CDF/PHYS/BOTTOM/CDFR/7559
version 2.0
April 24, 2005

The First Measurement of $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)/\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ at CDF II

Shin-Shan Eiko Yu, Joel Heinrich, Nigel Lockyer
University of Pennsylvania

Rick J. Tesarek, Dmitri Litvintsev
Fermilab

Abstract

We present measurements of the ratio of branching fractions of semileptonic decays of the B^0 and Λ_b relative to the fully reconstructed hadronic decay of similar topology. Using data taken with the CDF-II detector corresponding to a total sample size of 171.5 pb^{-1} we reconstruct , 579 ± 30, 106 ± 11, and 179 ± 19 of the statistically limiting $B^0 \to D^- \pi^+$, $B^0 \to D^{*-} \pi^+$ and $\Lambda_b \to \Lambda_c^+ \pi^-$ decays, respectively. After subtracting backgrounds we find:

$$\begin{aligned}
\frac{\mathcal{B}(\overline{B}^0 \to D^{*+} \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{*+} \pi^-)} &= 17.7 \ \pm 2.3\ (stat) \pm 0.6\ (syst) \pm 0.4\ (BR) \pm 1.1\ (UBR), \\
\frac{\mathcal{B}(\overline{B}^0 \to D^{+} \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{+} \pi^-)} &= 9.8 \ \pm 1.0\ (stat) \pm 0.6\ (syst) \pm 0.8\ (BR) \pm 0.9\ (UBR), \\
\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)} &= 20.0 \ \pm 3.0\ (stat) \pm 1.2\ (syst) \begin{smallmatrix}+0.7\\-2.1\end{smallmatrix}\ (BR) \pm 0.5\ (UBR).
\end{aligned}$$

where the uncertainties are from statistics, CDF internal systematics, external measured branching ratios and unmeasured branching ratios, respectively. The B^0 results are in good agreement with the world averages and will substantially improve our knowledge of these semileptonic branching ratios. The Λ_b results are the first of their kind and combining with additional information we determine the exclusive semileptonic branching fraction:

$$\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu) = \left(8.1 \ \pm 1.2\ (stat) \begin{smallmatrix}+1.1\\-1.6\end{smallmatrix}\ (syst) \pm 4.3\ (\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-))\right)$$

Contents

1 Difference between This Version and Version 2.0

Section 5.2 and Section 5.3

We now include a study of the acceptance using a more realistic decay model including form factors. We find the central value of the acceptance changes by $\sim$ 0.6%. Our Monte Carlo samples have a 2% statistical uncertainty for the acceptance.

Section 6.1 and 6.4

The latest $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ from CDF note 7558 [1] is used to normalize the semileptonic backgrounds to $\Lambda_b \to \Lambda_c^+ \pi^-$. The relative efficiency of each semileptonic background to the $\Lambda_b \to \Lambda_c^+ \pi^-$ decay in Tables 26–27 has reduced by 0.6%, after applying the scaling for the form factor decay model(see Section 5.3). The combination of the modified $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ and relative efficiencies reduces the total amount of physics background in the inclusive $\overline{B} \to \Lambda_c^+ \mu^- X$ events.

Section 7.1–7.5

The systematic uncertainties for the Λ_b relative branching ratios are re-calculated due to the change of $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$. In addition, we also include the systematic uncertainty from the scaling factor for the decay model.

Section 8–9

We re-evaluate $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)$ and compared our result with the recent DELPHI measurement.

Appendix A

We include the questions of the B group members we received and their answers.

2 Introduction

Hadrons containing a b-quark represent one of the most interesting topics in flavor physics. B mesons have been studied by several experiments and much is known about the lowsest lying mesons containing a b-quark. However, the situation is very different for the lowest lying baryon containing beauty, the Λ_b. According to the quark model the Λ_b contains u and d quarks in addition to the b-quark. Since the b-quark is much more massive than the the two lighter quarks one may use a heavy quark approximation to describe the system. The heavy quark approximation assumes the QCD properties of the a hadron are largely determined by the heavy quark. This means that in weak decays, the light quarks, the ud diquark in the Λ_b, act as spectators and the diquark properties do not change. More important, since the dominant decay for b-quarks is to a c-quark which is also considerably heavier than u and/or d quarks, the heavy symmetry is carried over to the daughter charm hadron. This heavy quark symmetry allows theorists to calculate the QCD properties of the b-hadrons, in particular, form factors and branching ratios. Specifically, the general description of the semileptonic decay of a baryond is described by six (6) form factors. Application of the heavy quark approximation reduces the number of form factor to two (2) and makes specific predictions on the nature of the form factors.

The Λ_b offers a unique test for theoretical models using heavy quark symmetry. In this paper we present a measurement of the ratio of branching fractions

$$\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \pi^-)}, \tag{1}$$

where we reconstruct the Λ_c^+ through its hadronic decay $\Lambda_c^+ \to pK^-\pi^+$. The careful reader will note that both the numerator and denominator modes contain four charge particles in the final state; three of those four particles originate from the Λ_c^+ which allows most systematic uncertainties to cancel in the final measurement.

Because such a measurement has never been performed, one would like to test our measurement technique on other systems. The B^0 system offers two such decays of a similar topology which have been measured previously. Specifically, we will test our techniques by measuring the additional ratios of branching fractions:

$$\frac{\mathcal{B}(B^0 \to D^- \mu^- \nu_\mu)}{\mathcal{B}(B^0 \to D^- \pi^+)} \tag{2}$$

$$\frac{\mathcal{B}(B^0 \to D^{*-} \mu^+ \nu_\mu)}{\mathcal{B}(B^0 \to D^{*-} \pi^+)}, \tag{3}$$

where we reconstruct the daughter decays $D^- \to K^+\pi^-\pi^-$ and $D^{*-} \to \overline{D}^0\pi^-$ with a subsequent decay $\overline{D}^0 \to K^+\pi^0$. Again, all of the decays under study contain four (4) tracks in the final state and three of the four tracks originate from a common parent for both the numerator and denominator decays.

In the following sections we describe the triggers and event selection used to obtain our data sample. We then describe how the final states are reconstructed and signals optimized. Once the number of events in each decay mode are determined, we describe the various backgrounds which must be subtracted from the semileptonic decays. We estimate systematic effects and finally present our final results.

3 Event Selection

Data used in this analysis are collected with the upgraded CDF detector from 9^{th} February 2002 to 6^{th} September 2003 and cover runs 138809 through 168889. This period corresponds to an integrated luminosity of $\sim$237 pb^{-1}. The data from `B_CHARM Scenario A` are processed with the `Production` executable, version 4.8.4, and compressed into the secondary datasets `hbot0h` and `hbot1i`. The total size of `hbot0h` and `hbot1i` is about 10 Terabytes (150M events), which is too big to be analyzed quickly multiple times. We apply loose selection cuts and reduce `hbot0h`, `hbot1i` to smaller, tertiary datasets. Yu [2] discusses the data skimming. Then we optimize the analysis cuts using the tertiary datasets in this section.

From the reduced datasets we reconstruct our signals:

- $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B} \to D^{*+}\mu^- X$, where $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$
- $\overline{B}^0 \to D^{+}\pi^-$ and $\overline{B} \to D^{+}\mu^- X$, where $D^{+} \to K^-\pi^+\pi^+$
- $\Lambda_b \to \Lambda_c^+\pi^-$ and $\overline{B} \to \Lambda_c^+\mu^- X$, where $\Lambda_c^+ \to pK^-\pi^+$

The reconstruction procedure is similar to that described in Yu [2]. The following cuts are studied more carefully and optimized :

- $\chi^2_{r-\phi}$ of B and charm vertex fit
- P_T of B and charm candidates
- $c\tau$ of B and charm candidates: $L_{xy} \times \frac{M}{P_T}$.

Our semileptonic signals are larger than the hadronic signals, and the statistical uncertainty of the relative branching fraction measurement is dominated by the uncertainty of the number of events in the hadronic signals. Therefore, we optimize the hadronic mode only and apply the optimized cuts to the semileptonic mode. The optimized quantity is the significance, $\frac{S}{\sqrt{S+B}}$, where "S" is the number of signal and "B" is the number of background events.

For our optimization, the amount of signal, "S" comes from a MC as described in Section 5.1. In order to scale the significance close to the true value measured from the data, we apply a normalization factor f_c on the signal MC,

$$f_c = \frac{S_{\rm data}}{S_{\rm MC}}, \tag{4}$$

where $S_{\rm data}$ and $S_{\rm MC}$ are the amount of the signal found in the data and MC after applying loose cuts, and

$$S = f_c \times S_{\rm MC}. \tag{5}$$

Figure 1 shows a comparison of the number of signal in the data and in the MC after applying the normalization factor.

We evaluate the background beneath the signal peak from the data. We first apply loose cuts on each mode to identify a clear B^0 or Λ_b peak;

- $c\tau(B) > 50$ μm
- each track $P_T > 0.5$ GeV/c
- π from the B hadron is CMU fiducial
- for $\overline{B}^0 \to D^{*+}\pi^-$:
  - $1.833 < M_{K\pi} < 1.893$ GeV/c^2
  - $0.143 < M_{K\pi\pi}$ - $M_{K\pi} < 0.148$ GeV/c^2

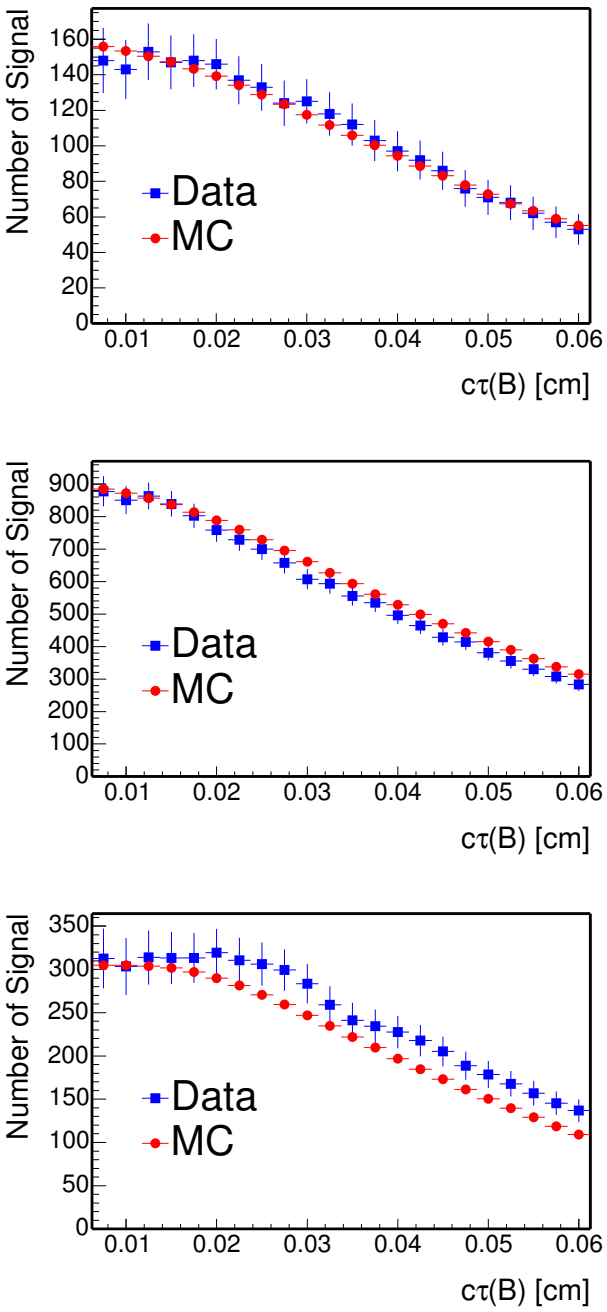


Figure 1: Signal optimization: number of signal in the MC and data as a function of the $c\tau(B)$ cut after applying the normalization factor for each hadronic mode. Top: $\overline{B}^0 \to D^{*+}\pi^-$. Middle: $\overline{B}^0 \to D^+\pi^-$. Bottom: $\Lambda_b \to \Lambda_c^+\pi^-$.

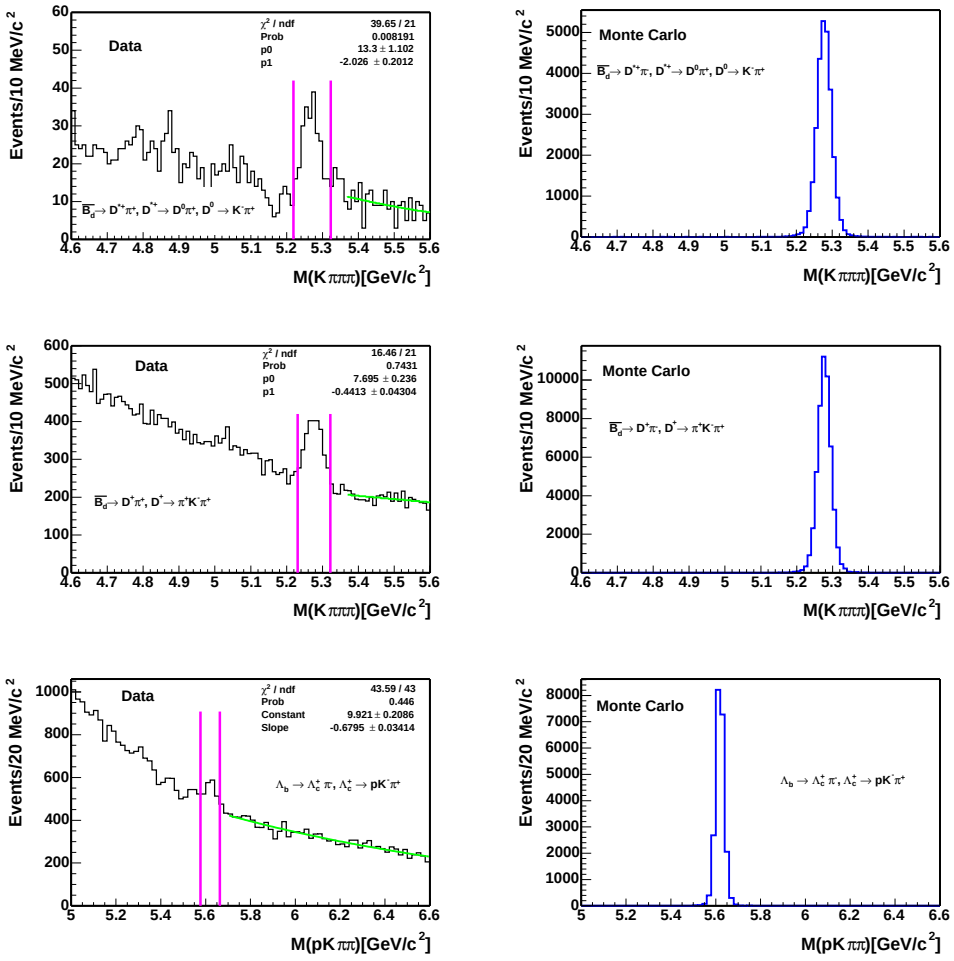


Figure 2: $D^*\pi$, $D\pi$ and $\Lambda_c\pi$ invariant mass for data and MC. From the top to the bottom rows are $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B}^0 \to D^+\pi^-$, and $\Lambda_b \to \Lambda_c^+\pi^-$ channels. Left: data, the pink solid lines indicate the signal region. Right: MC.

