbhd0d`	$8.03 \pm 0.81 \pm 0.22$	(0.66σ)	2.10 ± 0.26	(-0.15σ)	1.21 ± 0.13	(1.56σ)
`xbhd0h`	$10.43 \pm 0.81 \pm 0.28$	(1.95σ)	3.05 ± 0.25	(1.17σ)	0.93 ± 0.11	(0.46σ)
`xbhd0i`	$8.62 \pm 0.82 \pm 0.23$	(0.98σ)	2.41 ± 0.26	(0.29σ)	1.14 ± 0.12	(1.29σ)
comb	$9.21 \pm 0.47 \pm 0.25$	(1.38σ)	2.49 ± 0.15	(0.40σ)	1.11 ± 0.07	(1.26σ)
	$D^{*+}K^-/D^{*+}\pi^-$ (%)		$D^{*+}\rho^-/D^{*+}\pi^-$			
PDG	7.76 ± 0.45		2.46 ± 0.38			
`xbhd0d`	$9.59 \pm 1.49 \pm 0.26$	(1.16σ)	3.37 ± 0.31	(1.86σ)		
`xbhd0h`	$7.39 \pm 1.11 \pm 0.20$	(-0.31σ)	3.08 ± 0.25	(1.37σ)		
`xbhd0i`	$8.39 \pm 1.27 \pm 0.22$	(0.46σ)	3.02 ± 0.27	(1.21σ)		
comb	$8.39 \pm 0.73 \pm 0.22$	(0.71σ)	3.13 ± 0.16	(1.62σ)		

Table 5.9: Comparison with PDG for $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ fits. The results are corrected for kinematic and XFT efficiency. For the $DK/D\pi$ ratio, the first error is statistical and the second is due to the XFT efficiency correction. The numbers in parentheses indicate the deviation from the PDG value.

Channels		Sample size
$\overline{B}^0$	$\to D^{*+}\pi^-(n\gamma)$	50k
$\overline{B}^0/B^0$	$\to D^0\pi^+X$	50k
B^-/B^+	$\to D^0\pi^+X$	7k

Table 5.10: Monte Carlo samples used in the $\overline{B}^0 \to D^{*+}X$ fit

5.9 $\overline{B}^0 \to D^{*+}X$ fit configuration

The $\overline{B}^0 \to D^{*+}X$ fit is a non-extended fit using the likelihood function $(4 \cdot 3)$. There is very little combinatorial background in this fit; this frees us to use candidates reconstructed with a mass constraint on the D^0 candidate, unlike in the $\overline{B}^0 \to D^+X$ fit.

Table 5.3 lists the fit components. In addition to the $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^{*+}K^-$ signals, we include a $D^{*+}\rho^-$ component because of its overlap with the $D^{*+}K^-$. All remaining single-B modes are collected in a "remainder" template. No sizeable reflections occur in the D^{*+} sample.

Table 5.10 lists the single-B Monte Carlo samples from which the mass templates are generated. The mass distribution of the B decay modes contributing to these samples is shown in Figures 5.7–5.9.

Because the combinatorial background under the D^{*+} is small, only one combinatorial background template is needed. We cannot extract the background shape parameters from the data for lack of a sizeable background sample; we instead use the shape parameters of the fake-D background in the D^+ fit. These parameters are fixed in the $\overline{B}^0 \to D^{*+}\pi^-$ fit. Since the background is small, the fit results are insensitive to uncertainty on the background parameterization.

The Z templates used in this fit are generated in the same way as for the $\overline{B}^0 \to D^+X$ fit, with the exceptions that the species composition of the remainder template is accounted for through a parametric simulation of the type described in Section 4.9 and that the Z pdf of the combinatorial background, like the mass pdf, is taken from the D^+ sample.

5.10 $\overline{B}^0 \to D^{*+}X$ Monte Carlo–data agreement

As mentioned in Section 5.6, a relative scale factor of 1.00025 ± 0.00012 and a gaussian resolution function of width 6.43 ± 2.09 MeV are required to make the data and Monte Carlo B^0 mass peak agree. We crosscheck these numbers by scanning the likelihood using the $\overline{B}^0 \to D^{*+}X$ fit in the same way as we have done for the $\overline{B}^0 \to D^+X$ fit earlier. Figure 5.40 shows the scale factor scan, Figure 5.41 the smearing scan. The scan measures a scale factor of 1.00030 ± 0.00004 and a smearing width of 6.00 ± 0.75 MeV. Both values are comfortably within the errors on the default scale factor and

Parameter		Correlation coefficients				
	global	2	3	4	5	6
2	0.72894	1.000	−0.103	−0.134	0.204	−0.173
3	0.55815	−0.103	1.000	0.038	0.400	−0.035
4	0.92648	−0.134	0.038	1.000	0.199	−0.829
5	0.72548	−0.204	−0.400	−0.199	1.000	0.002
6	0.92304	−0.173	−0.035	−0.829	0.002	1.000

Table 5.11: Correlation coefficients for $\overline{B}^0 \to D^{*+}X$ (combined 1.2 fb^{-1} dataset)

smearing.

5.11 $\overline{B}^0 \to D^{*+}X$ results

Results of the likelihood fit are given in Table 5.5. Results are shown for each of the three data-taking periods and for the combined dataset. The correlation coefficients for the combined dataset are listed in Table 5.11. Figure 5.44 shows the likelihood profiles in all fit parameters. Mass projections are shown in Figures 5.26–5.29. Z projections are shown in Figures 5.34–5.37. We divide the Z projections into the same four mass regions as in the $\overline{B}^0 \to D^+X$ fit: the region below the $D^{*+}K^-$ (4.85 GeV $< m(B) <$ 5.17 GeV); the $D^{*+}K^-$ region (5.17 GeV $< m(B) <$ 5.26 GeV); the $D^{*+}\pi^-$ region (5.26 GeV $< m(B) <$ 5.4 GeV); and the region above the $D^{*+}\pi^-$ (5.4 GeV $< m(B) <$ 6.45 GeV).

Corrections on $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big)$ are determined in the same way as for the $\overline{B}^0 \to D^+X$ fit. We determine a kinematic efficiency correction $(\varepsilon_K\phi_K)/(\varepsilon_\pi\phi_\pi) = 1.023$ and an XFT correction $\eta_\pi/\eta_K = 1.0499 \pm 0.0281$. After corrections, we find

$$\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big) = 0.084 \pm 0.008(\text{stat}).$$

The denominator contains both the $\overline{B}^0 \to D^{*+}\pi^-$ main peak and the radiative tail, whose size we measure to be 7.3%±1.4% relative to the main peak; the statistical error on $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big)$ takes into account the correlations between the candidate yields. We compare this result to the world average in Section 5.13. (We do not estimate systematic uncertainties in this sample; since the control-sample measurements are intended as a crosscheck of our analysis method, repeating the exercise performed in the $\overline{B}^0 \to D^+X$ case has little benefit.)

Channels	Pull mean (%)	Pull width (%)
$D^{*+}\pi^-$	0.121 ± 1.015	100.638 ± 0.742
$D^{*+}K^-$	-5.114 ± 1.003	100.187 ± 0.713
$D^{*+}\rho^-$	-0.704 ± 1.005	100.472 ± 0.712
$D^{*+}\pi^-(n\gamma)$	-0.880 ± 1.011	100.989 ± 0.718
remainder	0.011 ± 1.003	100.258 ± 0.710
combinatorial background	-4.995 ± 1.003	100.130 ± 0.715

Table 5.12: Pull distributions for $\overline{B}^0 \to D^{*+}X$

5.12 $\overline{B}^0 \to D^{*+}X$ validation

We cross-check the fit by plotting the pull distributions of the fit parameters. Data is generated by throwing dice against the mass and dE/dx templates with the f_j determined in the fit as the "true" f_j. The data is then fit. Figure 5.45 shows that the fit provides unbiased results and that the error estimate is correct. Table 5.12 lists the pull means and widths for 10000 pseudoexperiments of 9664 events (the size of our dataset) each. The combinatorial background fraction is not a fit parameter, and so its error is not returned by the fit; instead we calculate it by propagating the errors on the fit parameters as follows:

$$\sigma_{\rm bg}^2 = \sum_{i,j}^{M-1} \sigma_{ij}, \qquad (5 \cdot 2)$$

where M is the number of fit components including the combinatorial background.

5.13 $\overline{B}^0 \to D^{*+}X$ comparison with world average

Table 5.9 shows our measurement of relative branching fractions for several modes after all corrections described in Section 4.10 have been applied. Like in the $\overline{B}^0 \to D^+X$ case, we find that the agreement between our measurement and the world average is satisfactory not only in $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big)$ but also in $D^{*+}\rho^-$.

5.14 Summary

Our control samples have fulfilled their purpose. We have applied our method to a clean $\overline{B}^0 \to D^{*+}X$ sample, which validates our treatment of the single-B backgrounds, and to a much dirtier $\overline{B}^0 \to D^+X$ sample, which shows that we can cope with combinatorial background. The analysis method gives good agreement between fit and data (or, in the case of the residual disagreement in the pion Z distribution, we have devel-

oped a method for assessing a systematic uncertainty). Having found good agreement in samples three or ten times the size of the signal sample, we expect good agreement there as well. Finally, the measurements of $\mathcal{B}\big(\overline{B}^0 \to D^+K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+\pi^-\big)$ and $\mathcal{B}\big(\overline{B}^0 \to D^{*+}K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+}\pi^-\big)$ are consistent with the world average. We can therefore proceed to the measurement of $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_s K^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s\pi^-\big)$.

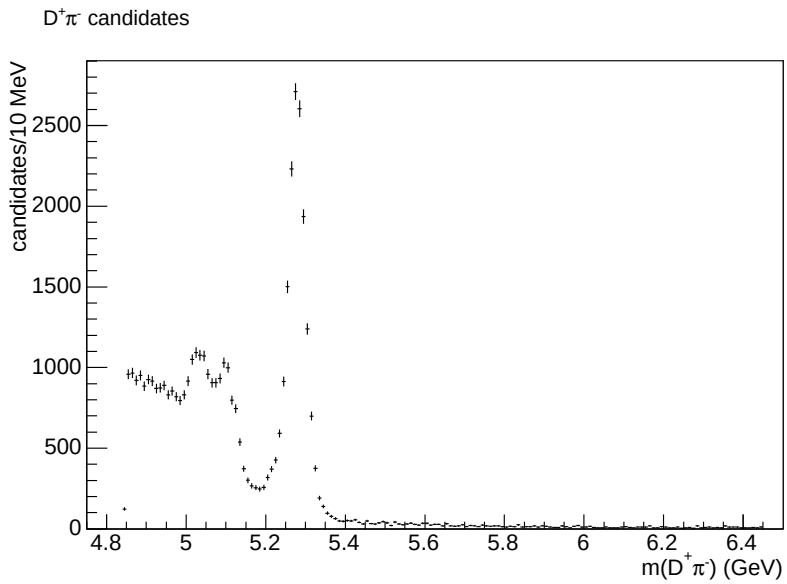

Figure 5.1: Mass distribution for $\overline{B}^0 \to D^+\pi^-$ candidates

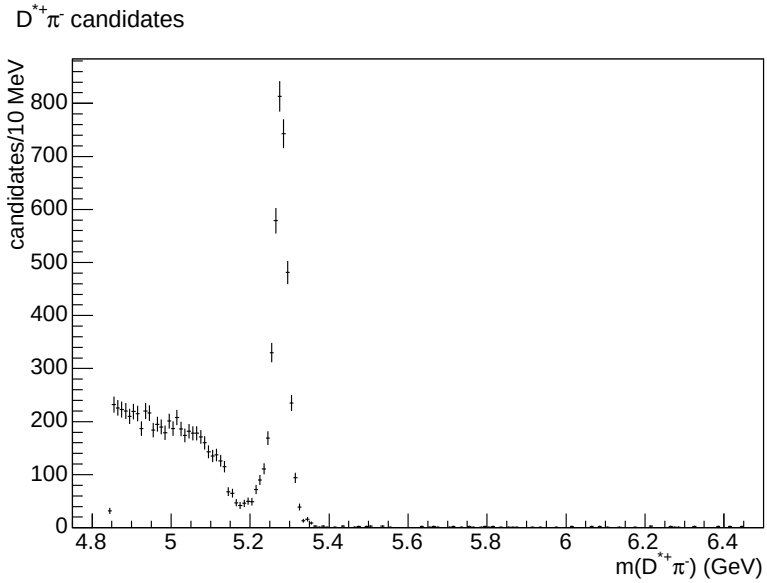

Figure 5.2: Mass distribution for $\overline{B}^0 \to D^{*+}\pi^-$ candidates

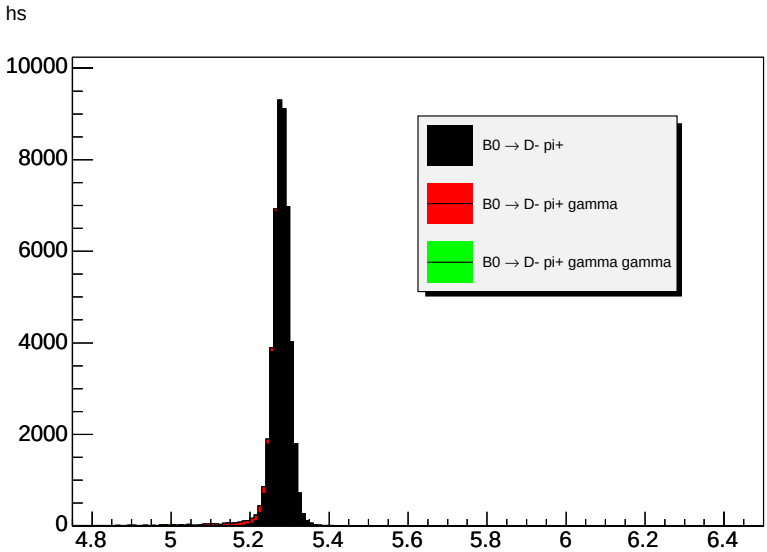

Figure 5.3: Composition of $\overline{B}^0 \to D^+\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$

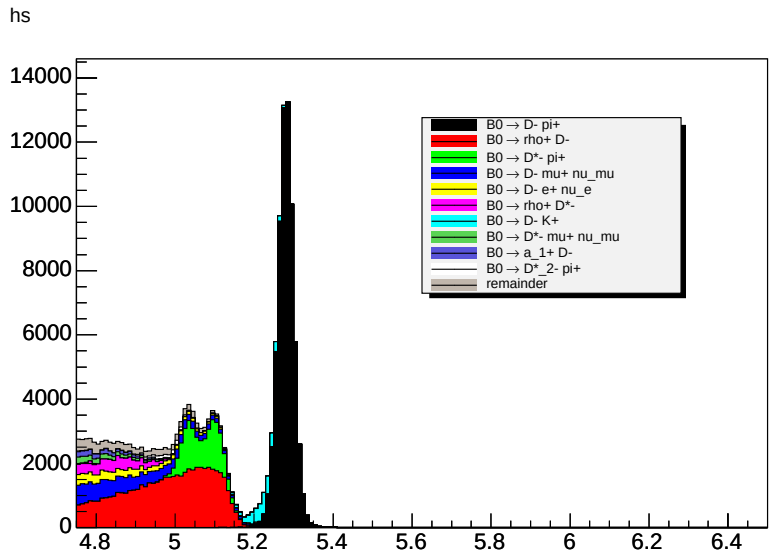

Figure 5.4: Composition of $\overline{B}^0 \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$

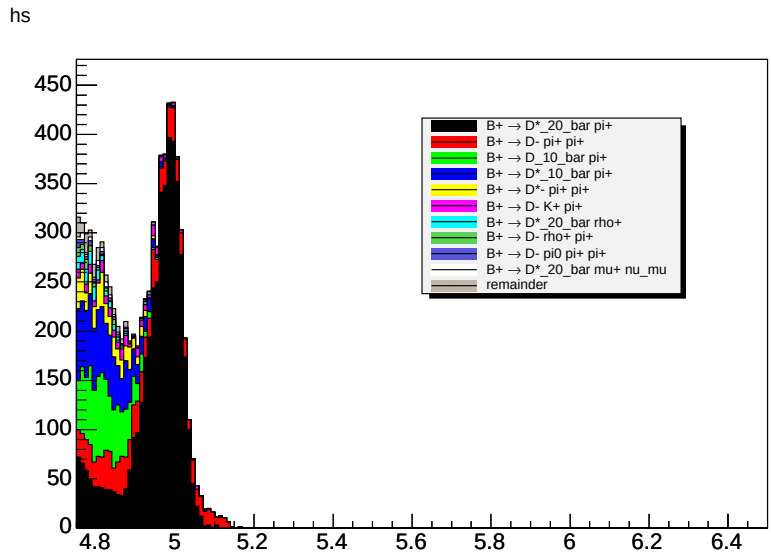

Figure 5.5: Composition of $B^- \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$

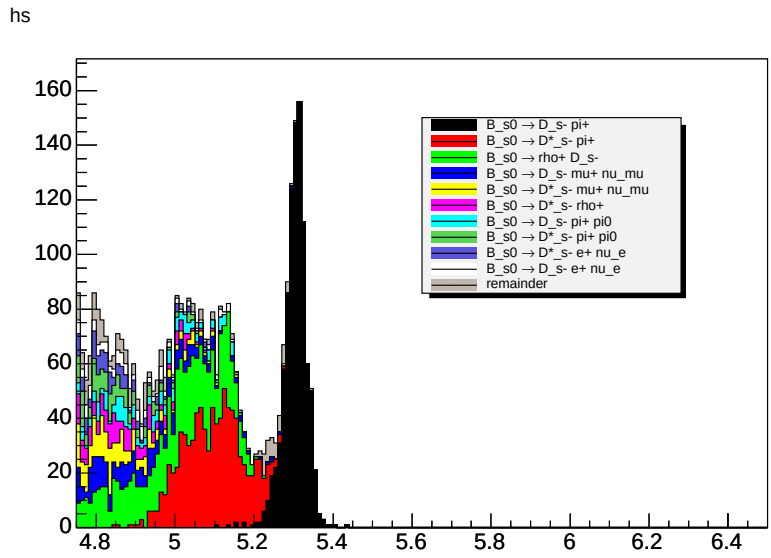

Figure 5.6: Composition of $B_s \to D_s(\phi\pi)X$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^+\pi^-$

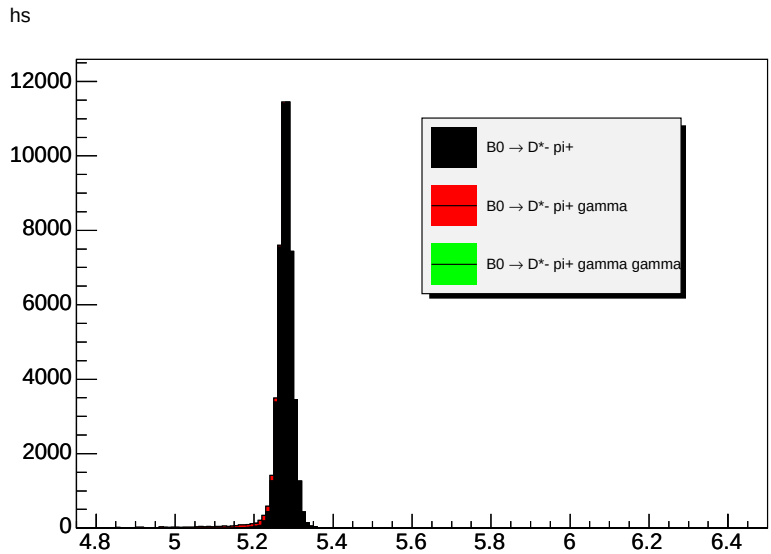

Figure 5.7: Composition of $\overline{B}^0 \to D^{*+}\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}^0 \to D^{*+}\pi^-$

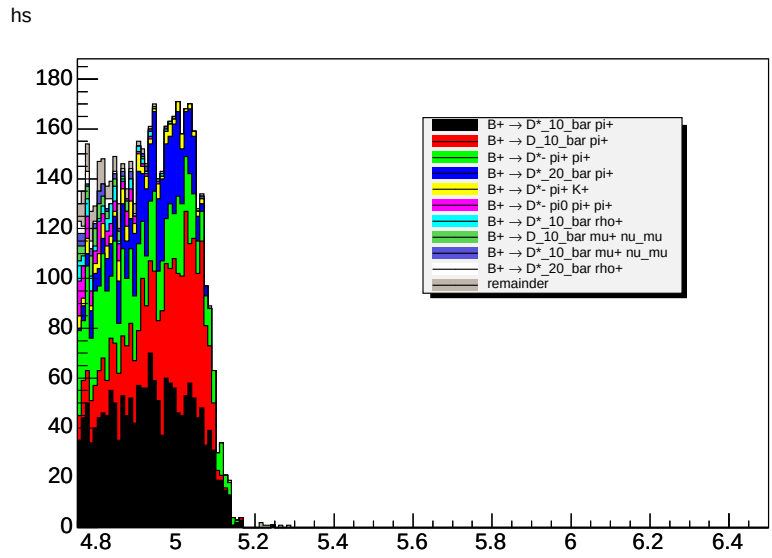

Figure 5.8: Composition of $B^- \rightarrow D^0\pi^- X$ Monte Carlo sample reconstructed as $\overline{B}^0 \rightarrow D^{*+}\pi^-$

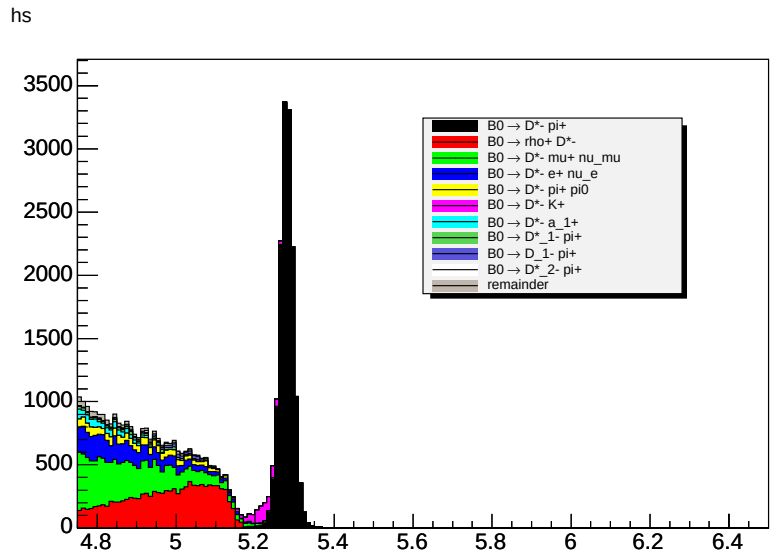

Figure 5.9: Composition of $\overline{B}^0 \rightarrow D^{*+}X$ Monte Carlo sample reconstructed as $\overline{B}^0 \rightarrow D^{*+}\pi^-$

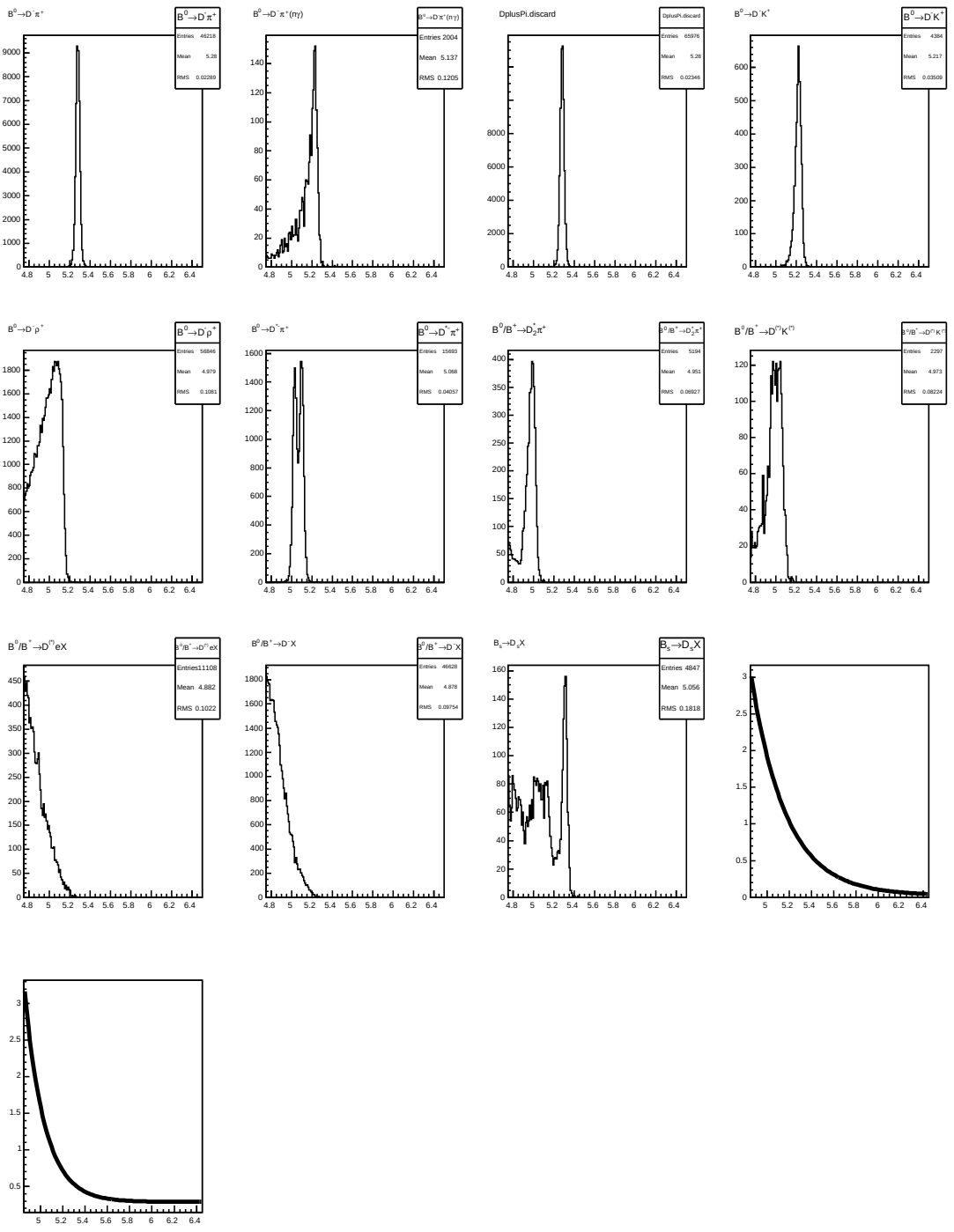

Figure 5.10: Mass templates used to model the $\overline{B}^0 \to D^+X$ fit components. The template marked "$\overline{B}^0 \to D^+\pi^-$.discard" is extracted from the non-PHOTOS Monte Carlo sample to determine the $D^+K^-/D^+\pi^-$ relative yield, but is not used as a fit template.

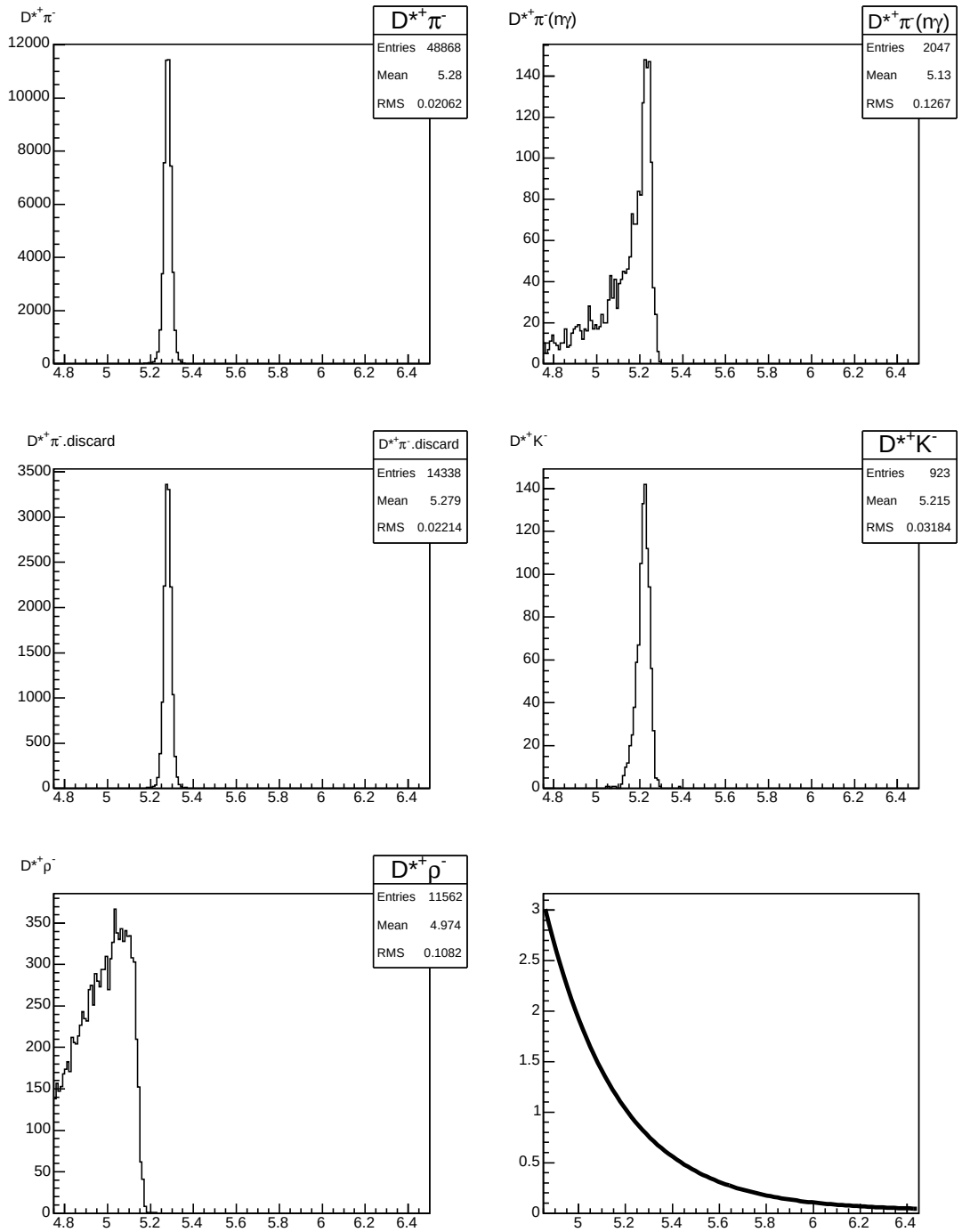

Figure 5.11: Mass templates used to model the $\overline{B}^0 \to D^{*+}X$ fit components. The template marked "$D^{*+}\pi^-$.discard" is extracted from the non-PHOTOS Monte Carlo sample to determine the $D^{*+}K^-/D^{*+}\pi^-$ relative yield, but is not used as a fit template.

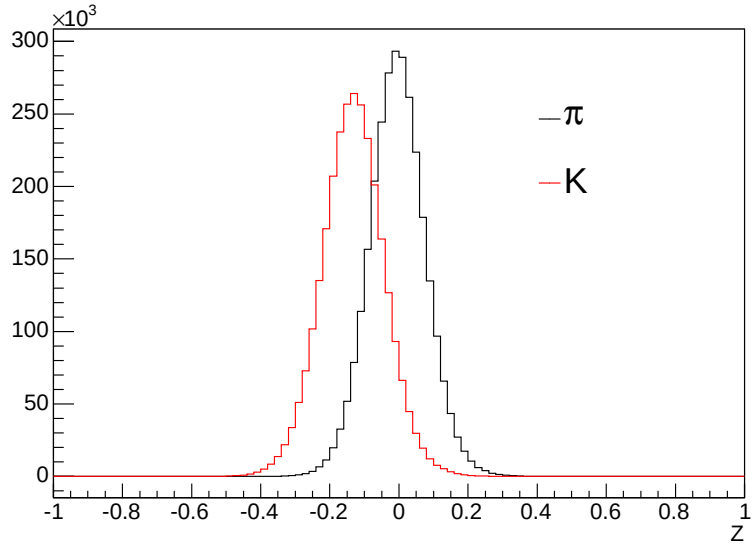

Figure 5.12: Z templates for K and π tracks reweighted for $\overline{B}^0 \to D^+X$

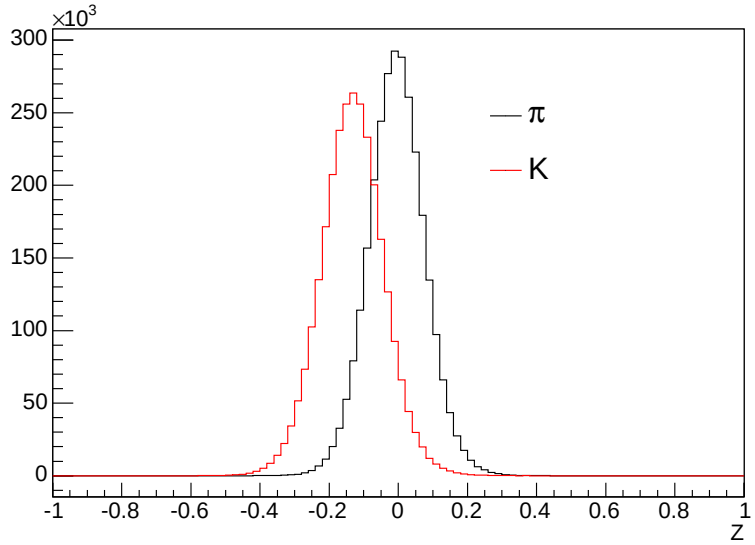

Figure 5.13: Z templates for K and π tracks in reweighted for $\overline{B}^0 \to D^{*+}X$

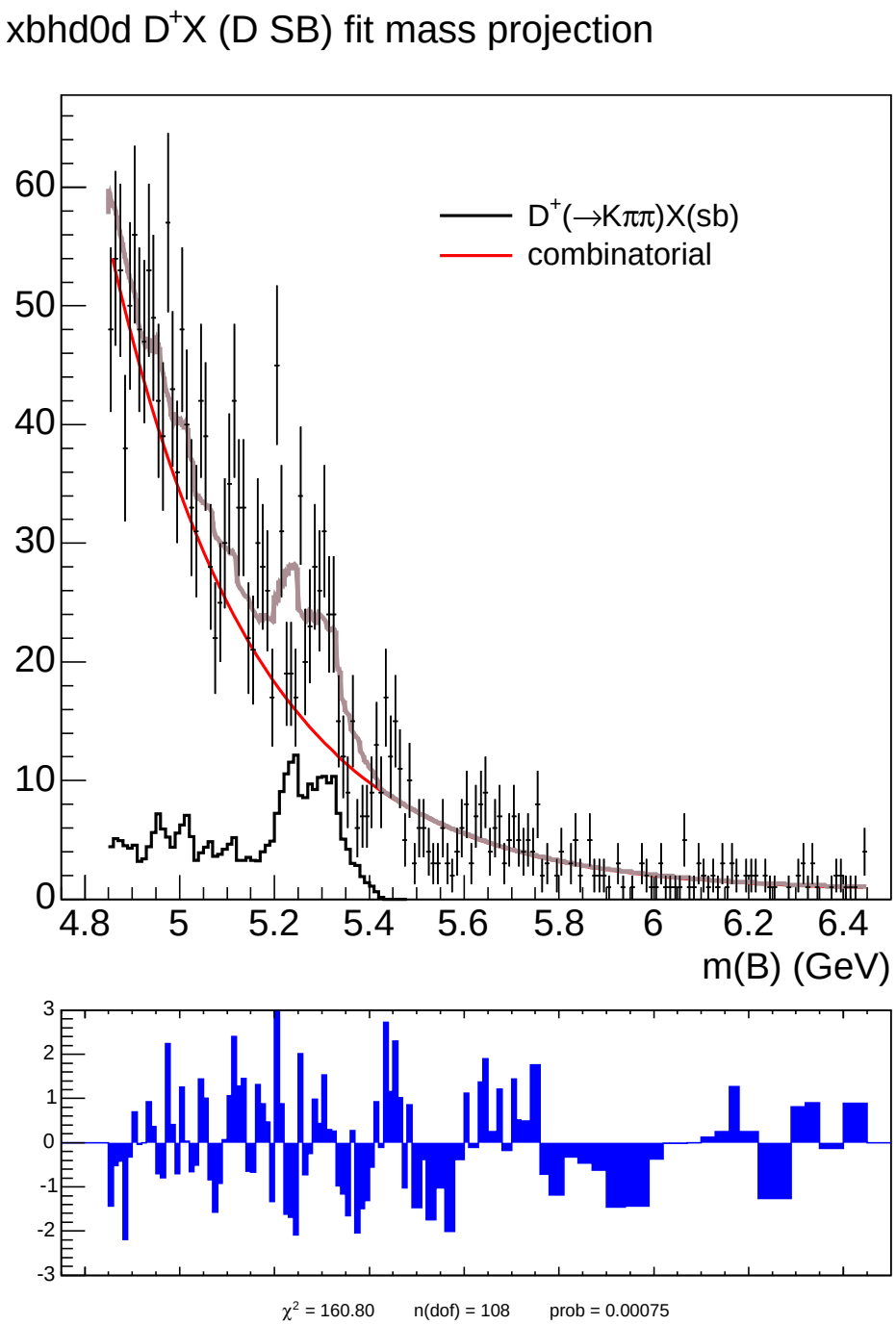

Figure 5.14: Fake-D^+ combinatorial template fit (`xbhd0d`). The data points show the $D^+\pi^-$ mass distribution of B candidates from the D sidebands. Signal subtraction is performed on this distribution to yield a combinatorial background of the form $(4 \cdot 6)$. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

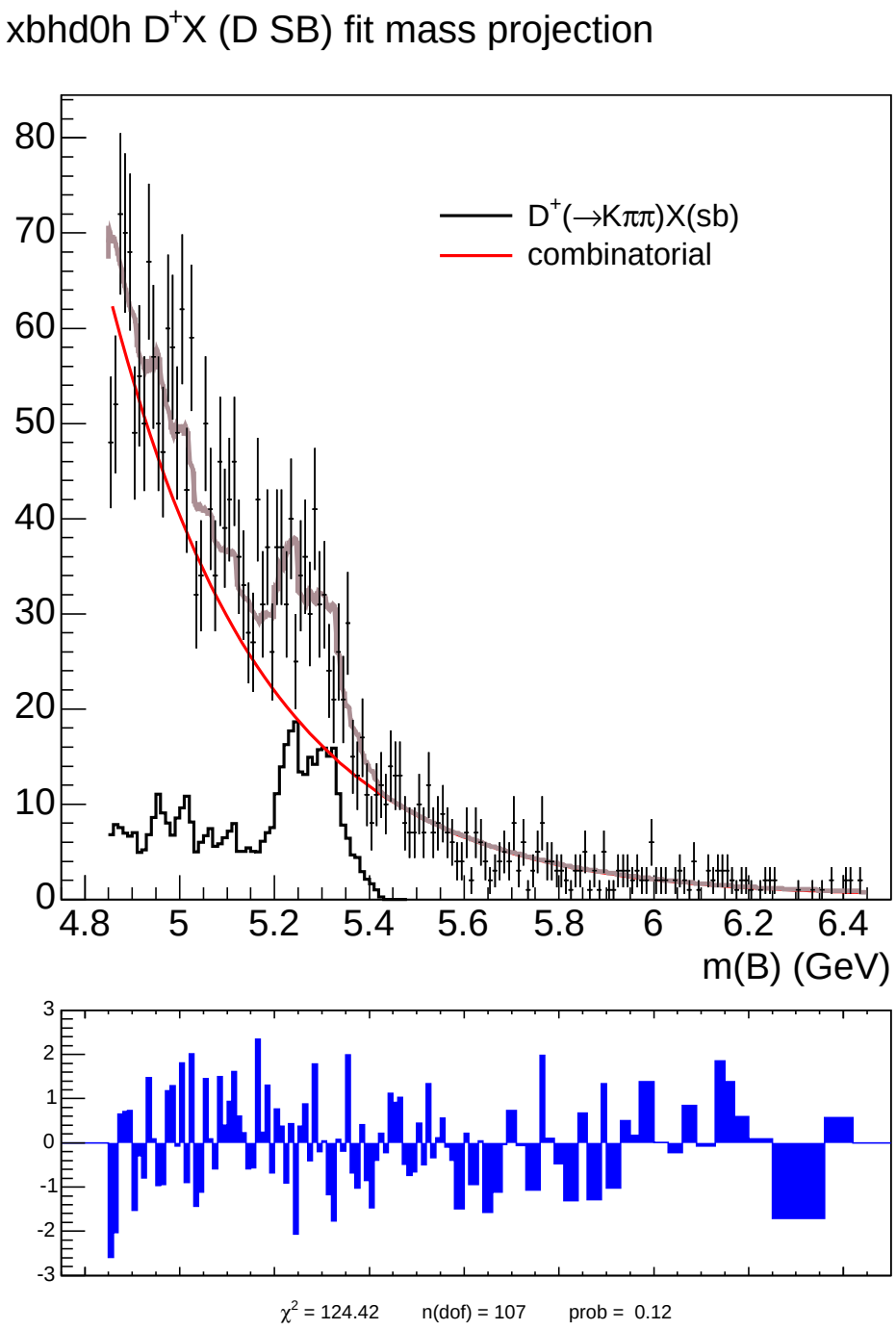

Figure 5.15: Fake-D^+ combinatorial template fit (`xbhd0h`). The data points show the $D^+\pi^-$ mass distribution of B candidates from the D sidebands. Signal subtraction is performed on this distribution to yield a combinatorial background of the form $(4 \cdot 6)$. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