- for $\overline{B}^0 \to D^+\pi^-$: $1.8517 < M_{K\pi\pi} < 1.8837$ GeV/c^2
- for $\Lambda_b \to \Lambda_c^+\pi^-$: $2.269 < M_{pK\pi} < 2.302$ GeV/c^2

We require that both the muon and pion from the B hadron point within CMU fiducial volume because we use the CMU only to identify the muons. CMU covers the region of pseudo-rapidity (η) less than 0.6. Making the same fiducial requirement for the hadronic mode allows the tracking efficiencies from both modes to cancel.

The backgrounds in the signal and in the upper mass regions are mainly combinatorial, and may be described by an exponential function, as we will see in Section 4.2. Therefore, we fit the upper mass region to an exponential function. Finally we extrapolate and integrate the exponential over the mass region of $\pm$ 3 σ around the signal peak to obtain "B". Figure 2 shows the B hadron mass distribution in the data and MC. The figure also shows the signal region we define and the upper mass region we fit to an exponential.

The optimization follows an iterative procedure which passes through the data multiple times. In the first pass, cuts on each variable are scanned and optimization points are found. In the second

Table 1: Final analysis cuts shared by all the modes.

All	
P_T for all tracks	> 0.5 GeV/c
π_B and μ_B P_T	> 2.0 GeV/c
P_T of 4 tracks	> 6.0 GeV/c
P_T of charm hadron	> 5.0 GeV/c
μ_B CMU χ^2_x	< 9
every track exits at COT layer 95	
π_B and μ_B matched to SVT tracks and CMU fiducial	

pass, we apply the optimized cuts for all but the variable which is being re-optimized. We iterate this process several times until the optimization points become stable; usually twice is enough. Figures 3– 5 show $\frac{S}{\sqrt{S+B}}$, $\frac{S}{B}$ and $\frac{S}{S_{\rm ref}}$ as a function of each cut variable, where $S_{\rm ref}$ is the number of signal events at the starting point. Tables 1– 2 list the final analysis cuts. Note that because the MC and the data $\chi^2_{r-\phi}$ do not agree well, as shown in Section 5.2, we choose to make a loose cut at the plateau region of the significance. The final analysis cuts for the P_T of charm hadrons are tighter than the optimization points. The tighter cuts arise from the 4 GeV/c P_T threshold applied to the c-quark in the MC sample for our semileptonic background study (see Section 6.3.2). This P_T threshold makes the reconstruction of charm hadrons below 4 GeV/c inefficient. The MC sample is produced by the CDF B group and it would take a prohibitive amount of CPU time to generate a new sample more suitable for our analysis. Therefore, we increase the P_T cut of our charm hadrons to 5 GeV/c. As the significance of the charm is a slowly varying curve, changing the cuts has little effect on the signal yield.

In addition to the cuts which are optimized above, we also require that the muon and pion from the B hadron each matches an SVT track. Finally, for the semileptonic modes, we make cuts on the four track invariant mass (eg: $M(\Lambda_c\mu)$) to reduce the backgrounds from the other B decays, see Section 6.1 for more details. Figures 6– 9 give the signal and sideband distribution of each optimized variable in the $\overline{B} \to D^{*+}X$ and $\overline{B} \to D^{+}X$ data after $N-1$ cuts. The signal distribution is sideband subtracted as described in the MC and data comparison presented in Section 5.2. Figures 10– 11 give the signal and sideband distribution of each optimized variable in the $\Lambda_b \to \Lambda_c^+ X$ data after $N-1$ cuts. The signal distribution is obtained by fitting the number of signal events in bins of the variable as described in Section 5.2. The sideband distribution is from the following mass region:

- For the $\Lambda_b \to \Lambda_c^+\pi^-$ mode: $4\ \sigma < M_{pK\pi\pi}\ - 5.6204 < 7\ \sigma$
- For the $\overline{B} \to \Lambda_c^+\mu^- X$ mode: $4\ \sigma < |M_{pK\pi}\ - 2.285| < 6\ \sigma$

The optimization yields a S/B of 37.6 and 62.8 for the $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B} \to D^{*+}\mu^- X$ modes, 2.6 and 1.3 for the $\overline{B}^0 \to D^{+}\pi^-$ and $\overline{B} \to D^{+}\mu^- X$ modes, 1.6 and 0.3 for the $\Lambda_b \to \Lambda_c^+\pi^-$ and $\overline{B} \to \Lambda_c^+\mu^- X$ modes. Figure 12 shows the charm+π (left) and charm (right) mass spectra from the hadronic and inclusive semileptonic signals in the data after applying the optimized analysis cuts. We have reconstructed our signals in the data collected from the `B_CHARM Scenario A` trigger path. We have optimized our analysis cuts. In the next section, we will present the fit to the charm and B hadron mass spectra to obtain the number of signal events.

Table 2: Final analysis cuts for each mode.

$\overline{B} \to D^{*+}X$	
D^0 VertexFit $\chi^2_{r-\phi}$	< 16
4 track VertexFit $\chi^2_{r-\phi}$	< 17
$c\tau(D^0 \to B)$	$>$ -70 μm
$c\tau(B \to$ beamspot$)$	> 200 μm
$1.833 < M_{K\pi} < 1.893$ GeV/c^2	
$3.0 < M_{K\pi\pi\mu} < 5.3$ GeV/c^2 for $\overline{B} \to D^{*+}\mu^- X$	
$0.143 < \Delta m < 0.148$ GeV/c^2 for $\overline{B}^0 \to D^{*+}\pi^-$	
$\overline{B} \to D^{+}X$	
D^+ VertexFit $\chi^2_{r-\phi}$	< 14
4 track VertexFit $\chi^2_{r-\phi}$	< 15
$c\tau(D^+ \to B)$	$>$ -30 μm
$c\tau(B \to$ beamspot$)$	> 200 μm
$3.0 < M_{K\pi\pi\mu} < 5.3$ GeV/c^2 for $\overline{B} \to D^{+}\mu^- X$	
$1.8517 < M_{K\pi\pi} < 1.8837$ GeV/c^2 for $\overline{B}^0 \to D^{+}\pi^-$	
$\Lambda_b \to \Lambda_c^+ X$	
P_T of proton	> 2 GeV/c
Λ_c^+ VertexFit $\chi^2_{r-\phi}$	< 14
4 track VertexFit $\chi^2_{r-\phi}$	< 15
$c\tau(\Lambda_c^+ \to \Lambda_b)$	$>$ -70 μm
$c\tau(\Lambda_b \to$ beamspot$)$	> 250 μm
$3.7 < M_{pK\pi\mu} < 5.64$ GeV/c^2 for $\overline{B} \to \Lambda_c^+\mu^- X$	
$2.269 < M_{pK\pi} < 2.302$ GeV/c^2 for $\Lambda_b \to \Lambda_c^+\pi^-$	

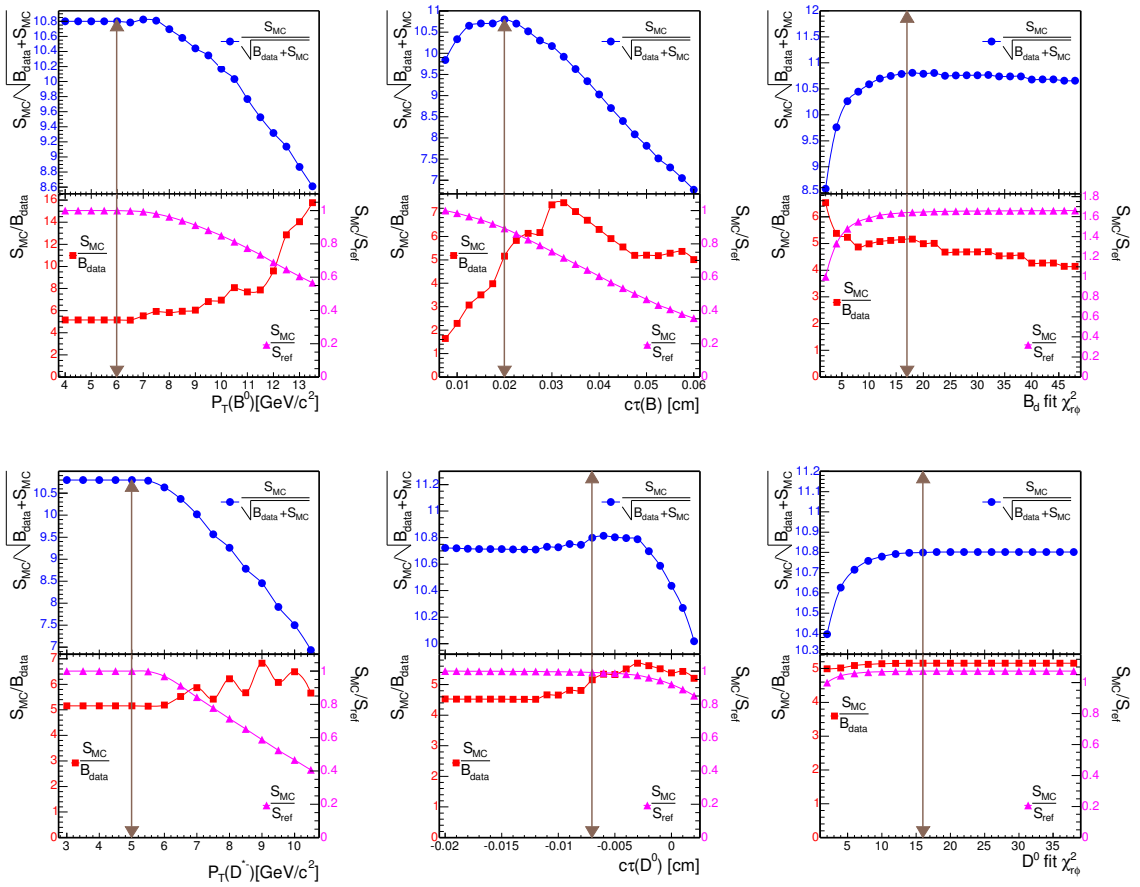


Figure 3: Significance ($\frac{S}{\sqrt{S+B}}$) marked by circles, signal/background ($\frac{S}{B}$) marked by squares and signal/reference ($\frac{S}{S_{\rm ref}}$) marked by triangles for cuts used in $\overline{B} \to D^{*+}X$ analysis.

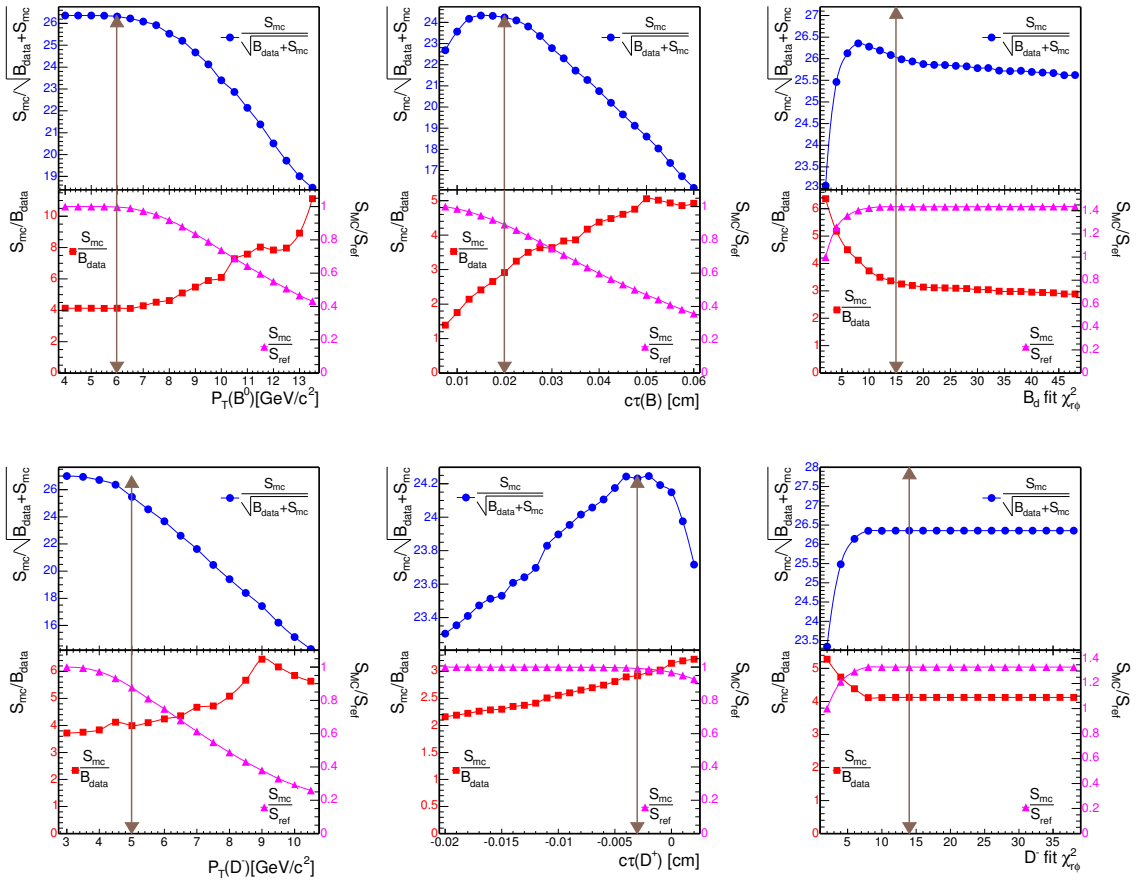


Figure 4: Significance ($\frac{S}{\sqrt{S+B}}$) marked by circles, signal/background ($\frac{S}{B}$) marked by squares and signal/reference ($\frac{S}{S_{\rm ref}}$) marked by triangles for cuts used in $\overline{B} \to D^+X$ analysis.

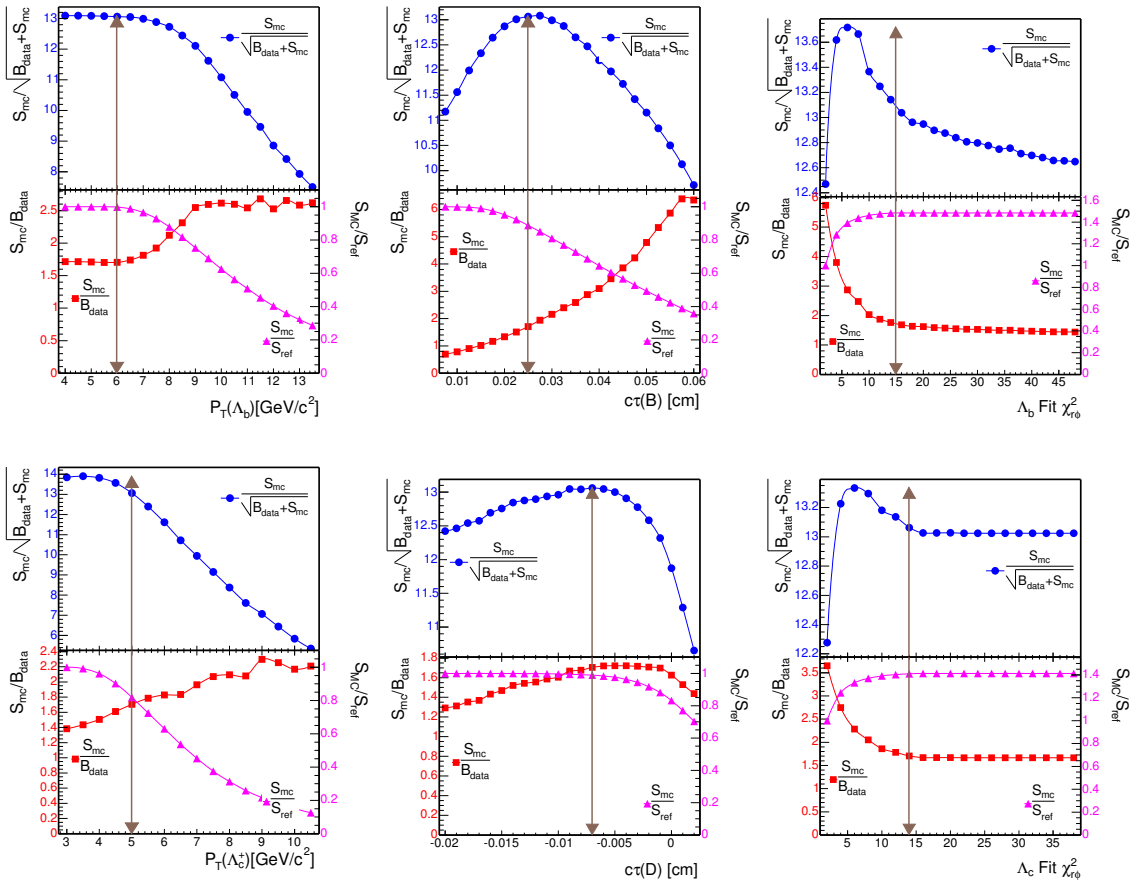


Figure 5: Significance ($\frac{S}{\sqrt{S+B}}$) marked by circles, signal/background ($\frac{S}{B}$) marked by squares and signal/reference ($\frac{S}{S_{\rm ref}}$) marked by triangles for cuts used in $\Lambda_b \to \Lambda_c^+ X$ analysis.

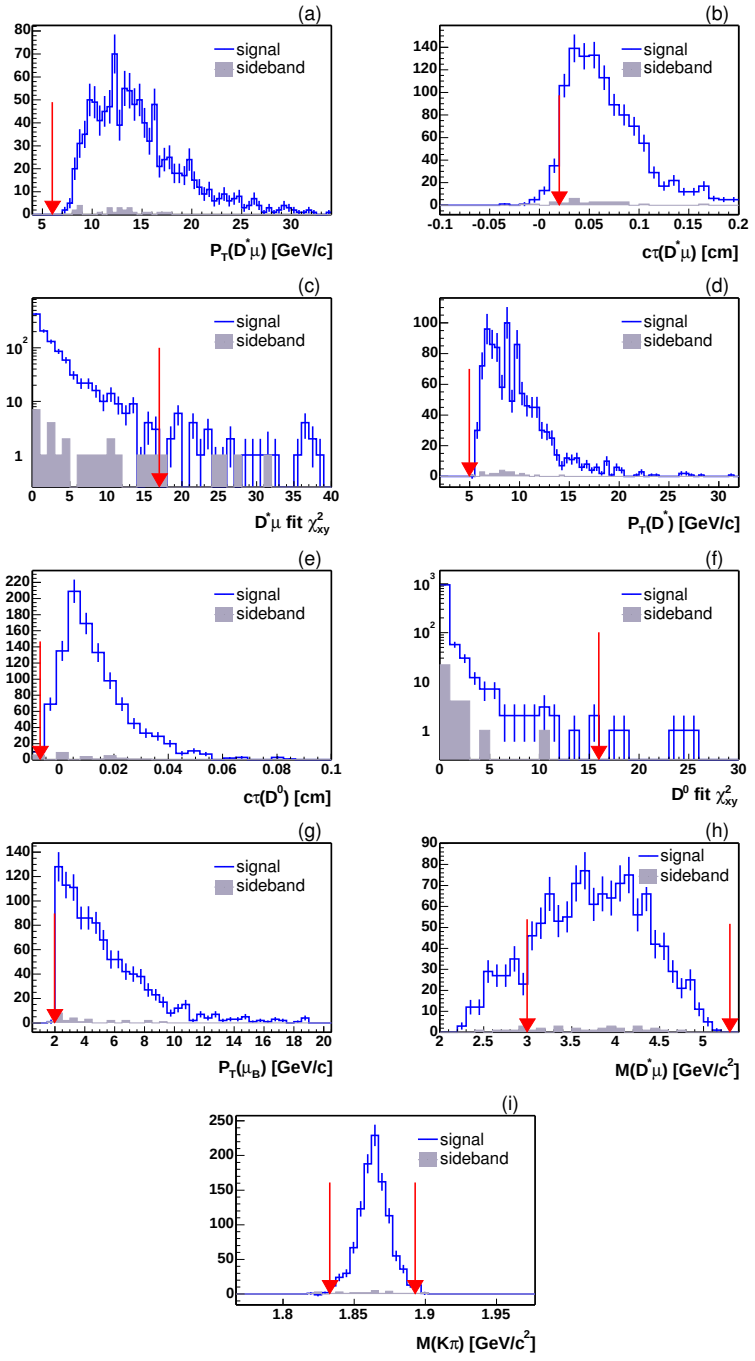


Figure 6: $\overline{B} \to D^{*+}\mu^- X$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(D^*\mu)$ $c\tau(D^*\mu)$, and the vertex fit $\chi^2_{r\phi}$ for the $D^*\mu$ vertex. (d)–(f): $P_T(D^{*+})$, $c\tau(D^0)$, and the vertex fit $\chi^2_{r\phi}$ for D^0 vertex. (g)–(i): $P_T(\mu_B)$, $M(D^*\mu)$ and $M_{K\pi}$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of the mass difference: $M_{K\pi\pi}$ - $M_{K\pi}$.

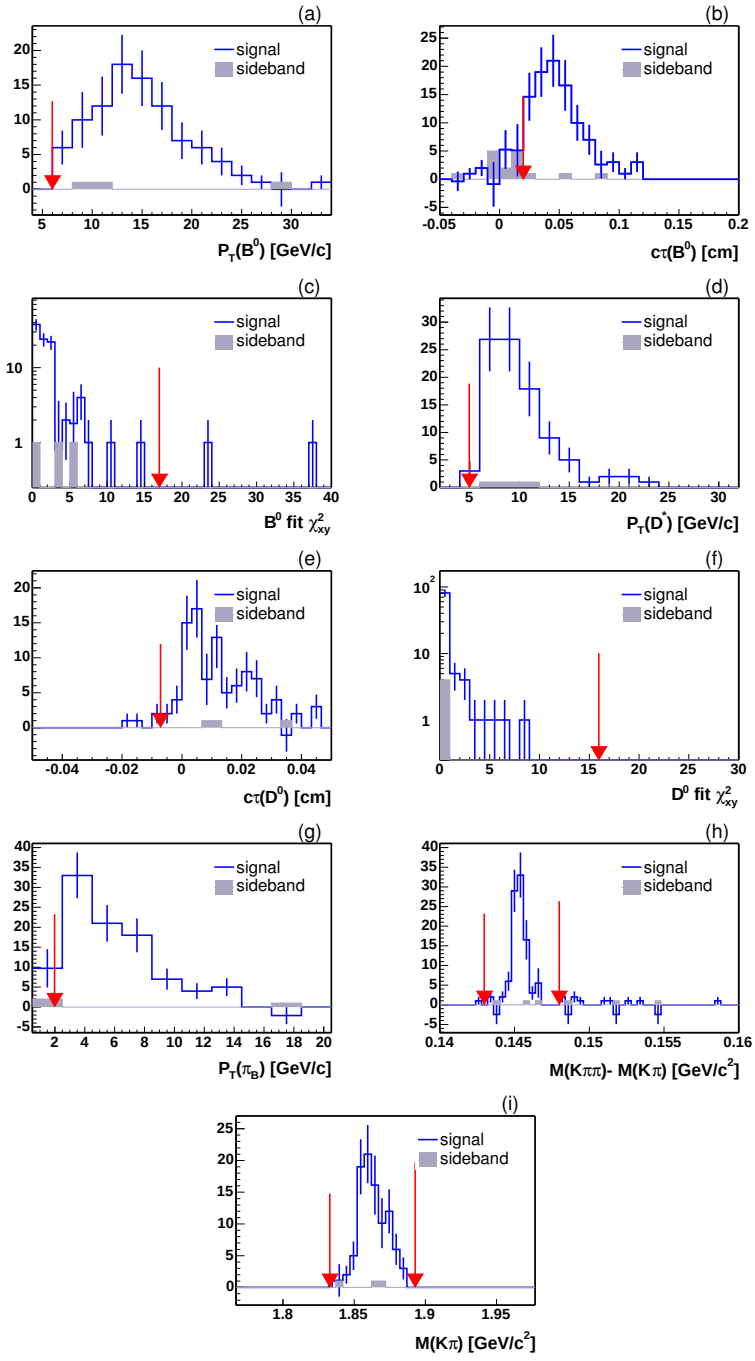


Figure 7: $\overline{B}^0 \to D^{*+}\pi^-$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(B^0)$, $c\tau(B^0)$, and the vertex fit $\chi^2_{r\phi}$ for the B^0 vertex. (d)–(f): $P_T(D^{*+})$, $c\tau(D^0)$, and the vertex fit $\chi^2_{r\phi}$ for D^0 vertex. (g)–(i): $P_T(\pi_B)$, $M_{K\pi\pi}$ - $M_{K\pi}$ and $M_{K\pi}$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of M_{B^0}.

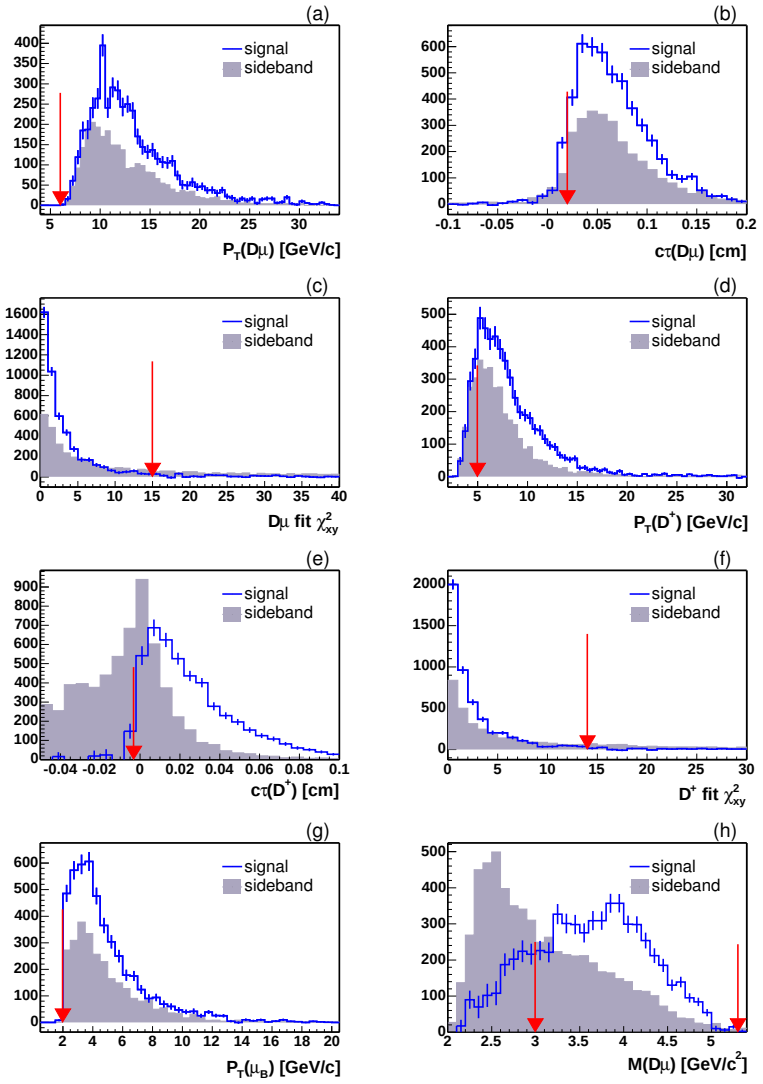


Figure 8: $\overline{B} \to D^+\mu^- X$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(D\mu)$, $c\tau(D\mu)$, and the vertex fit $\chi^2_{r\phi}$ for the $D\mu$ vertex. (d)–(f): $P_T(D^+)$, $c\tau(D^+)$, and the vertex fit $\chi^2_{r\phi}$ for D^+ vertex. (g)–(h): $P_T(\mu_B)$ and $M(D\mu)$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of $M_{K\pi\pi}$.