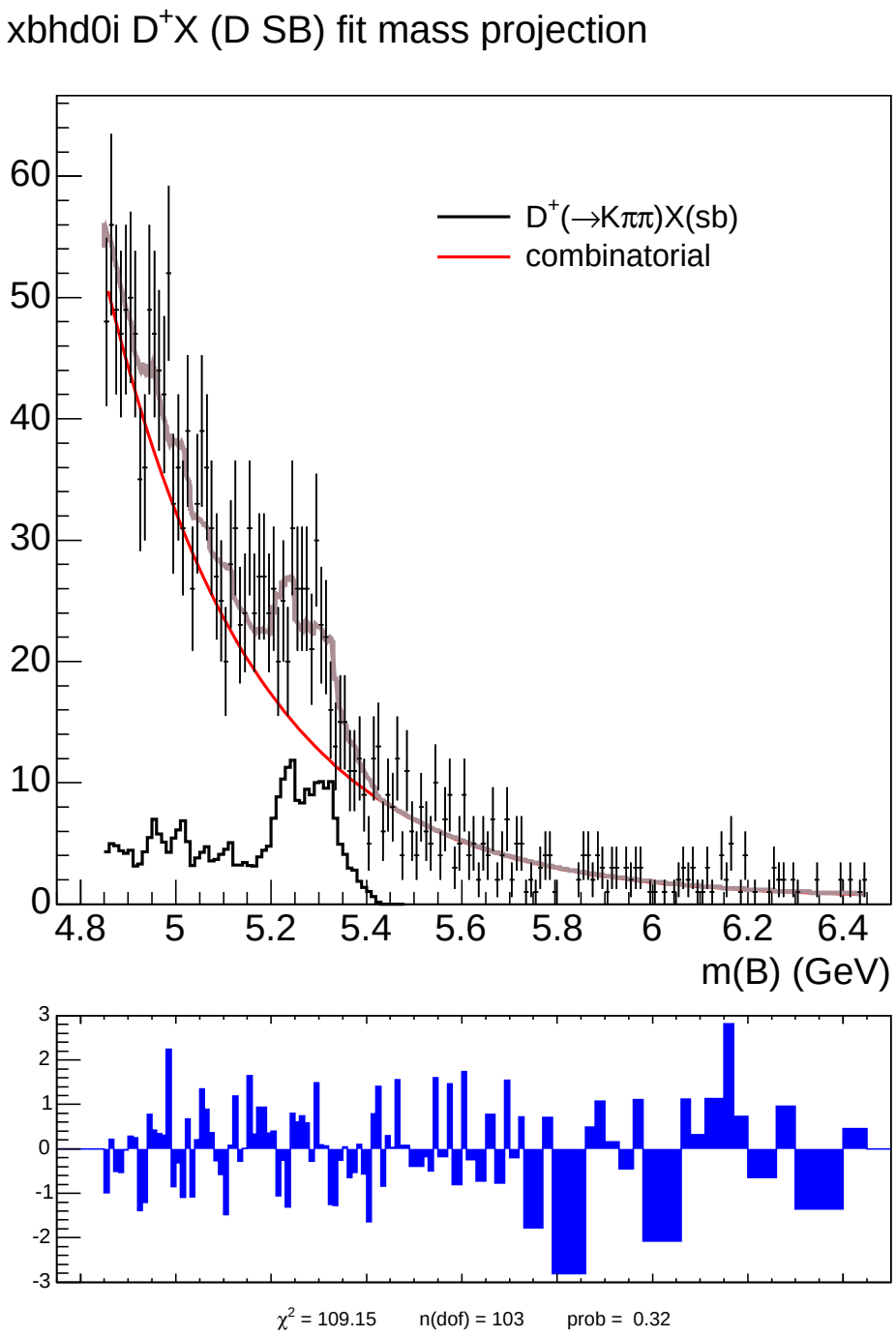

Figure 5.16: Fake-D^+ combinatorial template fit (`xbhd0i`). The data points show the $D^+\pi^-$ mass distribution of B candidates from the D sidebands. Signal subtraction is performed on this distribution to yield a combinatorial background of the form $(4 \cdot 6)$. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

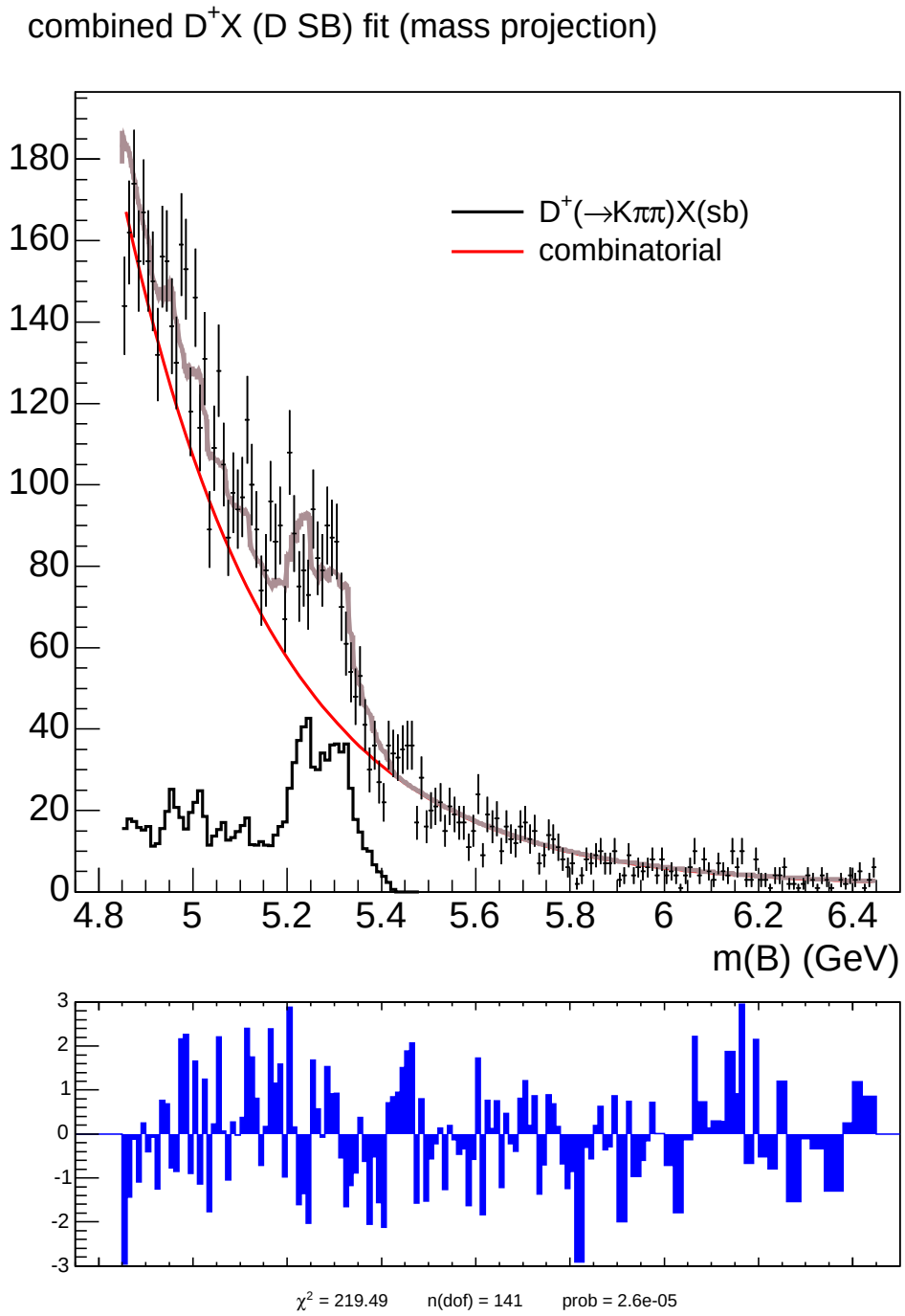

Figure 5.17: Fake-D^+ combinatorial template fit (combined). The data points show the $D^+\pi^-$ mass distribution of B candidates from the D sidebands. Signal subtraction is performed on this distribution to yield a combinatorial background of the form $(4 \cdot 6)$. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

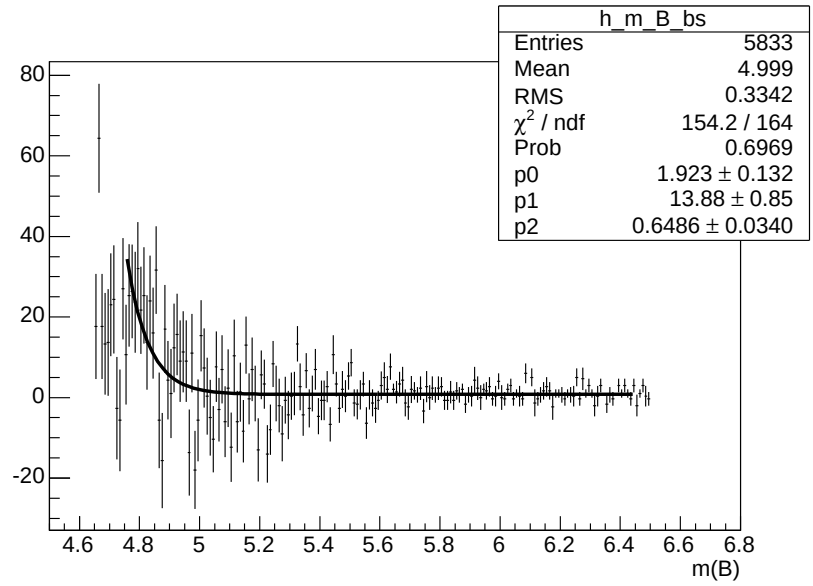

Figure 5.18: Real-D^+ combinatorial template fit (`xbhd0d`). The fit parameters are a normalization factor (p_0) and the shape parameters of the combinatorial-background function ($4 \cdot 6$).

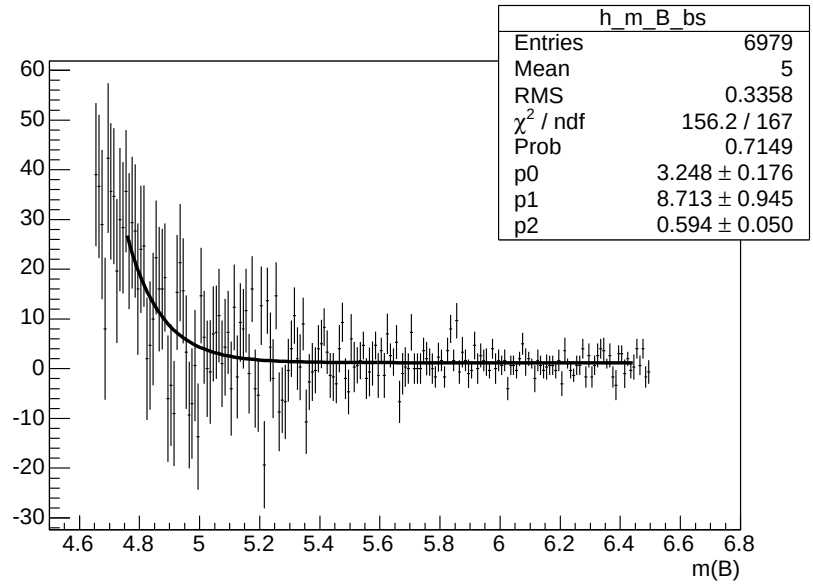

Figure 5.19: Real-D^+ combinatorial template fit (`xbhd0h`). The fit parameters are a normalization factor (p_0) and the shape parameters of the combinatorial-background function ($4 \cdot 6$).

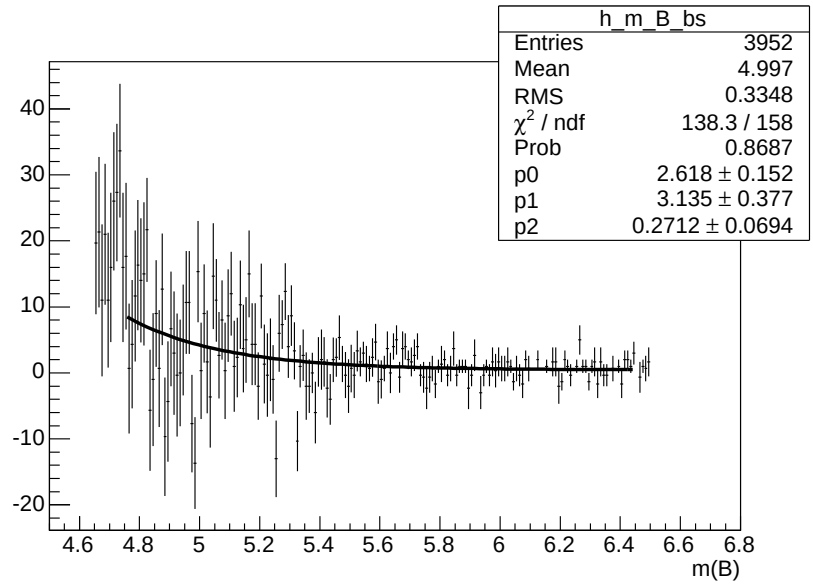

Figure 5.20: Real-D^+ combinatorial template fit (`xbhd0i`). The fit parameters are a normalization factor (p_0) and the shape parameters of the combinatorial-background function ($4 \cdot 6$).

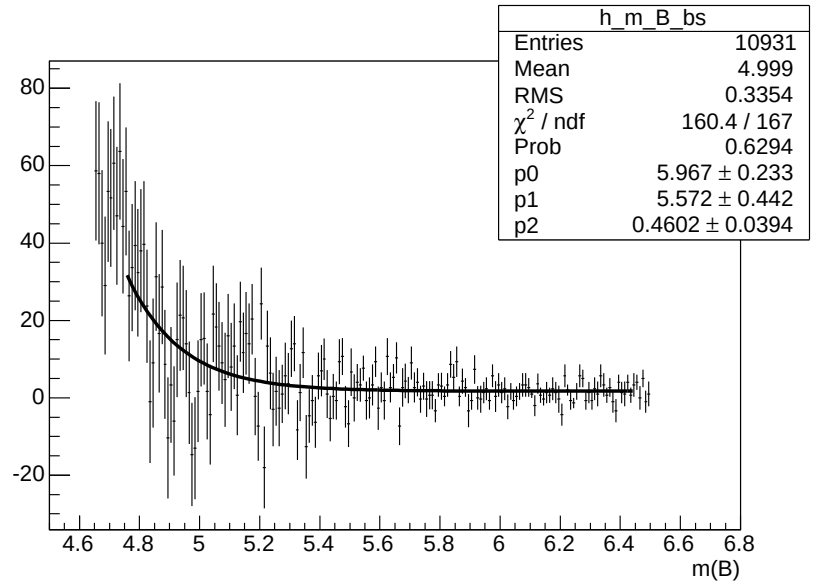

Figure 5.21: Real-D^+ combinatorial template fit (combined). The fit parameters are a normalization factor (p_0) and the shape parameters of the combinatorial-background function ($4 \cdot 6$).

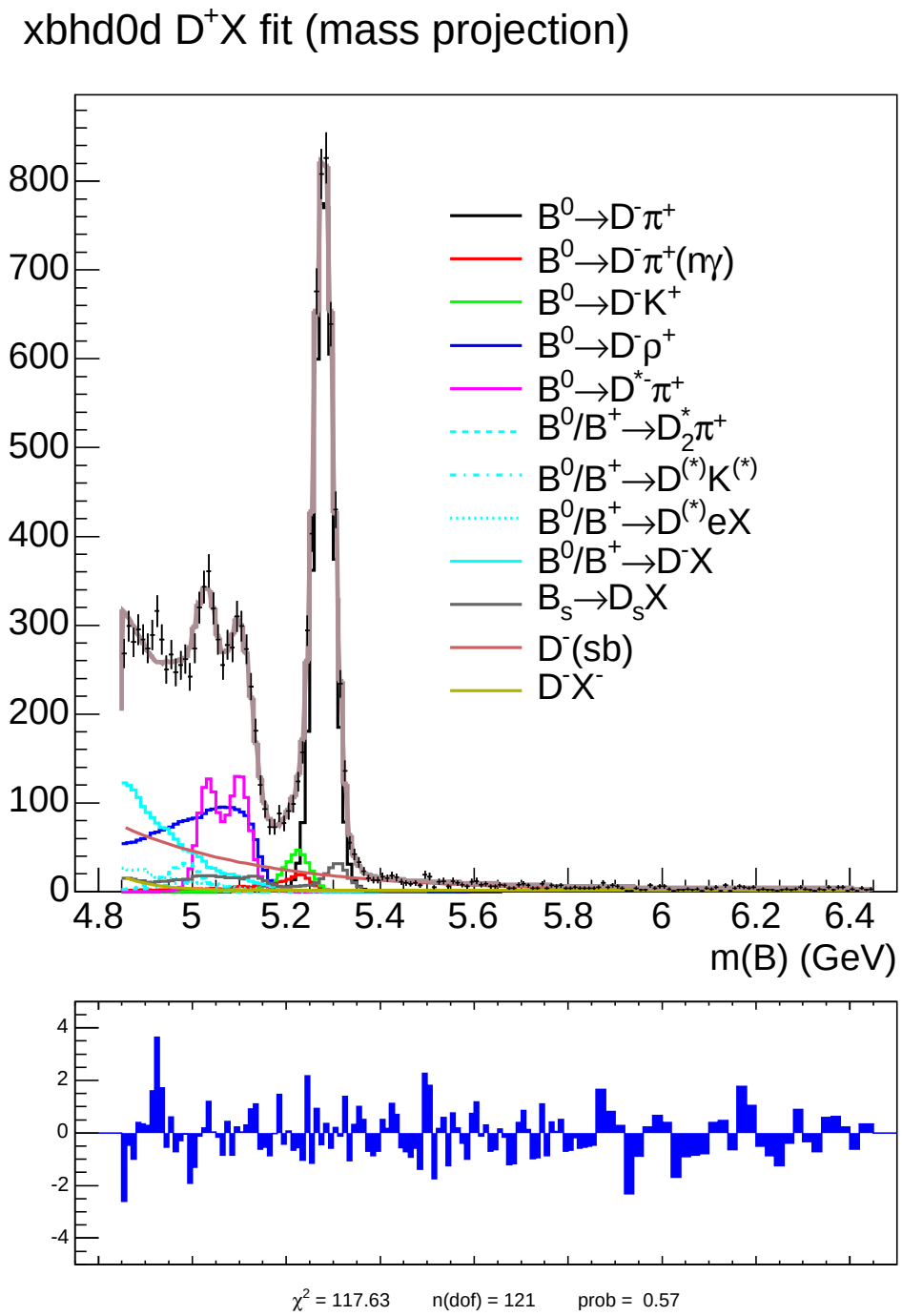

Figure 5.22: Mass projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0d`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

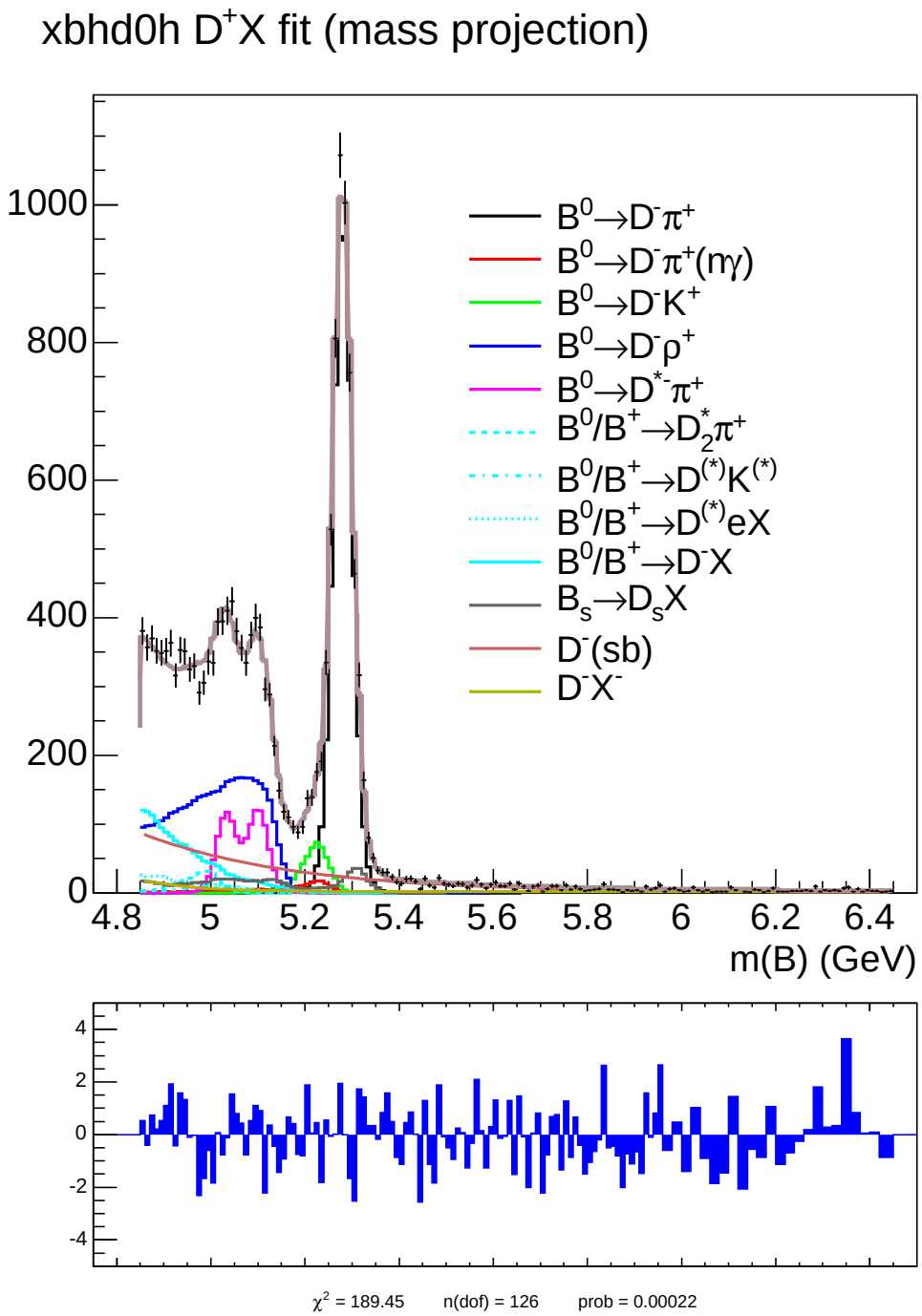

Figure 5.23: Mass projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0h`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

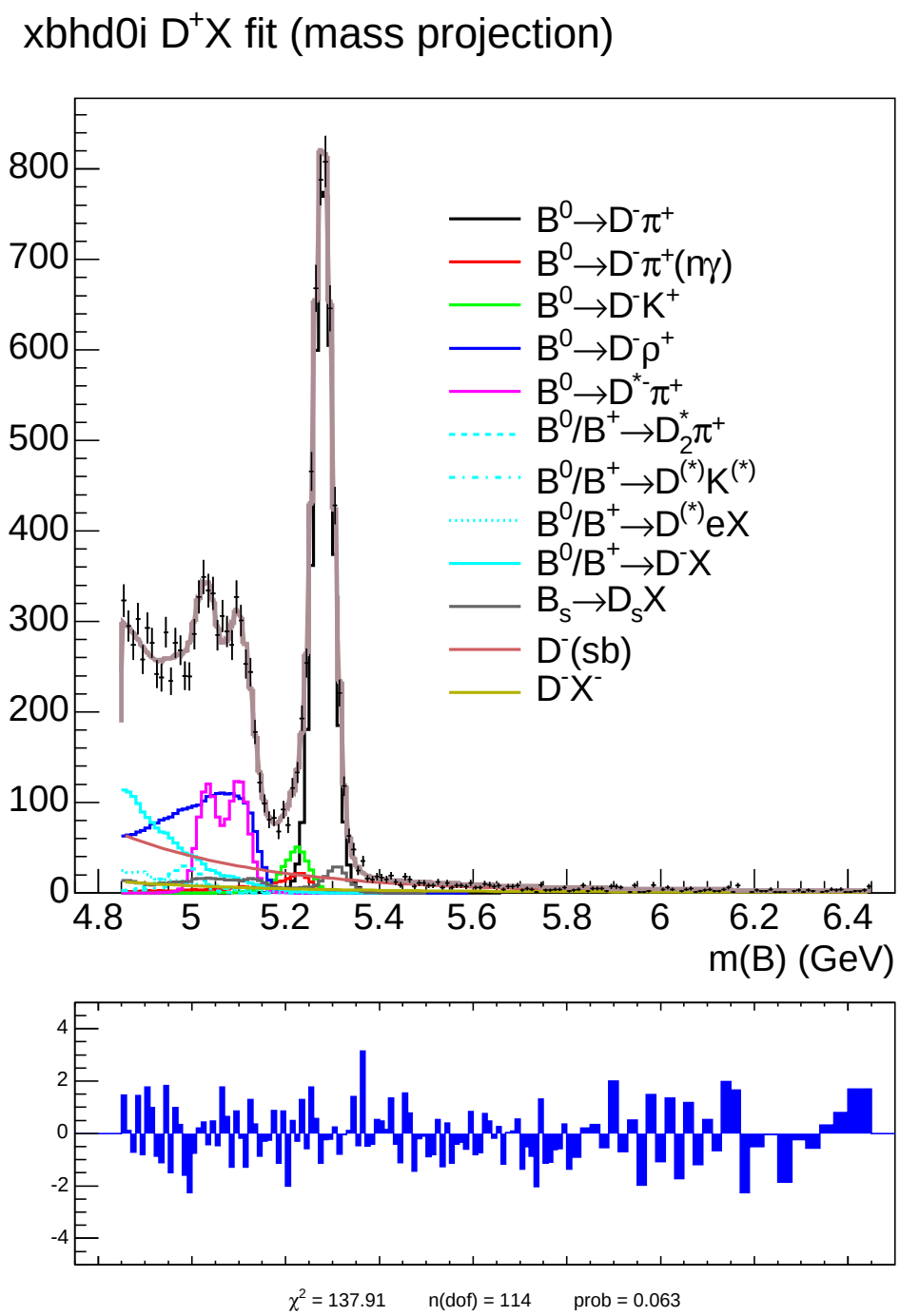

Figure 5.24: Mass projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0i`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

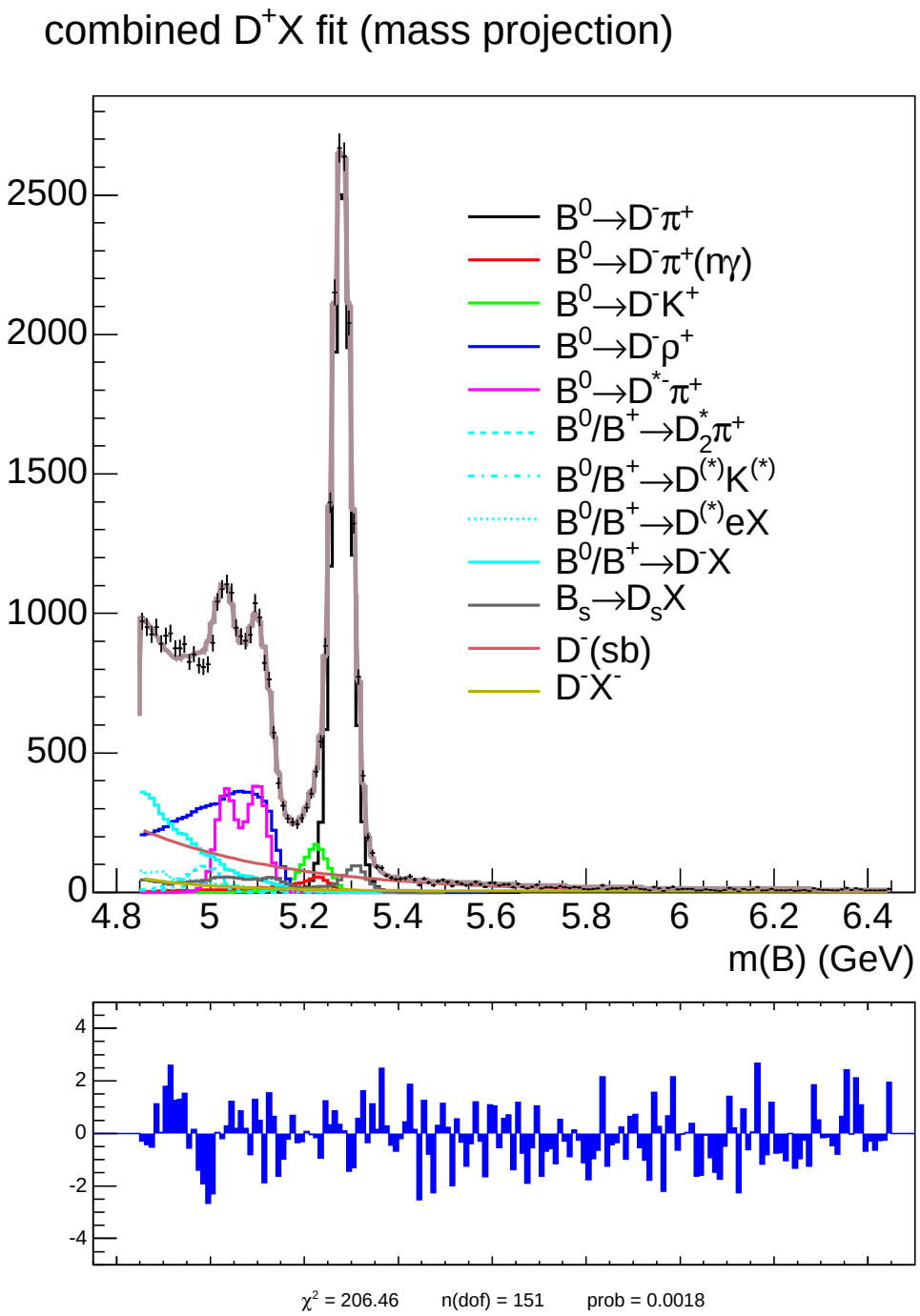

Figure 5.25: Mass projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in the combined 1.2 fb^{-1}. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

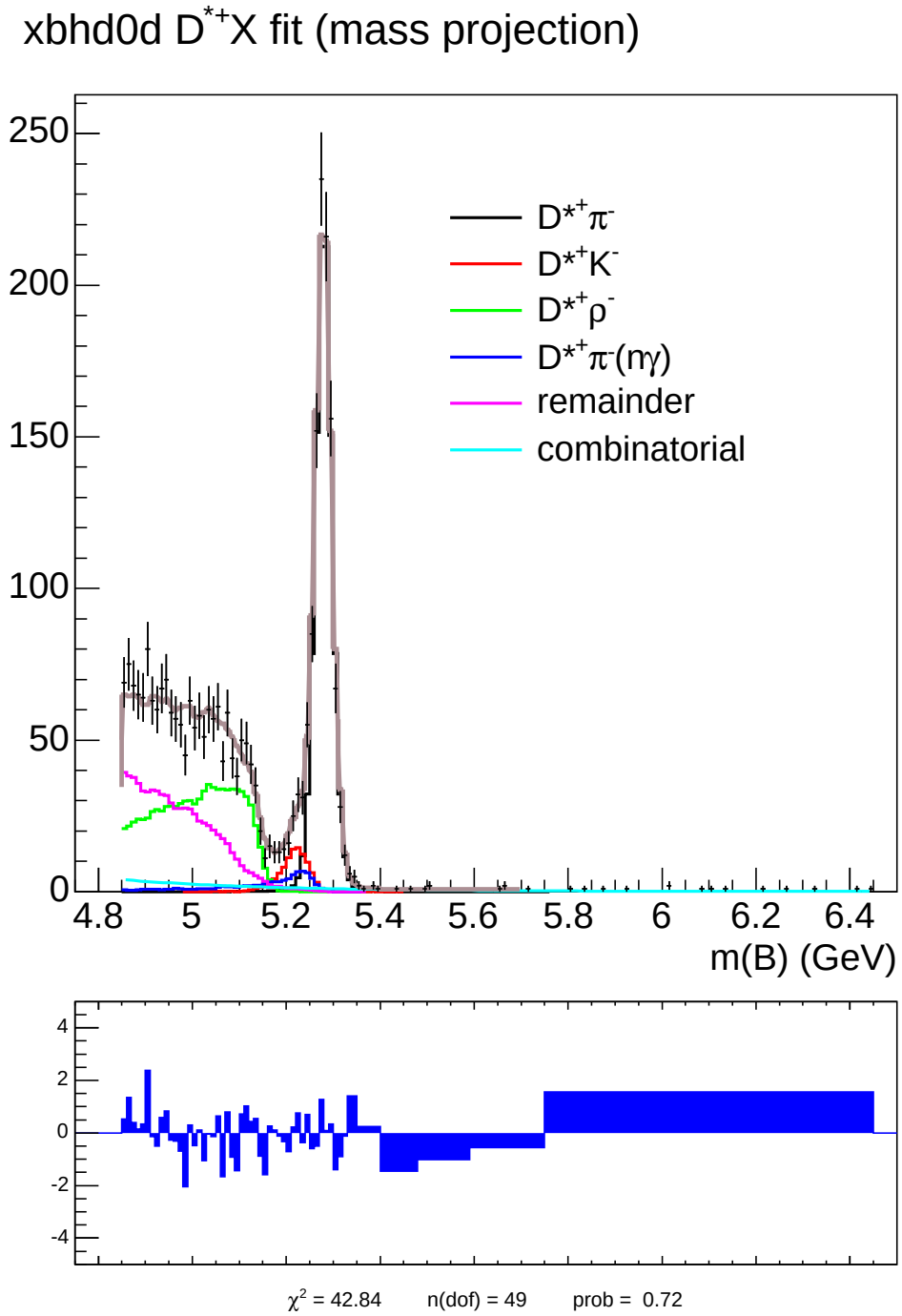

Figure 5.26: Mass projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0d`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

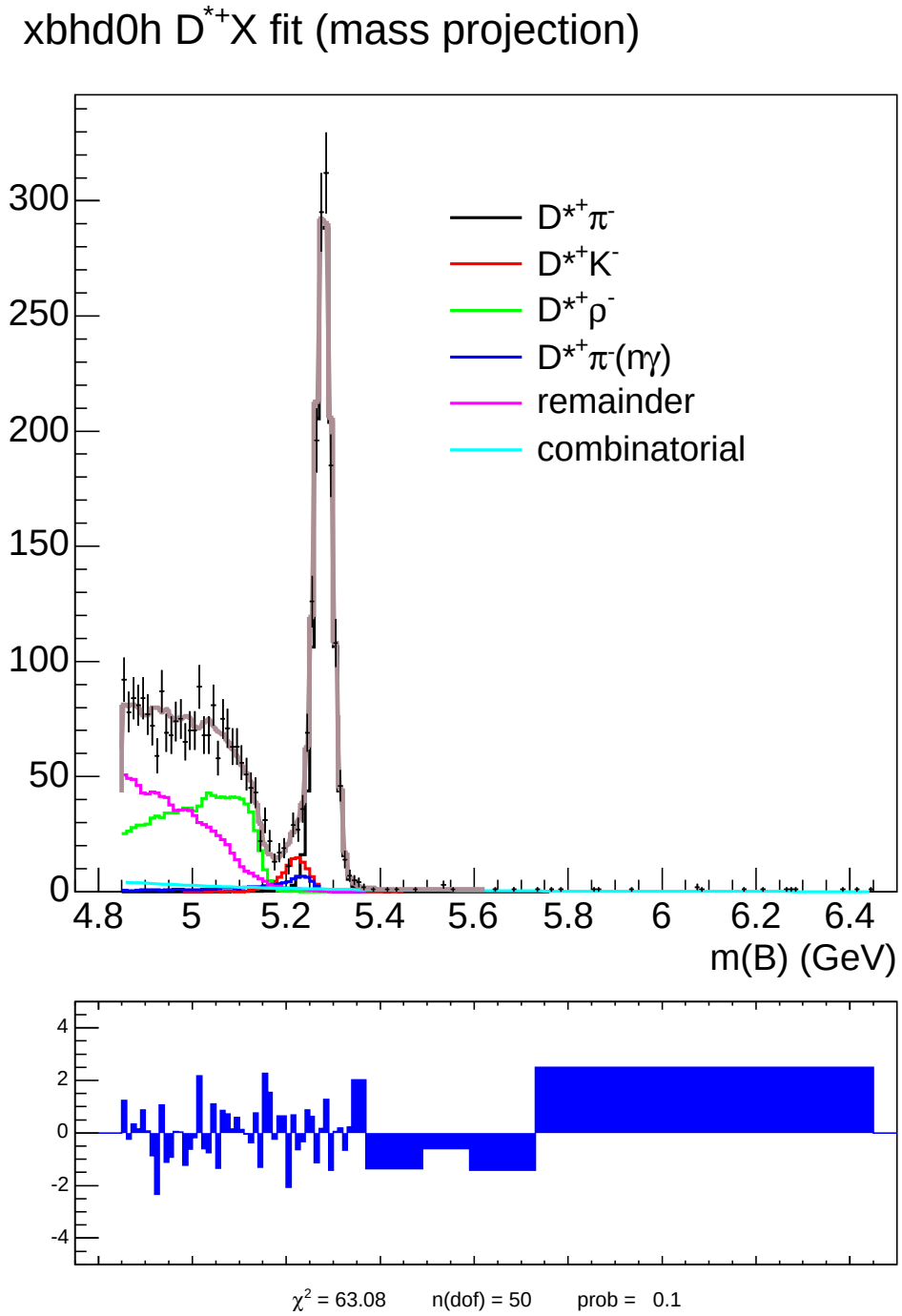

Figure 5.27: Mass projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0h`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

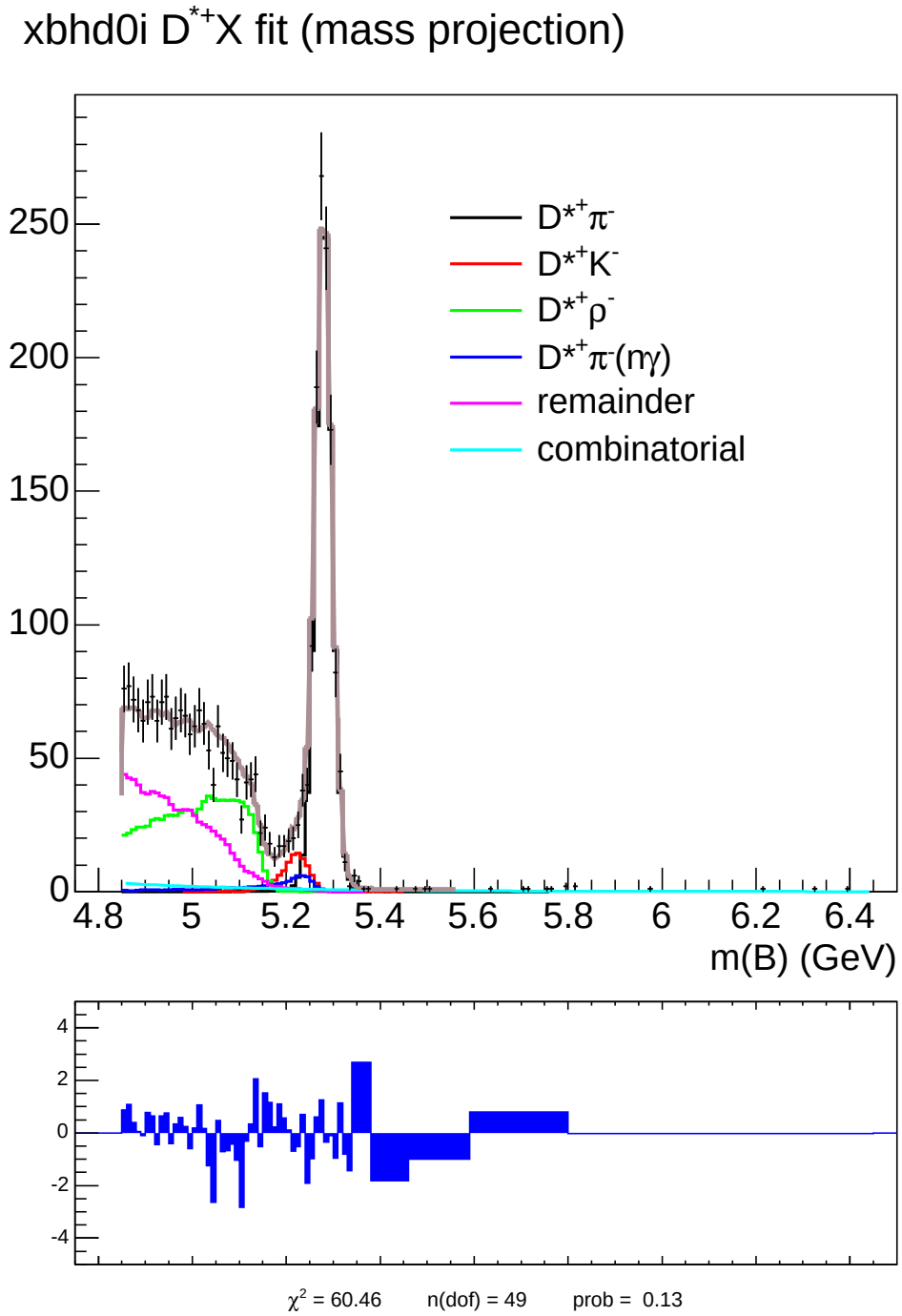

Figure 5.28: Mass projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0i`. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

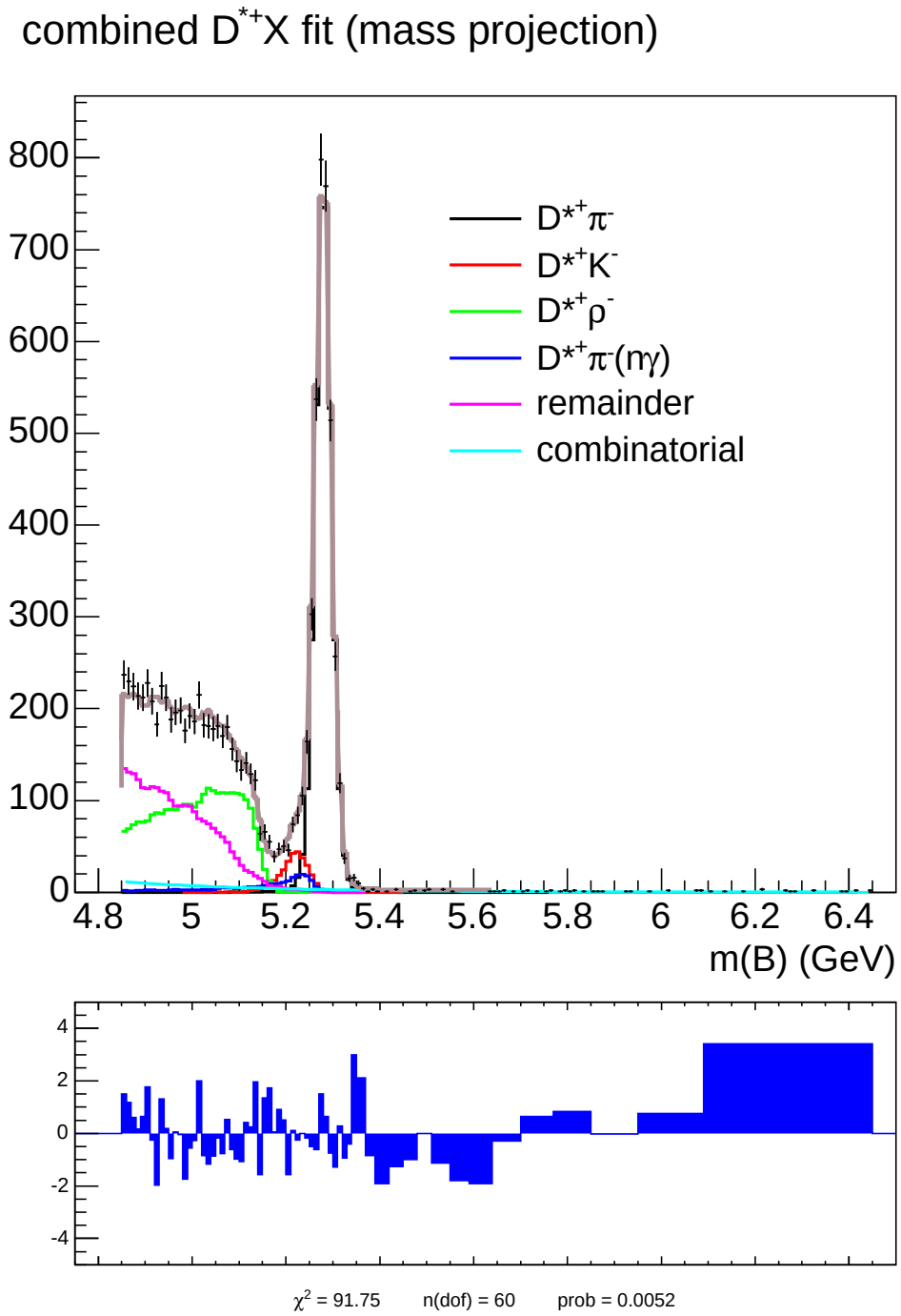

Figure 5.29: Mass projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in the combined 1.2 fb^{-1}. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

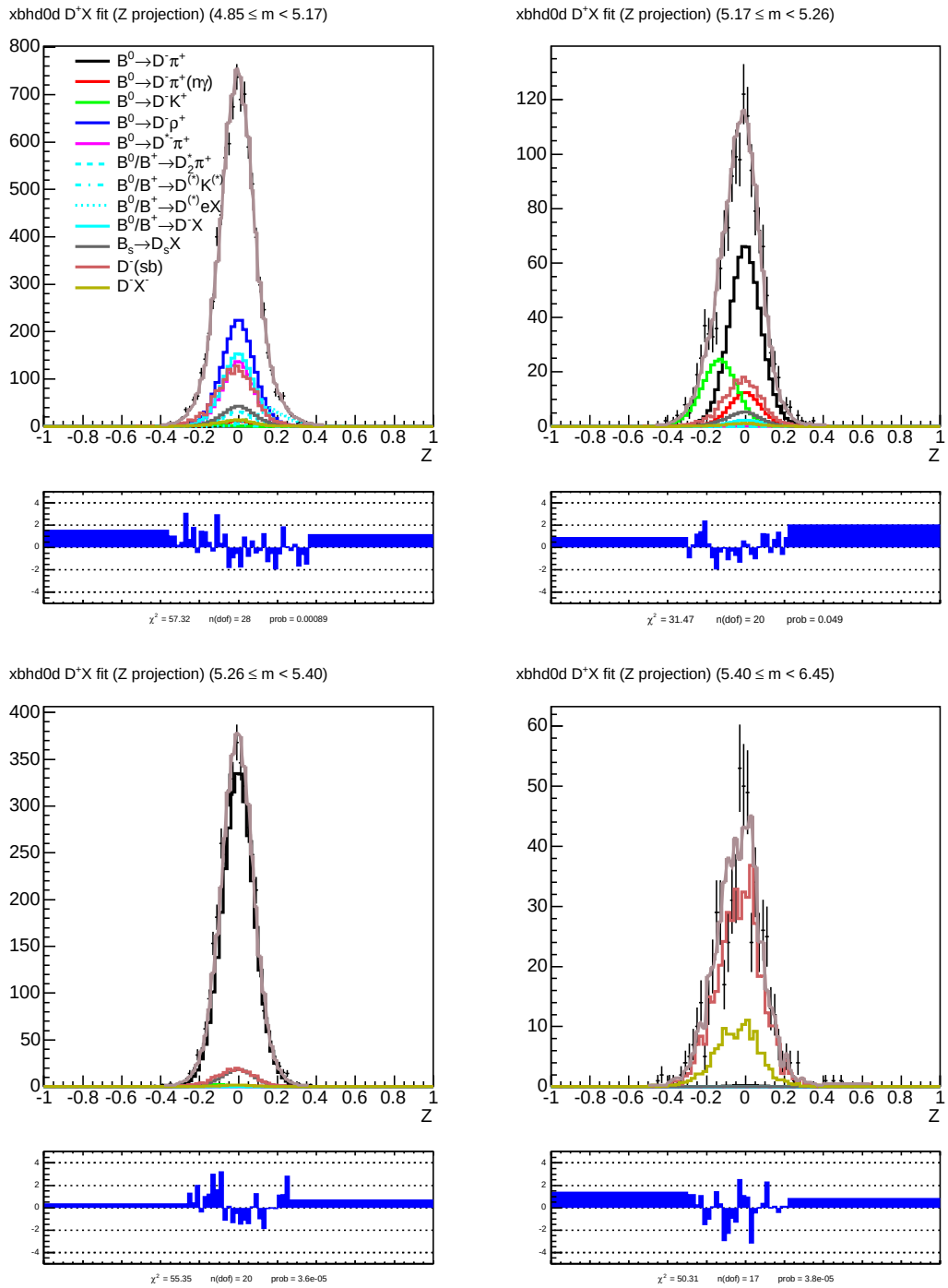

Figure 5.30: Z projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0d`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

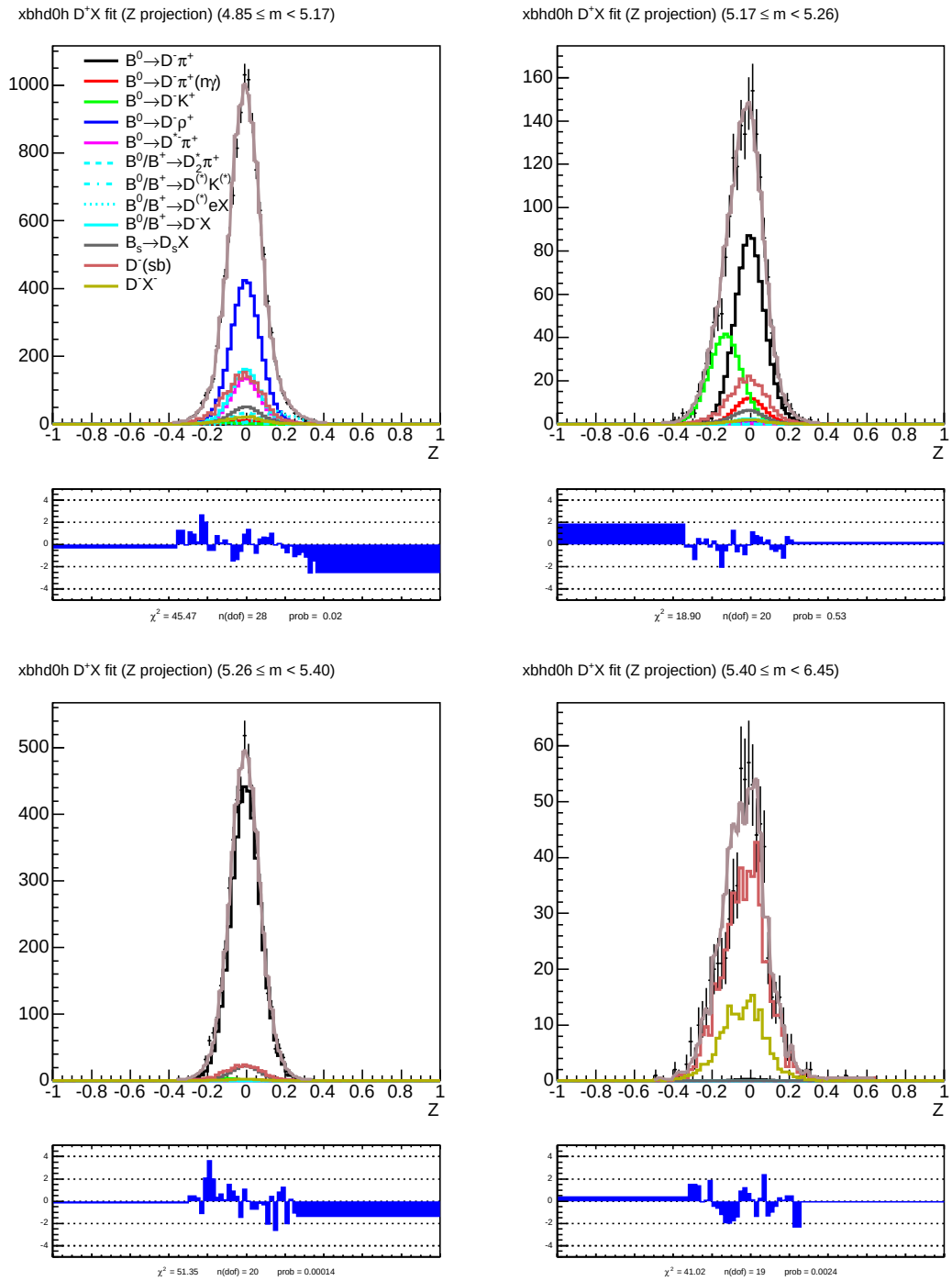

Figure 5.31: Z projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0h`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

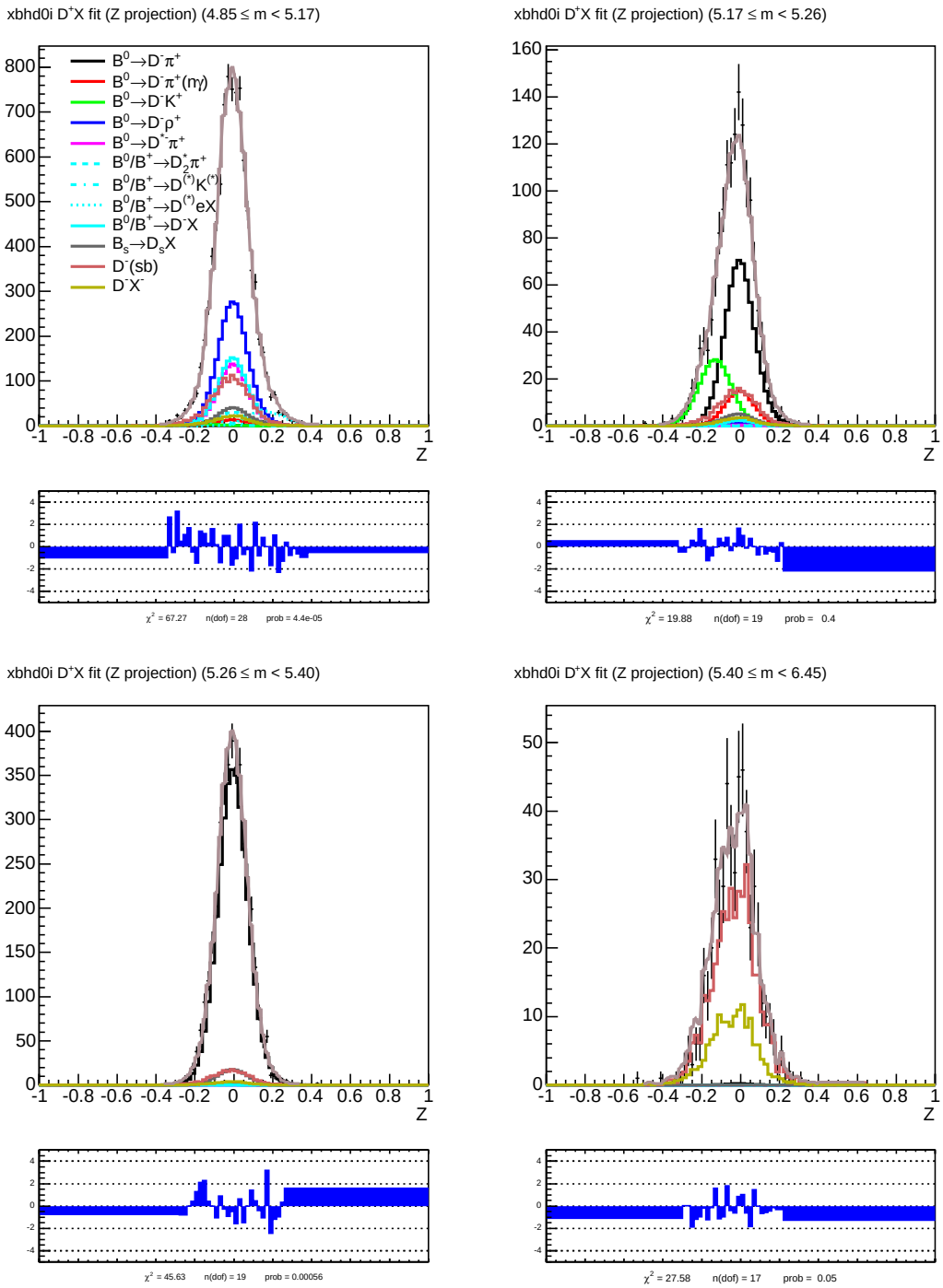

Figure 5.32: Z projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in `xbhd0i`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

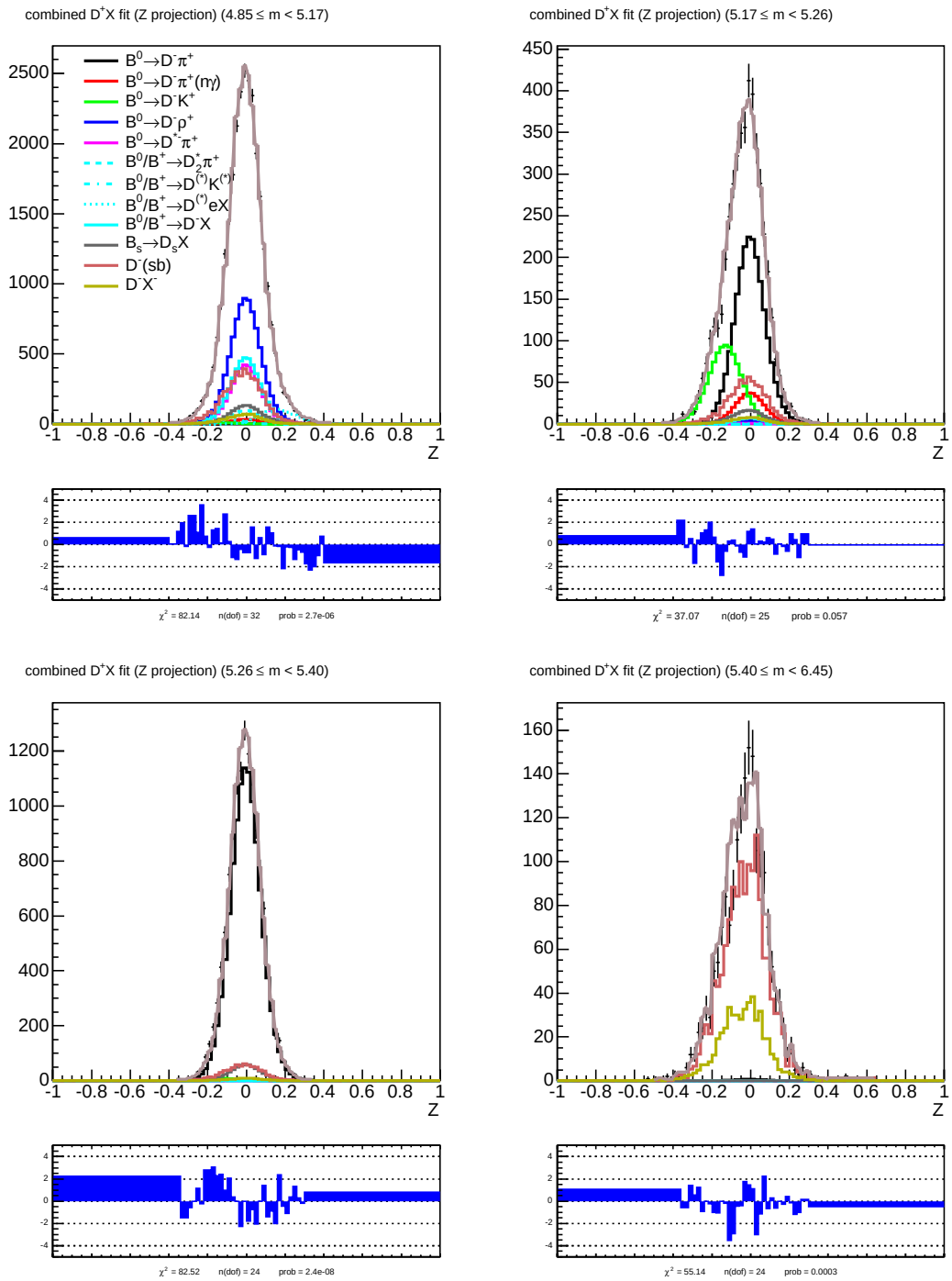

Figure 5.33: Z projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in the combined 1.2 fb^{-1}. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

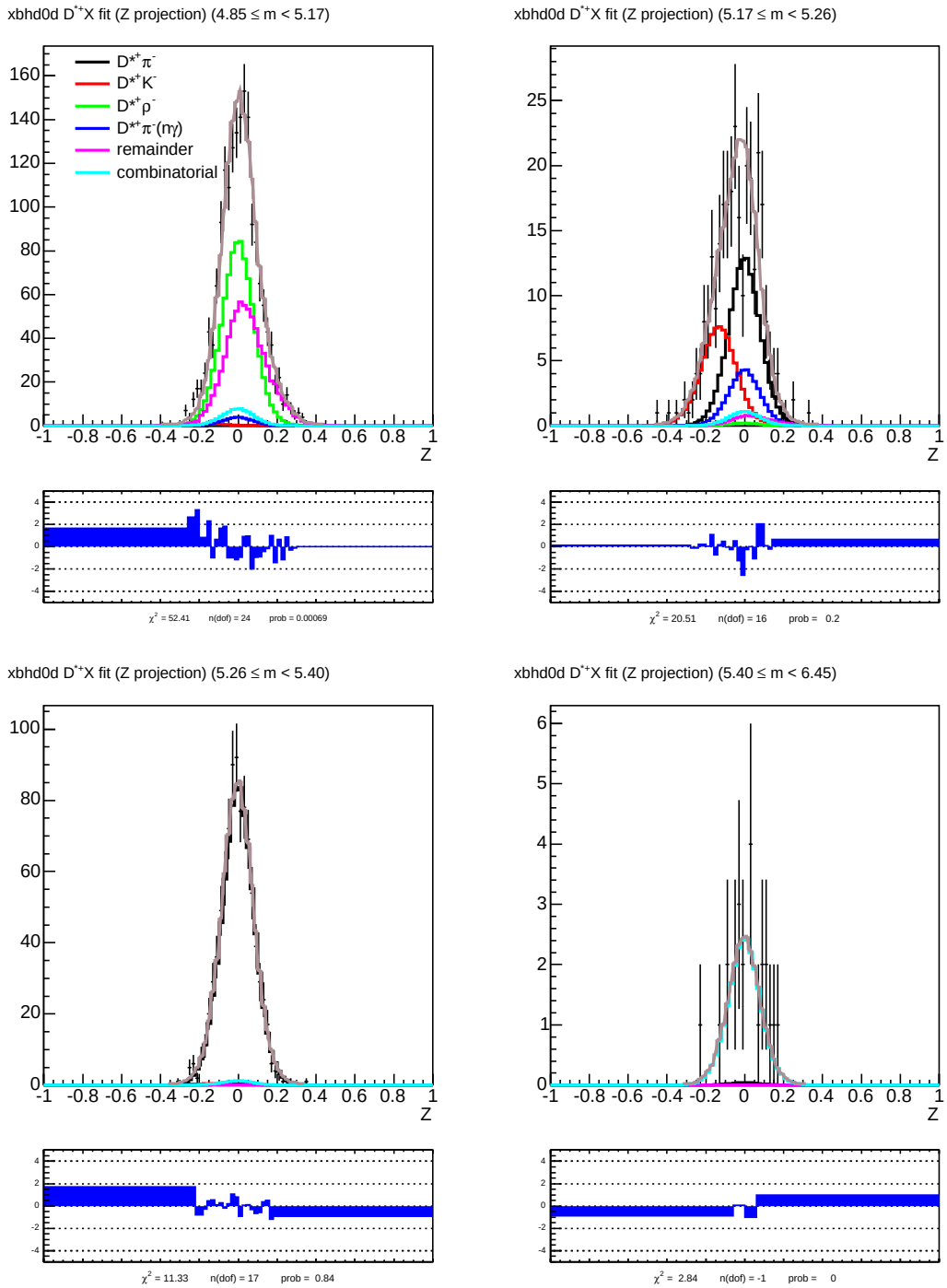

Figure 5.34: Z projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0d`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

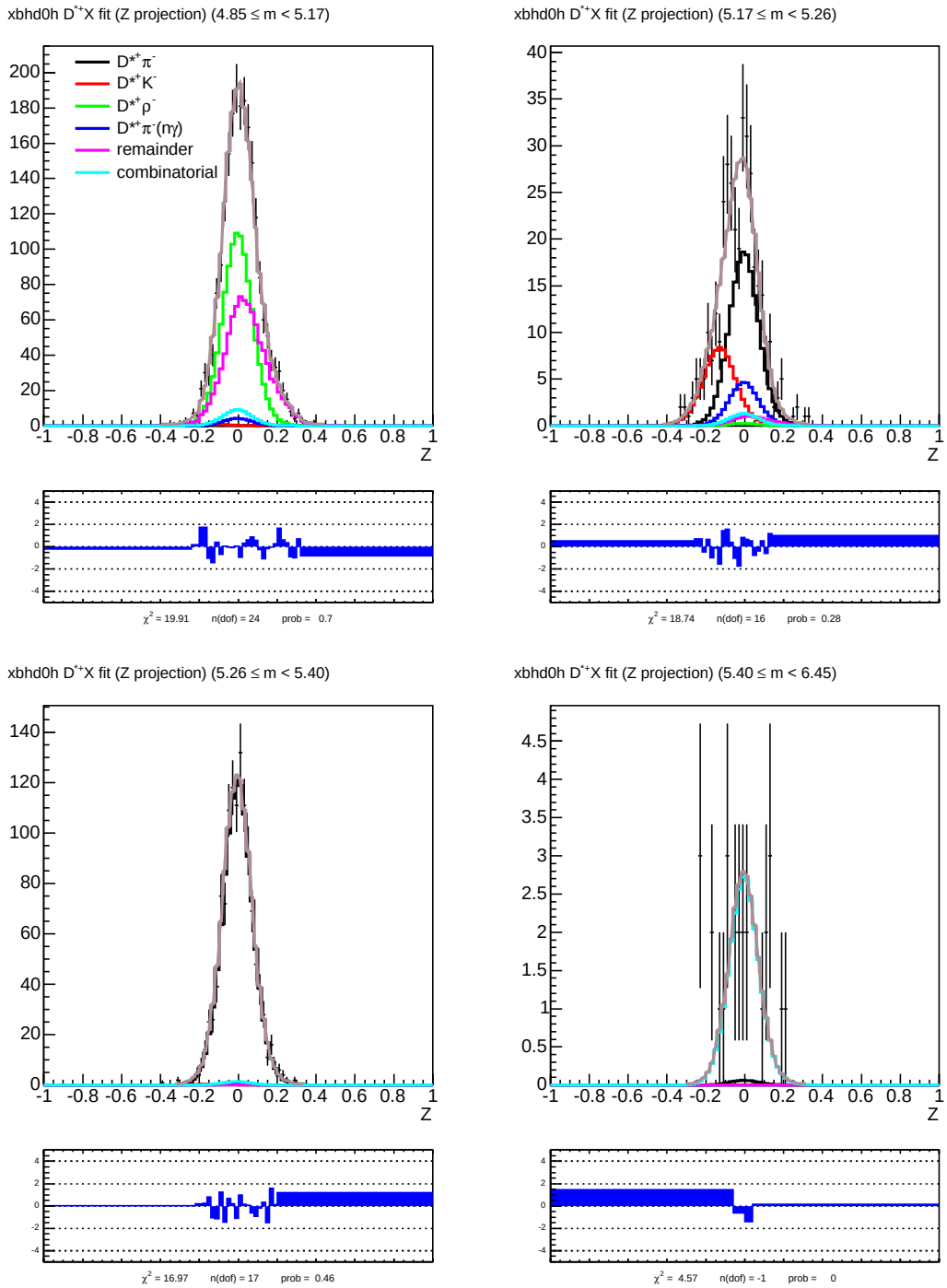

Figure 5.35: Z projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0h`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

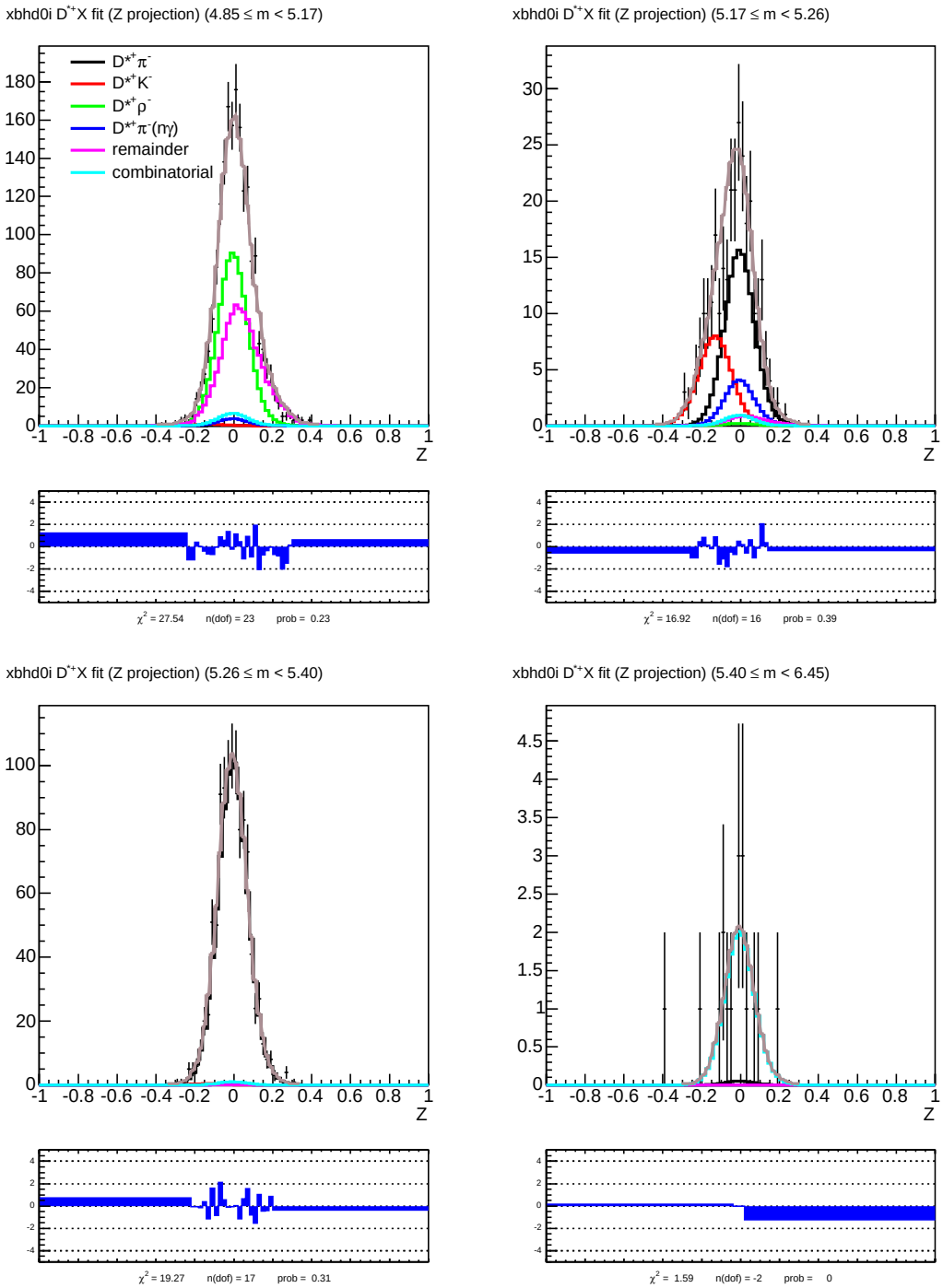

Figure 5.36: Z projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in `xbhd0i`. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

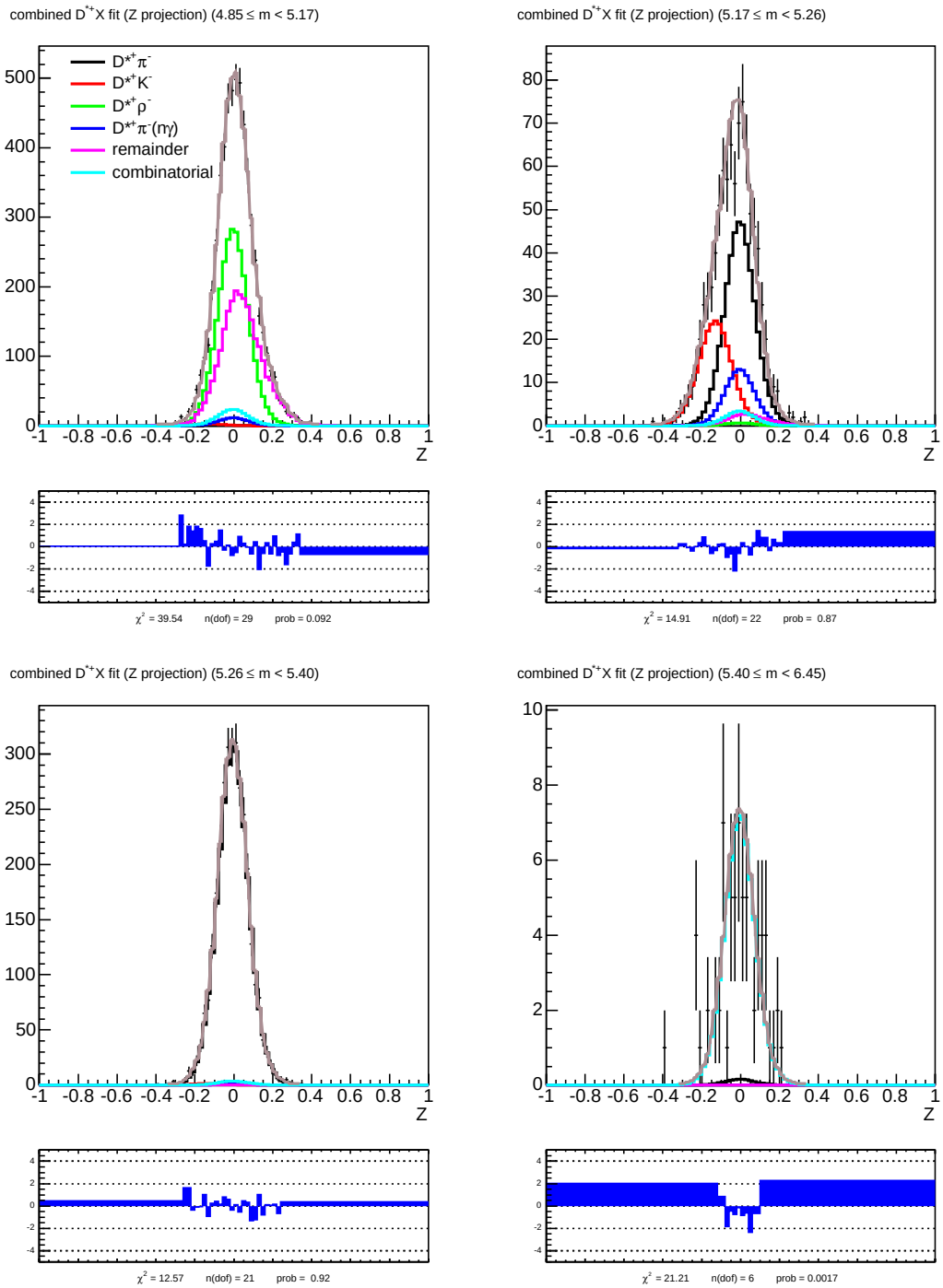

Figure 5.37: Z projection of the likelihood fit for $\overline{B}^0 \to D^{*+}X$ in the combined 1.2 fb^{-1}. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

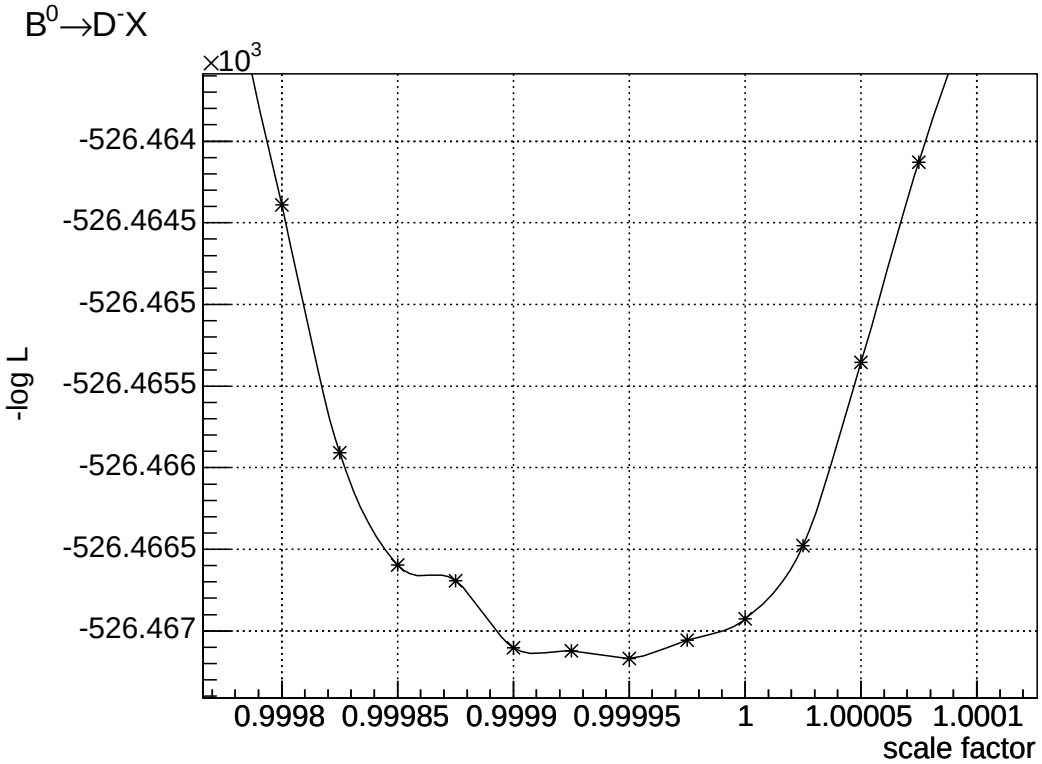

Figure 5.38: Scale factor scan for $\overline{B}^0 \to D^+X$. The scale factor plotted is relative to the default scale factor of 1.00025. The vertical grid cells correspond to $\Delta \log L = 0.5$.

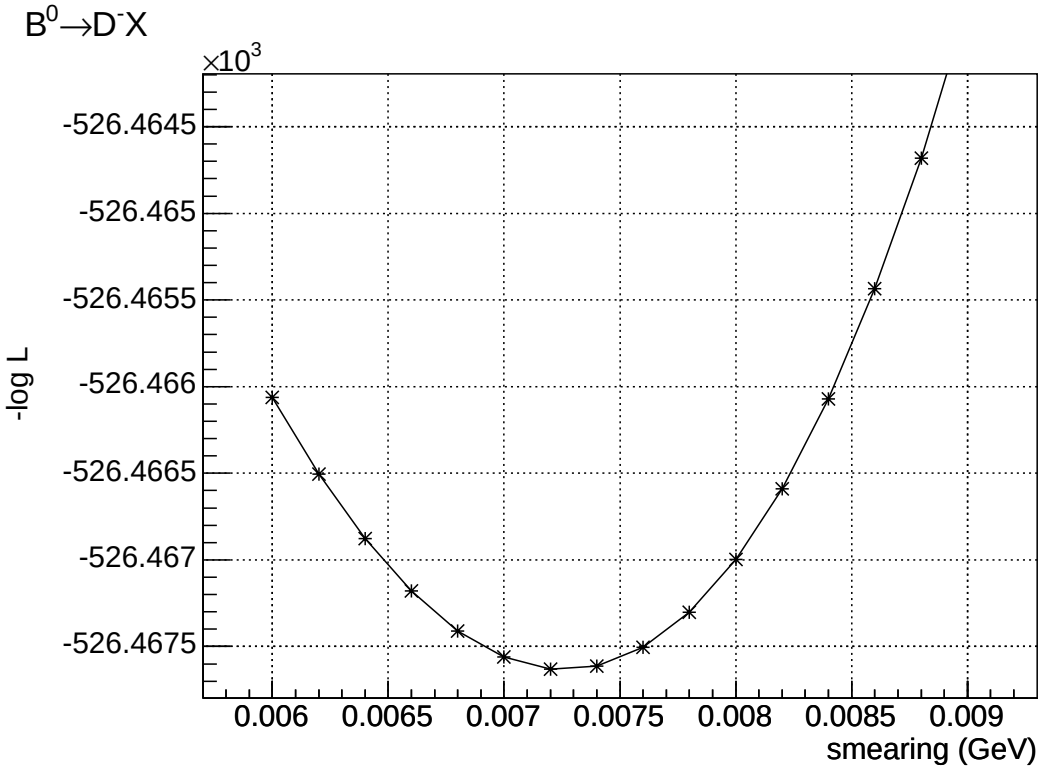

Figure 5.39: Smearing scan for $\overline{B}^0 \to D^+X$. The vertical grid cells correspond to $\Delta \log L = 0.5$.