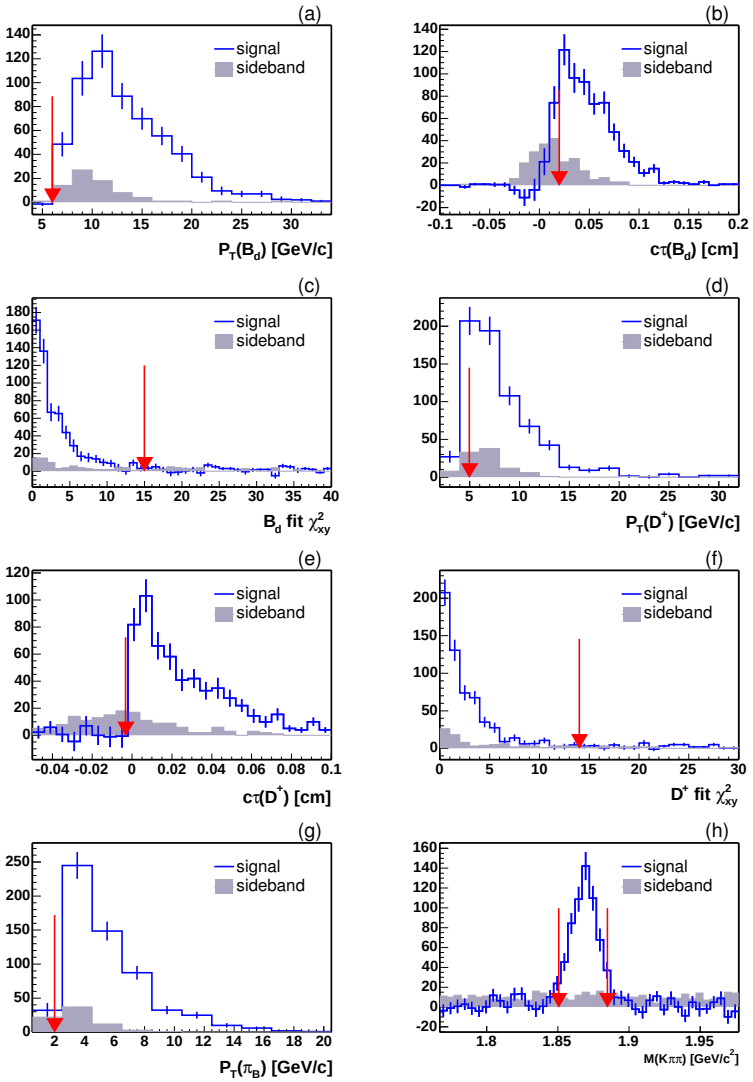


Figure 9: $\overline{B}^0 \to D^+\pi^-$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(B^0)$, $c\tau(B^0)$, and the vertex fit $\chi^2_{r\phi}$ for the B^0 vertex. (d)–(f): $P_T(D^+)$, $c\tau(D^+)$, and the vertex fit $\chi^2_{r\phi}$ for D^+ vertex. (g)–(h): $P_T(\pi_B)$, and $M_{K\pi\pi}$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of M_{B^0}.

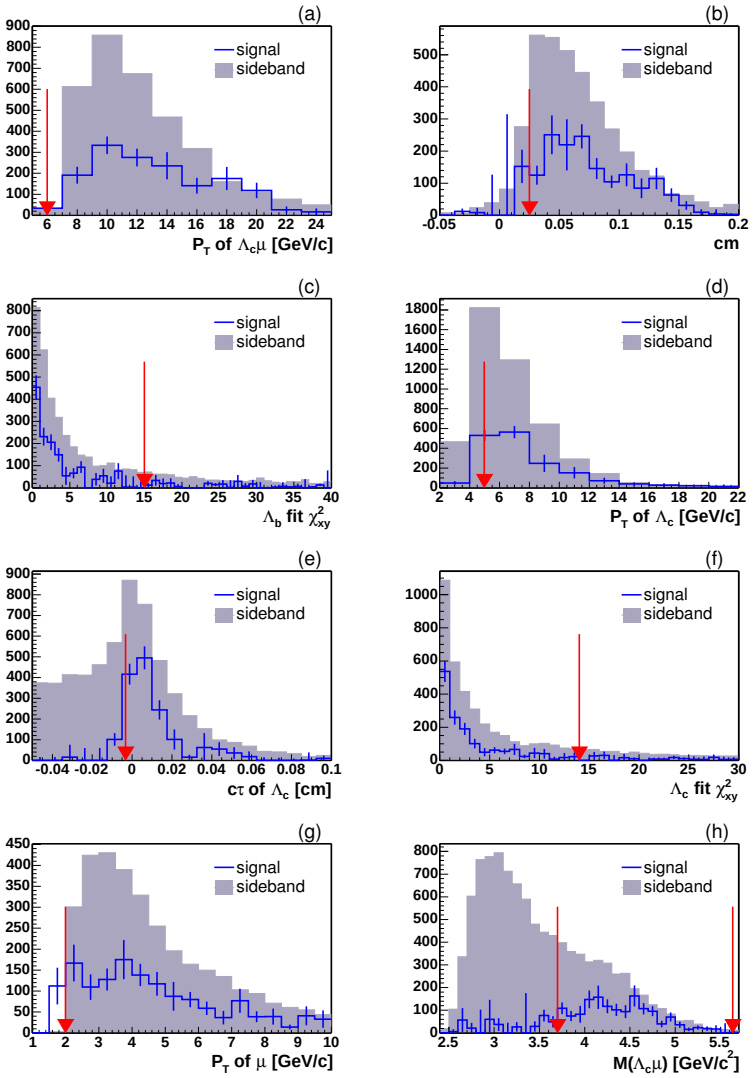


Figure 10: $\overline{B} \to \Lambda_c^+ \mu^- X$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(\Lambda_c\mu)$, $c\tau(\Lambda_c\mu)$, and the vertex fit $\chi^2_{r\phi}$ for the $\Lambda_c\mu$ vertex. (d)–(f): $P_T(\Lambda_c^+)$, $c\tau(\Lambda_c^+)$, and the vertex fit $\chi^2_{r\phi}$ for Λ_c^+ vertex. (g)–(h): $P_T(\mu_B)$ and $M(\Lambda_c\mu)$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of $M_{pK\pi}$.

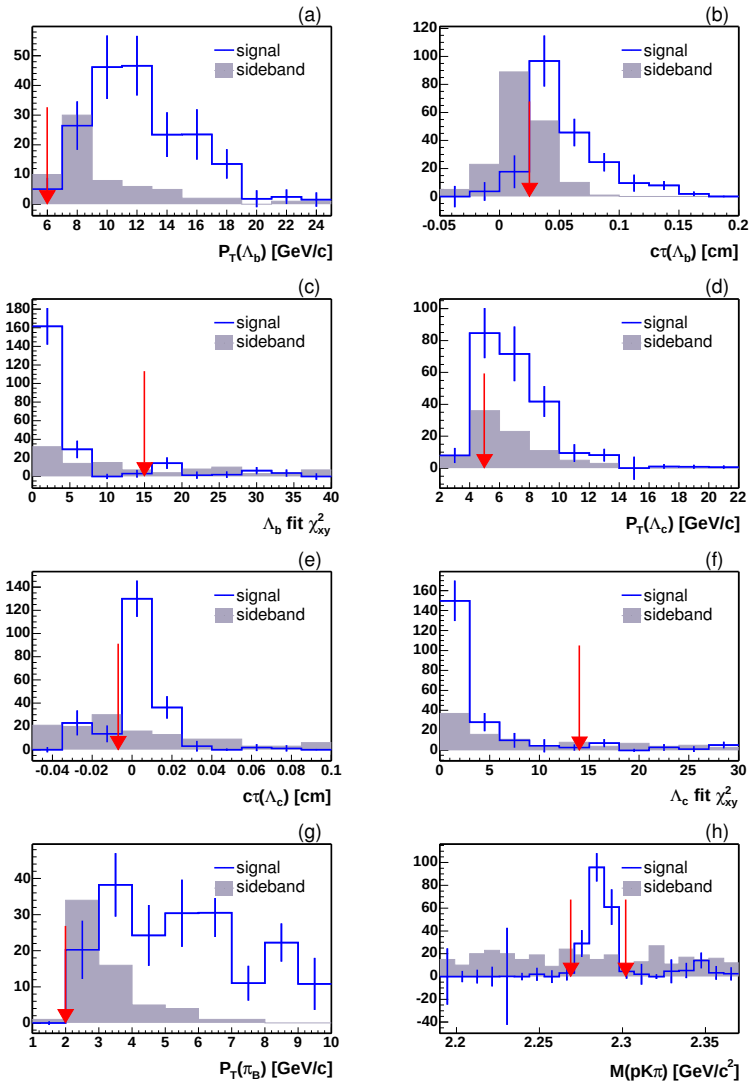


Figure 11: $\Lambda_b \to \Lambda_c^+ \pi^-$ cut variables from the data after sideband subtraction. (a)–(c): $P_T(\Lambda_b)$, $c\tau(\Lambda_b)$, and the vertex fit $\chi^2_{r\phi}$ for the Λ_b vertex. (d)–(f): $P_T(\Lambda_c^+)$, $c\tau(\Lambda_c^+)$, and the vertex fit $\chi^2_{r\phi}$ for Λ_c^+ vertex. (g)–(h): $P_T(\pi_B)$, and $M_{pK\pi}$. Arrows indicate the point of the final analysis cuts. The shaded histograms are the distribution from the sideband of M_{Λ_b}.

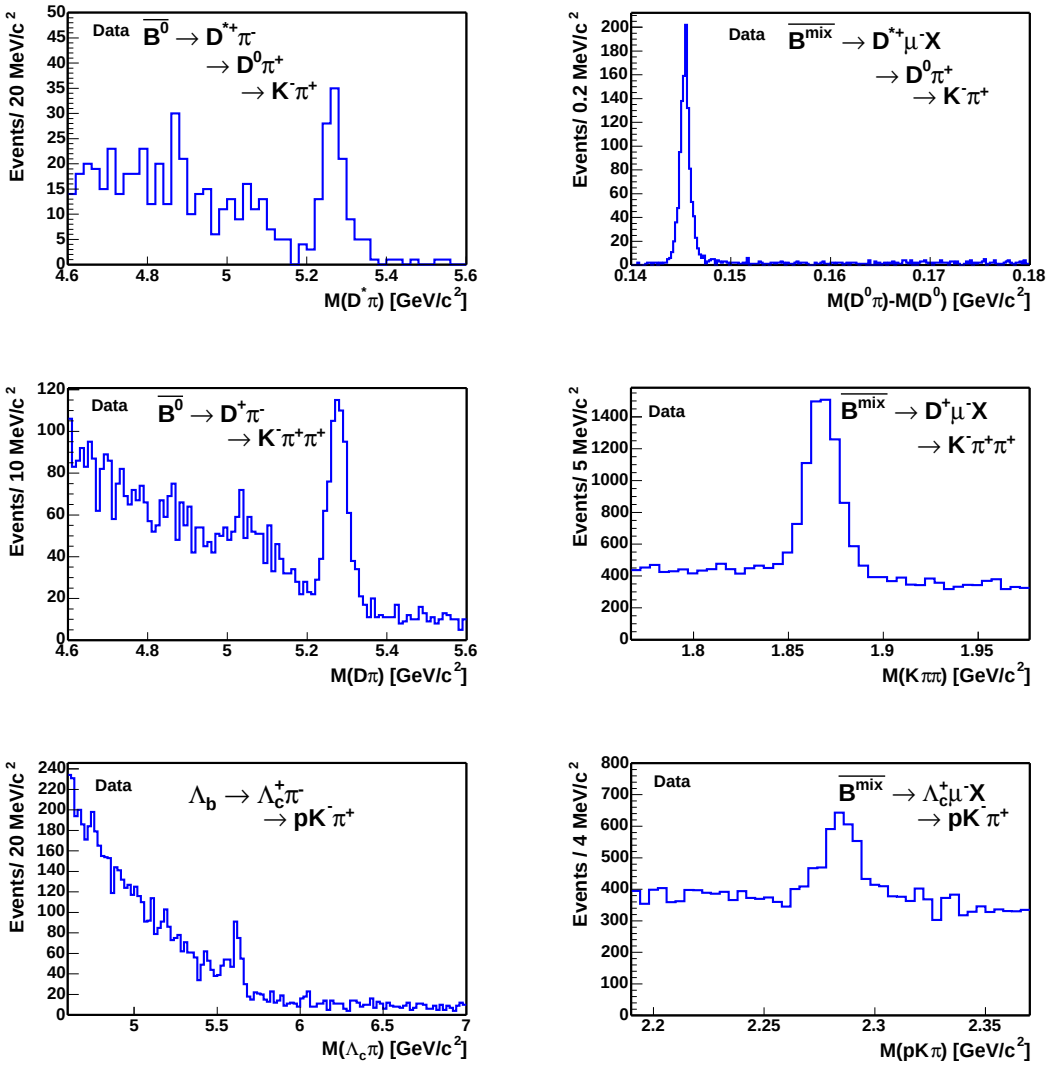


Figure 12: charm+π and charm mass spectra from our signals after all cuts. From the top left to the bottom right are: $M_{D^*\pi}$, $M_{D^0\pi} - M_{D^0}$, $M_{D\pi}$, $M_{K\pi\pi}$, $M_{\Lambda_c}\pi$ and $M_{pK\pi}$.