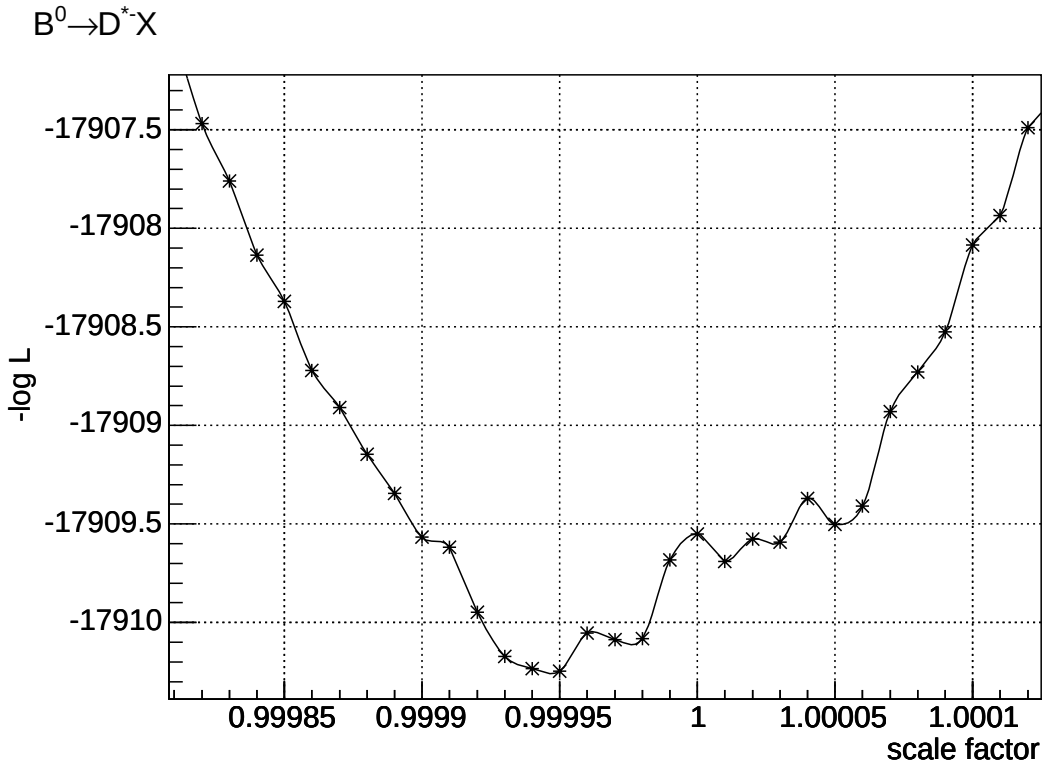

Figure 5.40: Scale factor scan for $\overline{B}^0 \to D^{*+}X$. The scale factor plotted is relative to the default scale factor of 1.00025. The vertical grid cells correspond to $\Delta \log L = 0.5$.

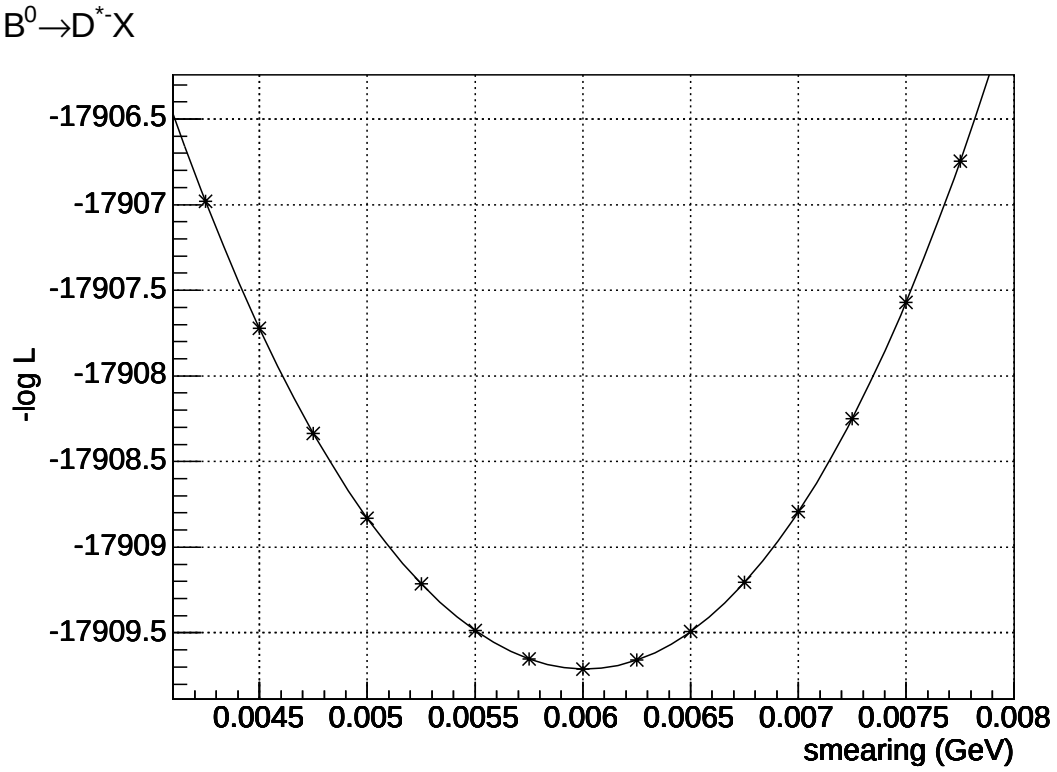

Figure 5.41: Smearing scan for $\overline{B}^0 \to D^{*+}X$. The vertical grid cells correspond to $\Delta \log L = 0.5$.

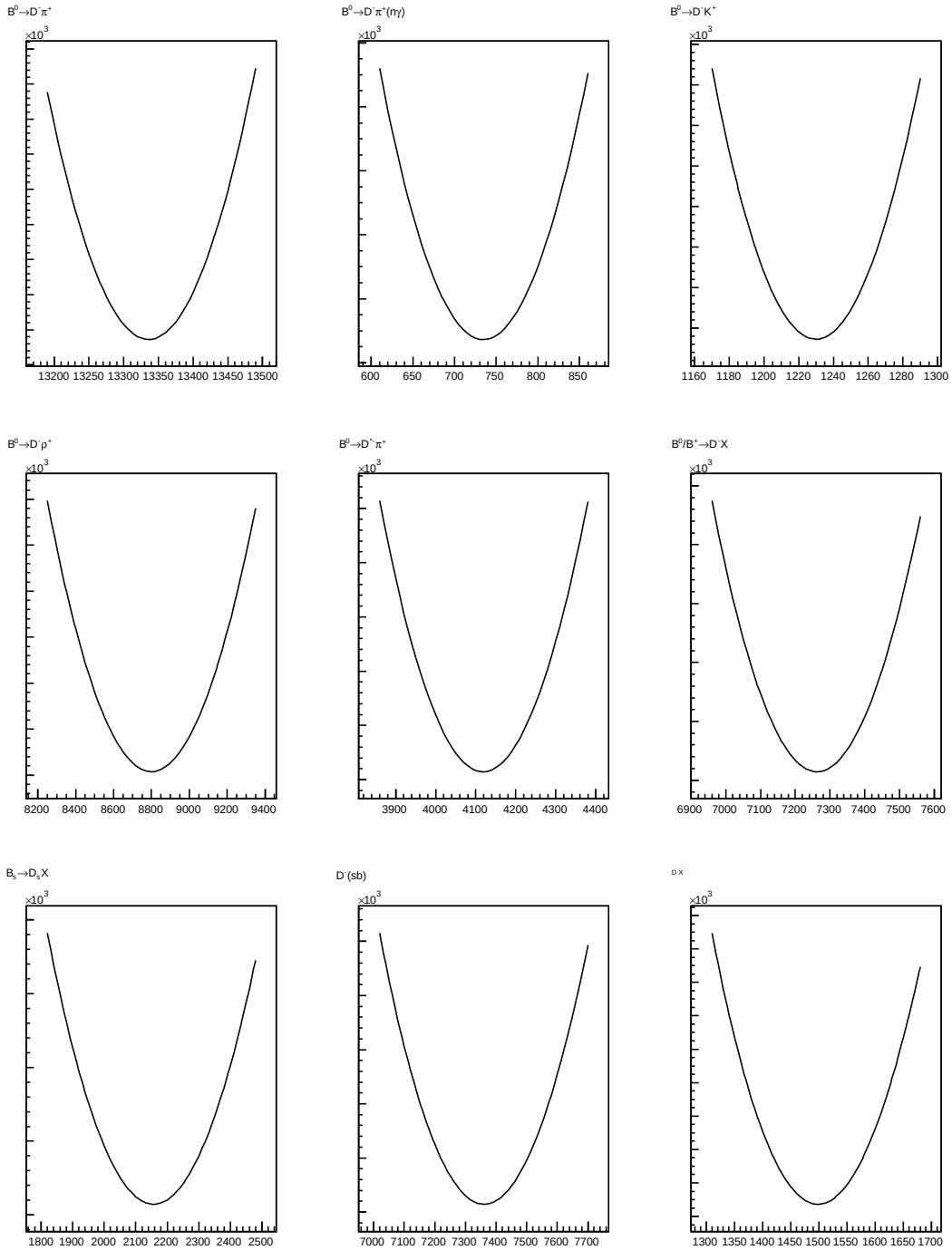

Figure 5.42: Likelihood scans for $\overline{B}^0 \to D^+\pi^-$. The candidate yield in each fit component is scanned.

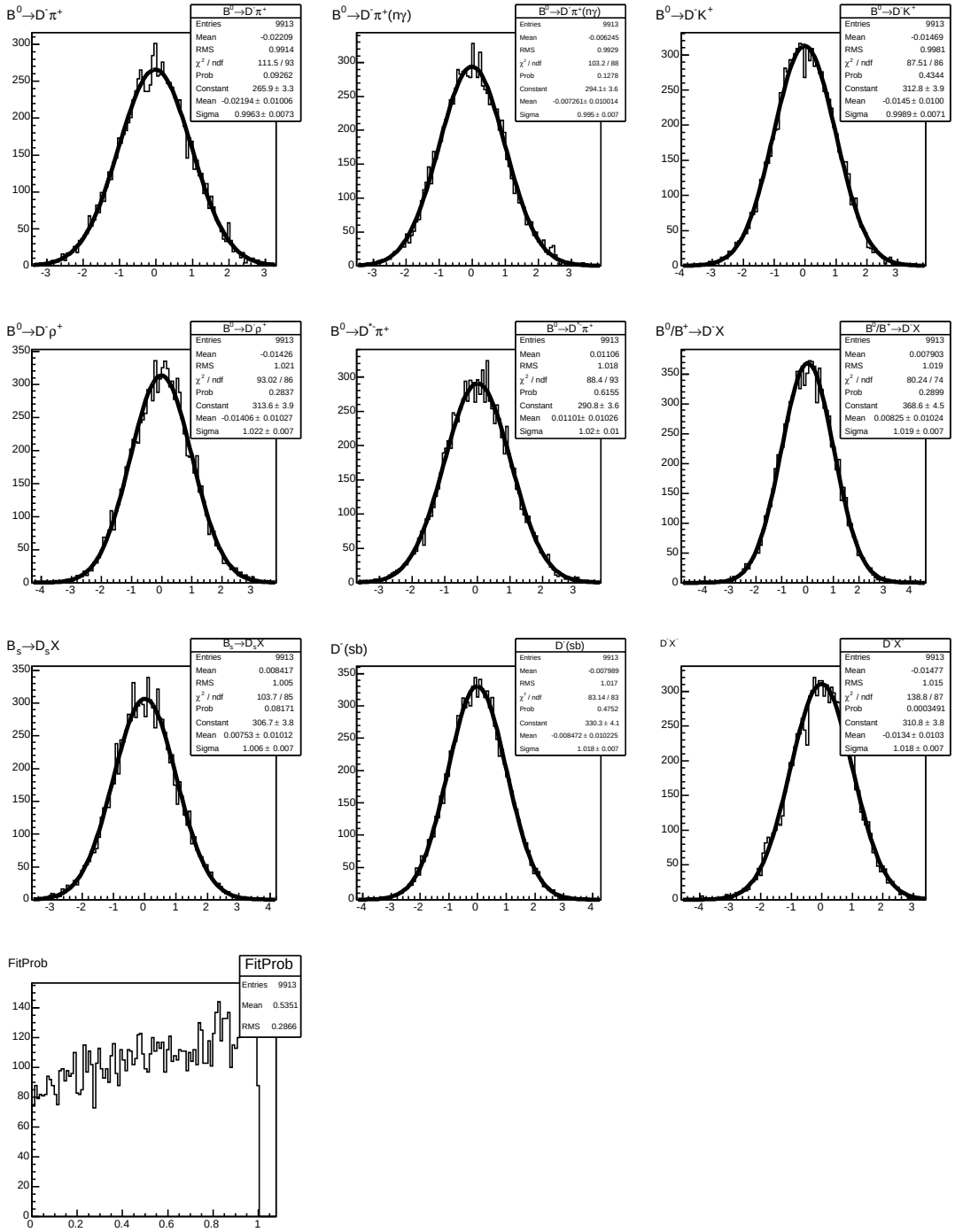

Figure 5.43: Pull distributions for $\overline{B}^0 \to D^+X$

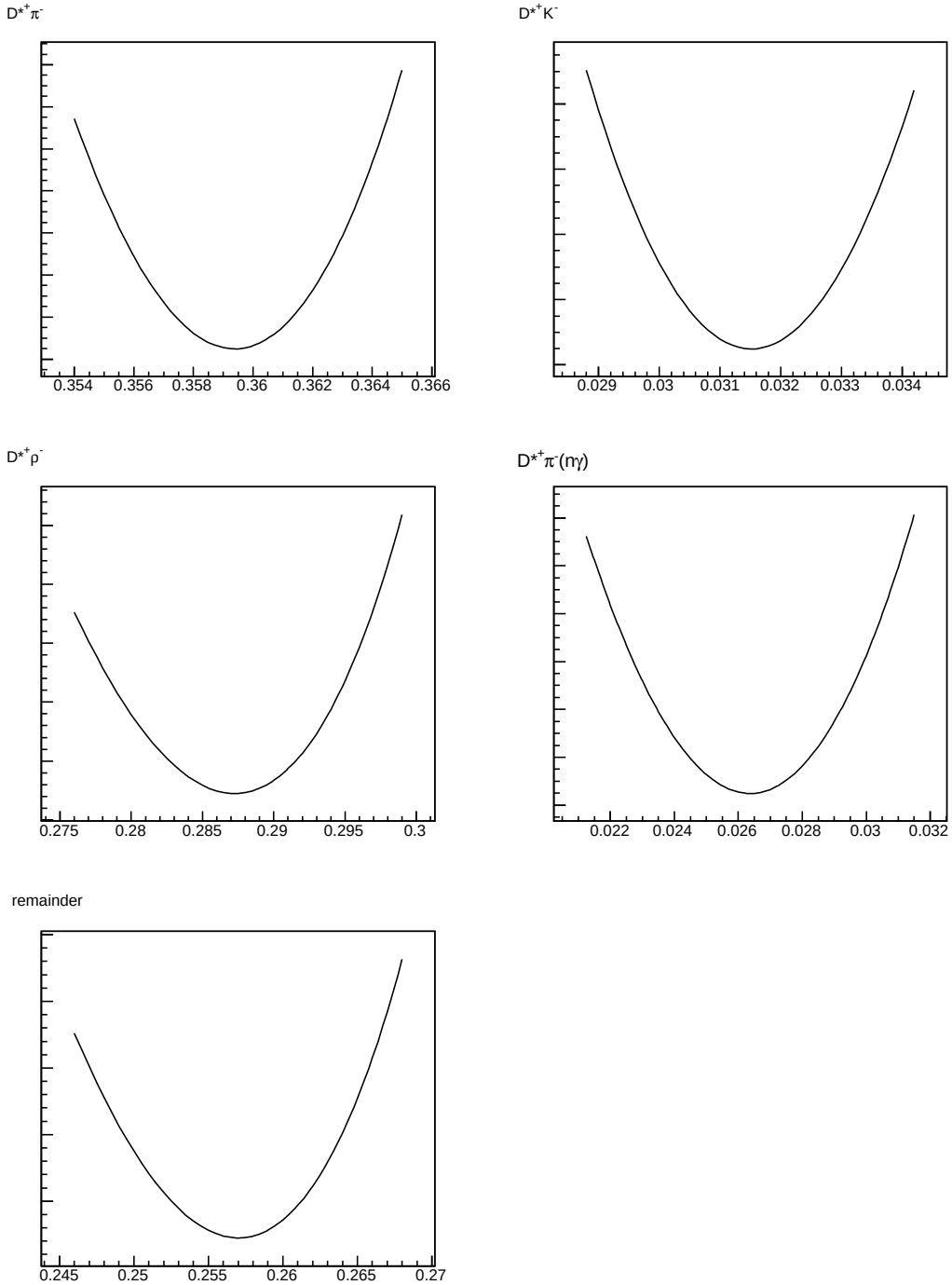

Figure 5.44: Likelihood scans for $\overline{B}^0 \to D^{*+}X$. The fractional candidate yield in each fit component is scanned.

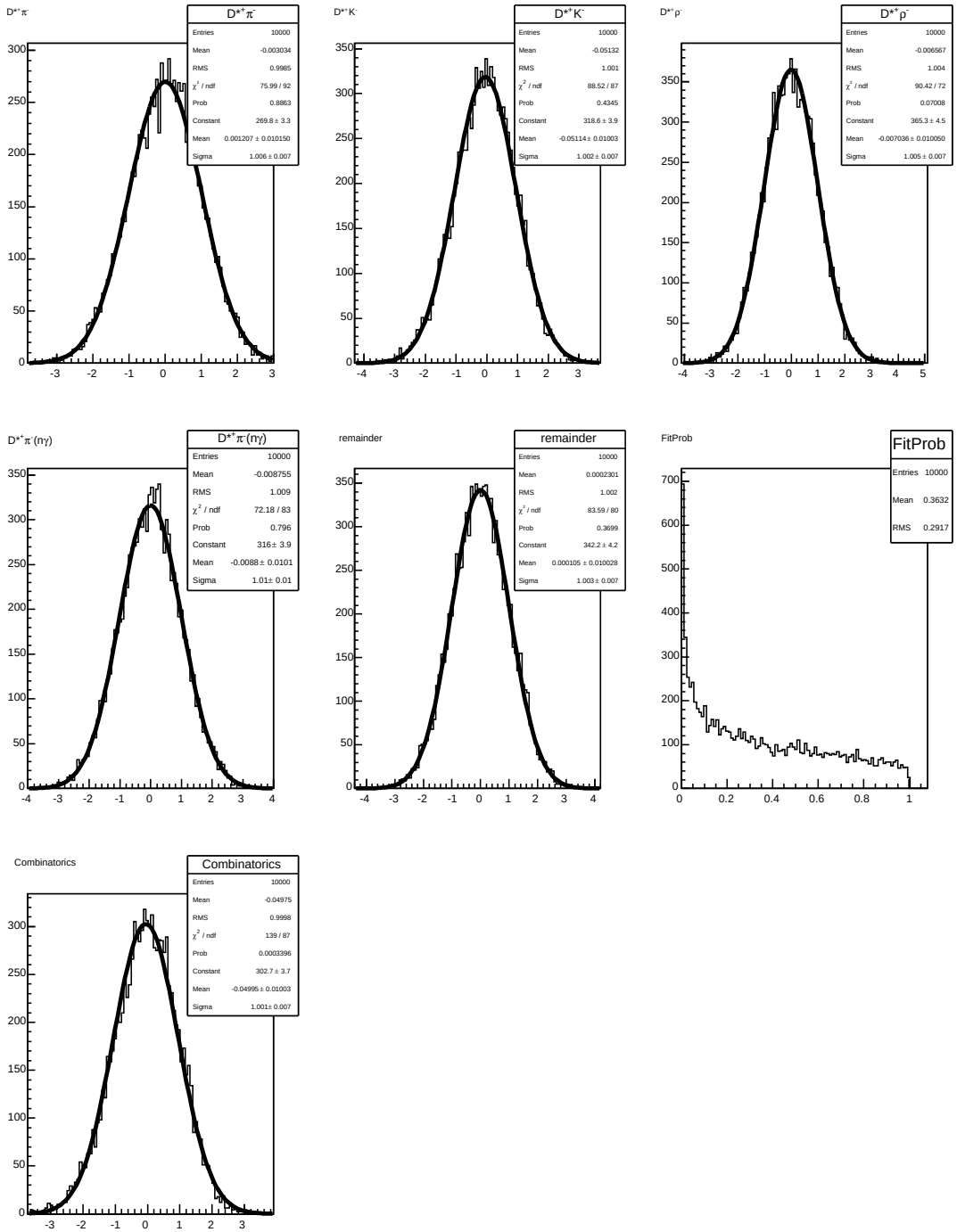

Figure 5.45: Pull distributions for $\overline{B}^0 \to D^{*+}X$

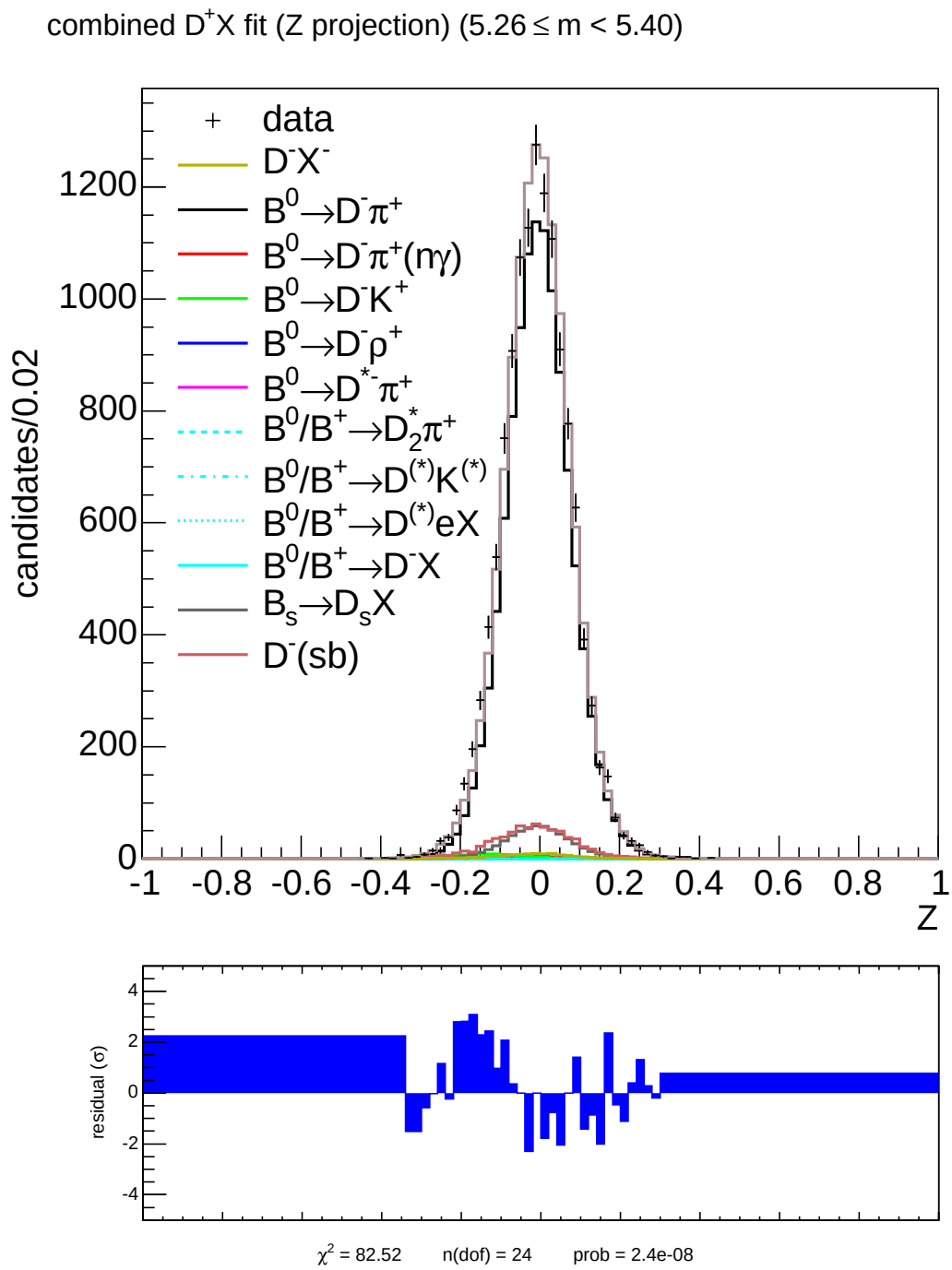

Figure 5.46: Z projection of the likelihood fit for $\overline{B}^0 \to D^+X$ in the combined 1.2 fb^{-1}, showing only the $D\pi$ mass region. (The remaining mass regions can be found in Figure 5.33.) The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

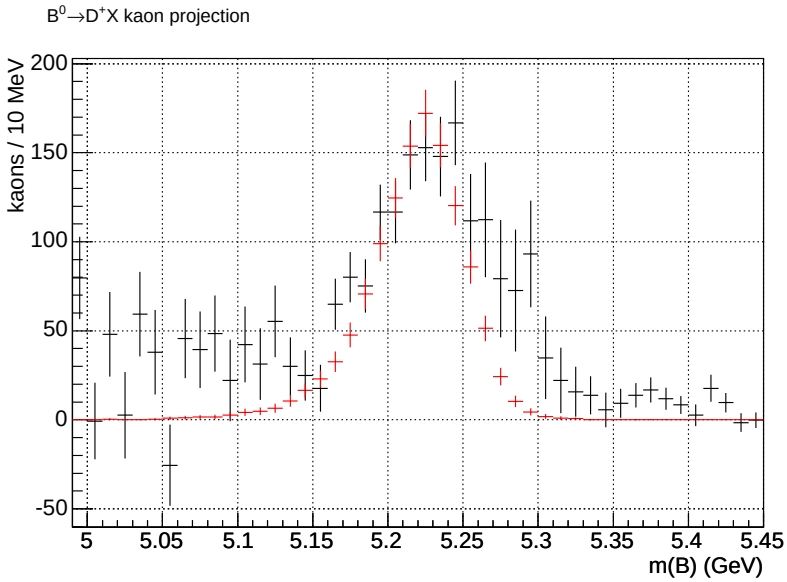

Figure 5.47: Kaon projection of the $\overline{B}^0 \to D^+X$ data sample (black) and $\overline{B}^0 \to D^+K^-$ mass shape scaled by the $\overline{B}^0 \to D^+K^-$ yield determined by the fit (red). (The kaon projection procedure is described in Section 5.5.) An excess of "kaon" tracks above the $\overline{B}^0 \to D^+K^-$ mass region can be seen.

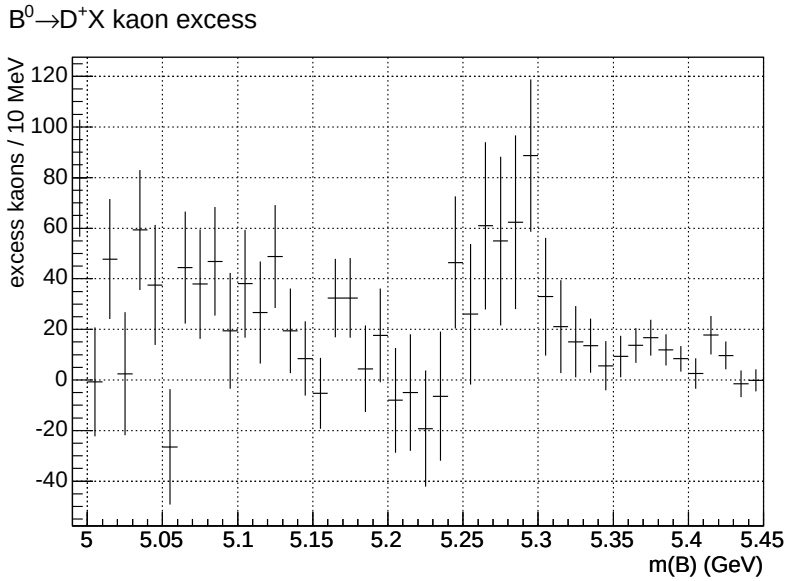

Figure 5.48: The excess of "kaon" tracks in $\overline{B}^0 \to D^+X$ data over the $\overline{B}^0 \to D^+X$ fit is shown by subtracting the data and fit distributions in the previous figure; the excess coincides with the $\overline{B}^0 \to D^+\pi^-$ mass.

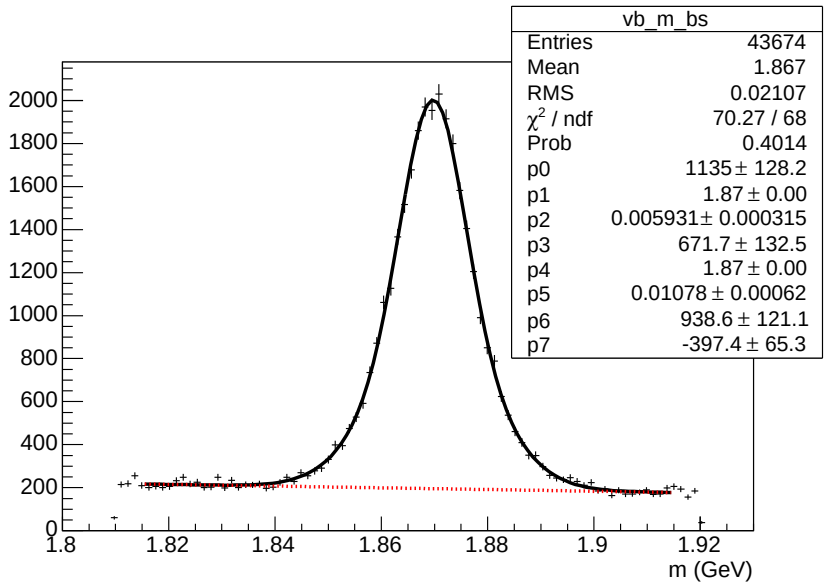

Figure 5.49: Fit to determine the fake-D^+ yield. The first six fit parameters are the amplitudes, mean and widths of the double-gaussian D^+ signal model; the remaining parameters are the intercept (p_6) and slope (p_7) of the straight-line combinatorial background model.

Chapter 6

$\overline{B}_s^0 \to D_s^+ X$ signal sample

In the previous chapter, we verified that our analysis method gives results consistent with the world average in our $\overline{B}^0 \to D^+ X$ and $\overline{B}^0 \to D^{*+} X$ control samples. We are now ready to apply the method to the $\overline{B}_s^0 \to D_s^+ X$ signal sample. This sample is reconstructed as $\overline{B}_s^0 \to D_s^+ \pi^-$, with $D_s^+ \to \phi(K^- K^+)\pi^+$. We will list the specifics of the $\overline{B}_s^0 \to D_s^+ X$ sample within the general analysis method described in Chapter 4.

6.1 Selection of candidates

Our starting point is the skimmed datasets described in Section 4.4. A list of skimming cuts is found in Table 6.1. We apply selection cuts to the $\overline{B}_s^0 \to D_s^+ \pi^-$ candidates that are almost identical to the B^0 analogue $\overline{B}^0 \to D^+ \pi^-$. There are two differences between the $\overline{B}_s^0 \to D_s^+ \pi^-$ cuts and the $\overline{B}^0 \to D^+ \pi^-$ cuts. The first is that a ϕ mass cut is applied, which provides powerful background rejection due to the narrowness of the ϕ resonance. The second difference is that in the D_s^+ case no Λ_c^+ veto is imposed. This is done in the interest of signal efficiency, because a large fraction of the D_s^+ signal reflects into the Λ_c^+ mass; because of the ϕ mass cut, the Λ_c^+ background in $D_s^+ \to \phi\pi^+$ is smaller than in $D^+ \to K^- \pi^+ \pi^-$. The complete set of analysis cuts is shown in Table 6.2; for comparison, the $\overline{B}^0 \to D^+ X$ cuts are also listed. The mass distribution for the $\overline{B}_s^0 \to D_s^\pm X$ candidates with these selections is shown in Figure 6.1. Consult Section 4.5 for a description of the cut variables.

6.2 Fit configuration

The fit is a non-extended maximum-likelihood fit using the likelihood function $(4 \cdot 3)$. The $\overline{B}_s^0 \to D_s^+ \pi^-$ candidate reconstruction is performed without a mass constraint on the D candidate. (This choice stems from our modeling of the fake-D background; see below.)

Table 6.3 lists the fit components. Apart from the $\overline{B}_s^0 \to D_s^+ \pi^-$ and $\overline{B}_s^0 \to D_s^\pm K^\mp$ signals, this list includes a number of single-B backgrounds from B_s^0 decays as well as reflections and crossfeeds from other B hadrons. With a few exceptions, the partially reconstructed B_s^0 modes are collected in the "remainder" $B_s \to D_s^{(*)} X$ template. $D_s^+ \rho^-$ and $D_s^{*+} \pi^-$ are split out because these components leak into the $D_s K$ mass region; the normalization of both modes floats in the fit.

	$\overline{B}_s^0 \to D_s^+ \pi^-$
$m(\phi)$	$980 < m/\text{MeV} < 1030$
$m(D)$	$1910 < m/\text{MeV} < 2030$
$\Delta\phi(D)$	< 1.5
$\Delta R(D)$	< 2
$\Delta z(D)$	< 5 cm
$\sum p_T(D)$	> 2.8 GeV
$p_T(D)$	> 2.5 GeV
$L_{xy}(D)$	> 200 µm
D SVT confirmation	one track
$\chi^2_{xy}(D)$	< 30
$m(B)$	$4650 < m/\text{MeV} < 6500$
$\Delta\phi(B)$	< 3
$\Delta R(B)$	< 2
$\Delta z(B)$	< 5 cm
$\sum p_T(B)$	> 5.5 GeV
$p_T(B)$	> 5 GeV
$L_{xy}(B)$	> 300 µm
$\lvert d_0(B)\rvert$	< 100 µm
B SVT confirmation	`SCEN_LOW`
$\chi^2_{xy}(B)$	< 30

Table 6.1: Skimming cuts for $\overline{B}_s^0 \to D_s^+ \pi^-$ candidates

	$\overline{B}^0 \to D^+\pi^-$	$\overline{B}^0_s \to D^+_s\pi^-$
$\chi^2_{xy}(B)$	< 10	< 10
$\chi^2_{xy}(D)$	< 15	< 15
$L_{xy}(B)/\sigma_{L_{xy}}(B)$	> 8	> 8
$L_{xy}(B \leftarrow D)$	> 0	> 0
$\lvert d_0(B)\rvert$	< 60 µm	< 60 µm
$p_T(B)$	> 5.5 GeV	> 5.5 GeV
$p_T(\pi_B)$	> 2.0 GeV	> 2.0 GeV
$\Delta R(D, \pi_B)$	< 1.5	< 1.5
$\lvert z_0(B)/\sigma_{z_0}(B)\rvert$	< 3	< 3
iso(B) $(\Delta R < 1)$	> 0.5	> 0.5
$m(D^+)$	$\lvert m/\text{MeV} - 1869.4\rvert < 20$	
$m(D^+_s)$		$\lvert m/\text{MeV} - 1968.3\rvert < 20$
$m(\phi)$		$1013 < m/\text{MeV} < 1028$
π_B SVT match distance	< 25	< 25
π_B SVT χ^2	< 15	< 15
π_B SVT p_T	> 2 GeV	> 2 GeV
$m(K, \pi \mapsto p, \pi)$	$\lvert m/\text{MeV} - 2285\rvert > 250$	
$m(K\pi\pi) - m(K\pi)$	> 180 MeV	> 180 MeV
Z	$\neq -999$	$\neq -999$
Good run list	B_s mixing version 14	B_s mixing version 14

Table 6.2: Selection cuts for $\overline{B}^0_s \to D^+_s\pi^-$ candidates

The $D_s^{(*)}K^{(*)}$ modes are the largest group of modes in which the track recoiling against the D_s^+ is a kaon; we split this group from the remainder template so we can give it a kaon dE/dx template. Likewise, we split out the $D_s^+e^\pm X$ decays and give them an electron dE/dx template. The $D_s^{(*)}K^{(*)}$ and $D_s^+e^\pm X$ components that are "split out" of the remainder template do not have floating normalizations but are instead fixed to the remainder template with ratios determined from the MC.

The fully reconstructed decays $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^-K^+$ give localized peaks in the mass plot. Since a measurement of the relative branching ratios of $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ has recently been published [42], we include templates for these decays. The templates are split according to the species recoiling against the $D_s^{(*)}$ with ratios fixed to the values in [42].

A template for $\overline{B}^0/B^0 \to D^+(K^-\pi^+\pi^+)X$ and $B^-/B^+ \to D^+(K^-\pi^+\pi^+)X$ reflections is also included, as is a template for $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)X$ reflections. The "parameter" column of Table 6.3 summarizes the connections between the various templates.

Mass and Z templates are determined in a way analogous to the $\overline{B}^0 \to D^+X$ fit. One exception to the analogy is that the $\overline{B}_s^0 \to D_s^+X$ fit does not use a real-D background component; the low statistics in the wrong-sign $D_s^+\pi^+$ sample makes it difficult to extract real-D pdf's, and the low statistics in the $\overline{B}_s^0 \to D_s^+\pi^-$ high mass region makes it hard for the fit to determine the relative normalization between two combinatorial backgrounds. A systematic uncertainty is assessed to cover the combinatorial-background shape modeling.

Single-B mass templates Table 6.4 lists the single-B Monte Carlo samples from which the mass templates are generated. These Monte Carlo samples are sorted into templates by the true Monte Carlo decay mode according the list in the previous paragraph. The mass distribution of the B decay modes contributing to these samples is shown in Figures 6.2–6.7. As in the control-sample case, it is necessary to scale and smear the Monte Carlo mass distributions to match the data. We use the same values as in the control samples, and we will assess a systematic uncertainty by varying these values within their errors.

Fake-D mass template The fake-D mass template is obtained from the D_s^+ sidebands, using the signal-subtraction procedure from Section 4.8. The invariant mass under the $D_s^+\pi^-$ mass assumption is calculated in the sidebands of the D_s^+ invariant mass plot: 1.924 GeV $< m(D_s^+) <$ 1.939 GeV or 1.999 GeV $<$ $m(D_s^+) <$ 2.014 GeV. Signal subtraction is performed on this data. To do this, the data is fit with two components, the first of which is a function of the form $(4 \cdot 6)$ and the second of which is a histogram obtained from the $\overline{B}_s^0 \to D_s^+(\phi\pi^+)X$ MC selecting on the D_s^+ sideband. Figure 6.10 shows the results of this fit. The shape parameters obtained from this auxiliary fit are fixed in the main $\overline{B}_s^0 \to D_s^+X$ fit. Because the fake-D_s fit gives better agreement with data when performed on candidates that are not mass-constrained, we use unconstrained candidates for all fit components.

Single-B Z templates Z templates marked "prompt D^*" are described in Section 4.9. Two-dimensional reweighting in secance (s) and momentum (p) is applied. The DeX template uses the parametric dE/dx simulation described in Section 4.9. The D^+ and Λ_c^+ reflection templates are composed from e, μ (using π as proxy), π and K templates according to species fractions determined from D^+ and Λ_c^+ Monte Carlo samples. The reflection templates are divided into four mass regions: the region below the D_sK (4.85 GeV $< m(B_s) <$ 5.26 GeV); the D_sK region (5.26 GeV $< m(B_s) <$ 5.35 GeV); the $D_s\pi$ region (5.35 GeV $< m(B_s) <$ 5.5 GeV); the region above the $D_s\pi$ (5.5 GeV $< m(B_s) <$ 6.45 GeV).

Fake-D Z template The Z template for the fake-D component is taken directly from the D_s^+ sideband data. Signal subtraction is performed on the fake-D template to account for the leakage of poorly measured real D_s^+ into the D_s^+ sidebands.

6.3 Results

Results of the likelihood fit are given in Table 6.5. The correlation coefficients for the combined dataset are listed in Table 5.6. Figure 6.13 shows the likelihood profiles in all fit parameters. The mass projection is shown in Figure 6.11; the Z projections are shown in Figure 6.12. We divide the Z projections into four mass regions: the region below the D_sK (4.85 GeV $< m(B_s) <$ 5.26 GeV); the D_sK region (5.26 GeV $< m(B_s) <$ 5.35 GeV); the $D_s\pi$ region (5.35 GeV $< m(B_s) <$ 5.5 GeV); the region above the $D_s\pi$ (5.5 GeV $< m(B_s) <$ 6.45 GeV). This is done so that the modeling of the combinatorial background Z distribution and the modeling of the pion Z distribution can be tested in the mass regions where these components dominate; we isolate the D_sK region to verify that the Z fit is good in the mass region we rely on for the final analysis result.

To calculate $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$ from the event yields returned by the fit, we apply the corrections discussed in Section 4.10. We determine a kinematic efficiency correction $(\varepsilon_K\phi_K)/(\varepsilon_\pi\phi_\pi) = 0.997$ and an XFT correction $\eta_\pi/\eta_K = 1.0483\pm 0.0283$. After corrections, we find

$$\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big) = 0.107 \pm 0.019(\text{stat}).$$

The denominator contains both the $\overline{B}_s^0 \to D_s^+\pi^-$ main peak and the radiative tail, whose size we measure to be 6.9%±7.0% relative to the main peak; the statistical error on $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$ takes into account the correlations between the candidate yields. We will estimate systematic uncertainties in Section 6.5.

To determine the statistical significance of the observation of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ signal, we perform a fit in which the $\overline{B}_s^0 \to D_s^\pm K^\mp$ fraction is fixed to zero. This fit is worse than the default fit by $\Delta \log L = -31.21$; the corresponding statistical significance of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ signal is 7.9σ.

Decay modes	Mass pdf	dE/dx pdf	Fit parameters
$B_s \to D_s\pi$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$B_s \to D_s\pi(n\gamma)$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$B_s \to D_sK$	MC	prompt $D^* \to K$, $s:p$ reweight	norm floating
$B_s \to D_s\rho$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$B_s \to D_s^*\pi$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$B^0/B^+ \to D^+X$	MC	species fractions from MC	norm floating
$B_s \to D_s^{(*)}K^{(*)}$	MC	prompt $D^* \to K$, $s:p$ reweight	norm fixed to $B_s \to D_s^{(*)}X$
$B_s \to D_s^{(*)}eX$	MC	parametric e simulation	norm fixed to $B_s \to D_s^{(*)}X$
$B_s \to D_s^{(*)}X$	MC	species fractions from MC	norm floating
$B^0 \to D_s^{(*)}K$	MC	prompt $D^* \to K$, $s:p$ reweight	norm fixed to $B^0 \to D_s^{(*)}\pi$
$B^0 \to D_s^{(*)}\pi$	MC	prompt $D^* \to \pi$, $s:p$ reweight	norm floating
$\Lambda_b \to \Lambda_cX$	MC	species fractions from MC	norm floating
combinatorial bg	D_s^+ SB	D_s^+ SB	norm $= 1 - \sum_i f_i$, shape fixed

Table 6.3: Fit components for $\overline{B}_s^0 \to D_s^+X$ fit. See the text for a description of the "mass pdf" and "dE/dx pdf" categories.

Channels	Sample size
$\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-(n\gamma)$	50k
$\overline{B}_s^0/B_s^0 \to D_s^+(\phi\pi^+)X$	100k
$\overline{B}^0/B^0 \to D^+X$	200
$B^-/B^+ \to D^+X$	20
$\overline{B}^0 \to D_s^{(*)+}(\phi\pi^+(\gamma/\pi^0))X$	10k
$\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)X$	800

Table 6.4: Monte Carlo samples used in this analysis. All samples are reconstructed as $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$.

Decay modes	Fit result
$B_s \to D_s\pi$	0.201 ± 0.008
$B_s \to D_s\pi(n\gamma)$	0.014 ± 0.014
$B_s \to D_sK$	0.021 ± 0.004
$B_s \to D_s\rho$	0.234 ± 0.036
$B_s \to D_s^*\pi$	0.096 ± 0.025
$B^0/B^+ \to D^+X$	0.051 ± 0.022
$B_s \to D_s^{(*)}K^{(*)}$	fixed to $B_s \to D_s^{(*)}X$
$B_s \to D_s^{(*)}eX$	fixed to $B_s \to D_s^{(*)}X$
$B_s \to D_s^{(*)}X$	0.180 ± 0.018
$B^0 \to D_s^{(*)}K$	fixed to $B^0 \to D_s^{(*)}\pi$
$B^0 \to D_s^{(*)}\pi$	0.039 ± 0.010
$\Lambda_b \to \Lambda_cX$	0.040 ± 0.017
combinatorial bg	norm $= 1 - \sum_i f_i$

Table 6.5: Fit parameters determined for $\overline{B}_s^0 \to D_s^+X$ by the combined mass and Z fit

Table 6.6: Correlation coefficients for $\overline{B}_s^0 \to D_s^+ X$

Parameter		Correlation coefficients								
	global	2	3	4	5	6	7	10	12	13
2	0.67160	1.000	0.149	0.003	0.131	−0.064	−0.510	0.113	−0.037	−0.157
3	0.89273	0.149	1.000	−0.196	0.587	−0.607	−0.628	−0.147	−0.220	−0.096
4	0.50211	0.003	−0.196	1.000	0.015	0.062	−0.117	0.059	−0.118	−0.001
5	0.96691	0.131	0.587	0.015	1.000	−0.860	−0.456	−0.621	−0.007	−0.140
6	0.95539	−0.064	−0.607	0.062	−0.860	1.000	0.336	0.575	−0.188	−0.085
7	0.91564	−0.510	−0.628	−0.117	−0.456	0.336	1.000	−0.124	0.167	0.007
10	0.86145	0.113	−0.147	0.059	−0.621	0.575	−0.124	1.000	−0.161	−0.161
12	0.72877	−0.037	−0.220	−0.118	−0.007	−0.188	0.167	−0.161	1.000	−0.067
13	0.79601	−0.157	−0.096	−0.001	−0.140	−0.085	0.007	−0.161	−0.067	1.000

Channels	Pull mean (%)	Pull width (%)
$B_s \to D_s\pi$	1.102 $\pm$ 1.006	99.354 $\pm$ 0.712
$B_s \to D_s\pi(n\gamma)$	1.788 $\pm$ 1.015	100.059 $\pm$ 0.726
$B_s \to D_sK$	-4.116 ± 1.021	100.712 $\pm$ 0.728
$B_s \to D_s\rho$	1.079 $\pm$ 1.000	98.613 $\pm$ 0.714
$B_s \to D_s^*\pi$	-1.287 ± 1.010	99.622 $\pm$ 0.720
$B^0/B^+ \to D^+X$	-3.409 ± 1.012	99.922 $\pm$ 0.719
$B_s \to D_s^{(*)}X$	-0.387 ± 1.019	100.124 $\pm$ 0.737
$B^0 \to D_s^{(*)}\pi$	-1.931 ± 1.025	101.128 $\pm$ 0.731
$\Lambda_b \to \Lambda_cX$	-5.179 ± 1.042	102.951 $\pm$ 0.739
combinatorial background	-3.195 ± 1.023	100.700 $\pm$ 0.737

Table 6.7: Pull distributions for $\overline{B}_s^0 \to D_s^+X$ (5102 events per pseudoexperiment)

6.4 Validation

To validate the fitter, we plot the pull distributions of the fit parameters. Data is generated by throwing dice against the mass and dE/dx templates with the f_j determined in the fit as the "true" f_j, yielding a generated number of candidates that is Poisson-distributed about f_j. The data is then fit. Figure 6.14 and Table 6.7 show the pulls for 10000 pseudoexperiments of 5102 events (the size of our dataset) each. The combinatorial background fraction is not a fit parameter, and so its error is not returned by the fit; instead we calculate it by propagating the errors on the fit parameters as follows:

$$\sigma_{\text{bg}}^2 = \sum_{i,j}^{M-1} \sigma_{ij}, \tag{6 · 1}$$

where M is the number of fit components including the combinatorial background.

Because of biases at the 5% level evident in these pull studies, we perform another study to show that the fitter behaves correctly in the asymptotic limit. Table 6.8 shows the pulls for 10000 pseudoexperiments of 51020 events each; no significant biases are observed in this study. The biases observed in the default-size Monte Carlo are treated as a systematic uncertainty below.

We also check that the fit result in the non-extended fit configuration gives the same result as an extended-likelihood fit (see Section 4.7). We find that the results differ at the 10^{-6} level (relative) in the $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$ ratio and at the $2 \cdot 10^{-3}$ level (also relative) in the error on $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$.

Channels	Pull mean (%)	Pull width (%)
$B_s \to D_s\pi$	0.608 ± 1.079	100.970 ± 0.773
$B_s \to D_s\pi(n\gamma)$	-0.943 ± 1.080	100.937 ± 0.779
$B_s \to D_sK$	-1.443 ± 1.072	100.364 ± 0.766
$B_s \to D_s\rho$	-0.480 ± 1.046	98.019 ± 0.745
$B_s \to D_s^*\pi$	-0.292 ± 1.045	97.942 ± 0.743
$B^0/B^+ \to D^+X$	-0.070 ± 1.067	99.878 ± 0.762
$B_s \to D_s^{(*)}X$	0.997 ± 1.045	97.986 ± 0.742
$B^0 \to D_s^{(*)}\pi$	0.998 ± 1.077	100.838 ± 0.769
$\Lambda_b \to \Lambda_cX$	-2.183 ± 1.066	99.967 ± 0.757
combinatorial background	-0.796 ± 1.058	98.953 ± 0.758

Table 6.8: Pull distributions for $\overline{B}_s^0 \to D_s^+X$ (51020 events per pseudoexperiment)

6.5 Systematic uncertainties

In this section we describe the various studies performed to determine the systematic uncertainties on $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$. The method for determining the systematic uncertainties is the one described in Section 4.12. Table 6.9 summarizes the systematic uncertainties.

Combinatorial background model From the high-statistics $\overline{B}^0 \to D^+X$ fit, we know that a combinatorial background model comprising two sources of background (fake D + track, real D + track) is a better model than a single-source background; because we are limited in the $D_s^+ \to \phi\pi^+$ case by the wrong-sign ($D_s^+\pi^+$) sample statistics, we fit without a real-D background component. We assess the systematic uncertainty associated with the background model by comparing different background models. The default model uses a single background of the form $(4 \cdot 6)$ with shape parameters determined from the fit of the D_s^+ sidebands described in Section 6.2. We compare this configuration to a two-background configuration in which the real-D background shape parameters and Z distribution from the $\overline{B}^0 \to D^+X$ fit are also used.

Monte Carlo smearing and scaling As explained in Section 5.6, it is necessary to scale the Monte Carlo templates by a factor of 1.00025 ± 0.00012 and smear them with gaussian resolution function of width 6.43 ± 2.09 MeV to achieve agreement with the data. We assess the systematic uncertainty associated with the errors on the scale factor by varying the scale factor up and down within its error. We use the greater of the changes in fit value resulting from an upward or downward variation as the systematic uncertainty due to the scale factor. The same method is employed to determine the systematic uncertainty due to the smearing.

$\Lambda_c X$, $D^+ X$ and fake-D_s species composition Three other sources of systematic uncertainty are a result of the species composition of the $\Lambda_c X$, $D^+ X$ and fake-D_s Z templates. We assess the systematic uncertainty due to the kaon content of these templates by comparing with pure-pion Z templates.

Choice of dE/dx reweighting variable As described in Section 4.9, the dE/dx templates need to be corrected for the chamber activity in the vicinity of the B daughter track. Appendix A describes in detail how we choose a proxy variable (secance) for the chamber activity and perform a secance-based reweighting. Secance is only one of several variables, all described in the appendix, that can be used for the reweighting. A variable which behaves almost as well is the total number of tracks N_{trks} in the event; this is also the most "global" proxy conceivable for the chamber activity. We assess a systematic uncertainty by fitting with an alternate model in which dE/dx reweighting in N_{trks} is used.

Z residuals in the $D_s^+\pi^-$-dominated region Finally we assess whether the residual Z-distribution discrepancies observed in the $\overline{B}^0 \to D^+ X$ sample exist in the $\overline{B}_s^0 \to D_s^+ X$ fit. We use the pion Z templates modified according to Section 5.7: the discrepancy between $\overline{B}^0 \to D^+ X$ data and fit in the $D^+\pi^-$-dominated region is added to the Z templates for $D_s^+\pi^-$ and $D_s^+\pi^-(n\gamma)$. When we perform a fit to the $\overline{B}_s^0 \to D_s^+(\phi\pi^+)X$ data, we find that the χ^2 of the Z projection in the $D_s^+\pi^-$-dominated region (5.35 GeV $< m(B_s^0) <$ 5.5 GeV) decreases from 21.04 to 19.59 (with 14 degrees of freedom). On the basis of this procedure, we cannot rule out the existence of a residual discrepancy like the one observed in the D^+ fit. Therefore we choose to asses the size of the systematic uncertainty that would be caused by the presence of such a discrepancy. We do so by generating toy MC with the modified templates and fitting with the default templates.

Fitter bias due to finite statistics As we have seen in Section 6.4, the fitter exhibits a small bias due to the finite statistics of the data sample. We take the systematic uncertainty due to this bias to be the difference in mean $\mathcal{B}\big(\overline{B}_s^0 \to D_s^\pm K^\mp\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^+\pi^-\big)$ between the toy Monte Carlo sample that exhibits the bias and the larger toy Monte Carlo sample that does not exhibit the bias.

Statistical fluctuations in the mass templates We derive mass pdf's from Monte Carlo samples of finite size. The statistical fluctuations in the mass pdf's are not taken into account by the likelihood function; instead we assess a systematic uncertainty here. We perform 10000 toy Monte Carlo experiments; before each experiment, we regenerate the mass templates by letting each bin vary within its Poisson error. From the width of the pull distribution in these 10000 pseudoexperiments, which is 17.4% of the expected statistical uncertainty of 2%, we determine that we need to add a systematic uncertainty of 0.35% to account for statistical fluctuations in the mass templates.

Effect	$D_s^{\pm}K^{\mp}/D_s^{+}\pi^{-}$ uncertainty (%)
Z reweighting in $N_{\text{trks}} : p$	0.45
Residual discrepancy in π Z template	0.38
Statistical fluctuations in the mass templates	0.35
Signal subtraction in fake-D Z template	0.29
Composition of $D^{+}X$ Z template	0.21
Combinatorial bg model	0.18
MC mass scale factor	0.18
MC mass smearing	0.13
Composition of $\Lambda_c X$ Z template	0.11
Fitter bias due to finite statistics	0.09
Sum in quadrature	0.84

Table 6.9: Systematic uncertainties on $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm}K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+}\pi^{-}\big)$

6.6 Summary

We have presented the first observation of the decay $\overline{B}_s^0 \to D_s^{\pm}K^{\mp}$; the statistical significance of the observation is 7.9σ. The $\overline{B}_s^0 \to D_s^{\pm}K^{\mp}$ event yield is 109 ± 19 (statistical error only). We use this sample to measure $\mathcal{B}\big(\overline{B}_s^0 \to D_s^{\pm}K^{\mp}\big)/\mathcal{B}\big(\overline{B}_s^0 \to D_s^{+}\pi^{-}\big) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$. This result is statistically consistent with the analogous relative B^0 branching fractions. Our result is statistics-limited because of the size of the $\overline{B}_s^0 \to D_s^{+}\pi^{-}$ sample but also because many of the reflection (Λ_c^{+}, D^{+}) and crossfeed ($\overline{B}^0 \to D_s^{\pm}X$) backgrounds overlap with our signal both in mass and in Z.

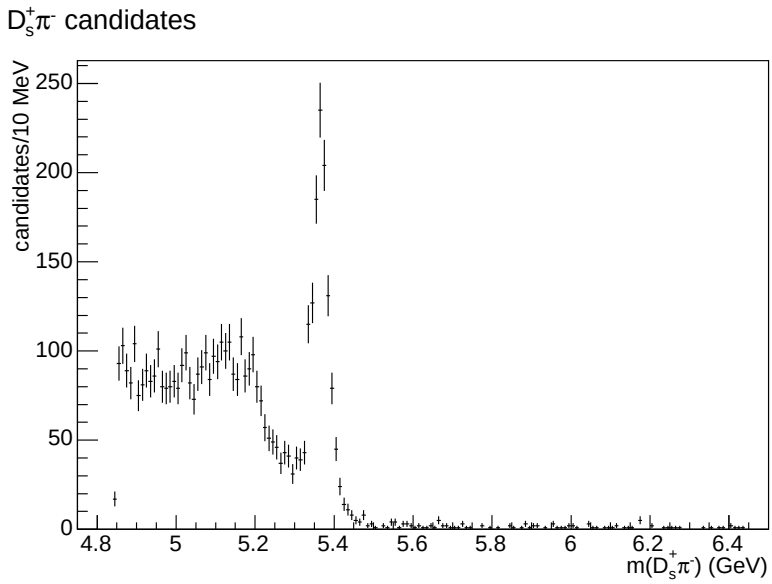

Figure 6.1: Mass distribution for $\overline{B}_s^0 \to D_s^+\pi^-$ candidates

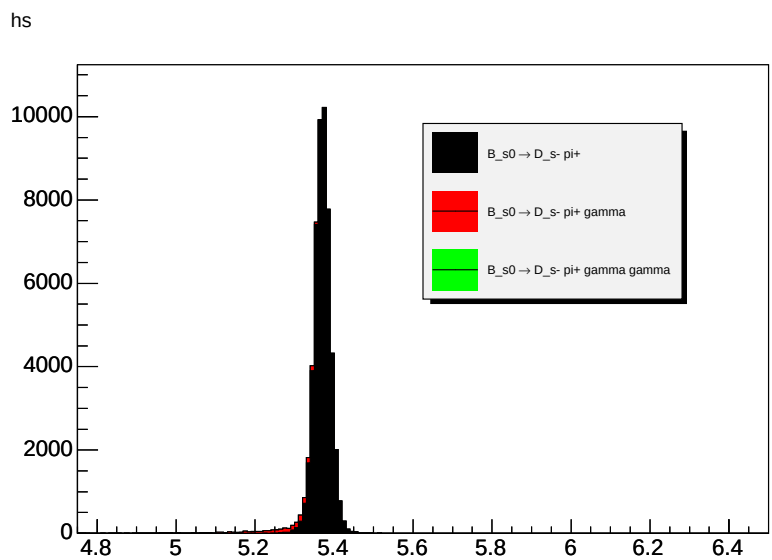

Figure 6.2: Composition of $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-(n\gamma)$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

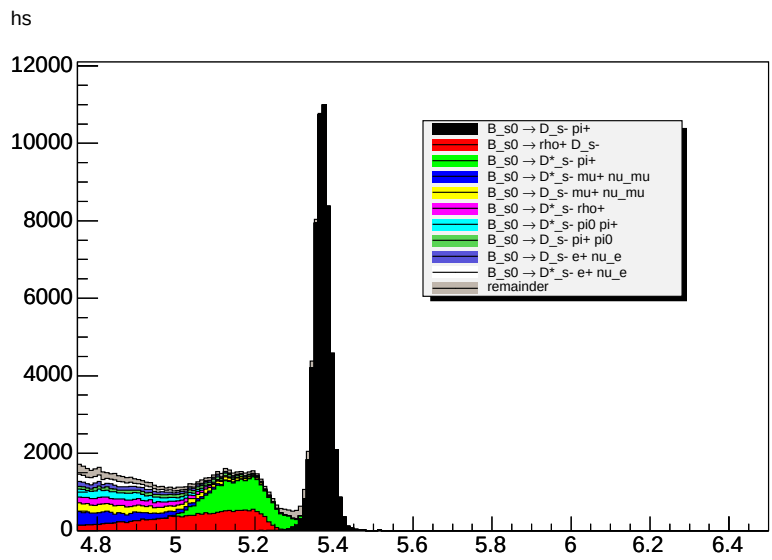

Figure 6.3: Composition of $\overline{B}_s^0/B_s^0 \to D_s^+(\phi\pi^+)X$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

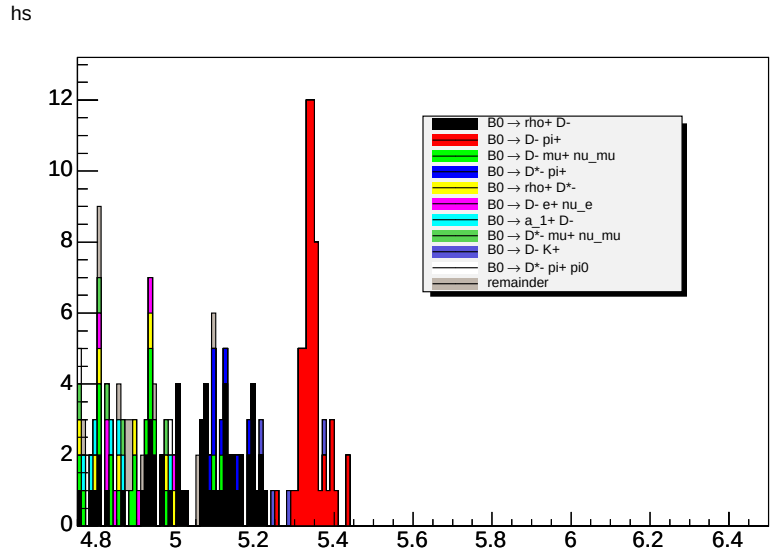

Figure 6.4: Composition of $\overline{B}^0/B^0 \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

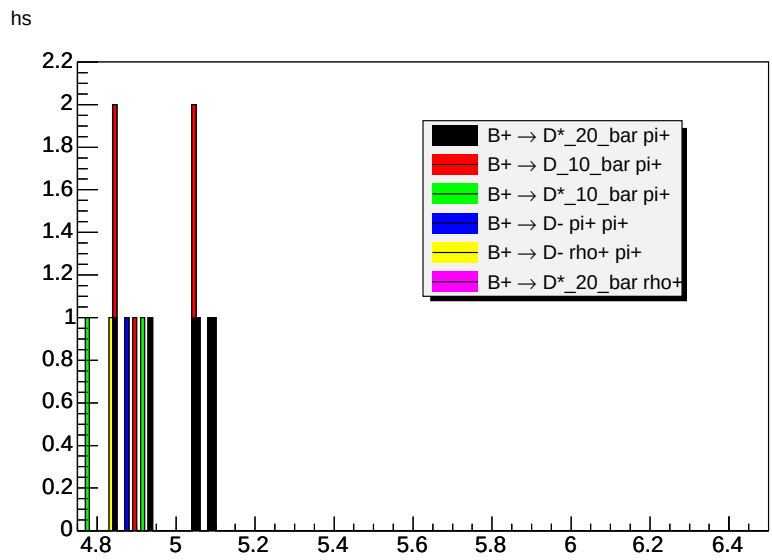

Figure 6.5: Composition of $B^-/B^+ \to D^+X$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

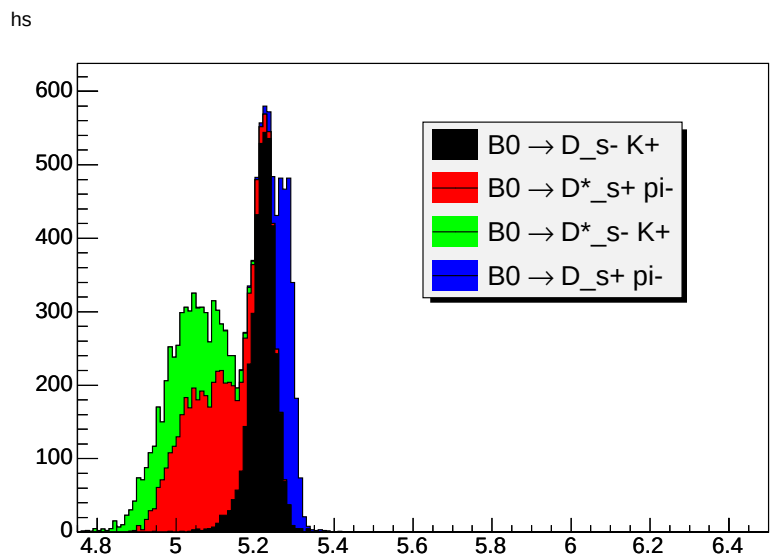

Figure 6.6: Composition of $\overline{B}^0 \to D_s^{(*)+}(\phi\pi^+(\gamma/\pi^0))X$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

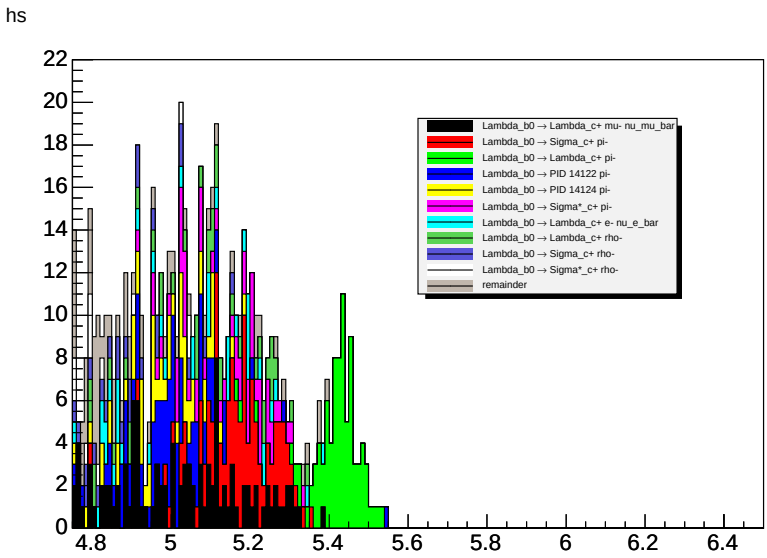

Figure 6.7: Composition of $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)X$ Monte Carlo sample reconstructed as $\overline{B}_s^0 \to D_s^+\pi^-$

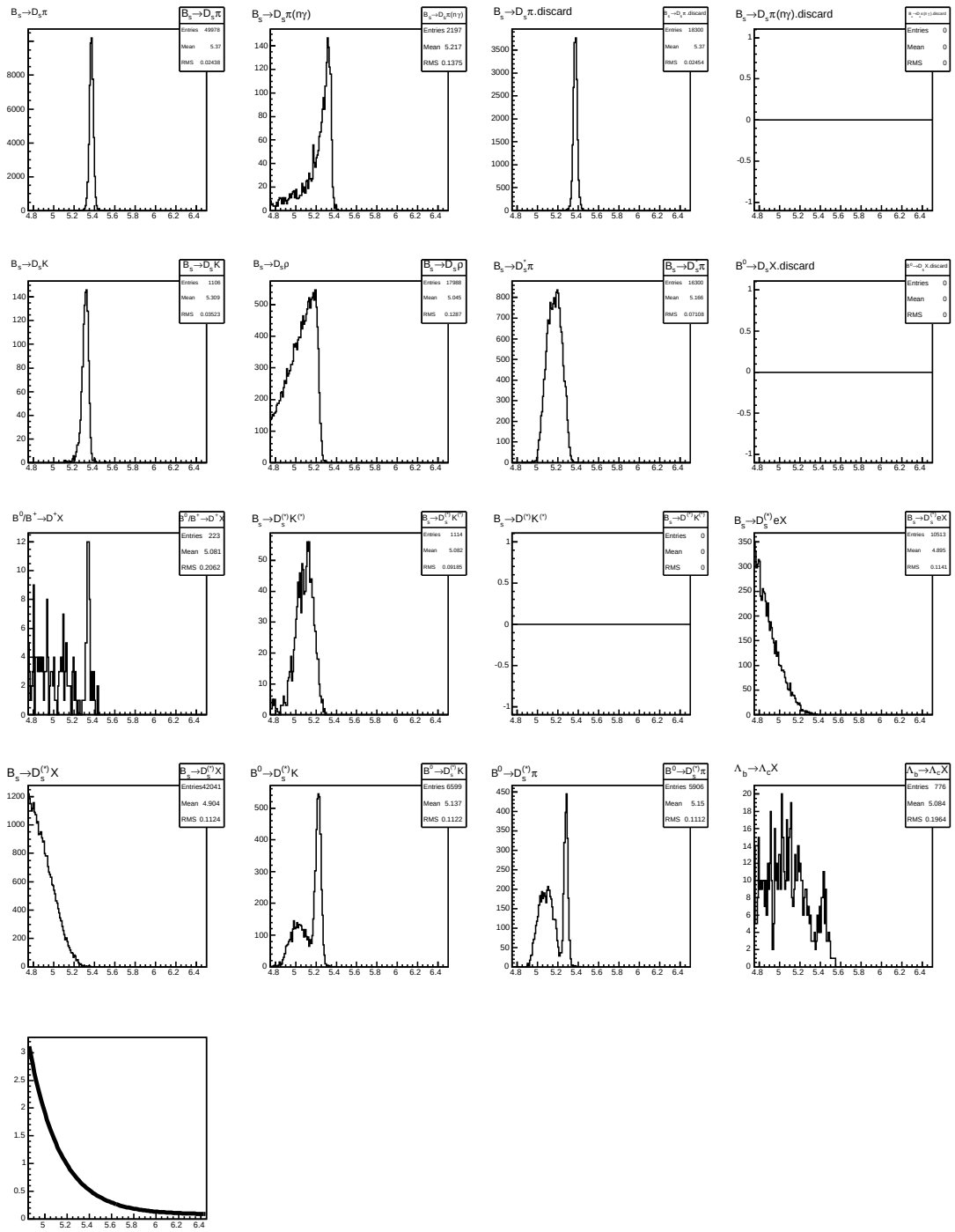

Figure 6.8: Mass templates used to model the $\overline{B}_s^0 \to D_s^+X$ fit components. The templates marked "discard" or left empty are artifacts of the fit technology and are not used as fit templates. The last template is the combinatorial background function.

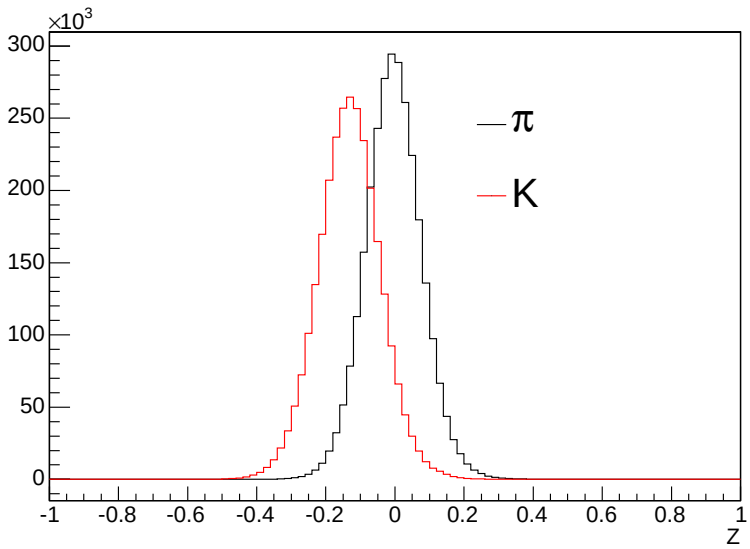

Figure 6.9: Z templates for K and π tracks reweighted for $\overline{B}_s^0 \to D_s^+ X$

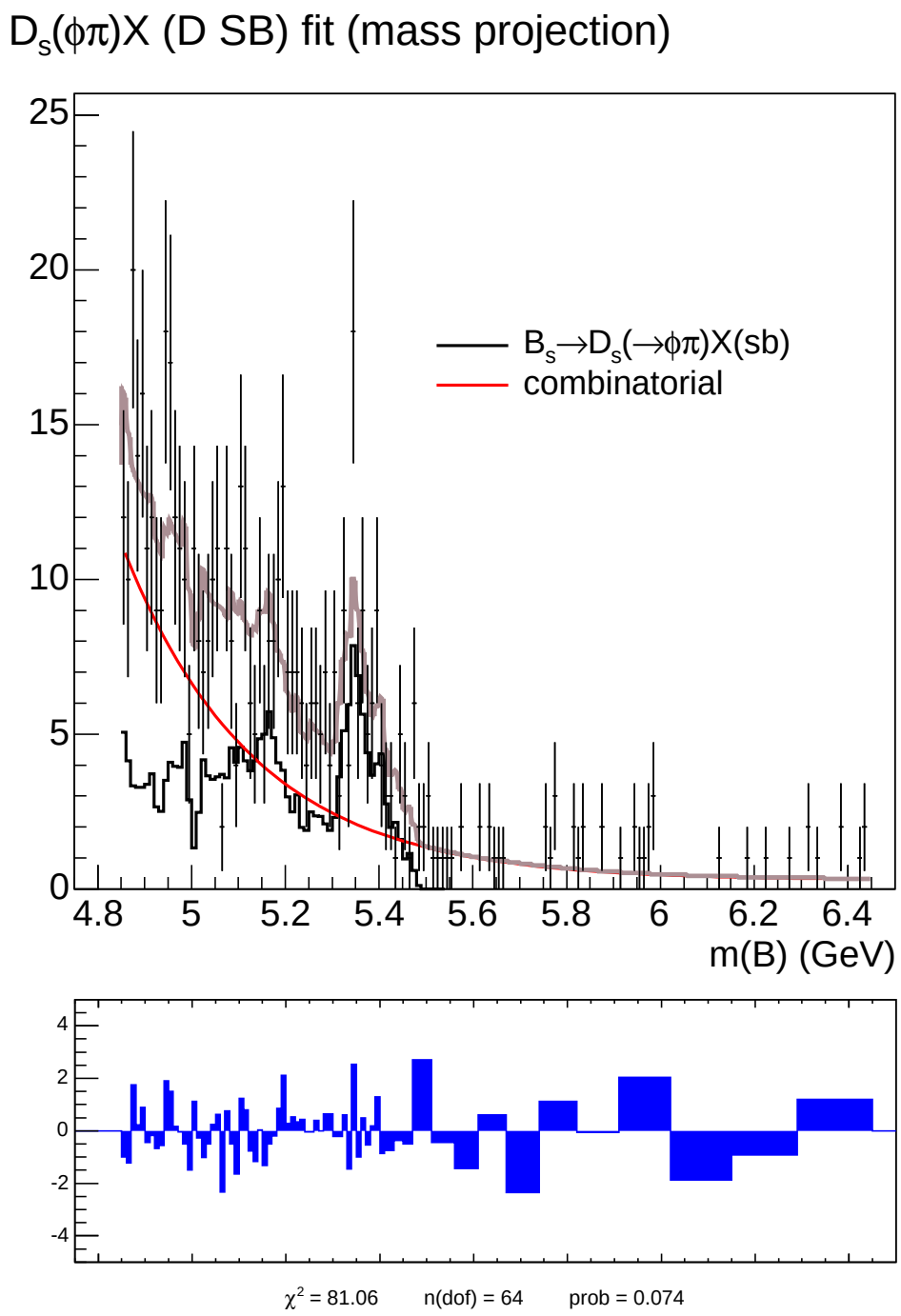

Figure 6.10: Fake-D_s^+ combinatorial template fit. The data points show the $D_s^+\pi^-$ mass distribution of B candidates from the D_s^+ sidebands. Signal subtraction is performed on this distribution to yield a combinatorial background of the form $(4 \cdot 6)$. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

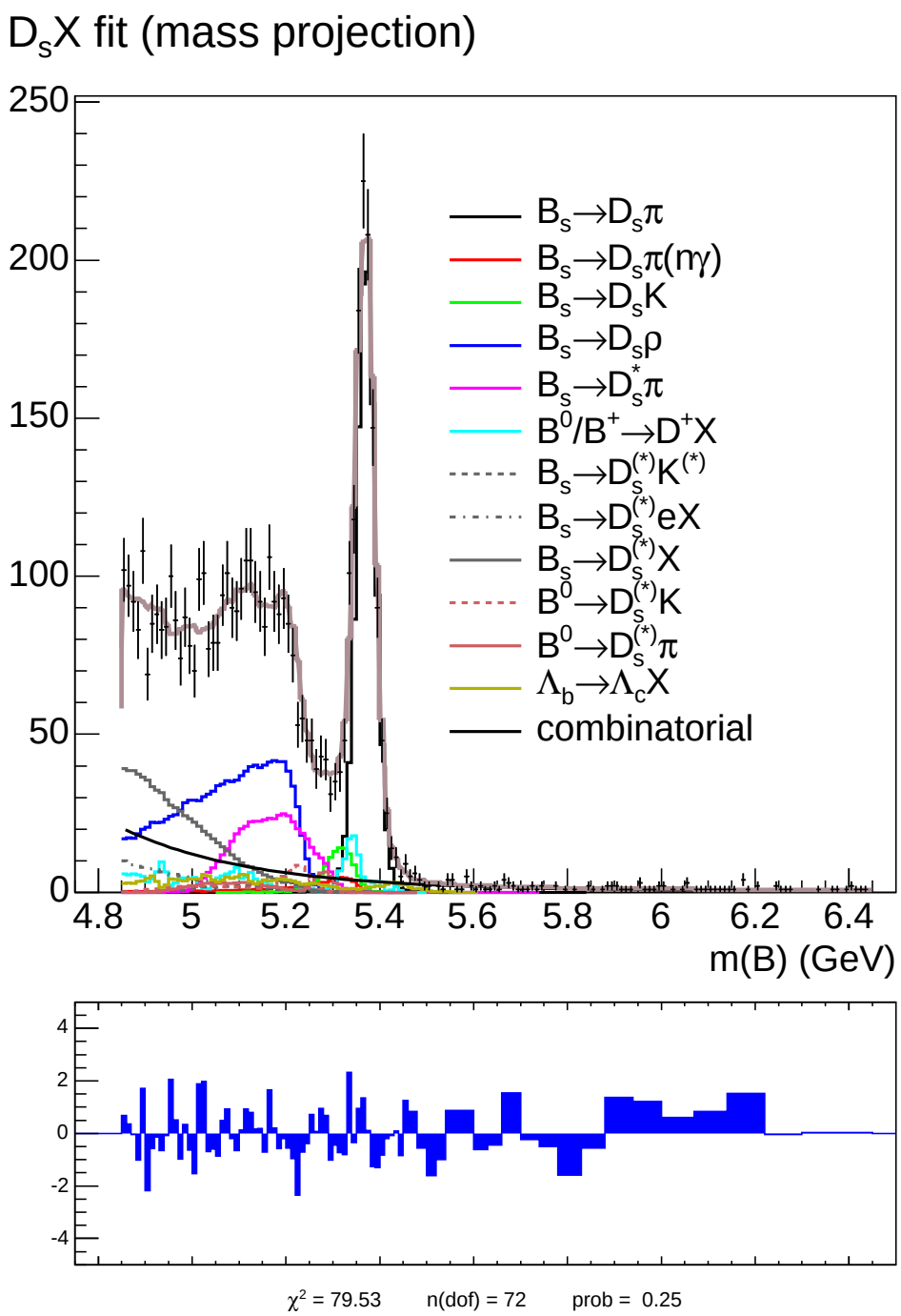

Figure 6.11: Mass projection of the likelihood fit for $\overline{B}_s^0 \to D_s^+X$ in the combined 1.2 fb^{-1}. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). For the bins with low statistics, neighboring bins are combined until the predicted number of events is > 5.

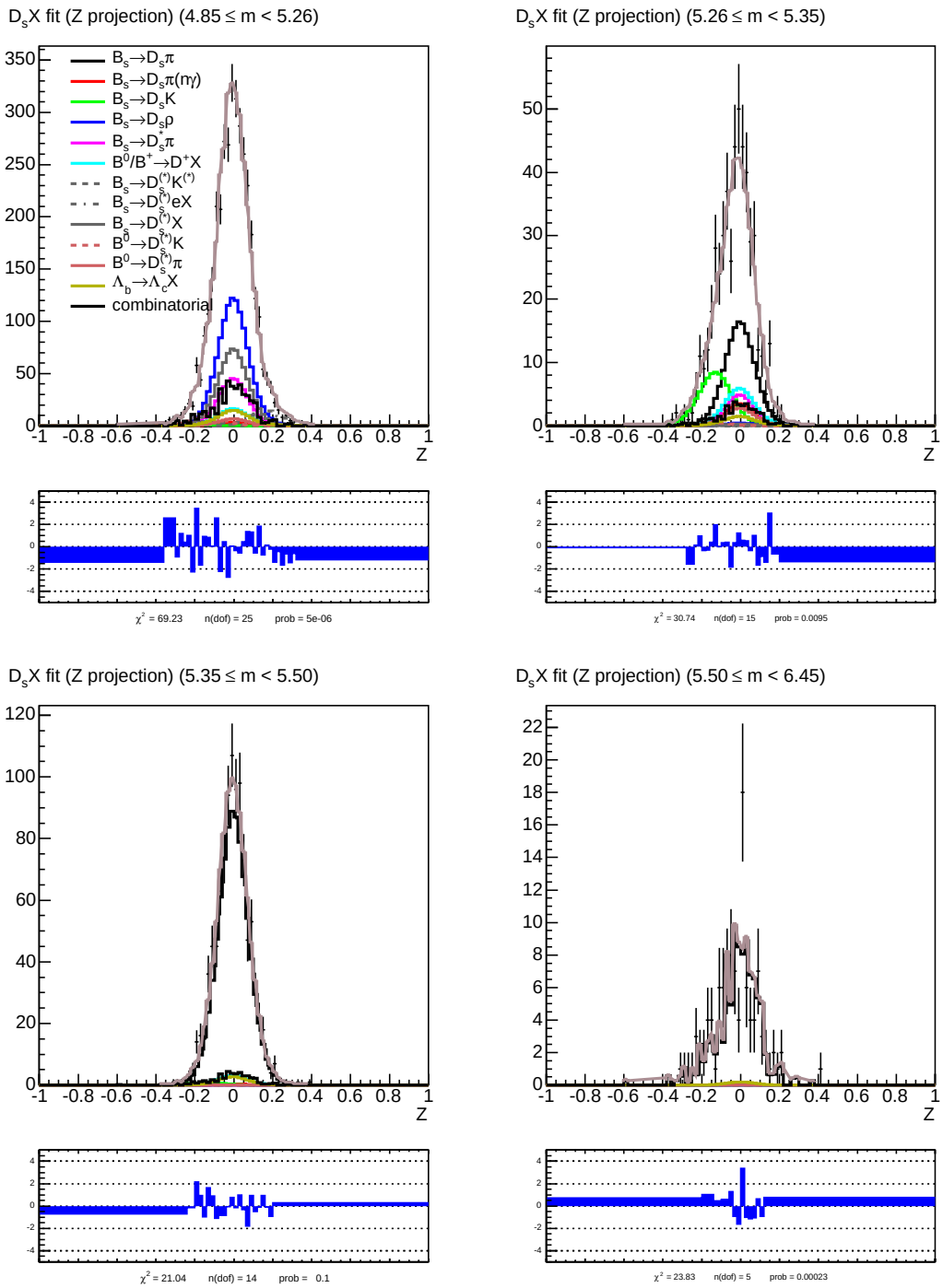

Figure 6.12: Z projection of the likelihood fit for $\overline{B}_s^0 \to D_s^+ X$ in the combined 1.2 fb^{-1}. The projections are shown in four mass regions indicated in the plot titles. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). For the bins with low statistics, neighboring bins are combined until the predicted number of events is > 5.