4 Signal Yield in the Data

In this section, we explain how the signal yield in the data is extracted. We extract the yield by fitting the charm+π (or charm) mass spectra in Figure 12 to a function which describes both the signal and the background. We integrate the signal function to obtain the yield. The signal function for all modes is a Gaussian or double-Gaussians. The background function varies with the decay mode. All our fits use an unbinned, extended likelihood technique. The general extended likelihood function ($\mathcal{L}$) is expressed as:

$$\log \mathcal{L} = \sum_i \log\{N_{\rm sig} \cdot \mathcal{S}(m_i) + N_{\rm bg} \cdot B(m_i)\} - N_{\rm sig} - N_{\rm bg} + \log \mathcal{C}, \tag{6}$$

where i represents i^{th} event, m represents the reconstructed charm+π (charm) mass. The amounts of signal and background are denoted as $N_{\rm sig}$ and $N_{\rm bg}$, respectively, while $\mathcal{S}(m)$ ($\mathcal{B}(m)$) are the functions which describe the signal (background) mass spectrum. The last term in Equation 6, $\mathcal{C}$, is a Gaussian constraint on one fit parameter, x:

$$\mathcal{C} = \mathcal{G}(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{(x-\mu)}{\sigma})^2}, \tag{7}$$

where we constrain the variable x around the mean μ. The difference of x and μ follows a Gaussian distribution with an uncertainty σ. The unbinned likelihood fitter calls the `MINUIT` package. `MINUIT` varies the fit parameters to minimizes $-2 \cdot \log \mathcal{L}$.

The performance of the fitter was checked on 1000 toy MC samples similar to the data distribution. We plot the pull distribution for each parameter. For a large number of toy MC tests, the pull is expected to follow a Gaussian distribution. We examine if the fitter returns an output consistent with the input, i.e. if the mean of the pull distribution is consistent with zero and if the width is consistent with one. Note that the μ and σ of the Gaussian constraint in Equation 7 are determined from a subsidiary measurement using the data and the MC. Therefore, we simulate this measurement in the toy MC test, by smearing the mean of the constraint with a Gaussian distribution of mean μ and sigma σ in Equation 7. In order to evaluate the quality of the fit, we also superimpose the fit result on the data histograms and compute a χ^2. Remark that as the B hadrons are fully reconstructed in the hadronic channels, the yields we extract are the true amount of signal for this analysis. The yields we extract for the inclusive semileptonic channels include the exclusive signals and indistinguishable backgrounds: such as muon fakes, decays from $b\bar{b}$, $c\bar{c}$, or other B hadrons. These backgrounds will be estimated in Section 6 and subtracted in the calculation of the relative branching ratios.

4.1 Mass Fit of the Semileptonic Modes

4.1.1 $D^*\mu$ Yield

As seen in Figure 12 (top right), the events with $D^*\mu$ in the final state have almost no combinatorial background. The combinatorial background is reduced largely by requiring $M_{K\pi}$ be consistent with the world average D^0 mass and cutting on the variable $M_{D^*\mu}$. The $M_{D^0\pi} - M_{D^0}$ distribution is fitted to a double Gaussian signal and a constant background. The extended log likelihood function is expressed as:

$$\begin{aligned} \log \mathcal{L} \;=\; & \sum_i \log\{N_{\rm sig} \cdot [(1-f_2) \cdot \mathcal{G}_1(m_i, \mu, \sigma_1) + f_2 \cdot \mathcal{G}_2(m_i, \mu, \sigma_2)] \\ & + N_{\rm bg} \cdot \frac{1}{M_{\rm max} - M_{\rm min}}\} - N_{\rm sig} - N_{\rm bg}, \end{aligned} \tag{8}$$

where f_2 is the fraction of the second Gaussian, The mass window $0.14 < M_{D^0\pi} - M_{D^0} < 0.18$ GeV/c^2 is specified by $M_{\rm min}$ and $M_{\rm max}$. Both Gaussians have the same mean but different sigmas. Table 3

Table 3: $D^*\mu$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	$N_{\rm sig}$		-0.023 ± 0.031	1.006 ± 0.023	1059 ± 33
2	f_2		0.002 ± 0.034	1.072 ± 0.024	0.56 ± 0.10
3	μ	[GeV/c^2]	0.049 ± 0.033	1.044 ± 0.024	0.145410 ± 0.000016
4	σ_1	[GeV/c^2]	-0.048 ± 0.033	1.052 ± 0.024	0.00031 ± 0.00004
5	σ_2	[GeV/c^2]	0.011 ± 0.032	1.031 ± 0.023	0.00071 ± 0.00006
6	$N_{\rm bg}$		0.010 ± 0.031	1.000 ± 0.022	321 ± 19

Table 4: Correlation coefficients returned from the fit ($D^*\mu$ data)

	1	2	3	4	5	6
1	1.000					
2	-0.050	1.000				
3	0.002	0.064	1.000			
4	0.040	-0.853	-0.031	1.000		
5	0.085	-0.859	-0.058	0.686	1.000	
6	-0.070	0.087	-0.004	-0.070	-0.147	1.000

lists the mean, width of the pull distribution from 1000 toy MC test and the fit value of each parameter from the unbinned likelihood fit to the data. Figure 13 shows the fit to each pull distribution. The toy MC test result indicates that the fitter returns a fit value consistent with the input. Table 4 gives the correlation coefficients returned from the likelihood fit to the data, where the index of each parameter follows that in Table 3. Figure 14 shows the fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\overline{B}\to D^{*+}\mu^- X} = 1059 \ \pm 33.$$

4.1.2 $D\mu$ Yield

A first glance of $M_{K\pi\pi}$ in Figure 12 (middle right) might suggest that we could fit $M_{K\pi\pi}$ to a Gaussian signal and a first-order polynomial background. But, since we do not apply particle identification (PID) in this analysis, the background under the signal contains not only the combinatorial background, but also contamination from the D_s decays. Not using PID means that a pion mass might be assigned to a kaon, and D_s^+ may be reconstructed as D^+. Figure 15 shows the mis-reconstructed D^+ mass spectrum from the $\overline{B}_s \to D_s^+\mu^-\overline{\nu}_\mu$ MC, where D_s are forced to decay into the final states listed in Table 5. These final states are selected after a study to identify the dominant D_s decays reconstructed in the D^+ mass window. The MC used to assess D_s background is produced as described in Section 5.1.

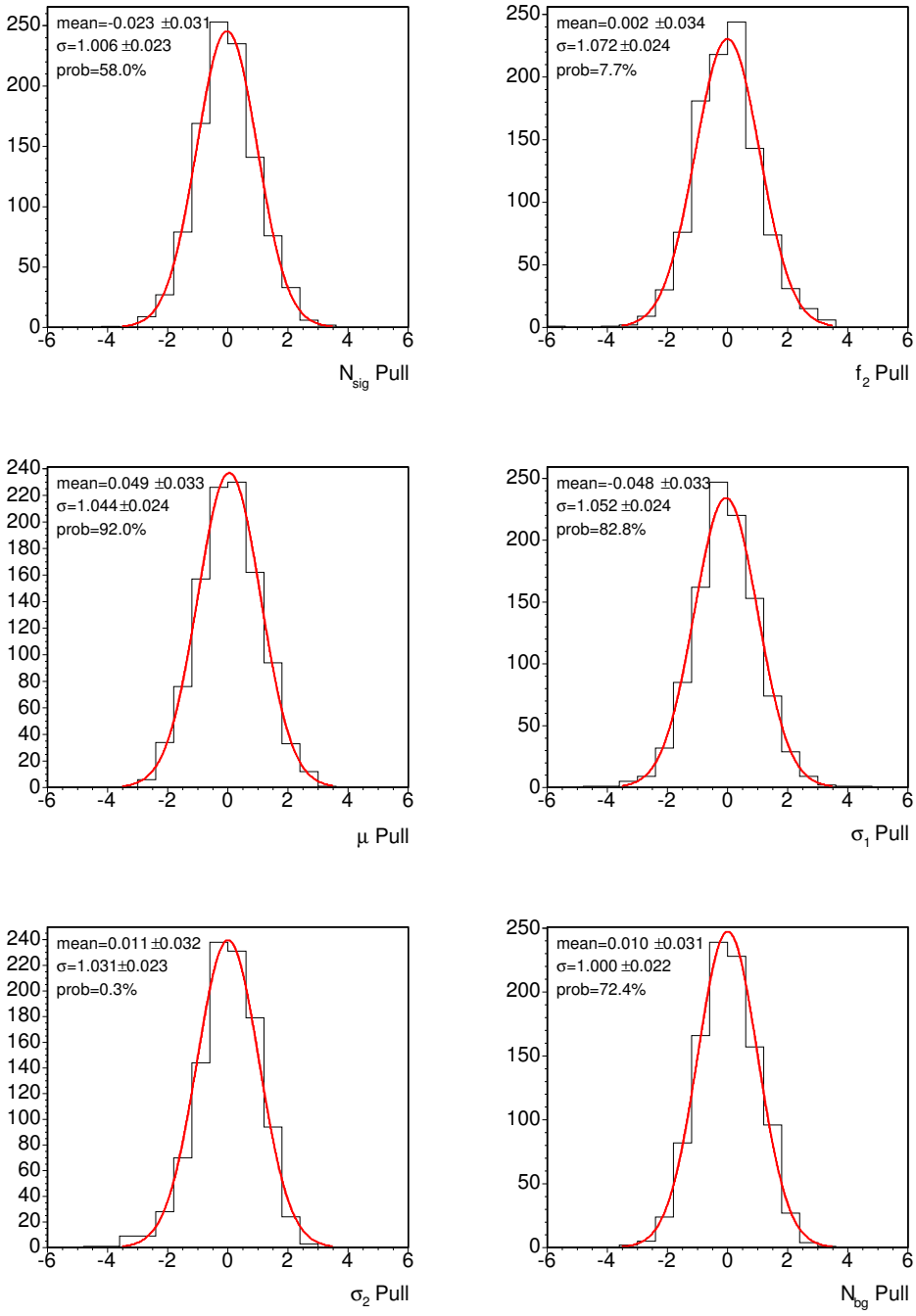


Figure 13: Pull of each fit parameter in the unbinned likelihood fit ($D^*\mu$)

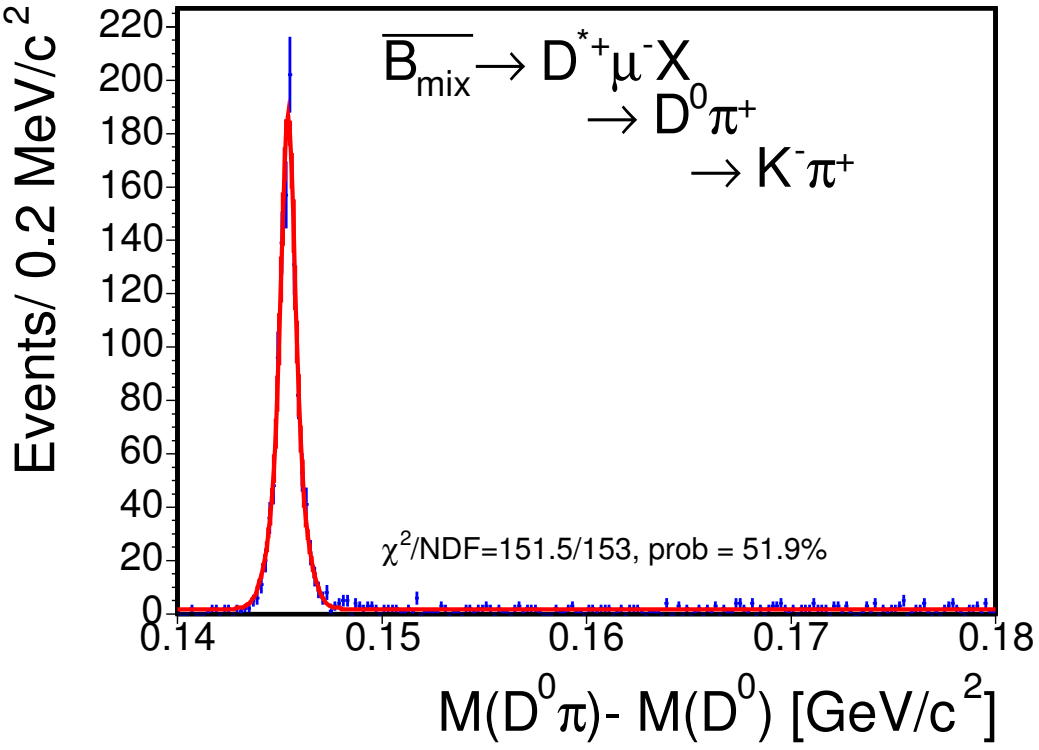


Figure 14: $M_{D^0\pi} - M_{D^0}$ from the $D^*\mu$ events fit to a double- Gaussian signal and a constant background. The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated.

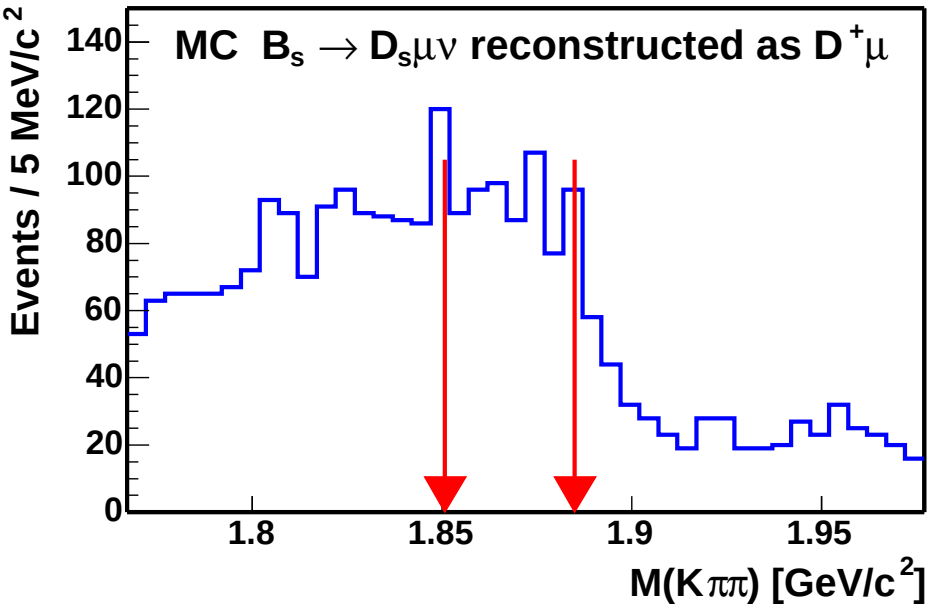


Figure 15: $\overline{B}_s \to D_s^+ \mu^- \overline{\nu}_\mu$ MC reconstructed as $D\mu$ final state. Here D_s are forced to decay into the modes listed in Table 5. The arrows indicate the 3 σ D^+ signal region.