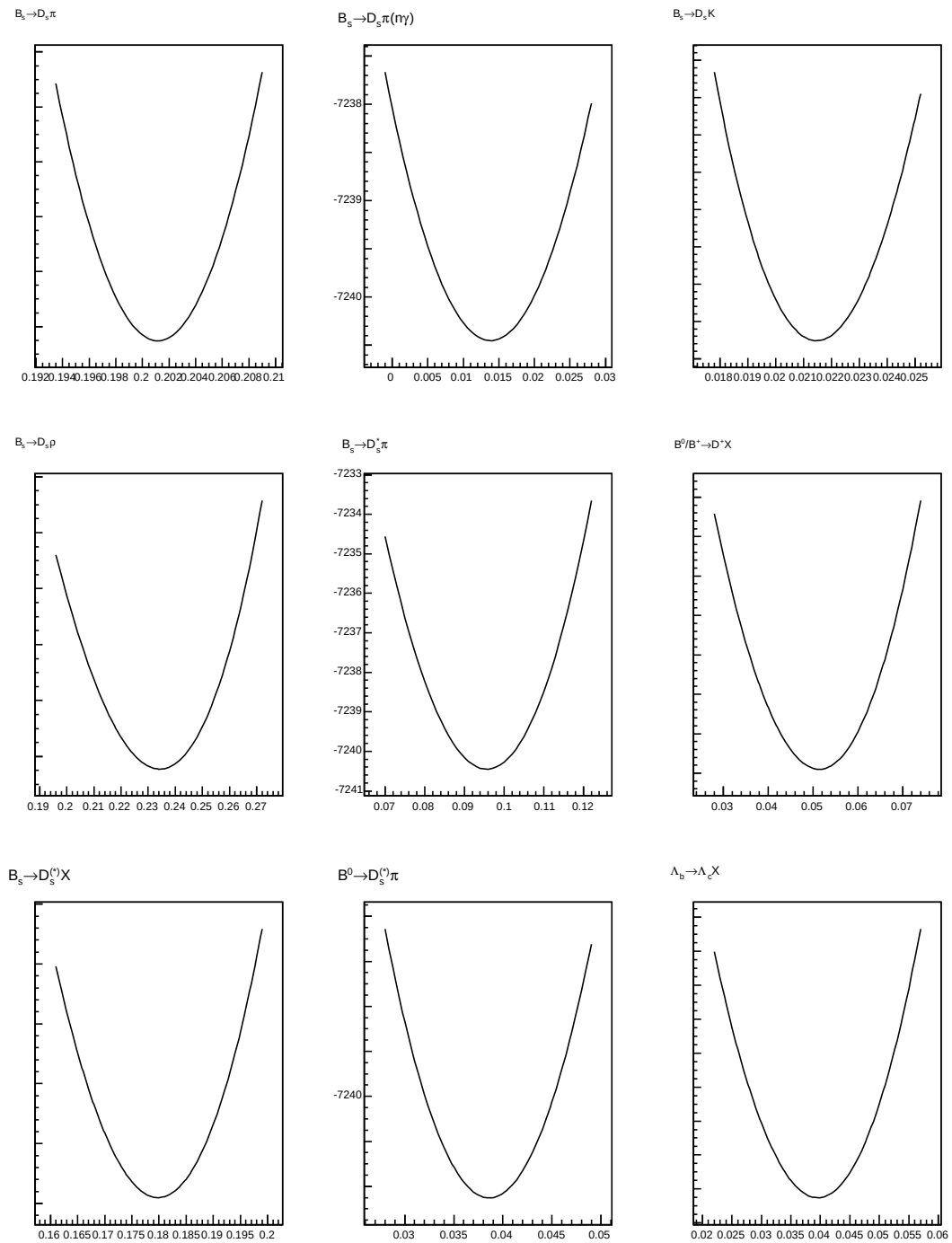

Figure 6.13: Likelihood scans for $\overline{B}_s^0 \to D_s^+ X$. The fractional candidate yield in each fit component is scanned.

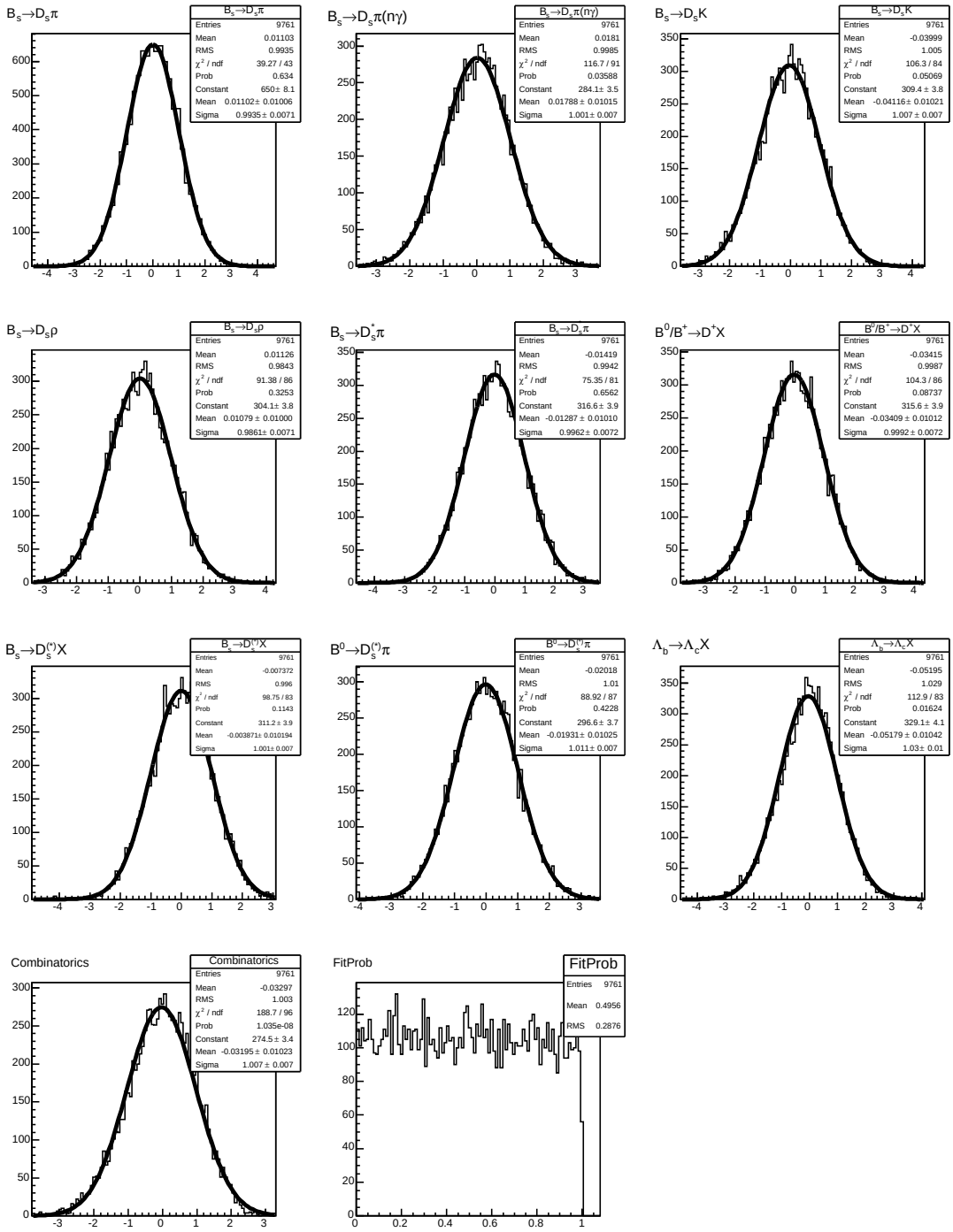

Figure 6.14: Pull distributions for $\overline{B}_s^0 \to D_s^+ X$

Chapter 7

Summary and conclusions

We have reported the first observation of the decay $\overline{B}^0_s \to D^\pm_s K^\mp$. The statistical significance of this observation is 7.9σ. We have also reported the first measurement of the relative branching fractions $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_s K^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s \pi^-\big) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$; our measurement can be combined with $\mathcal{B}\big(\overline{B}^0_s \to D^+_s \pi^-\big)$ [43] to get an absolute branching fraction. Our measurement on the $\overline{B}^0_s \to D^+_s X$ signal sample was performed after we had re-measured the analogous relative branching fractions $\mathcal{B}\big(\overline{B}^0 \to D^+ K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+ \pi^-\big) = 0.092 \pm 0.005(\text{stat}) \pm 0.007(\text{sys})$ and $\mathcal{B}\big(\overline{B}^0 \to D^{*+} K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+} \pi^-\big) = 0.084 \pm 0.008(\text{stat})$ in the $\overline{B}^0 \to D^+ X$ and $\overline{B}^0 \to D^{*+} X$ control samples and verified agreement with the world average. Figures 7.1–7.6 show the fit projections in the three samples in a summary representation.

Because the decay $\overline{B}^0_s \to D^\pm_s K^\mp$ is an important avenue towards a measurement of the CKM angle γ, the observation and measurement of $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_s K^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s \pi^-\big)$ answers two important questions. First, there is the question of whether strong destructive interference between the diagrams contributing to $\overline{B}^0_s \to D^\pm_s K^\mp$ reduces the $\overline{B}^0_s \to D^\pm_s K^\mp$ yield to levels that make the γ measurement impossible. Our result shows that there is no unpleasant surprise in the yield; $\mathcal{B}\big(\overline{B}^0_s \to D^\pm_s K^\mp\big)/\mathcal{B}\big(\overline{B}^0_s \to D^+_s \pi^-\big)$ is consistent with the B^0 analogues $\mathcal{B}\big(\overline{B}^0 \to D^+ K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^+ \pi^-\big)$ and $\mathcal{B}\big(\overline{B}^0 \to D^{*+} K^-\big)/\mathcal{B}\big(\overline{B}^0 \to D^{*+} \pi^-\big)$. Second, planning for future analyses that attempt to measure γ using $\overline{B}^0_s \to D^\pm_s K^\mp$ decays can now use yield predictions based on a measured branching fraction.

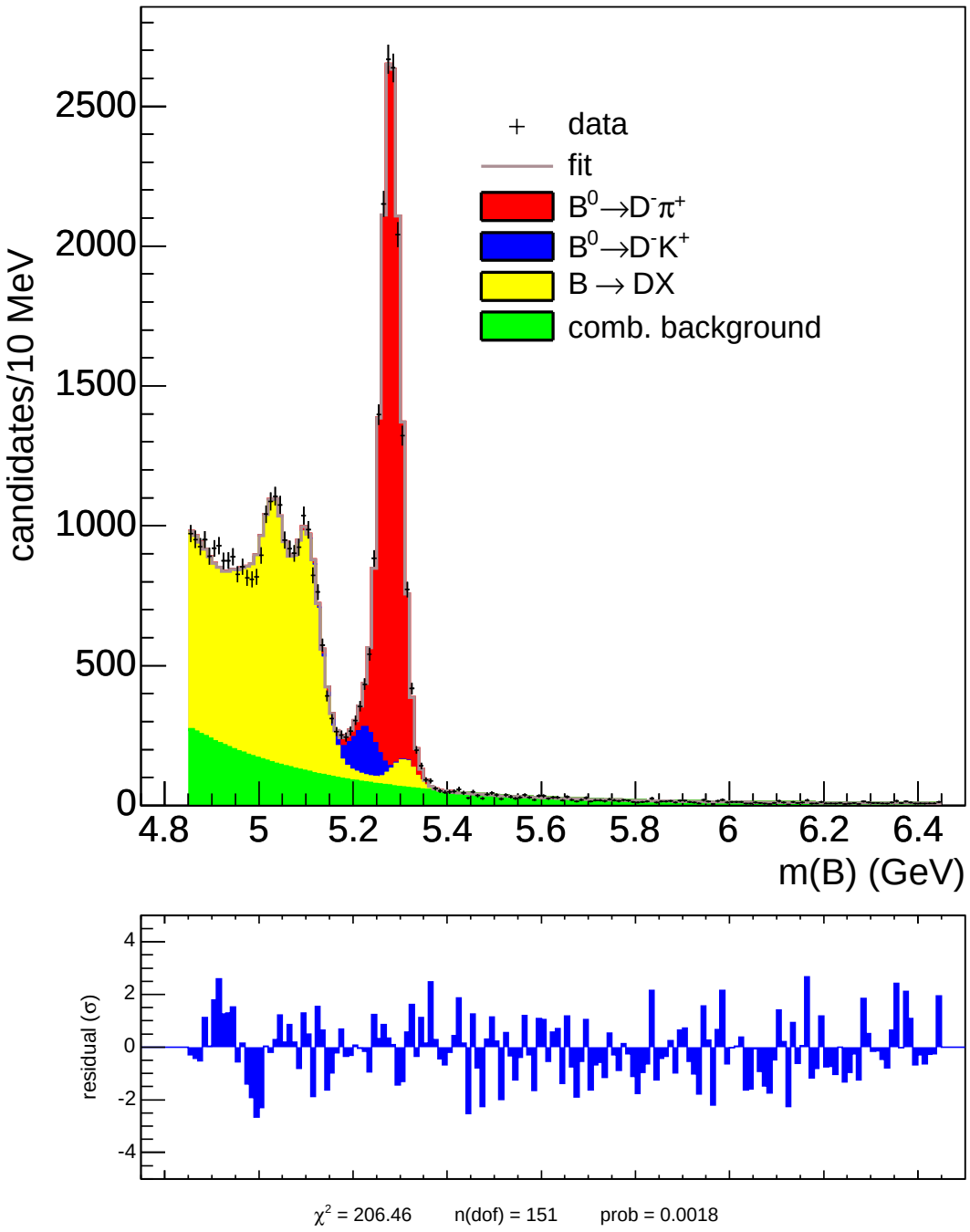

Figure 7.1: Mass projection for the $\overline{B}^0 \to D^+X$ fit (summary view)

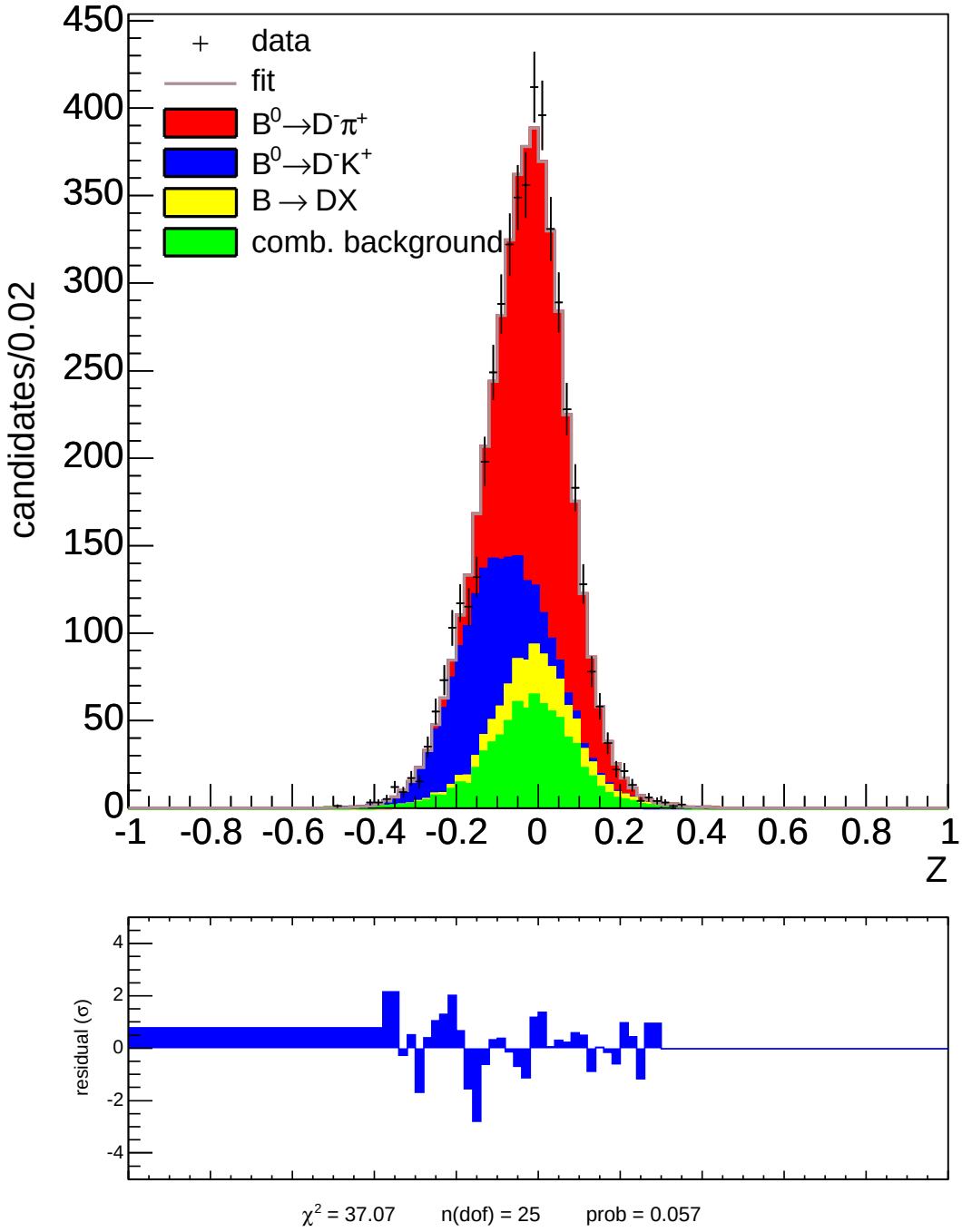

Figure 7.2: Z projection for the $\overline{B}^0 \to D^+X$ fit (summary view)

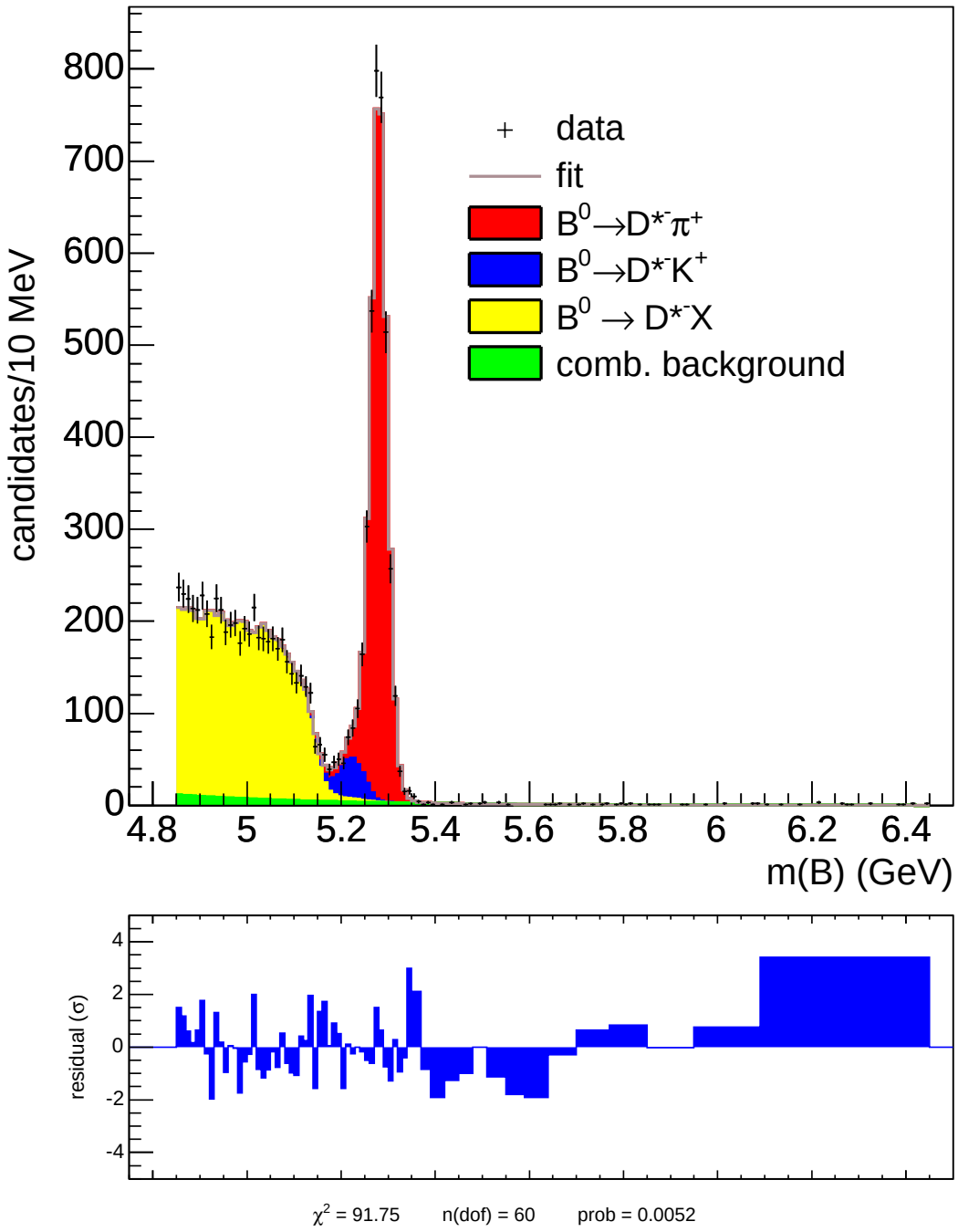

Figure 7.3: Mass projection for the $\overline{B}^0 \to D^{*+}X$ fit (summary view)

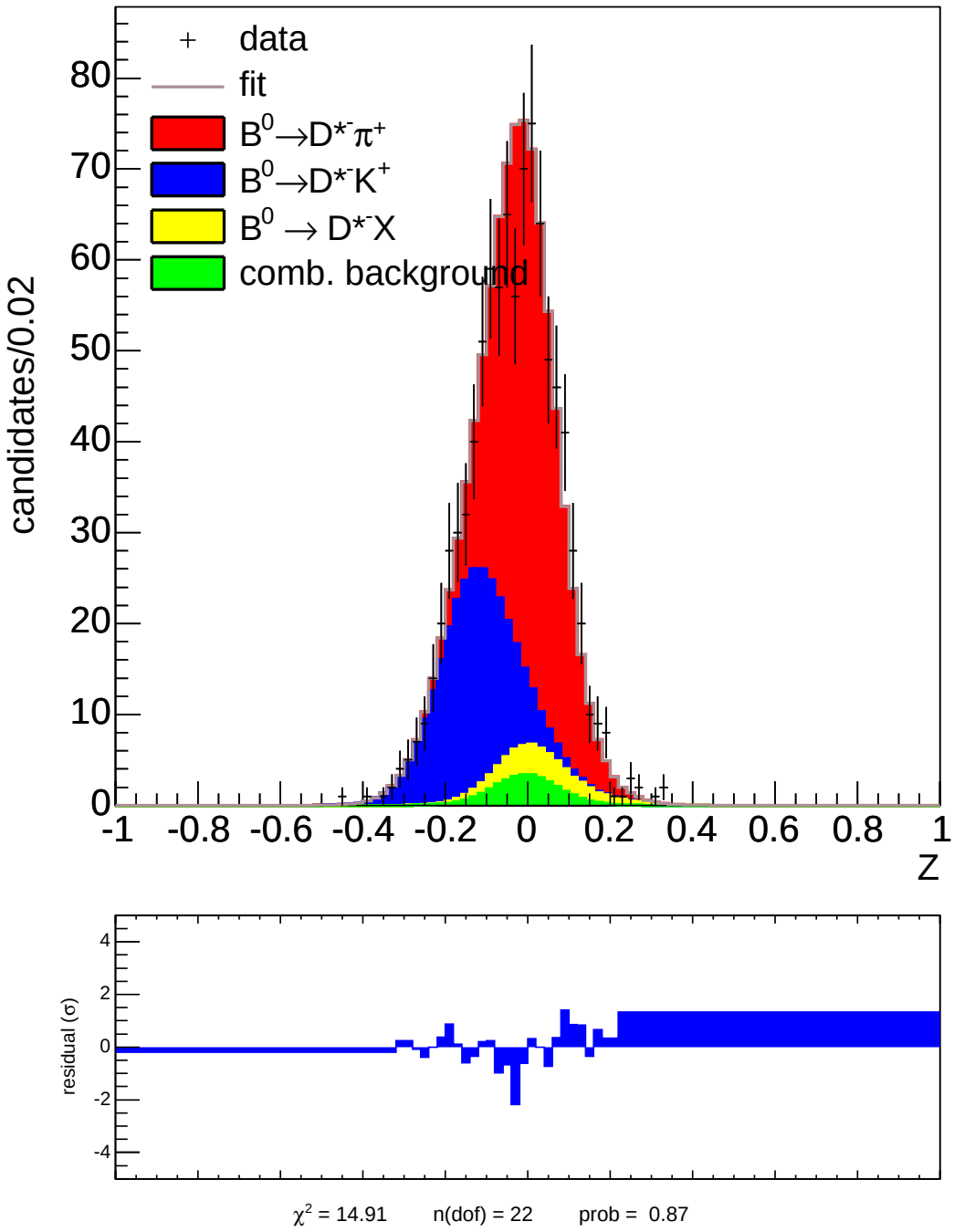

Figure 7.4: Z projection for the $\overline{B}^0 \to D^{*+}X$ fit (summary view)

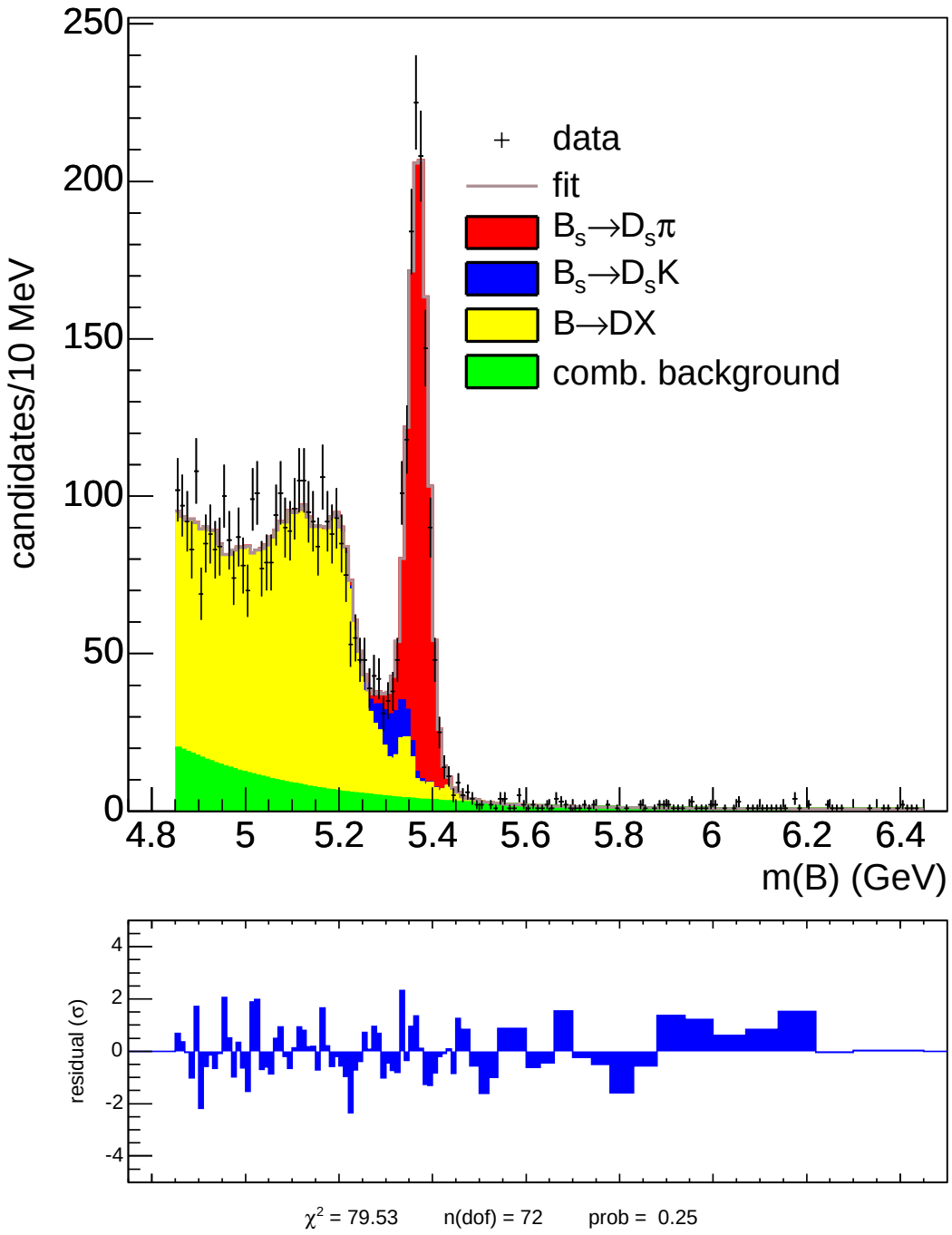

Figure 7.5: Mass projection for the $\overline{B}_s^0 \to D_s^+ X$ fit (summary view)

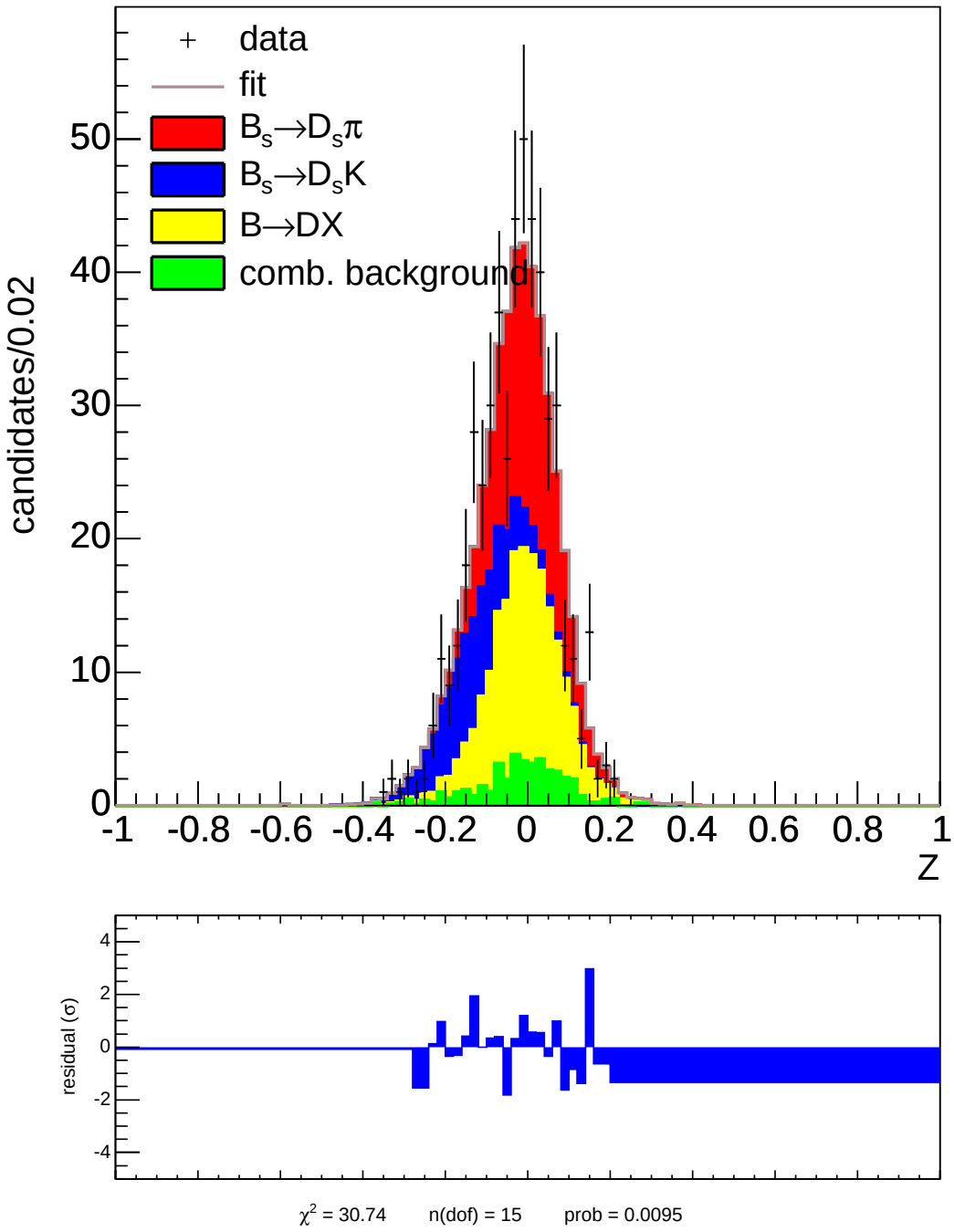

Figure 7.6: Z projection for the $\overline{B}_s^0 \to D_s^+ X$ fit (summary view)

Appendix A

dE/dx corrections

This appendix describes the procedure for extracting the pion and kaon dE/dx templates used in the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurements of Chapters 5 and 6. The pions and kaons we use come from $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. These decays have the desirable properties that the flavor of the daughter tracks of the D^0 is tagged by the charge of the pion from the D^{*+} and that the narrow $\Delta m = m(D^{*+}) - m(D^0)$ peak makes it possible to reject most background. We describe the D^{*+} samples (a sample of prompt D^{*+} and a sample of D^{*+} from B decays) in Section A.1. Then we briefly describe the hit-level and track-level dE/dx corrections in Section A.2; these are corrections that had already been developed before the present analysis. After applying these "default" corrections, we find a remaining difference between the two D^{*+} samples. The description of the sample dependence is the subject of Section A.1. Finally, Section A.4 describes the "sample-level" correction that removes that sample dependence.

A.1 D^{*+} samples

Two D^{*+} samples are available to us. The first comes mainly from decays of B mesons to D^{*+} decays; the second is a "prompt" D^{*+} sample, coming mainly from $c\bar{c}$ production. The prompt sample has much larger statistics and is the sample we use to extract the pion and kaon Z pdf's. We use the sample of D^{*+} from B decays to check Z for sample dependence; we do indeed find a sample dependence, the correction of which is necessitates this appendix.

The selection cuts for candidates at the skimming level are shown in Table A.1. In addition to the skimming-level cuts, which are different for the prompt D^{*+} and the D^{*+} from B decays, we make analysis-level cuts on the $D^0 \to K^-\pi^+$ candidates and the D^0 daughter tracks; these cuts are listed in Table A.2. A description of the cut variables can be found in Section 4.5. The requirements on the D^0 daughter tracks track cuts are identical with the requirements on the B daughter tracks in the samples used for the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurements. In particular, the tracks are required to be trigger tracks, since the track-level dE/dx corrections were developed on trigger tracks. After cuts, $2.8 \cdot 10^6$ tracks in each species are available in the prompt D^{*+} sample; the statistical uncertainty on the pdf histograms derived from this sample is tiny compared to the data statistics in the $B \to DX$ fits.

We perform sideband subtraction in Δm to reduce the already small background

	$\overline{B}^0 \to D^{*+}(D^0\pi^+)\pi^-$	prompt $D^{*+} \to D^0\pi^+$
$m(D^0)$	$1810 < m/\text{MeV} < 1920$	$1700 < m/\text{MeV} < 2000$
$\Delta\phi(D^0)$	< 1.5	
$\Delta R(D^0)$	< 2	
$\Delta z(D^0)$	< 5 cm	< 5 cm
$\sum p_T(D^0)$	> 2.4 GeV	> 5.5 GeV
$p_T(D^0)$	> 2.4 GeV	> 4.5 GeV
$L_{xy}(D^0)$	> 200 µm	> 200 µm
D^0 SVT confirmation	one track	`SCEN_LOW`
$\chi^2_{xy}(D^0)$	< 30	
$m(B)$	$4650 < m/\text{MeV} < 6500$	
$\Delta\phi(B)$	< 3	
$\Delta R(B)$	< 2	
$\Delta z(B)$	< 5 cm	
$p_T(\pi_B)$	> 1 GeV	> 1 GeV
$\Delta R(D^0, \pi_B)$	< 1.8	< 1.8
$\sum p_T(B)$	> 5.5 GeV	> 5.5 GeV
$p_T(B)$	> 5 GeV	
$\Delta z(D^{*+})$		< 5 cm
$p_T(D^0) + p_T(\pi_*)$		> 5.5 GeV
$p_T(D^{*+})$		> 5 GeV
$L_{xy}(B)$	> 300 µm	
$\lvert d_0(B)\rvert$	< 100 µm	
$\chi^2_{xy}(B)$	< 30	
B SVT confirmation	`SCEN_LOW`	
$m(D^{*+}) - m(D^0)$	< 200 MeV	< 200 MeV

Table A.1: Skimming cuts for $\overline{B}^0 \to D^{*+}(D^0\pi^+)\pi^-$ and prompt $D^{*+} \to D^0\pi^+$ candidates

under the D^{*+} even further. The fit function for the signal is a double gaussian; for the background, we use

$$p_{\text{bg}}(m) = a_0 \left(1 - \exp\left(-\frac{\Delta m - 0.140\ \text{GeV}}{a_1}\right)\right). \qquad (\text{A}\cdot 1)$$

The fit is performed in the range $0.140 < \Delta m < 0.170$ GeV. The D^* signal is modeled as a sum of two Gaussians. The signal region is $0.144 < \Delta m < 0.147$ GeV, the sideband region $0.155 < \Delta m < 0.165$ GeV. Figures A.1–A.3 show the fits in the prompt D^{*+} samples.

$m(D^0)$	$1844.6 < m/\text{MeV} < 1884.6$
Offline track p_T	> 2 GeV
Track–SVT match distance	< 25
χ^2 of matching SVT track	< 15
p_T of matching SVT track	> 2 GeV
Z	$\neq -999$
Good run list	B_s mixing version 14

Table A.2: Analysis cuts for $D^0 \to K^-\pi^+$ candidates and D^0 daughter tracks

A.2 Hit-level and track-level dE/dx corrections

As a preambulatory remark, we remind the reader that the variable used for particle identification in the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurements is

$$Z = \log\left(\frac{dE/dx(\text{measured})}{dE/dx(\text{expected for } \pi)}\right), \tag{A·2}$$

which is chosen because, in contrast to the dE/dx itself, it is normally distributed.

CDF analyses that use dE/dx information typically apply two levels of corrections. The first, "hit-level" corrections [44], correct the measured energy of each hit for dip angle, aspect angle, drift distance, superlayer, wire, high voltage and pressure in the COT. "Track-level" corrections [37] are applied to the 80% truncated mean of the track hit energies; these corrections are based on the parameters of an entire track: η, ϕ and number of hits. In addition, the track-level corrections are run-number-dependent to account for changes in the trigger configuration, which can affect the dE/dx through the track quality requirements on a trigger track.

Track-level corrections are not available for the latest part of the data used in our analysis. For this part of the data sample, we assume that the last existing η, ϕ and hit-number corrections can simply be extended; we determine a correction only for run-dependent offsets. This correction is additive in Z; it is calculated by histogramming the pion Z in groups of 1000 runs, then merging bins until each bin contains at least 10000 events after sideband subtraction. For consistency, this correction is applied to the entire dataset; its effects are small in the `xbhd0d` and `xbhd0h` run periods, but is as high as $\sim 20 \cdot 10^{-3}$ in `xbhd0i`. Figure A.4 shows the correction by run number.

A.3 Sample dependence

Are the "default" corrections for hit-level and track-level effects sufficient to allow us to extract templates for the Z of B daughter tracks from the D^{*+} daughter tracks? What we observe instead is that there are effects due to the environment of the track

that introduce a sample dependence. The difference between the prompt D^{*+} and the D^{*+} from B is listed in the "before" columns of Tables A.3–A.6. Whatever effect causes a difference between D^{*+} daughter tracks will also cause the B daughter tracks to behave differently from the templates extracted from the D^{*+} daughters; if we want to avoid a bias in our measurement from this effect, we need to find a way to correct for it.

To understand what the source of the sample dependence is, we make profile histograms of Z for kaons and pions in the prompt D^{*+} sample as a function of a number of variables. These variables are the p; p_T; η; ϕ_0; the number of hits N_{hits} used in calculating the dE/dx truncated mean; run number; min(ΔR), the ΔR of the closest track in ΔR; min($\Delta\eta$), the $\Delta\eta$ of the closest track in η; min($\Delta\phi$), the $\Delta\phi$ of the closest track in ϕ_0; iso(0.3), iso(0.5), and iso(0.7), the track isolation calculated for the specified cone size, where iso(ΔR_{max}) for a given track i is defined as

$$\text{iso}(\Delta R_{\text{max}}) = \frac{p_T^i}{p_T^i + \sum\limits_{j \neq i} p_T^j} \qquad (\Delta R(i,j) < \Delta R_{\text{max}}); \qquad (\text{A} \cdot 3)$$

$N_{\text{trks}}(\Delta R < 0.3)$, $N_{\text{trks}}(\Delta R < 0.5)$, $N_{\text{trks}}(\Delta R < 0.7)$, the number of other tracks within the specified ΔR; the total number of tracks N_{trks} in the event; and "secance", defined as the number of r–ϕ intersections (of the track in question with other tracks) that occur within the nominal COT radii of $40 < r < 137$ cm.