Table 5: Dominant mis-identified D_s sequential decays in $D\mu$ signal. Branching fractions without uncertainties have an upper limit in the PDG.

Selected final states of D_s decays		
Mode	$\mathcal{B}$ (%)	relative to $\mathcal{B}(D_s \to \phi\pi)$
$D_s^+ \to \phi\pi^+$	3.6 $\pm$ 0.9	1
$D_s^+ \to \phi K^+$	0.03 $\pm$?	0.008 $\pm$?
$D_s^+ \to \eta\pi^+$	1.7 $\pm$ 0.5	0.48 $\pm$ 0.05
$D_s^+ \to \eta'\pi^+$	3.9 $\pm$ 1.0	1.08 $\pm$ 0.09
$D_s^+ \to \omega\pi^+$	0.28 $\pm$ 0.11	0.077 $\pm$ 0.025
$D_s^+ \to \rho^0\pi^+$	0.04 $\pm$?	0.011 $\pm$?
$D_s^+ \to \rho^0 K^+$	0.15 $\pm$?	0.042 $\pm$?
$D_s^+ \to f_0\pi^+$	0.57 $\pm$ 0.17	0.16 $\pm$ 0.03
$D_s^+ \to f_2\pi^+$	0.35 $\pm$ 0.12	0.098 $\pm$ 0.022
$D_s^+ \to \rho^+\eta$	10.8 $\pm$ 3.1	2.98 $\pm$ 0.44
$D_s^+ \to \rho^+\eta'$	10.1 $\pm$ 2.8	2.78 $\pm$ 0.41
$D_s^+ \to \overline{K^0}\pi^+$	0.4 $\pm$?	0.11 $\pm$?
$D_s^+ \to \overline{K^{*0}}\pi^+$	0.65 $\pm$ 0.28	0.18 $\pm$ 0.06
$D_s^+ \to K^0 K^+$	3.6 $\pm$ 1.1	1.01 $\pm$ 0.16
$D_s^+ \to K^{*0} K^+$	3.3 $\pm$ 0.9	0.92 $\pm$ 0.09
$D_s^+ \to \pi^+\pi^+\pi^-$	0.005 $\pm^{+0.022}_{-0.005}$	0.0014$\pm$ 0.0007
$D_s^+ \to K^+K^-\pi^+$	0.9 $\pm$ 0.4	0.25 $\pm$ 0.09
$D_s^+ \to K^+K^+K^-$	0.02 $\pm$?	0.0056$\pm$?

We need to include the mis-identified D_s mass shape in our likelihood fit so to properly estimate the number of $D\mu$ events in the data. Assuming that $\overline{B} \to D_s^+ \mu^- X$ has a similar mass spectrum as $\overline{B}_s \to D_s^+ \mu^- \overline{\nu}_\mu$, we could use the MC for Figure 15 to obtain the function which describes the line-shape of mis-reconstructed D_s mass spectrum. We find the D_s spectrum ($\mathcal{F}$) could be described by a constant and a triangular function convoluted with a Gaussian ($\mathcal{T}$):

$$\mathcal{F}(m) = (1 - f_{\rm trg}) \cdot \frac{1}{M_{\rm max} - M_{\rm min}} + f_{\rm trg} \cdot \mathcal{T}(m), \tag{9}$$

where $f_{\rm trg}$ is the fraction of triangular function, $M_{\rm max}$ and $M_{\rm min}$ specify the mass window, $1.767 < M_{K\pi\pi} < 1.977$ GeV/c^2, and

$$\mathcal{T}(m) = \frac{2(m - M_0)}{(M_{\rm off} - M_0)^2} \otimes \mathcal{G}(m, M_0, \sigma_{\rm trg}). \tag{10}$$

Here, $\otimes$ represents convolution, $\mathcal{G}$ is the Gaussian and $\sigma_{\rm trg}$ is the width of $\mathcal{G}$. The triangular function value starts from zero at M_0 and increases as the mass increases. When the mass reaches $M_{\rm off}$, the function values is at its maximum and drops precipitously to zero. A graphical representation of $M_{\rm off}$ and M_0 may be found in Figure 16. The exact form of $\mathcal{T}(m)$ is found in Appendix B.1 derived by Heinrich. Figure 17 shows the result of the fit to the MC.

Now with the function form of the $M_{K\pi\pi}$ spectrum from the D_s decays, we have to normalize the MC yield to the data. The D_s yield may be obtained by reconstructing one of the D_s final states in the data: $\overline{B} \to D_s^+ \mu^- X$, where $D_s^+ \to \phi\pi^+$, $\phi \to K^+K^-$, then using MC to determine the ratio of this D_s decay to that of all the D_s decays in Table 5, $R_{\phi\pi}$:

$$R_{\phi\pi} = \frac{N_{\phi\pi}^{\rm MC}}{N_{\rm all}^{\rm MC}}. \tag{11}$$

The normalization of D_s is then expressed as:

$$N_{\overline{B} \to D_s^+ \mu^- X} = \frac{N_{\overline{B} \to D_s^+ \mu^- X, D_s \to \phi\pi, \phi \to KK}}{R_{\phi\pi}}. \tag{12}$$

In order to obtain $N_{\overline{B} \to D_s^+ \mu^- X, D_s \to \phi\pi, \phi \to KK}$ in the data, the same analysis cuts for $D\mu$ are applied, except that we assign kaon mass to one of the same-sign charged tracks and pion mass to the other. We still assign kaon mass to the track which has the opposite charge of the other two. In addition, the candidates are required to pass the following cuts:

- $1.767 < M_{K\pi\pi} < 1.977$ GeV/c^2
- $|M_{KK} - 1.019| < 0.01$ GeV/c^2

The cut on M_{KK} guarantees that there is no mis-identified D^+ in the D_s^+ signal we reconstruct. We confirm this by reconstructing D_s^+ from the $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$ MC and no D_s candidate is found. See Figure 18 for the $\overline{B} \to D_s^+ \mu^- X$ signal in the data, we find:

$$N_{\overline{B} \to D_s^+ \mu^- X, D_s \to \phi\pi, \phi \to KK} = 237 \pm 17.$$

We then reconstruct the same D_s decay chain in the MC as in the data and obtain

$$R_{\phi\pi} = 0.131 \pm 0.007. \tag{13}$$

See Figure 19 for the reconstructed $D_s^+ \to \phi\pi^+$ in the semi-inclusive $\overline{B}_s \to D_s^+ \mu^- \overline{\nu}_\mu$ MC. Inserting the result of $N_{\overline{B} \to D_s^+ \mu^- X, D_s \to \phi\pi, \phi \to KK}$ and $R_{\phi\pi}$ into Equation 12, we have:

$$N_{\overline{B} \to D_s^+ \mu^- X} = 1812 \pm 160, \tag{14}$$

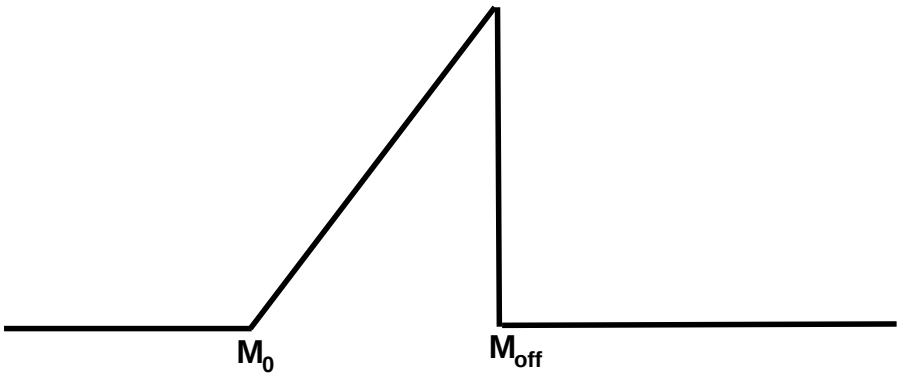


Figure 16: Graphical representation of the triangular function

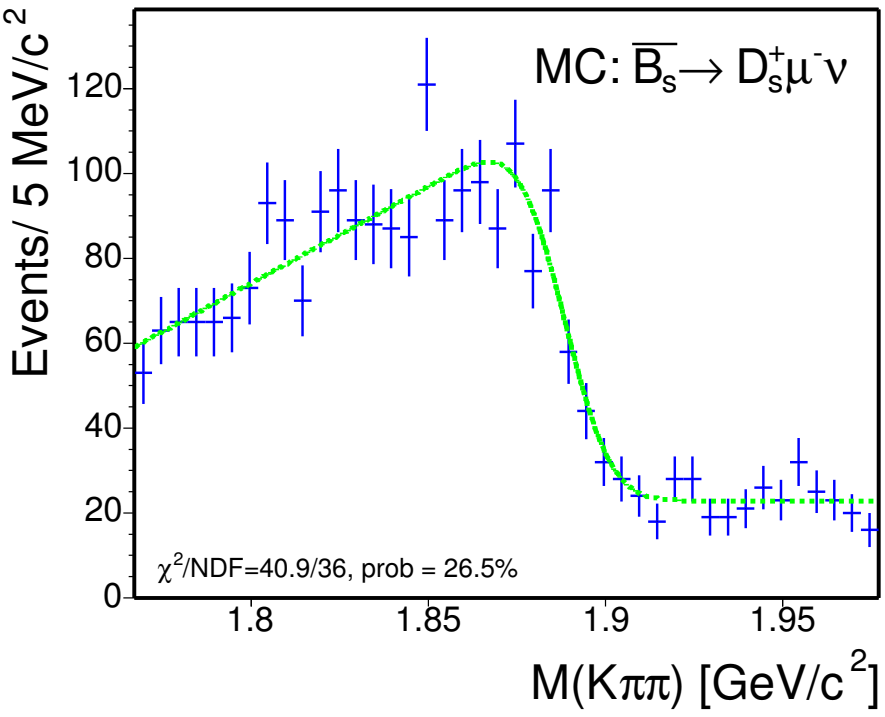


Figure 17: Mis-reconstructed D_s from the $\overline{B}_s \to D_s^+ \mu^- \overline{\nu}_\mu$ MC fit to a constant and a triangular function convoluted with a Gaussian. The dashed curve indicates the result of the unbinned likelihood fit.

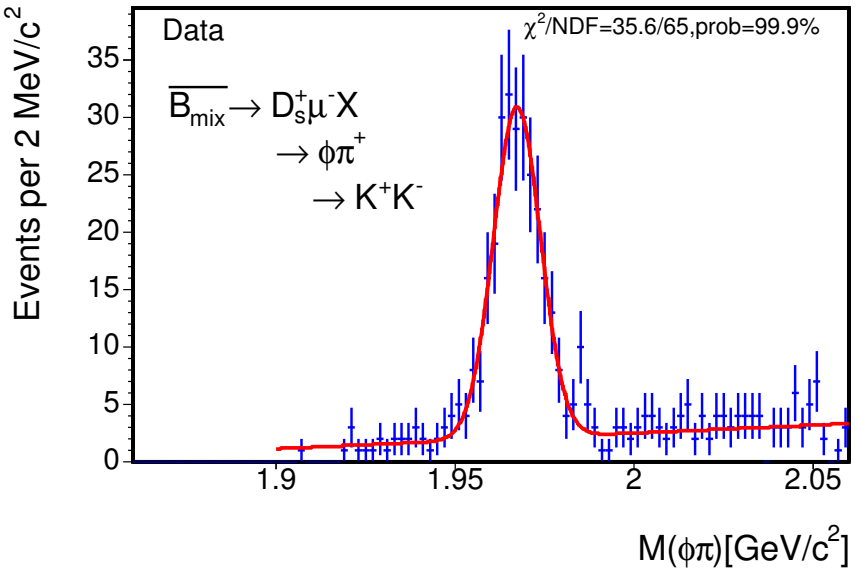


Figure 18: Data: Reconstructed $\overline{B} \rightarrow D_s^+ \mu^- X$, where $D_s^+ \rightarrow \phi\pi^+$, and $\phi \rightarrow K^+K^-$. $M_{K\pi\pi}$ is required to be between 1.767 and 1.977 GeV/c^2. There are 237 $\pm$ 17 events in the peak.

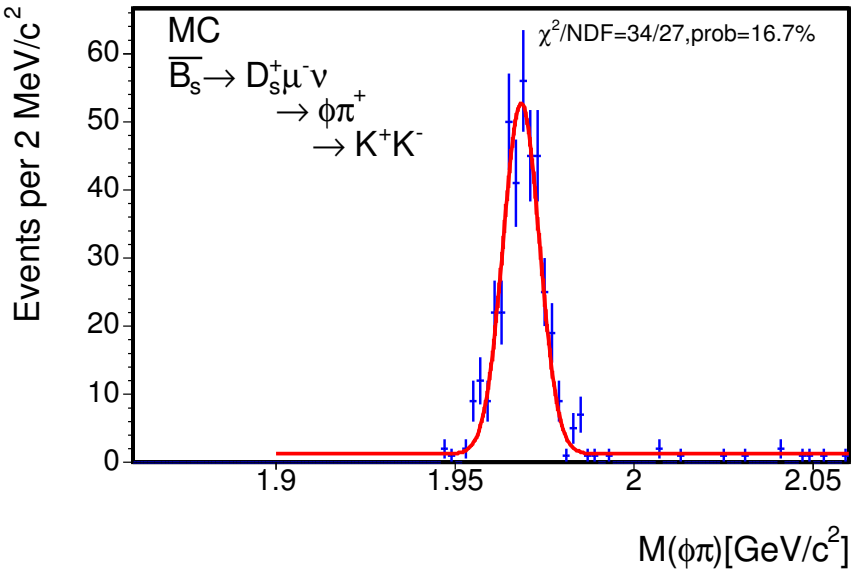


Figure 19: MC: Reconstructed $\overline{B}_s \rightarrow D_s^+ \mu^- \overline{\nu}_\mu$, where $D_s^+ \rightarrow \phi\pi^+$, and $\phi \rightarrow K^+K^-$. $M_{K\pi\pi}$ is required to be between 1.767 and 1.977 GeV/c^2. There are 326 $\pm$ 19 events, out of 2493 events from selected D_s decays.