The pion Z dependence on these variables is shown for `xbhd0d`, `xbhd0h` and `xbhd0i` in Figures A.5–A.7. The kaon Z profiles are shown in Figures A.8–A.10. A large dependence of Z on many of the variables, in particular in $N_{\text{trks}}(\Delta R < 0.5)$, $N_{\text{trks}}(\Delta R < 0.7)$, N_{trks} and secance, is evident; the variation in Z over the range plotted is $60 \cdot 10^{-3}$, corresponding to half the K–π separation. (For kaons, there is also a strong additional dependence on momentum, which is a consequence calculating the expected dE/dx in the wrong mass hypothesis; we will return to this point in Section A.4 below.)

The likely source of the Z dependence on these variables is the activity in the chamber in the vicinity of the track whose dE/dx is being measured. If the track passes through a region of the tracking chamber in which many other particles are depositing hits, it is more likely that energy depositions not belonging to the track will be merged with the hits belonging to the track. The result is a higher measured dE/dx. A high degree of correlation exists between the variables on which Z strongly depends and the density of hits in the vicinity of the track, whether we are considering a very "global" measure of the hit density like N_{trks} or a very "local" measure like secance.

A.4 Reweighting

If the D^{*+} daughter tracks in the dE/dx template samples inhabit a denser hit environment than the B daughter tracks in the $\mathcal{B}(B \to DK)/\mathcal{B}(B \to D\pi)$ measurement

samples, the effect described in the previous section will introduce a bias in the branching fraction measurements. Figure A.17 shows that the prompt D^{*+} tracks do indeed traverse a much denser environment than the B daughter tracks, presumably due to the greater isolation of b jets compared to c jets.

We remedy this situation by reweighting the D^{*+} daughter tracks so that their distributions in the density proxy variables ($N_{\text{trks}}(\Delta R < 0.5)$, $N_{\text{trks}}(\Delta R < 0.7)$, N_{trks} and secance) match those of the B daughter tracks. Since all these variables are proxies for the same effect, they are highly correlated amongst each other; it is therefore sensible to restrict the reweighting to only one of these variables.

The most suitable choice of reweighting variable is determined as follows. $N_{\text{trks}}(\Delta R < 0.5)$ and $N_{\text{trks}}(\Delta R < 0.7)$ are not considered because the signal populates the zero bin of these distributions, while the D^{*+} daughters do not. The decision between N_{trks} and secance is made by comparing the effect of reweighting in each variable on the difference in $\langle Z \rangle$ in the prompt D^{*+} and D^{*+} from B samples. Tables A.3–A.6 show that in almost all cases, reweighting in secance results in a smaller sample dependence than reweighting in N_{trks}. We therefore choose secance reweighting.

In the case of the kaon Z templates, an additional effect must be corrected for. As mentioned earlier, the kaon Z depends strongly on momentum because the expected dE/dx is calculated in the pion hypothesis. Therefore, we reweight the D^{*+} daughter tracks (both for kaon and pion templates) in momentum to match the B daughter tracks. The two reweightings (in secance and momentum) are performed simultaneously; we call this "two-dimensional" reweighting. Figures A.18–A.20 show the Z distributions before and after reweighting for the `xbhd0d`, `xbhd0h` and `xbhd0i` run periods. (A consequence of reweighting the D^{*+} daughter tracks to match the $B \to DX$ samples is that each $B \to DX$ sample will have slightly different Z templates; the figures show the templates for the two control samples $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$.) To obtain the Z templates for the combined run period, we add the templates from the sub-periods; since $B \to DX$ samples and prompt D^{*+} samples are both collected by the two-track trigger, we expect them to have the same relative yields in a given run range.

		$\langle Z \rangle_\pi$			$\langle Z \rangle_K$			$\langle Z \rangle_\pi - \langle Z \rangle_K$		
		$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.
0d	before	+2.02	+0.00	+2.02	−130.44	−137.24	+6.80	132.47	137.24	−4.78
0d	after	−2.16	−3.01	+0.85	−135.80	−135.54	−0.26	133.64	132.52	+1.11
0h	before	+2.30	+0.03	+2.27	−125.88	−132.66	+6.77	128.18	132.69	−4.51
0h	after	−4.48	−3.87	−0.62	−131.45	−131.53	+0.08	126.97	127.67	−0.70
0i	before	+1.57	−0.07	+1.64	−125.26	−131.99	+6.73	126.83	131.92	−5.09
0i	after	−4.93	−4.72	−0.21	−130.82	−131.29	+0.47	125.89	126.57	−0.68

Table A.3: $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in secance and momentum to the $\overline{B}^0 \to D^{*+} X$ sample

Table A.4: $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in N_{trks} and momentum to the $\overline{B}^0 \to D^{*+}X$ sample

		$\langle Z \rangle_\pi$			$\langle Z \rangle_K$			$\langle Z \rangle_\pi - \langle Z \rangle_K$		
		$B \to D^*X$	prompt	diff.	$B \to D^*X$	prompt	diff.	$B \to D^*X$	prompt	diff.
0d	before	+2.02	+0.00	+2.02	−130.44	−137.24	+6.80	132.47	137.24	−4.78
0d	after	−4.00	−1.73	−2.27	−135.50	−132.99	−2.51	131.50	131.26	+0.24
0h	before	+2.30	+0.03	+2.27	−125.88	−132.66	+6.77	128.18	132.69	−4.51
0h	after	−3.69	−2.59	−1.10	−131.60	−129.27	−2.33	127.91	126.68	+1.22
0i	before	+1.57	−0.07	+1.64	−125.26	−131.99	+6.73	126.83	131.92	−5.09
0i	after	−5.28	−3.40	−1.88	−130.31	−129.49	−0.82	125.03	126.09	−1.06

Table A.5: $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in secance and momentum to the $\overline{B}^0 \to D^+ X$ sample

		$\langle Z \rangle_\pi$			$\langle Z \rangle_K$			$\langle Z \rangle_\pi - \langle Z \rangle_K$		
		$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.
0d	before	+2.02	+0.00	+2.02	−130.44	−137.24	+6.80	132.47	137.24	−4.78
0d	after	−2.98	−3.16	+0.18	−136.09	−135.47	−0.62	133.11	132.31	+0.80
0h	before	+2.30	+0.03	+2.27	−125.88	−132.66	+6.77	128.18	132.69	−4.51
0h	after	−4.15	−3.72	−0.43	−130.73	−131.31	+0.58	126.59	127.59	−1.00
0i	before	+1.57	−0.07	+1.64	−125.26	−131.99	+6.73	126.83	131.92	−5.09
0i	after	−5.29	−4.65	−0.63	−130.87	−130.76	−0.11	125.58	126.11	−0.53

		$\langle Z \rangle_\pi$			$\langle Z \rangle_K$			$\langle Z \rangle_\pi - \langle Z \rangle_K$		
		$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.	$B \to D^* X$	prompt	diff.
0d	before	+2.02	+0.00	+2.02	−130.44	−137.24	+6.80	132.47	137.24	−4.78
0d	after	−3.51	−1.45	−2.06	−135.67	−132.75	−2.92	132.16	131.30	+0.86
0h	before	+2.30	+0.03	+2.27	−125.88	−132.66	+6.77	128.18	132.69	−4.51
0h	after	−3.87	−2.30	−1.57	−132.15	−129.18	−2.97	128.28	126.88	+1.40
0i	before	+1.57	−0.07	+1.64	−125.26	−131.99	+6.73	126.83	131.92	−5.09
0i	after	−4.97	−3.27	−1.70	−130.47	−128.60	−1.86	125.49	125.33	+0.16

Table A.6: $\langle Z \rangle$ for pions and kaons and kaon–pion separation before and after reweighting in N_{trks} and momentum to the $\overline{B}^0 \to D^+ X$ sample

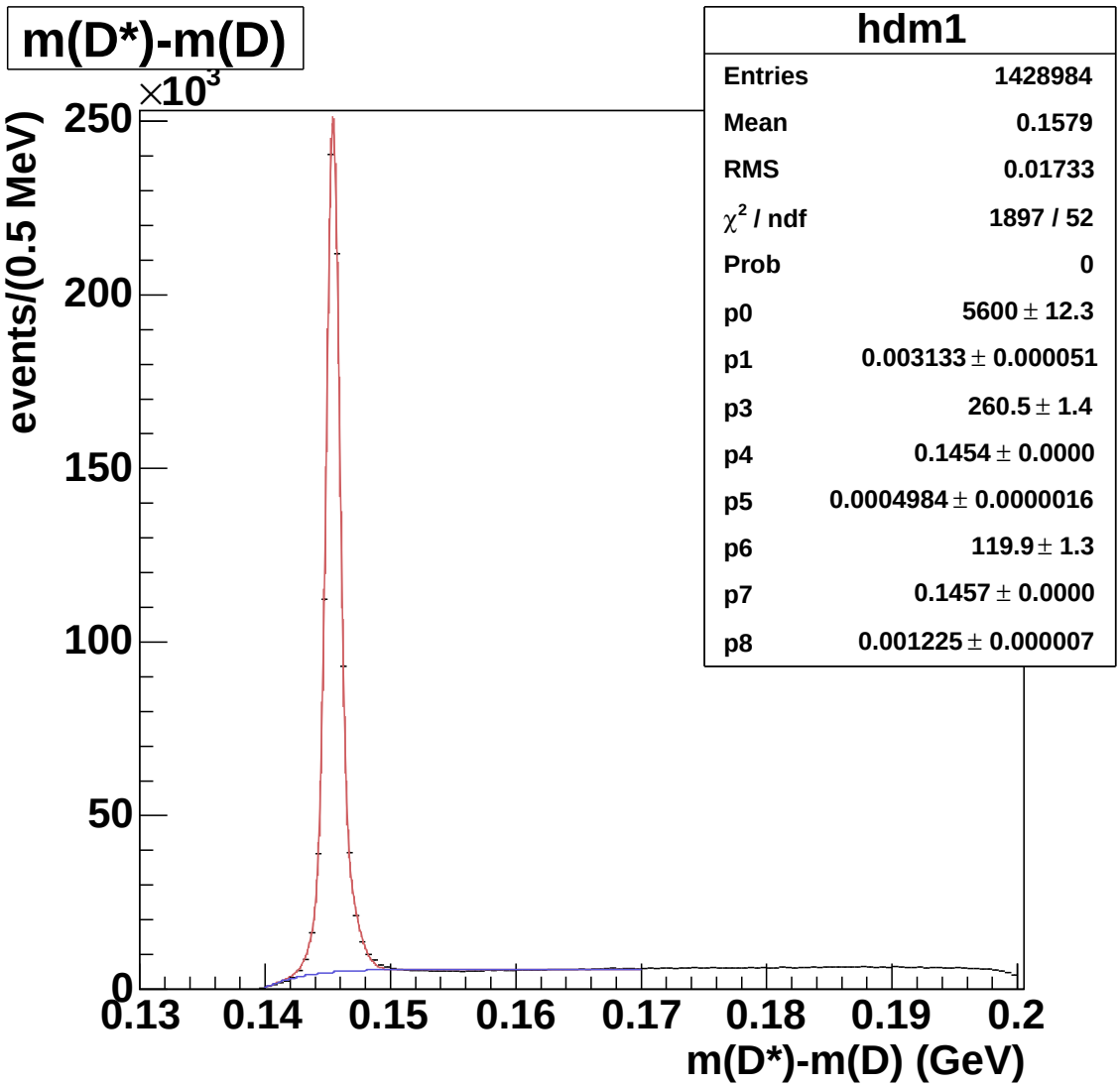

Figure A.1: Δm distribution in the `xbhd0d` prompt D^{*+} sample. The fit parameters p_0 and p_1 correspond to a_0 and a_1 of the background function $(\mathrm{A} \cdot 1)$; the remaining parameters are amplitudes, means and widths for the double-gaussian signal function.

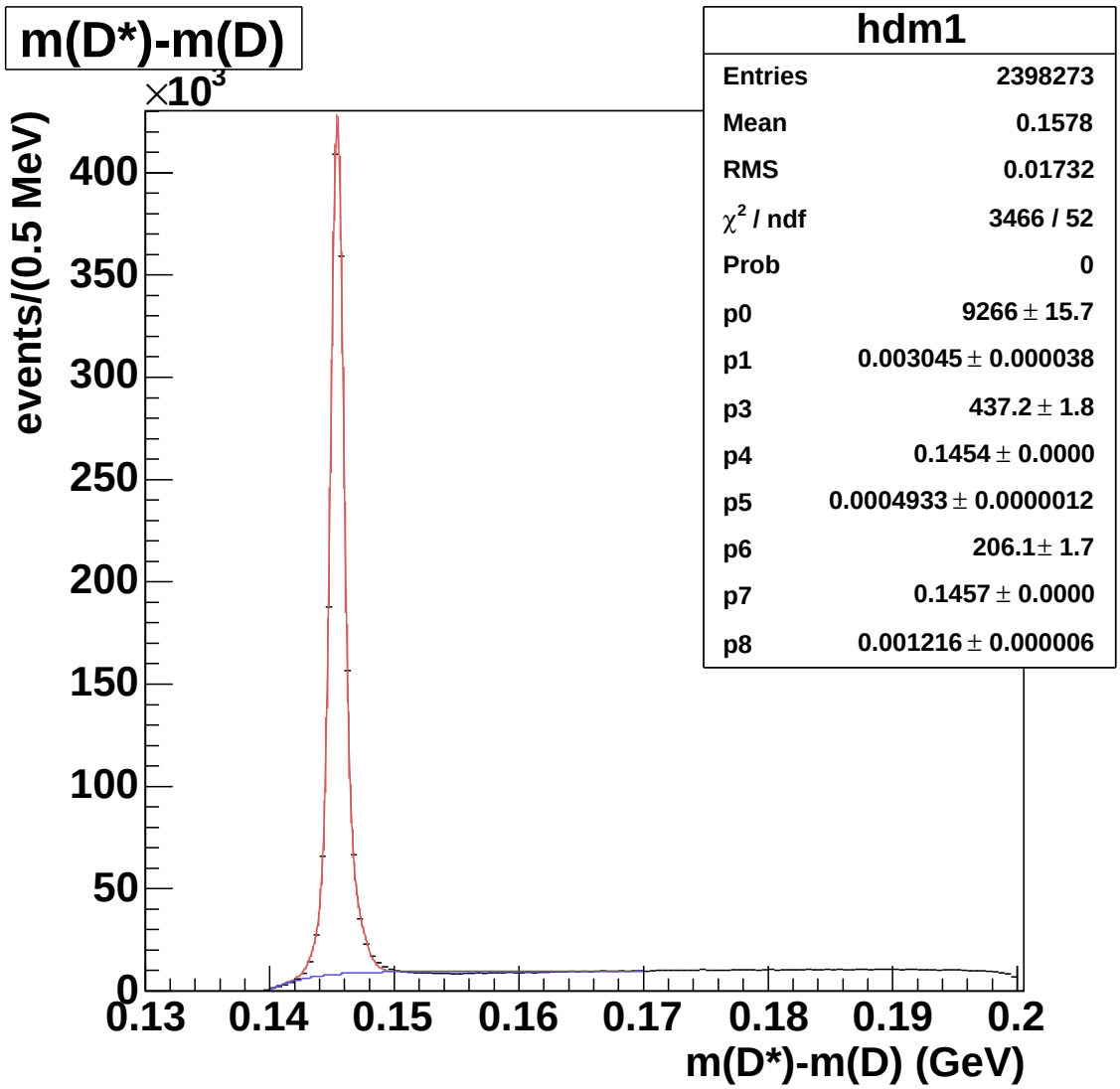

Figure A.2: Δm distribution in the `xbhd0h` prompt D^{*+} sample. The fit parameters p_0 and p_1 correspond to a_0 and a_1 of the background function $(\mathrm{A} \cdot 1)$; the remaining parameters are amplitudes, means and widths for the double-gaussian signal function.

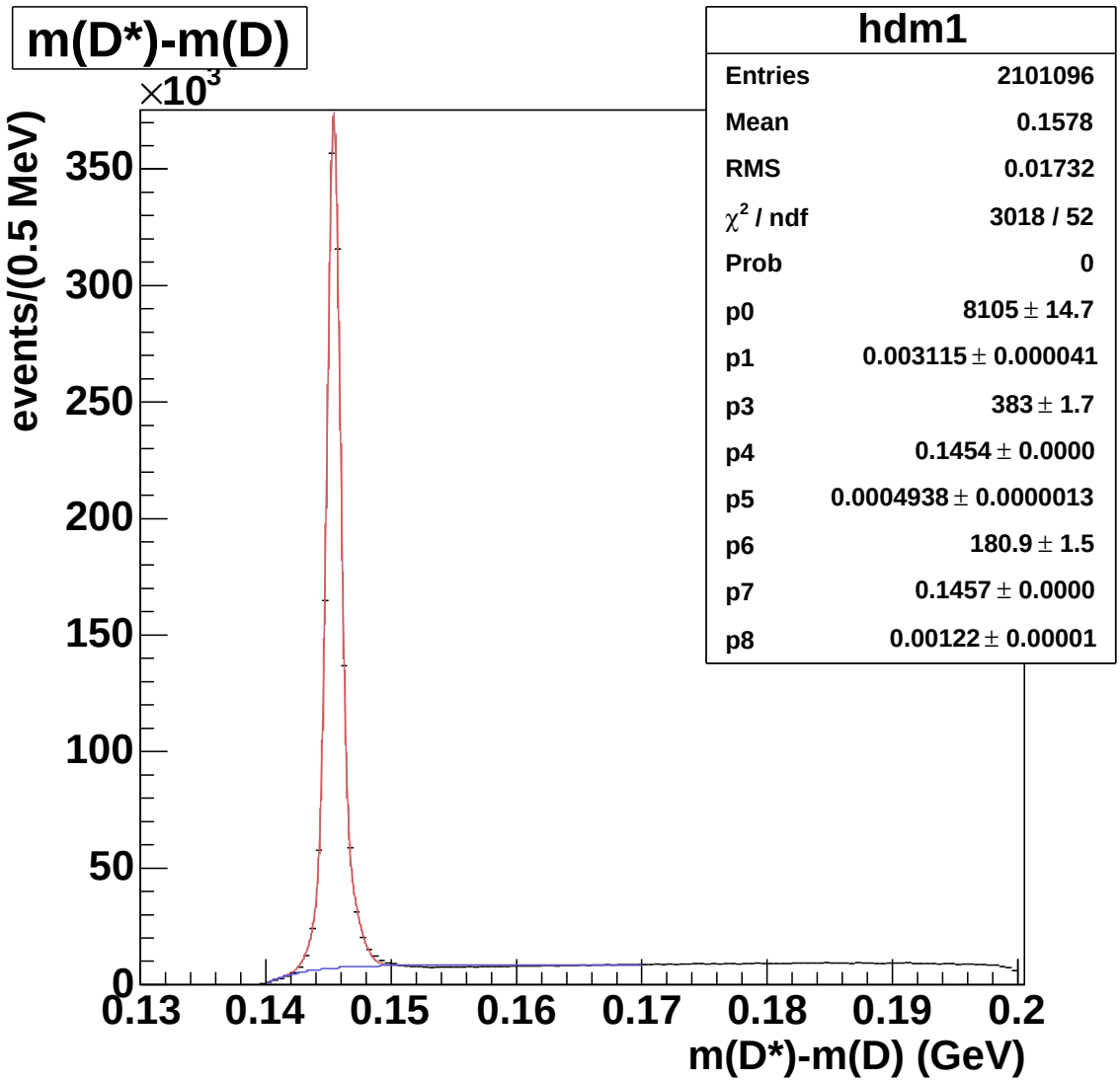

Figure A.3: Δm distribution in the `xbhd0i` prompt D^{*+} sample. The fit parameters p_0 and p_1 correspond to a_0 and a_1 of the background function $(\mathrm{A} \cdot 1)$; the remaining parameters are amplitudes, means and widths for the double-gaussian signal function.

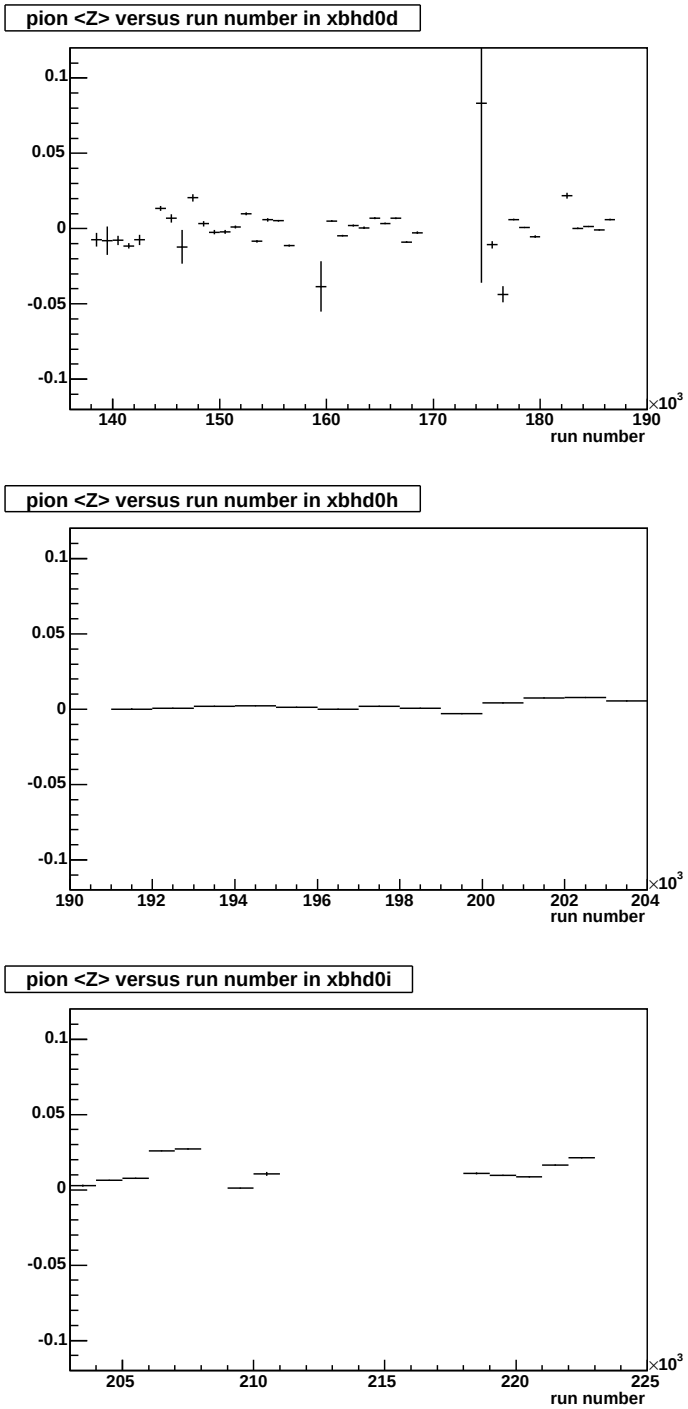

Figure A.4: Z versus run number in the prompt D^* sample for the three datasets `xbhd0d`, `xbhd0h` and `xbhd0i`

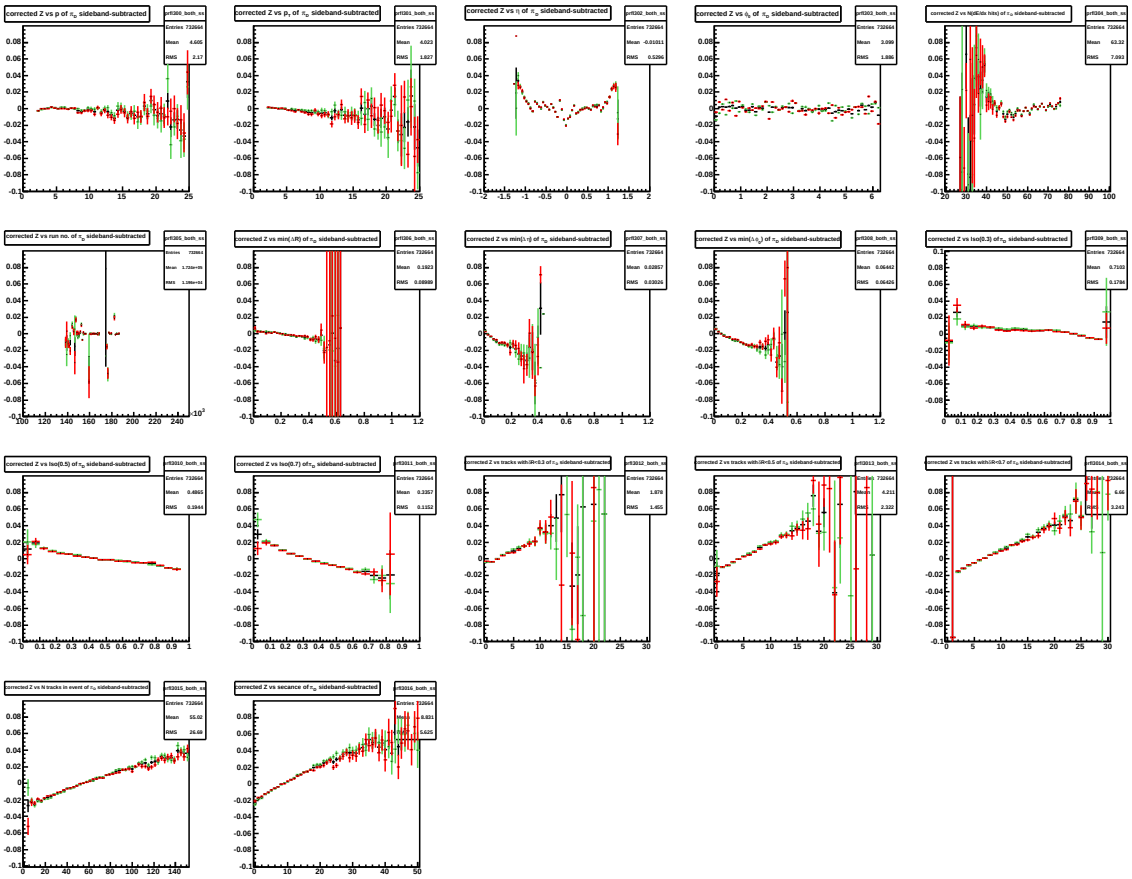

Figure A.5: Pion Z profiles for the `xbhd0d` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

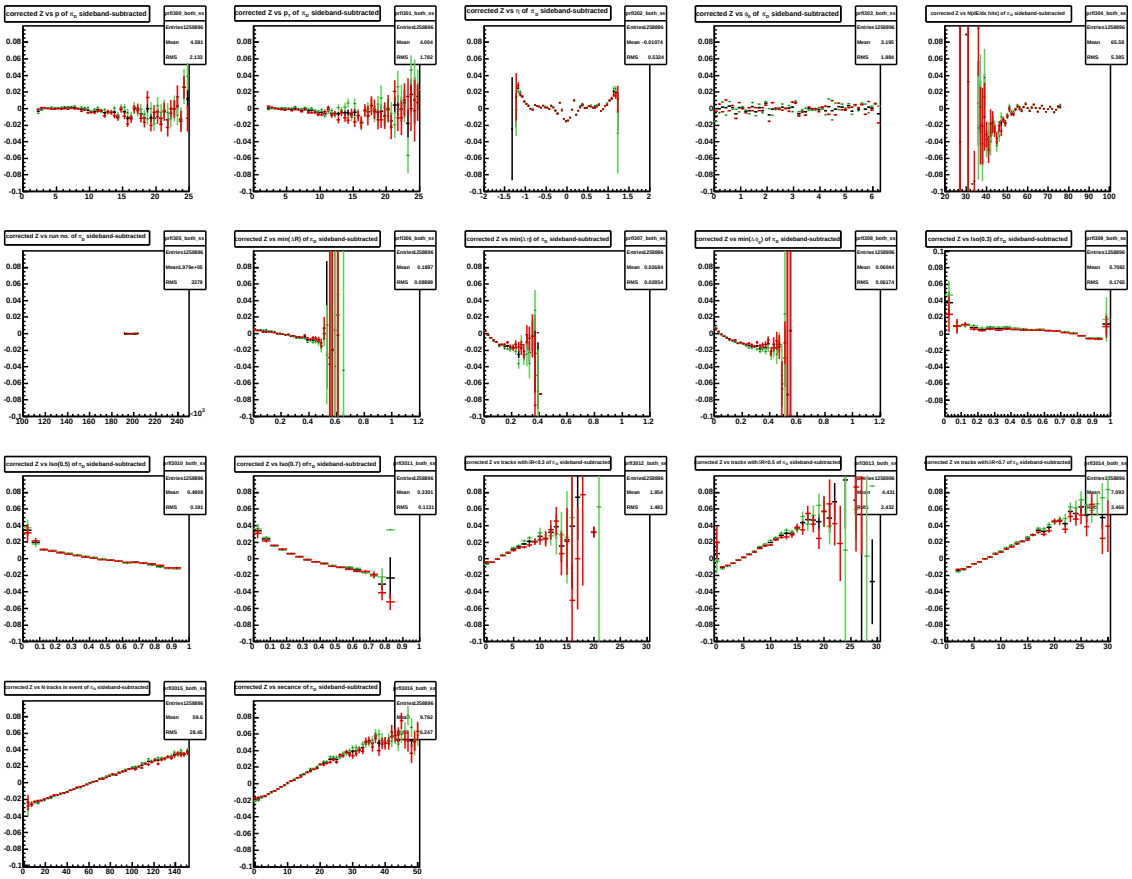

Figure A.6: Pion Z profiles for the `xbhd0h` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

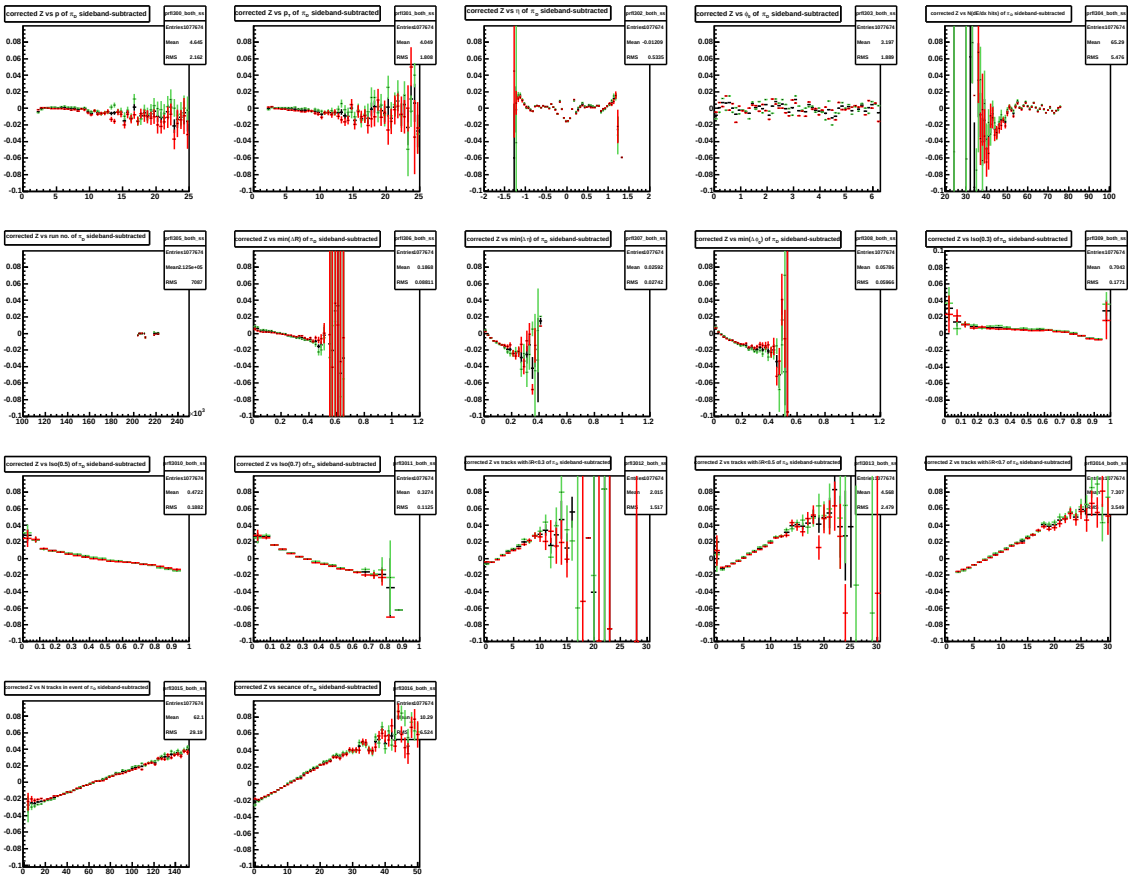

Figure A.7: Pion Z profiles for the `xbhd0i` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

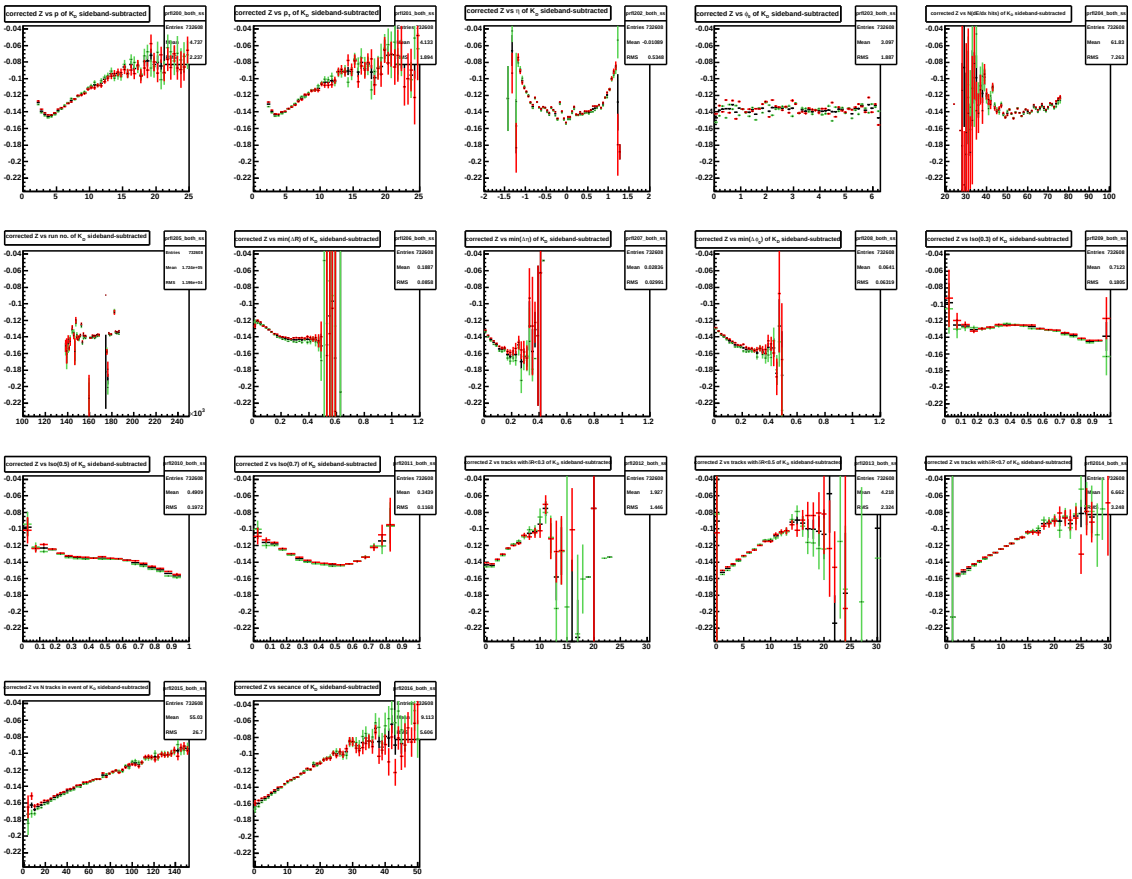

Figure A.8: Kaon Z profiles for the `xbhd0d` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

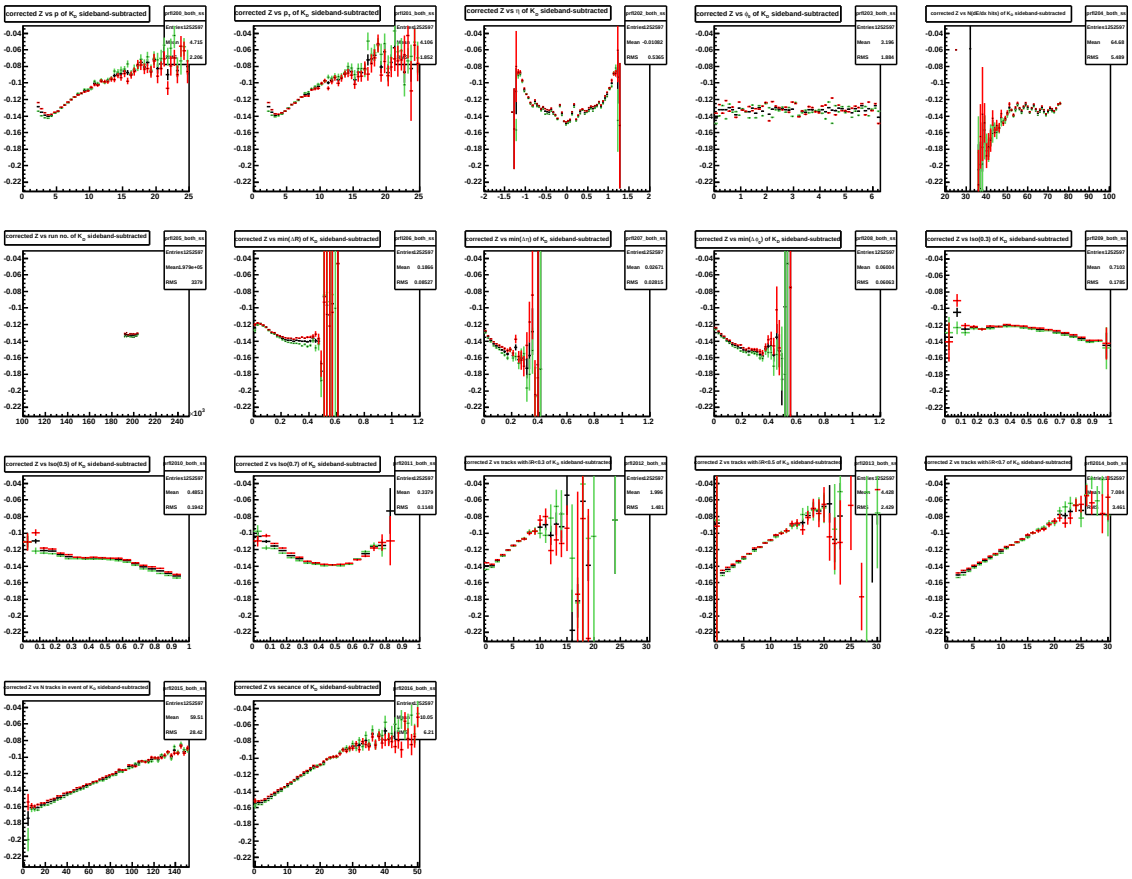

Figure A.9: Kaon Z profiles for the `xbhd0h` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

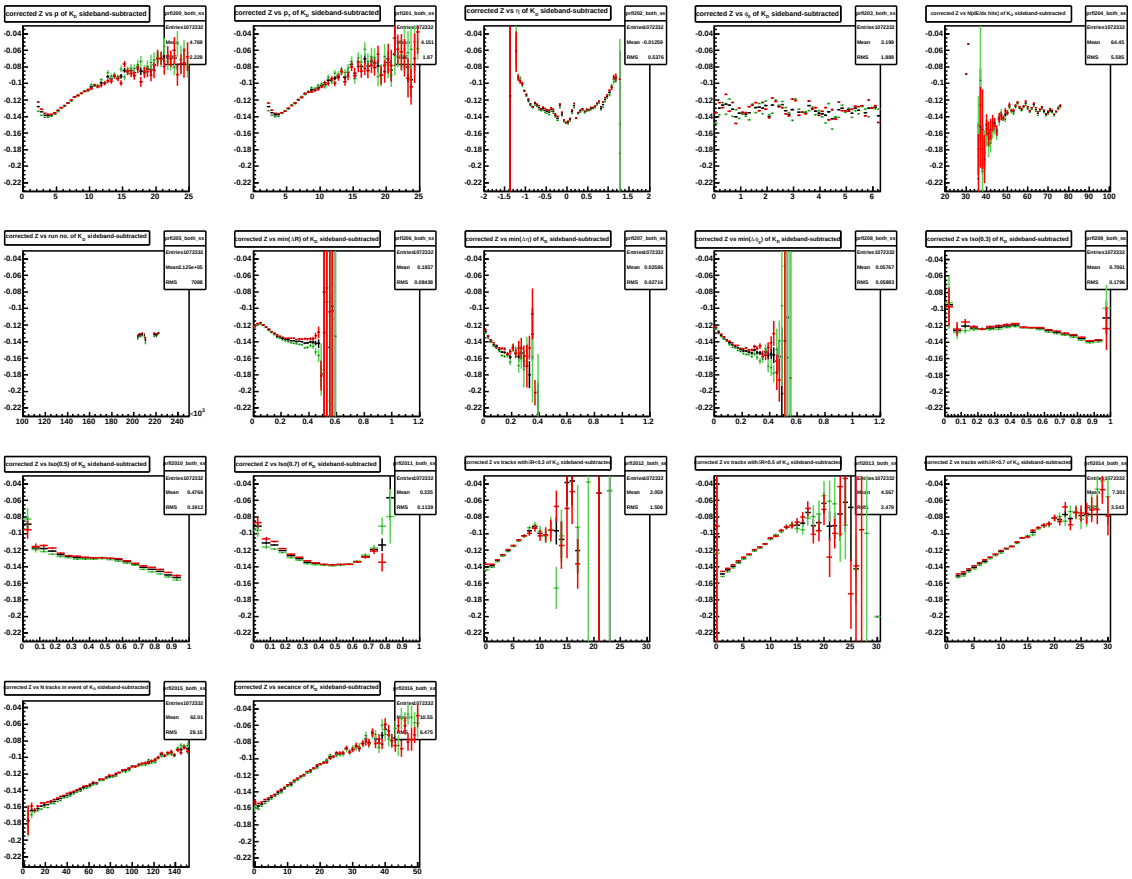

Figure A.10: Kaon Z profiles for the `xbhd0i` prompt D^{*+} sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

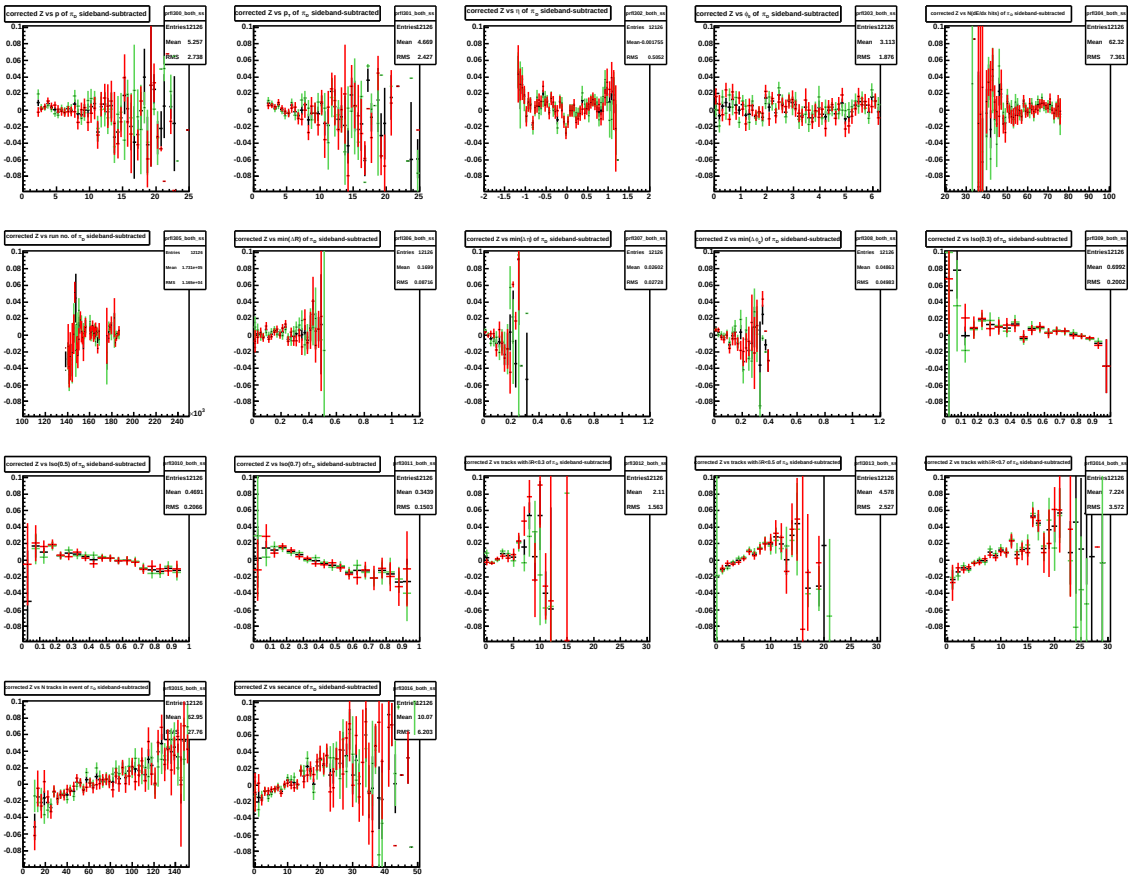

Figure A.11: Pion Z profiles for the `xbhd0d` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

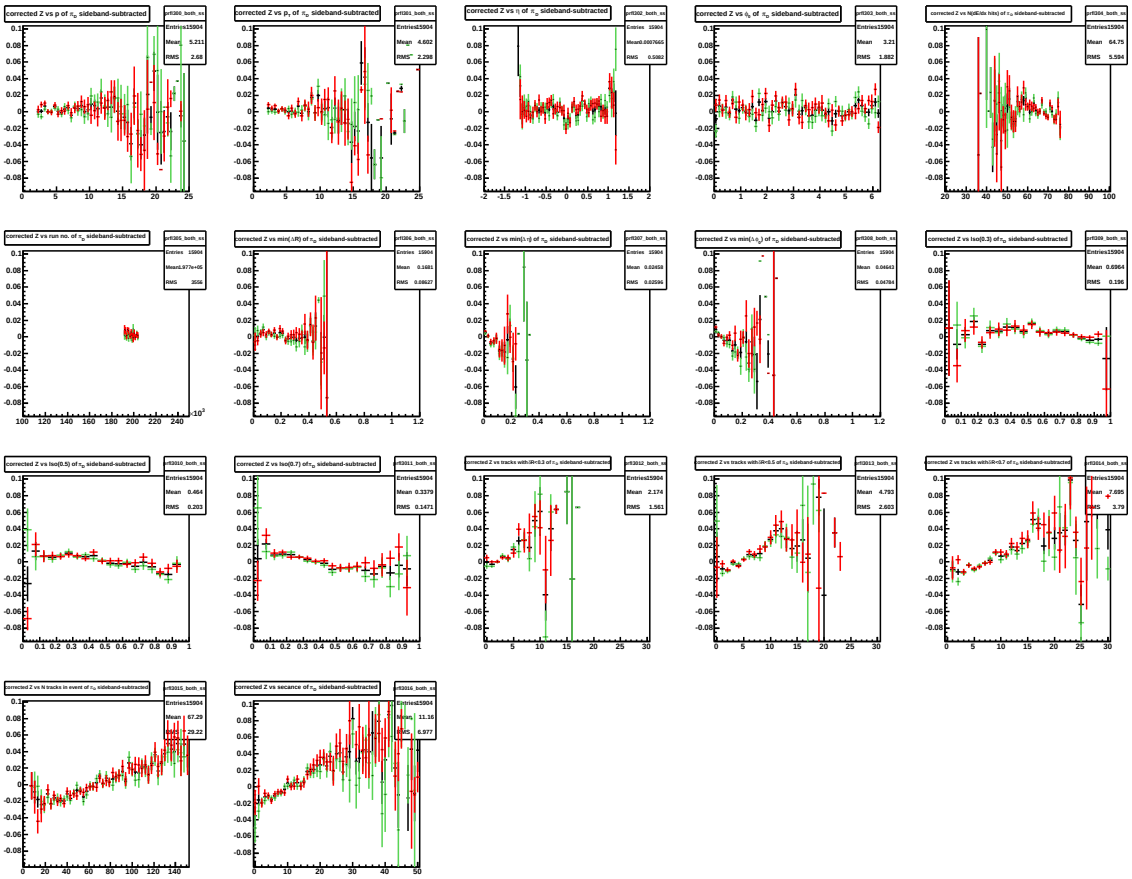

Figure A.12: Pion Z profiles for the `xbhd0h` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, $\min(\Delta R)$, $\min(\Delta\eta)$, $\min(\Delta\phi)$, iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

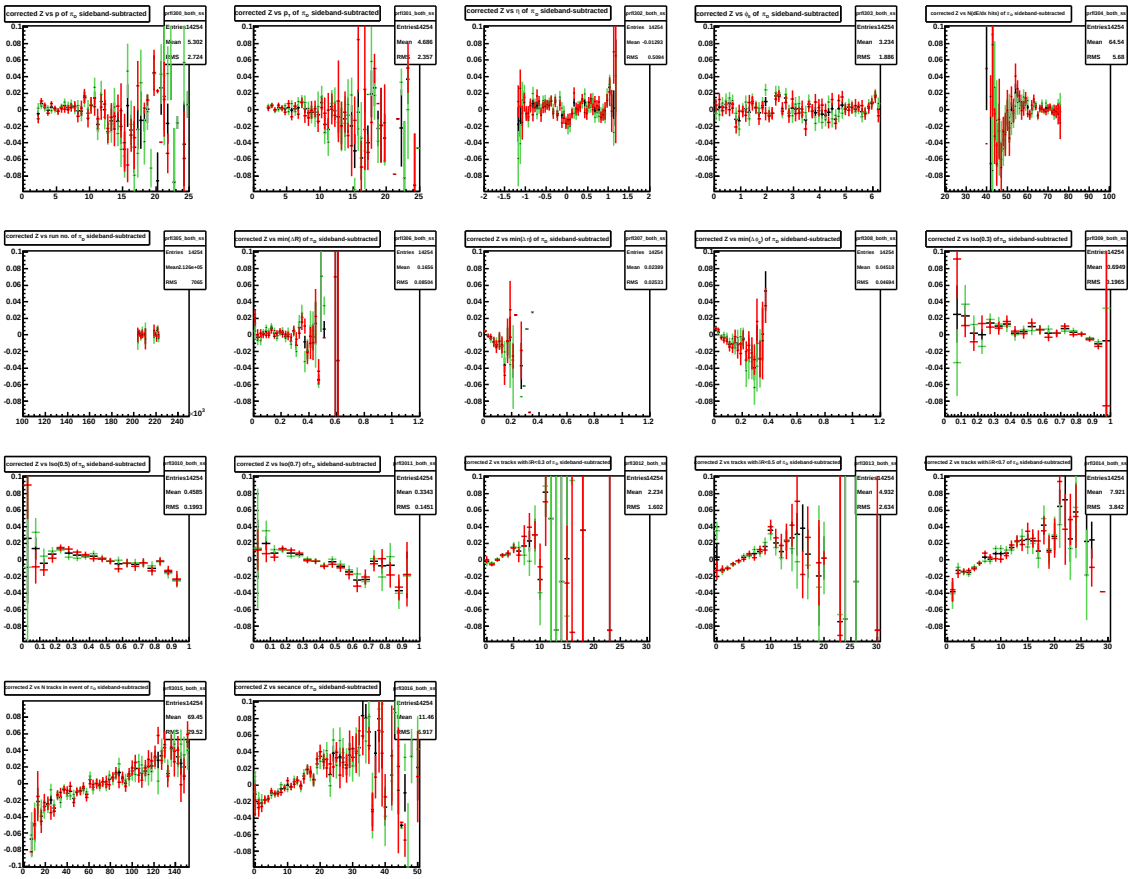

Figure A.13: Pion Z profiles for the `xbhd0i` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

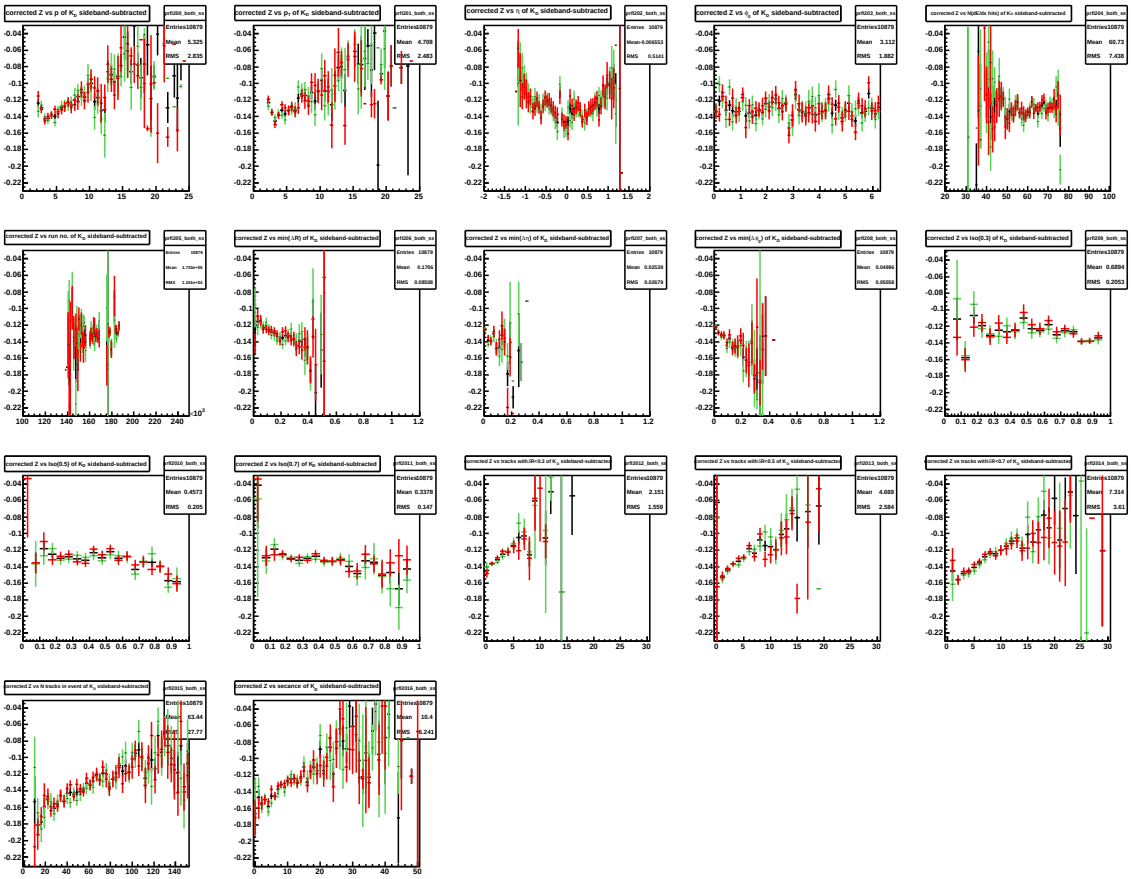

Figure A.14: Kaon Z profiles for the `xbhd0d` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, $\min(\Delta R)$, $\min(\Delta\eta)$, $\min(\Delta\phi)$, iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

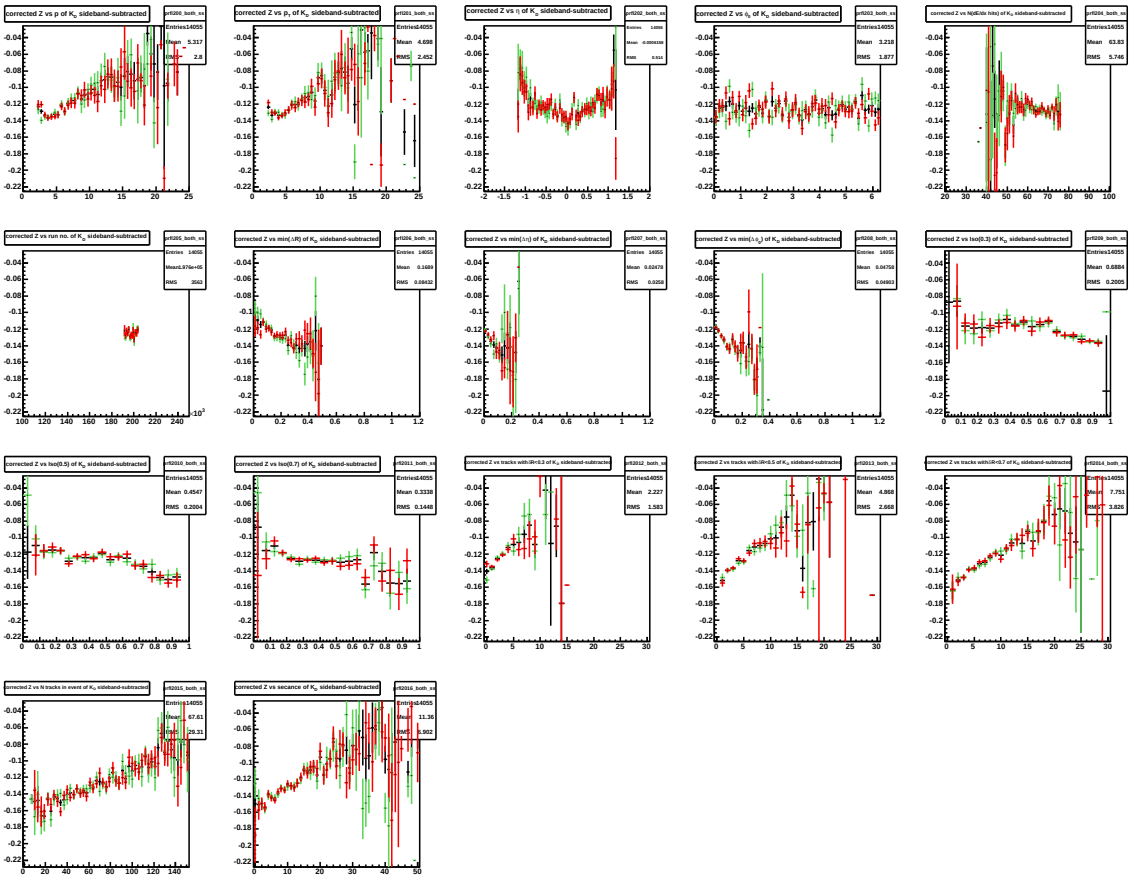

Figure A.15: Kaon Z profiles for the `xbhd0h` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

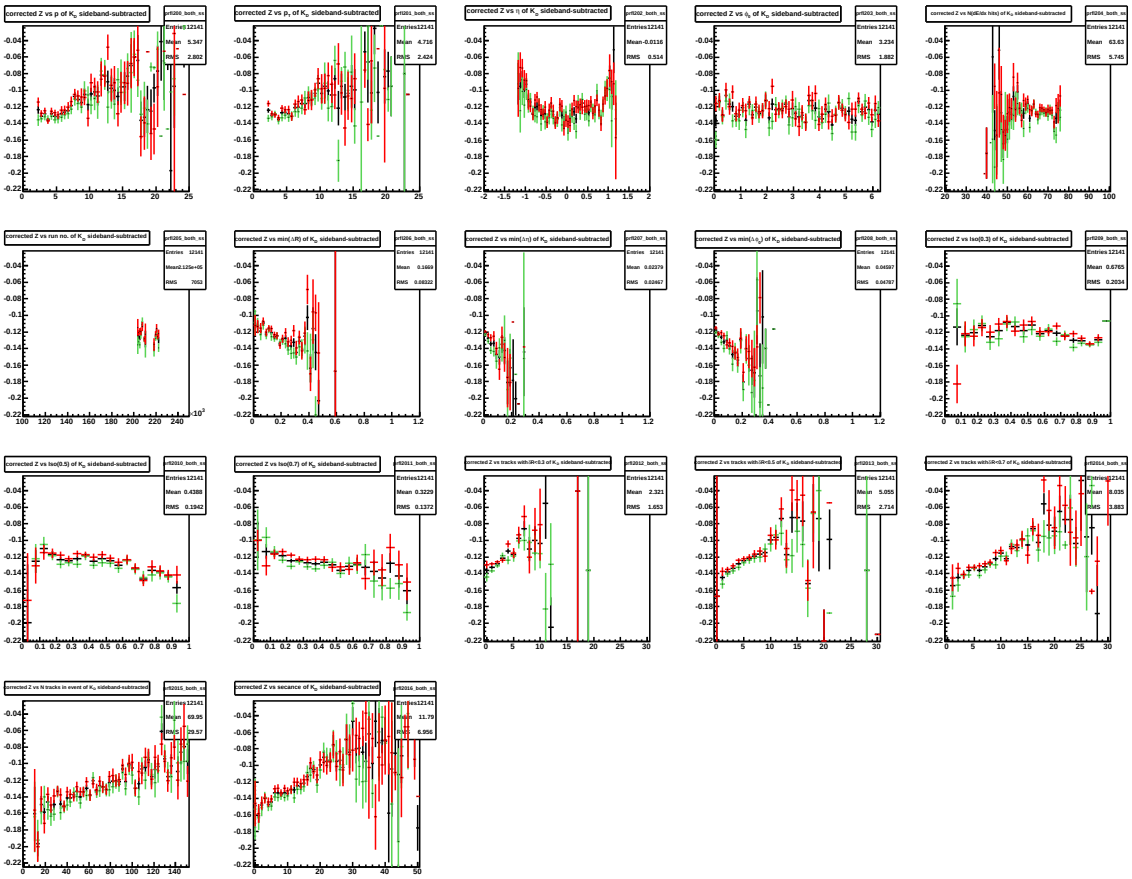

Figure A.16: Kaon Z profiles for the `xbhd0i` $\overline{B}^0 \to D^{*+}X$ sample after sideband subtraction in Δm. The variables against which Z is plotted are (top row) p, p_T, η, ϕ_0, N_{hits}; (second row) run number, min(ΔR), min($\Delta\eta$), min($\Delta\phi$), iso(0.3); (third row) iso(0.5), iso(0.7), $N_{\text{trks}}(0.3)$, $N_{\text{trks}}(0.5)$, $N_{\text{trks}}(0.7)$; (bottom row) N_{trks}, secance. The full y-axis range is roughly 1.5 times the typical K–π separation.

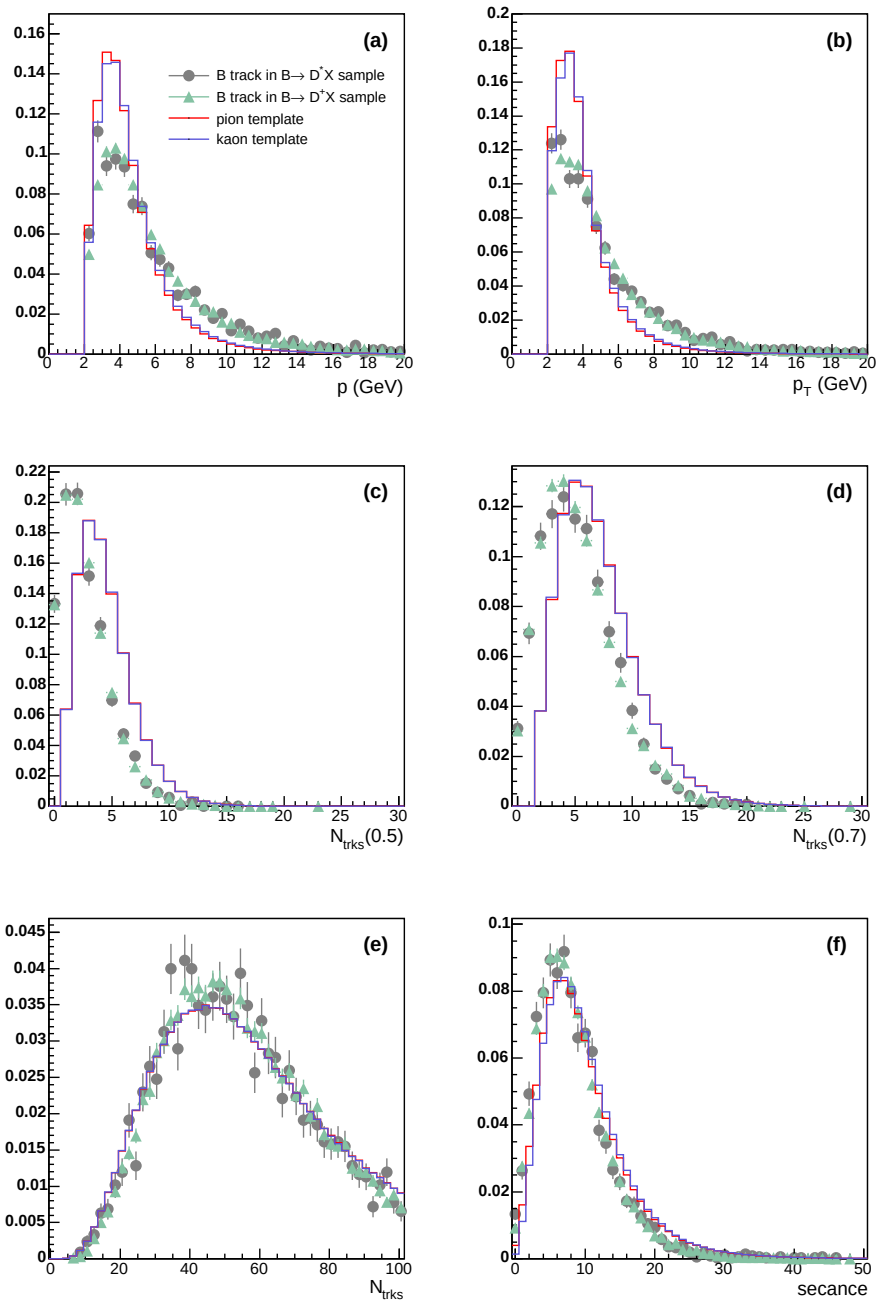

Figure A.17: Kinematic variable distributions of tracks from four samples. We plot (a) p, (b) p_T, (c) $N_{\text{trks}}(0.5)$, (d) $N_{\text{trks}}(0.7)$, (e) N_{trks} and (f) secance.

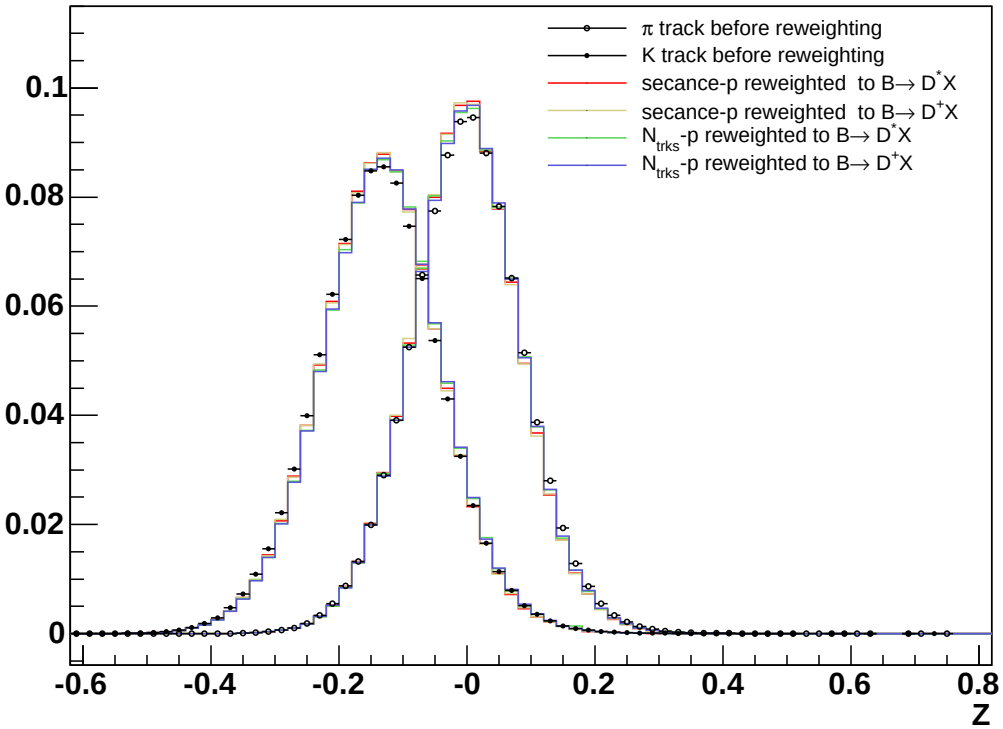

Figure A.18: Z templates in `xbhd0d` before and after reweighting. Histograms are shown in pairs of the same color; the left and right histograms are the kaon and pion templates.

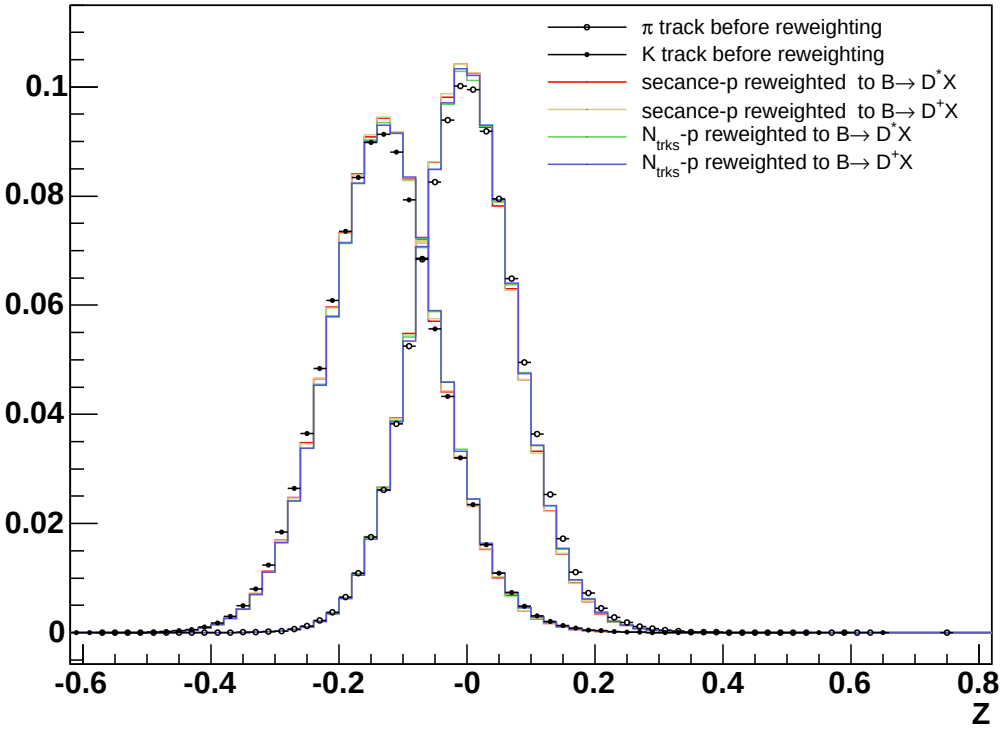

Figure A.19: Z templates in `xbhd0h` before and after reweighting. Histograms are shown in pairs of the same color; the left and right histograms are the kaon and pion templates.

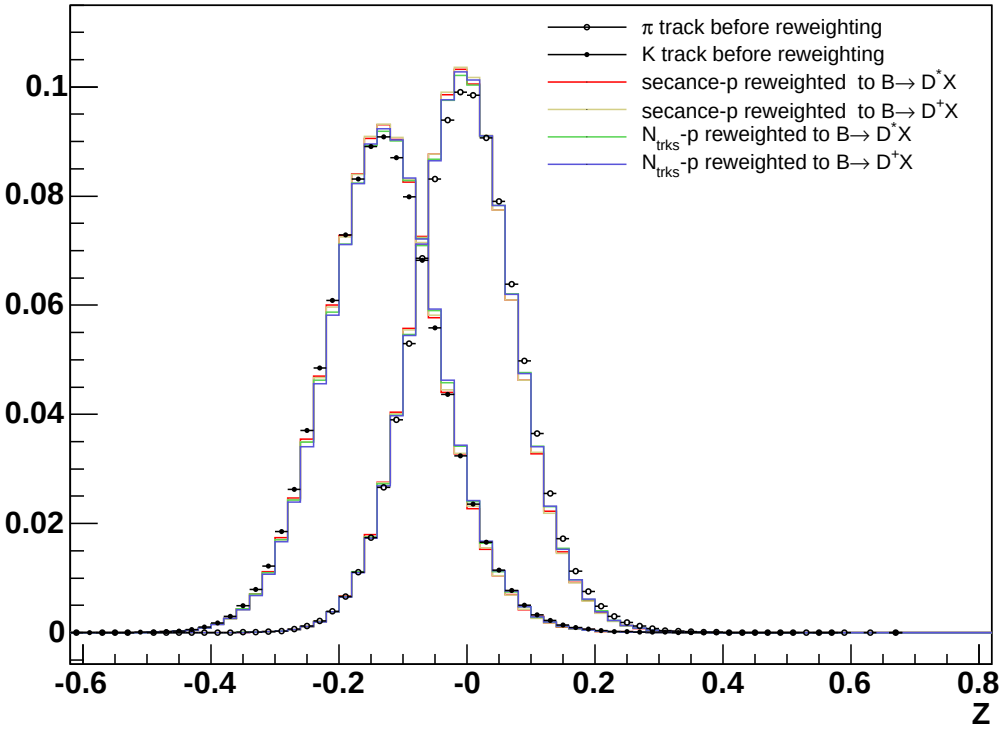

Figure A.20: Z templates in `xbhd0i` before and after reweighting. Histograms are shown in pairs of the same color; the left and right histograms are the kaon and pion templates.

Bibliography

[1] T. D. Lee and C.-N. Yang, Phys. Rev. **104**, 254 (1956).

[2] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, Phys. Rev. **105**, 1413 (1957).

[3] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

[4] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[5] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[6] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[7] J. Charles *et al.*, Eur. Phys. J. **C41**, 1 (2005), hep-ph/0406184, updated results and plots available at: http://ckmfitter.in2p3.fr.

[8] M. Bona *et al.*, JHEP **07**, 028 (2005), hep-ph/0501199, updated results and plots available at: http://utfit.roma1.infn.it.

[9] M. Bona *et al.*, JHEP **10**, 081 (2006), hep-ph/0606167.

[10] B. Aubert *et al.*, (2006), hep-ex/0607104.

[11] A. Poluektov *et al.*, Phys. Rev. **D73**, 112009 (2006), hep-ex/0604054.

[12] M. Gronau and D. London., Phys. Lett. **B253**, 483 (1991).

[13] M. Gronau and D. Wyler, Phys. Lett. **B265**, 172 (1991).

[14] M. Gronau, Phys. Rev. **D58**, 037301 (1998), hep-ph/9802315.

[15] D. Atwood, I. Dunietz and A. Soni, Phys. Rev. **D63**, 036005 (2001), hep-ph/0008090.

[16] A. Giri, Y. Grossman, A. Soffer and J. Zupan, Phys. Rev. **D68**, 054018 (2003), hep-ph/0303187.

[17] R. Aleksan, I. Dunietz and B. Kayser, Z. Phys. **C54**, 653 (1992).

[18] I. Dunietz, Phys. Rev. **D52**, 3048 (1995), hep-ph/9501287.

[19] A. Sill, Nucl. Instrum. Meth. **A447**, 1 (2000).

[20] A. A. Affolder *et al.*, Nucl. Instrum. Meth. **A526**, 249 (2004).

[21] D. Allspach *et al.*, IEEE Trans. Nucl. Sci. **52**, 2956 (2005).

[22] C. Hays, P. Tamburello, A. Kotwal, P. Wittich and R. Snider, The COT pattern recognition algorithm and offline code, CDF note 6992.

[23] C. Chen *et al.*, Measurement of the track reconstruction efficiency in the COT using a MC track embedding technique, CDF note 6394.

[24] K. Bloom and W.-M. Yao, "Outside-in" silicon tracking at CDF, CDF note 5991.

[25] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).

[26] B. Ashmanskas *et al.*, Nucl. Instrum. Meth. **A518**, 532 (2004), physics/0306169.

[27] G. Gomez-Ceballos *et al.*, Nucl. Instrum. Meth. **A518**, 522 (2004).

[28] W. M. Yao *et al.*, J. Phys. **G33**, 1 (2006), and 2007 partial update for the 2008 edition.

[29] J. Marriner, Secondary vertex fit with mass and pointing constraints (CTVMFT), CDF note 1996.

[30] A. Cerri, M. Shapiro, A. Heijboer and J. Kroll, Study of proper time scale factors with event by event primary vertexing, CDF note 7673.

[31] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).

[32] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. **B327**, 49 (1989).

[33] D. J. Lange, Nucl. Instrum. Meth. **A462**, 152 (2001).

[34] R. Brun, R. Hagelberg, M. Hansroul and J. C. Lassalle, CERN-DD-78-2-REV.

[35] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[36] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).

[37] S. D'Auria *et al.*, Track-based calibration of the COT specific ionization, CDF note 6932.

[38] S. Giagu *et al.*, BR ratios and direct *CP* violation in Cabibbo suppressed decays of D^0, CDF note 6391.

[39] B_s mixing group, Updated measurement of Δm_d and calibration of flavor tagging in fully reconstructed decays, CDF note 7722.

[40] K. Abe *et al.*, Phys. Rev. **D69**, 112002 (2004), hep-ex/0307021.

[41] A. Cerri, A. Deisher, H.-C. Fang, J. Mülmenstädt and M.D. Shapiro, Measurement of B^+ branching fractions using combined mass and dE/dx fits, CDF note 8777.

[42] B. Aubert *et al.*, Phys. Rev. Lett. **98**, 081801 (2007), hep-ex/0604012.

[43] A. Abulencia *et al.*, Phys. Rev. Lett. **98**, 061802 (2007), hep-ex/0610045.

[44] S.-S. Yu, J. Heinrich, N. Lockyer, D. Ambrose and P. Wittich, COT dE/dx Measurement and Corrections, CDF note 6361.

Index

B_CHARM trigger family, 22
Bgenerator, 36

candidate selection, 35
charged-particle tracking, 18
χ^2_{xy}, 35
CKM matrix, 6
CKMFitter, 8
coordinate system, 17
COT, 18
    gain drop, 18

D^{*+} satellite, 48
d_0 (impact parameter), 35
dE/dx corrections, 135
Δm
    sideband subtraction for Z pdf's, 135
    in D^{*+} reconstruction, 34
ΔR, 34

efficiency corrections, 40
η (pseudorapidity), 17
$\overline{\eta}$ (unitarity triangle parameter), 8
EvtGen, 36

fake-D templates, 37
f_j (fit parameters), 36

γ
    CKM angle, 8
    measurement in B^0_s decays, 9
    measurement in B^+ decays, 8

isolation
    of track, 138
    of B candidate, 35

kinematic efficiency, 40
Kobayashi-Maskawa mechanism, 4

λ (Cabibbo suppression factor), 7

Λ^+_c reflection
    into D^+, 48
    into D^+_s, 105
likelihood function, 36
L_{xy}, 34

mass constraint, 34
mass templates
    in $\overline{B}^0 \to D^{*+}X$ fit, 61
    in $\overline{B}^0 \to D^+X$ fit, 50
    in $\overline{B}^0_s \to D^+_sX$ fit, 108
    in general $B \to DX$ fits, 38
μ_j (fit parameters), 37

N_{trks}, 138
ntuples, 35

PHOTOS, 38
prescaled trigger, 22
primary vertex reconstruction, 34

radiative tail of $B \to D\pi$, 38
real-D templates, 37
reconstruction of B candidates, 34
remainder template
    in $\overline{B}^0 \to D^+X$ fit, 50
    in $\overline{B}^0 \to D^{*+}X$ fit, 61
    in $\overline{B}^0_s \to D^+_sX$ fit, 105
    in general $B \to DX$ fits, 38
residual discrepancies in $\overline{B}^0 \to D^+X$ fit, 53
$\overline{\rho}$ (unitarity triangle parameter), 8

scale factor
    in $\overline{B}^0 \to D^{*+}X$ fit, 61
    in $\overline{B}^0 \to D^+X$ fit, 57
    in $\overline{B}^0_s \to D^+_sX$ fit, 108
secance, 40, 138
semi-inclusive Monte Carlo samples, 38
signal subtraction, 39

single-B templates, 37
skimmed datasets, 35
smearing
    in $\overline{B}^0 \to D^{*+}X$ fit, 61
    in $\overline{B}^0 \to D^{+}X$ fit, 57
    in $\overline{B}^0_s \to D^{+}_s X$ fit, 108
superlayer, 18
SVT, 21
SVT χ^2, 35
SVT match distance, 35
SVX II, 17

templates
    fake-D, 37
    real-D, 37
    single-B, 37
θ_C (weak mixing angle), 7
tracking, 18
trigger system, 20
trigger scenario, 22
two-track trigger, 21

unitarity triangle, 8
UTFit, 8

vertex detector, 17
volunteer event, 21

weak mixing matrix, 6
Wolfenstein parameterization, 7
wrong-sign B candidates, 38

`xbhd0d`, `xbhd0h`, `xbhd0i` (datasets), 47
XFT, 21
XFT efficiency, 41

y (rapidity), 17

Z, 36
Z corrections, 135
Z templates
    in $\overline{B}^0 \to D^{*+}X$ fit, 61
    in $\overline{B}^0 \to D^{+}X$ fit, 51
    in $\overline{B}^0_s \to D^{+}_s X$ fit, 108
    in general $B \to DX$ fits, 39
z_0 (impact parameter), 35