The uncertainty in Equation 14 comes from the fractional uncertainties on: $N_{\overline{B}\to D_s^+\mu^- X, D_s\to\phi\pi, \phi\to KK}$ (7.2%) and $R_{\phi\pi}$ (5%).

In the unbinned fit, the extended log likelihood function is expressed by the sum of two likelihoods: one describing the data and the other describing the $\overline{B}_s \to D_s^+\mu^-\overline{\nu}_\mu$ MC since we fit the data and MC simultaneously;

$$\log\mathcal{L} = \log\mathcal{L}^{\rm data} + \log\mathcal{L}^{\rm MC}, \tag{15}$$

The likelihood function for the data, $\log\mathcal{L}^{data}$, is a sum of a signal Gaussian, a first-order polynomial for the combinatorial background ($\mathcal{H}$), and the function for the D_s ($\mathcal{F}$, see Equation 9). A Gaussian constraint on the amount of D_s, $\mathcal{C}_{D_s}$, is employed.

$$\begin{aligned}\log\mathcal{L}^{\rm data} &= \sum_i \log\{N_{\rm sig}\cdot\mathcal{G}(m_i,\mu,\sigma) + N_{\rm combg}\cdot\mathcal{H}(m_i) + N_{D_s}\cdot\mathcal{F}(m_i)\} \\ &- N_{\rm sig} - N_{\rm combg} - N_{D_s} + \log\mathcal{C}_{D_s},\end{aligned} \tag{16}$$

where

$$\begin{aligned}\mathcal{H}(m_i) &= \frac{1}{M_{\rm max}-M_{\rm min}} + p_1\cdot(m_i - \frac{M_{\rm max}+M_{\rm min}}{2}), \\ \mathcal{C}_{D_s} &= \mathcal{G}(N_{D_s},\mu^p,\sigma^p).\end{aligned}$$

From the prediction of Equation 14, we have $\mu^p = 1812$, and $\sigma^p = 160$.

The likelihood function $\log\mathcal{L}^{\rm MC}$ is used to fit $\overline{B}_s \to D_s^+\mu^-\overline{\nu}_\mu$ MC and obtain the parameterization of $\mathcal{F}(m)$. Here the normalization does not matter.

$$\log\mathcal{L}^{\rm MC} = \sum_i \log\{\mathcal{F}(m_i)\}. \tag{17}$$

Table 6 lists the mean, width of the pulls from 1000 toy MC test and the result returned from the unbinned likelihood fit to the data. Figure 20 shows the fit to each pull distribution. All the pull means are consistent with zero and the pull widths are consistent with one. Table 7 gives the correlation coefficients returned from the likelihood fit to the data. Figure 21 shows the fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\overline{B}\to D^+\mu^- X} = 4721 \ \pm 104.$$

We also perform a cross-check by removing the constraint on N_{D_s} and obtain $N_{\overline{B}\to D^+\mu^- X} = 4667\pm139$, $N_{D_s} = 2184\pm620$, which are consistent with the result of the constrained fit. The fit without constraint has a χ^2/NDF=197.0/199 and probability is 52.7%.

4.1.3 $\Lambda_c\mu$ Yield

When a proton mass is assigned to a kaon or pion, numerous B meson to D meson semileptonic decays could be mis-reconstructed as a $\Lambda_c\mu$ final state. In order to estimate the B meson background shape under our signal, we use generator level MC and generate the semileptonic decays (μ channel) of each B meson flavor separately. After applying analysis cuts, we add up the mis-reconstructed mass spectrum from each kind of B meson according to the production fractions:

$$\begin{aligned}b\to B_d &= (39.7\pm1.3)\ \%, \\ b\to B_u &= (39.7\pm1.3)\ \%, \\ b\to B_s &= (10.7\pm1.1)\ \%.\end{aligned}$$

Figure 22 shows a smooth mass spectrum from the generator MC. The shape is best described by a second-order polynomial, with χ^2/NDF = 36.6/42, prob = 70%. A first-order polynomial fit yields χ^2/NDF = 56.6/43, prob = 8%. Because the combinatorial background may be parameterized by a

Table 6: $D\mu$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	N_{sig}		-0.012 ± 0.035	1.004 ± 0.025	4721 ± 104
2	μ	[GeV/c^2]	0.027 ± 0.037	1.048 ± 0.027	1.8680 ± 0.0002
3	σ	[GeV/c^2]	0.007 ± 0.035	0.992 ± 0.025	0.0084 ± 0.0002
4	N_{combg}		-0.076 ± 0.038	1.073 ± 0.027	15178 ± 197
5	p_1		0.018 ± 0.036	1.027 ± 0.026	-5.2 ± 0.7
6	N_{D_s}		0.042 ± 0.037	1.065 ± 0.027	1832 ± 155
7	f_{trg}		0.022 ± 0.036	1.023 ± 0.026	0.617 ± 0.021
8	M_0	[GeV/c^2]	0.055 ± 0.035	1.007 ± 0.025	1.69 ± 0.02
9	M_{off}	[GeV/c^2]	-0.025 ± 0.036	1.019 ± 0.026	1.888 ± 0.002
10	σ_{trg}	[GeV/c^2]	-0.035 ± 0.037	1.056 ± 0.027	0.010 ± 0.002

Table 7: Correlation coefficients returned from the fit ($D\mu$ data)

	1	2	3	4	5	6	7	8	9	10
1	1.000									
2	0.011	1.000								
3	0.444	0.026	1.000							
4	-0.032	-0.011	-0.065	1.000						
5	-0.066	-0.050	-0.036	-0.316	1.000					
6	-0.277	0.007	-0.168	-0.727	0.384	1.000				
7	-0.041	0.019	-0.038	0.025	0.148	-0.005	1.000			
8	-0.083	0.107	-0.029	0.070	-0.134	-0.033	-0.090	1.000		
9	-0.054	-0.193	-0.090	-0.013	0.022	0.044	0.094	-0.542	1.000	
10	0.079	0.175	0.058	-0.032	0.010	-0.007	0.272	0.376	-0.534	1.000

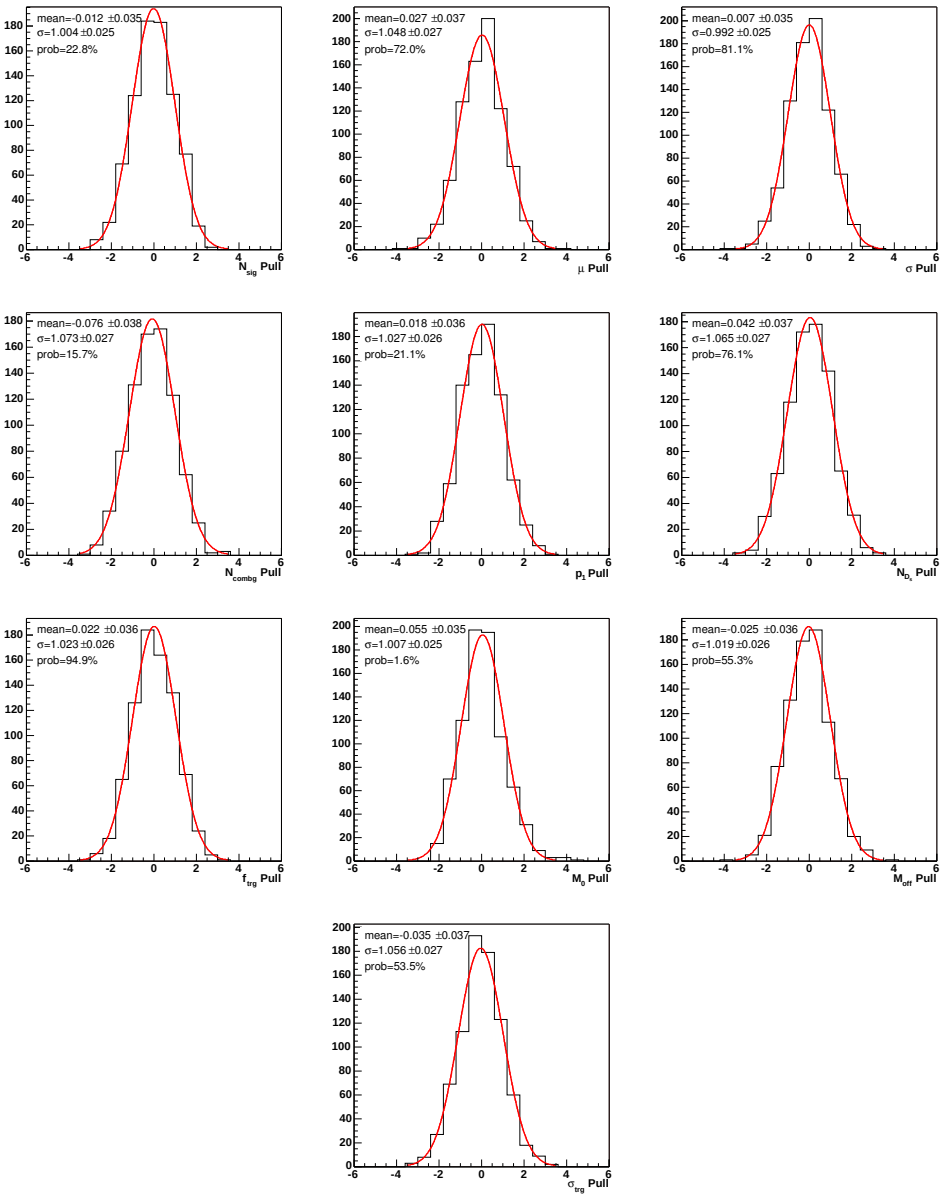


Figure 20: Pull of each fit parameter in the unbinned likelihood fit ($D\mu$)

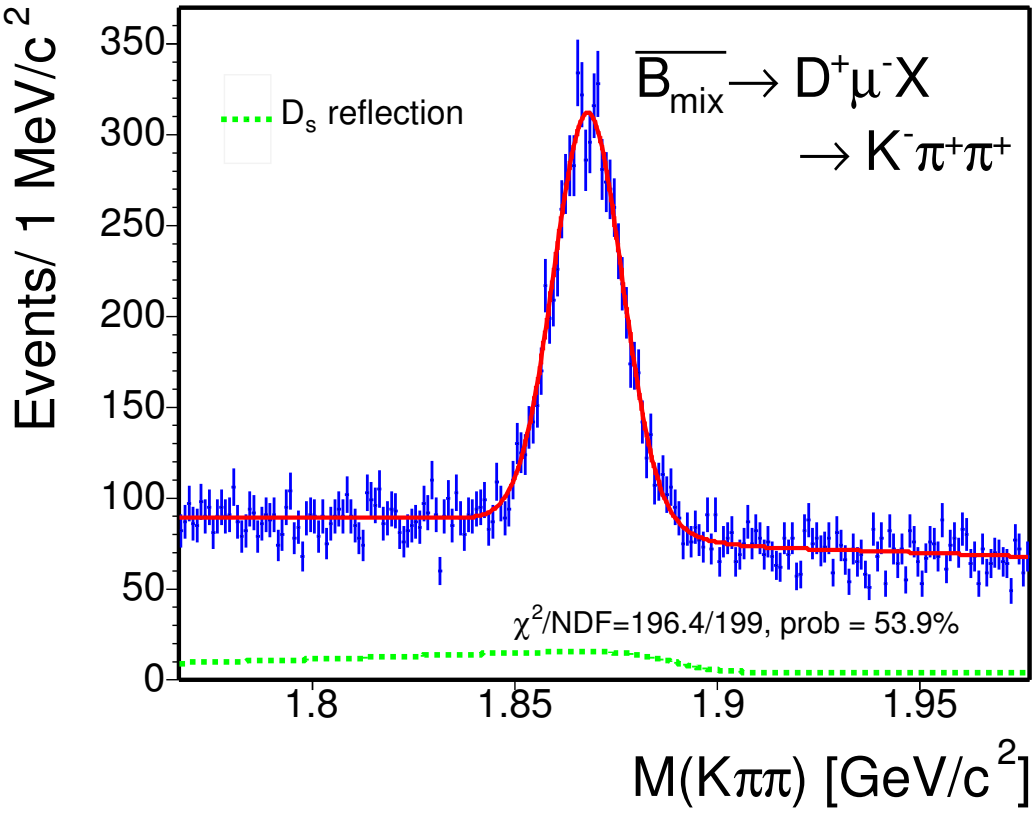


Figure 21: $M_{K\pi\pi}$ from the $D\mu$ events fit to a Gaussian (signal), a first-order polynomial (combinatorial background), and a constant plus a triangular function convoluted with a Gaussian (marked by the dashed line). The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated.

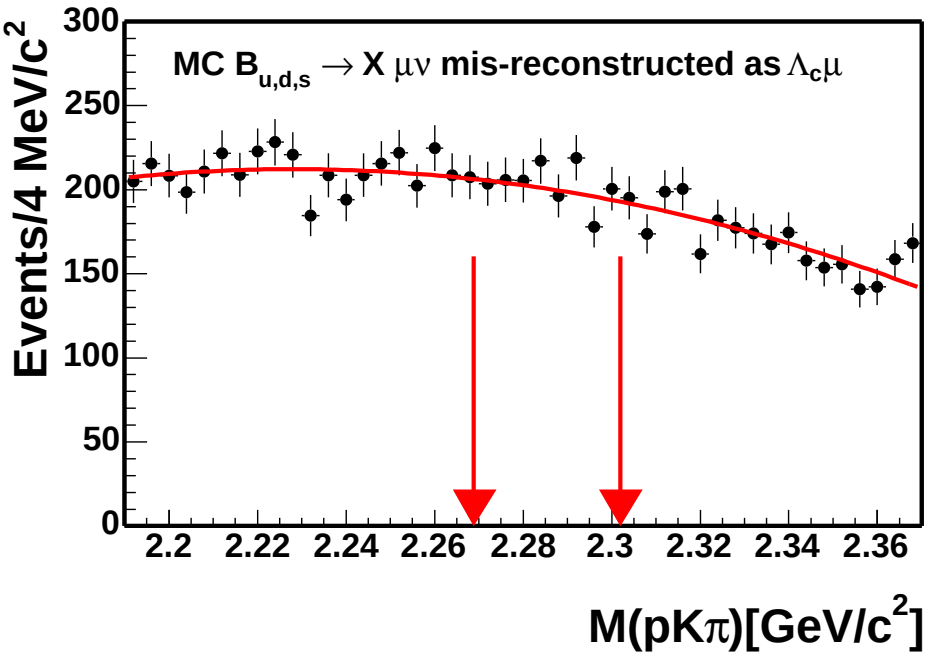


Figure 22: $B \rightarrow X\mu\nu_\mu$ MC mis-reconstructed as $\Lambda_c\mu$. The mass spectrum fit to a second-order polynomial. χ^2/NDF = 36.6/42, prob= 70.7%. Two arrows indicate the signal region of Λ_c^+.

first-order polynomial, and adding a first- to a second-order polynomial gives a second-order polynomial, we fit the combinatorial and the B meson background together to a second-order polynomial ($\mathcal{H}$). The extended log likelihood function could be expressed as:

$$\begin{aligned} \log \mathcal{L} &= \sum_i \log\{N_{\rm sig} \cdot \mathcal{G}(m_i, \mu, \sigma) + N_{\rm bg} \cdot \mathcal{H}(m_i)\} \\ &- N_{\rm sig} - N_{\rm bg}, \end{aligned} \tag{18}$$

where

$$\mathcal{H}(m_i) = \frac{1}{M_{\rm max} - M_{\rm min}} + p_1 \cdot (m_i - M_{\rm mid}) + p_2 \cdot (12 \cdot (m_i - M_{\rm mid})^2 - M_{\rm diff}^2).$$

Here, $M_{\rm max}$ and $M_{\rm min}$ specify the Λ_c^+ mass window: $2.19 < M_{pK\pi} < 2.37$ GeV/c^2. The average of $M_{\rm max}$ and $M_{\rm min}$, or the mid point in the mass window is $M_{\rm mid}$. The difference of $M_{\rm max}$ and $M_{\rm min}$ is $M_{\rm diff}$.

Table 8 lists the mean, width of the pulls from the toy MC test and the parameter value from the fit to the data. Figure 23 gives the pull of each fit parameter. The pull mean of each fit parameter is consistent with zero and pull width is consistent with one. Table 9 gives the correlation coefficients returned from the likelihood fit to the data. Figure 24 shows fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\overline{B} \rightarrow \Lambda_c^+ \mu^- X} = 1237 \ \pm 97.$$

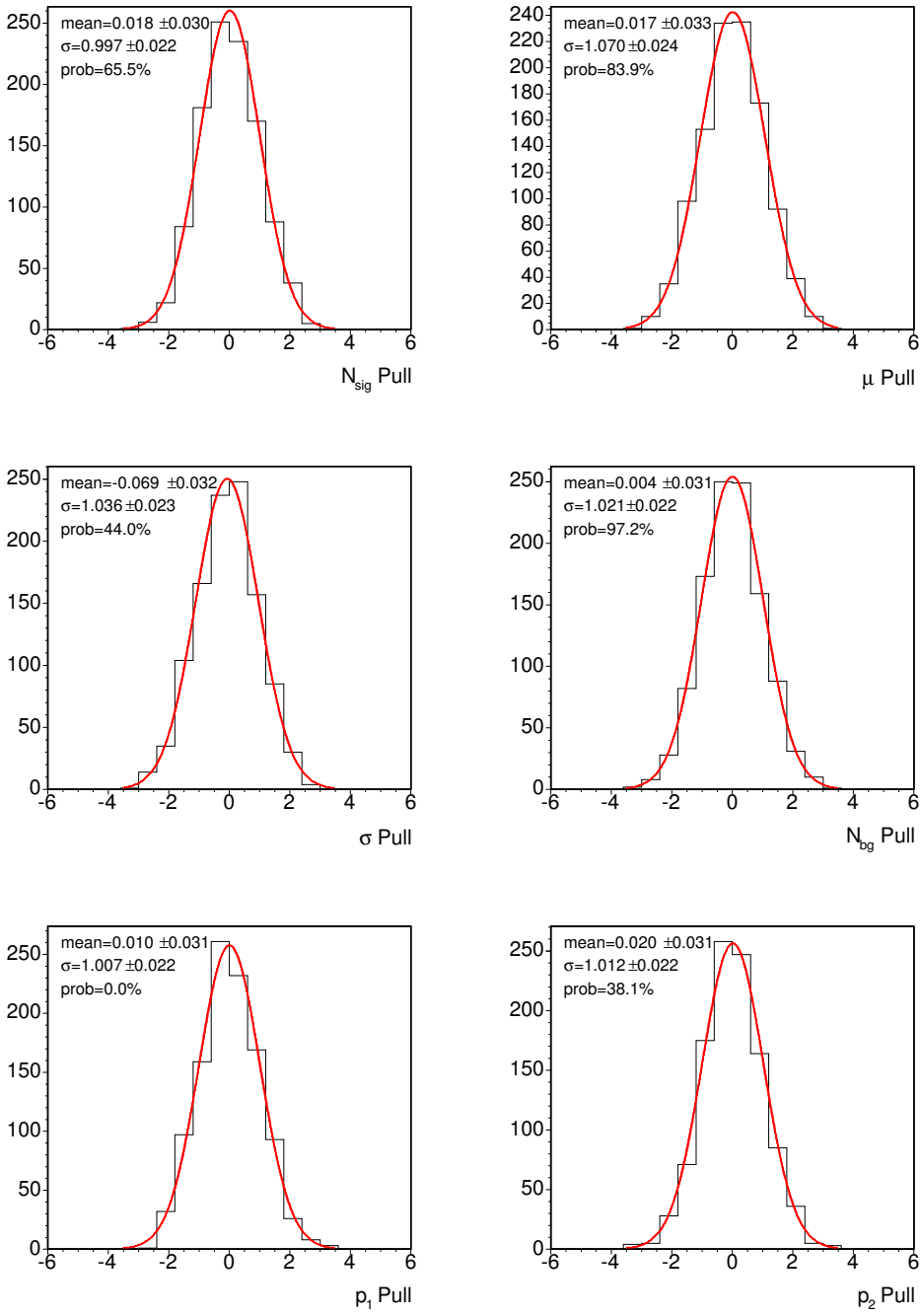


Figure 23: Pull of each fit parameter in the unbinned likelihood fit ($\Lambda_c\mu$)

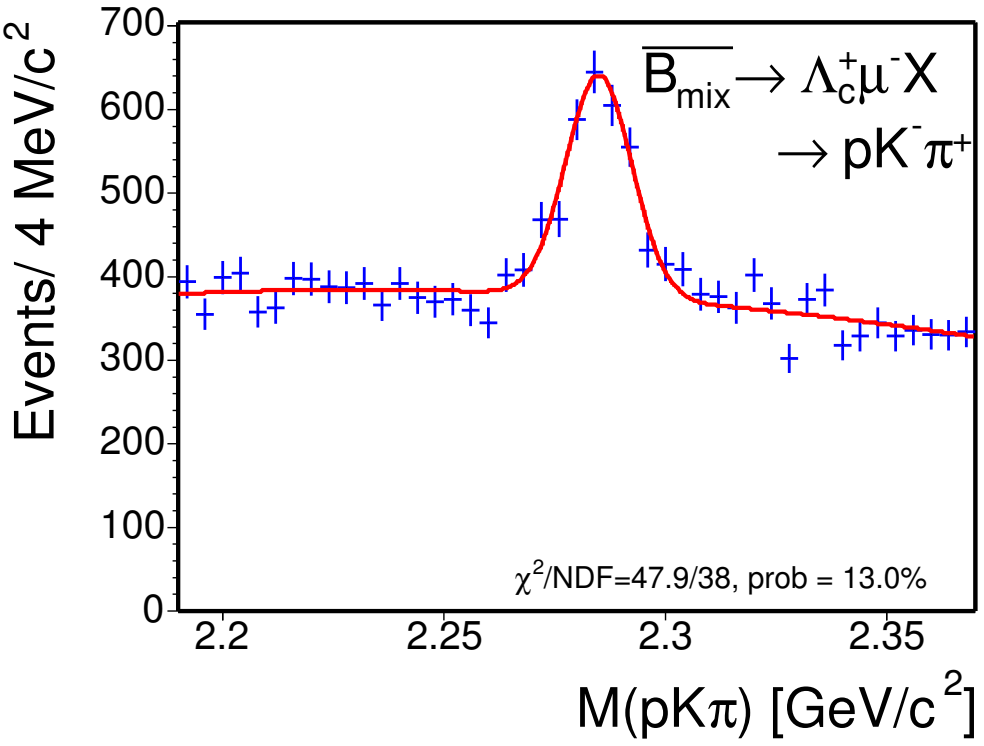


Figure 24: $M_{pK\pi}$ from the $\Lambda_c\mu$ events fit to a Gaussian (signal), a second-order polynomial (combinatorial + B meson). The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated.

Table 8: $\Lambda_c\mu$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	$N_{\rm sig}$		0.018 ± 0.030	0.997 ± 0.022	1237 ± 97
2	μ	[GeV/c^2]	0.017 ± 0.033	1.070 ± 0.024	2.2850 ± 0.0005
3	σ	[GeV/c^2]	-0.069 ± 0.032	1.036 ± 0.023	0.0074 ± 0.0006
4	$N_{\rm bg}$		0.004 ± 0.031	1.021 ± 0.022	16576 ± 157
5	p_1		0.010 ± 0.031	1.007 ± 0.022	-4.3 ± 0.8
6	p_2		0.020 ± 0.031	1.012 ± 0.022	-3.7 ± 1.8

Table 9: Correlation coefficients returned from the fit ($\Lambda_c\mu$ data)

	1	2	3	4	5	6
1	1.000					
2	-0.024	1.000				
3	0.460	-0.020	1.000			
4	-0.429	0.013	-0.245	1.000		
5	0.443	-0.022	0.243	-0.236	1.000	
6	-0.060	-0.067	-0.033	0.032	-0.058	1.000

4.2 Mass Fit of the Hadronic Modes

Figure 12 shows that to the left of the hadronic signal peak, the "charm+π" mass spectrum exhibits an interesting structure. If we use an exponential or a first-order polynomial to fit the the lower and upper mass regions separately, we find that the slope of the lower mass region is steep but seem to turn off just below the peak, while the slope of the upper mass region is shallow and approaches a constant. In order to extract a correct number of events observed in the hadronic channels, we have to take into account the background structure when fitting the charm+π mass spectrum.

For the $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\pi^-$ modes, we import the B^0 and Λ_b mass functions derived in the analyses of Furic [3], and Martin and Maksimović [4], respectively. Several parameters that describe the background shapes or normalizations are fixed. We find small modifications are needed for the numerical values of the fixed parameters in the $\overline{B}^0 \to D^+\pi^-$ mode, as a few variables we apply cut on are different from those in Furic's analysis. We apply our cuts on the MC used in Furic's analysis and refit the MC to extract the numbers for our analysis. For the $\overline{B}^0 \to D^{*+}\pi^-$ mode, we produce an inclusive $B \to D^{*+}X$ MC sample to study the background composition. The decay modes with distinguished mass shape are separated from the other modes. The decays with similar mass spectra are lumped together and fit to the same background function. Figure 25 shows the B^0 and Λ_b mass spectra from the contributions of different decays.

Our hadronic mass spectra share several common features: It is clear that the background from the B hadron decays only contribute to the mass region below the signal, while in the data, the upper mass region is composed of combinatorial background, which may be described by an exponential or a constant. The combinatorial background extends down to the lower B mass region as well. In the region 40 to 70 MeV/c^2 below the signal peak, Cabibbo suppressed decays, $\overline{B}^0 \to D^{*+}K^-$, $\overline{B}^0 \to D^+K^-$, $\Lambda_b \to \Lambda_c^+K^-$, with a branching ratio about 8% of our Cabibbo favored signals, produce a small contamination. Going further down in the charm+π mass, we have partially reconstructed B decays from the semileptonic modes, and other mis-identified B hadronic decays.

Note that since both $\overline{B}^0 \to D^{*+}\pi^-$ and $\Lambda_b \to \Lambda_c^+\pi^-$ have low statistics, we constrain the widths of their signal Gaussians in the following way: We first fit the width of $M_{D\pi}$ ($\sigma_{D\pi}^{\rm data}$) from the high statistics $\overline{B}^0 \to D^+\pi^-$ sample (~600 events) in the data. Then we multiply $\sigma_{D\pi}^{\rm data}$ with the MC width ratio: $\sigma_{\Lambda_c\pi,D^*\pi}^{\rm MC}/\sigma_{D\pi}^{\rm MC}$ and predict $\sigma_{\Lambda_c\pi,D^*\pi}^{\rm data}$.

4.2.1 $\overline{B}^0 \to D^{*+}\pi^-$ Yield

The study from the $\overline{B} \to D^{*+}X$ MC shows that the background in the lower mass region is dominated by the following decays: Cabibbo suppressed decay $\overline{B}^0 \to D^{*+}K^-$, $\overline{B}^0 \to D^{*+}\rho^-$, and the remaining $\overline{B} \to D^{*+}X$. See the texts below for the detailed descriptions.

1. $\overline{B}^0 \to D^{*+}K^-$: fully reconstructed Cabibbo suppressed decays. The mass spectrum is a peak about 40 MeV/c^2 below the $\overline{B}^0 \to D^{*+}\pi^-$ signal, with small tails on the lower mass side. The shape is modeled by a lifetime function;

$$\mathcal{E}(m) = \exp(m, \tau_{D^*K}) \otimes \mathcal{G}(m, \mu_{D^*K}, \sigma_{D^*K}), \tag{19}$$

   where τ_{D^*K} is the lifetime, μ_{D^*K} is the zero point of the lifetime function also the mean of the Gaussian. The width of the Gaussian also the resolution of the lifetime function is σ_{D^*K}. The exact form of $\mathcal{E}(m)$ is found in Appendix B.2. See Figure 26 (top) for the fit to $\overline{B}^0 \to D^{*+}K^-$ MC.

2. $\overline{B}^0 \to D^{*+}\rho^-$, where $\rho^- \to \pi^0\pi^-$: modeled by a triangular function convoluted with a Gaussian;

$$\mathcal{T}(m) = \frac{2(m - M_0^{D^*\rho})}{(M_{\rm off}^{D^*\rho} - M_0^{D^*\rho})^2} \otimes \mathcal{G}(m, M_0^{D^*\rho}, \sigma_{D^*\rho}). \tag{20}$$

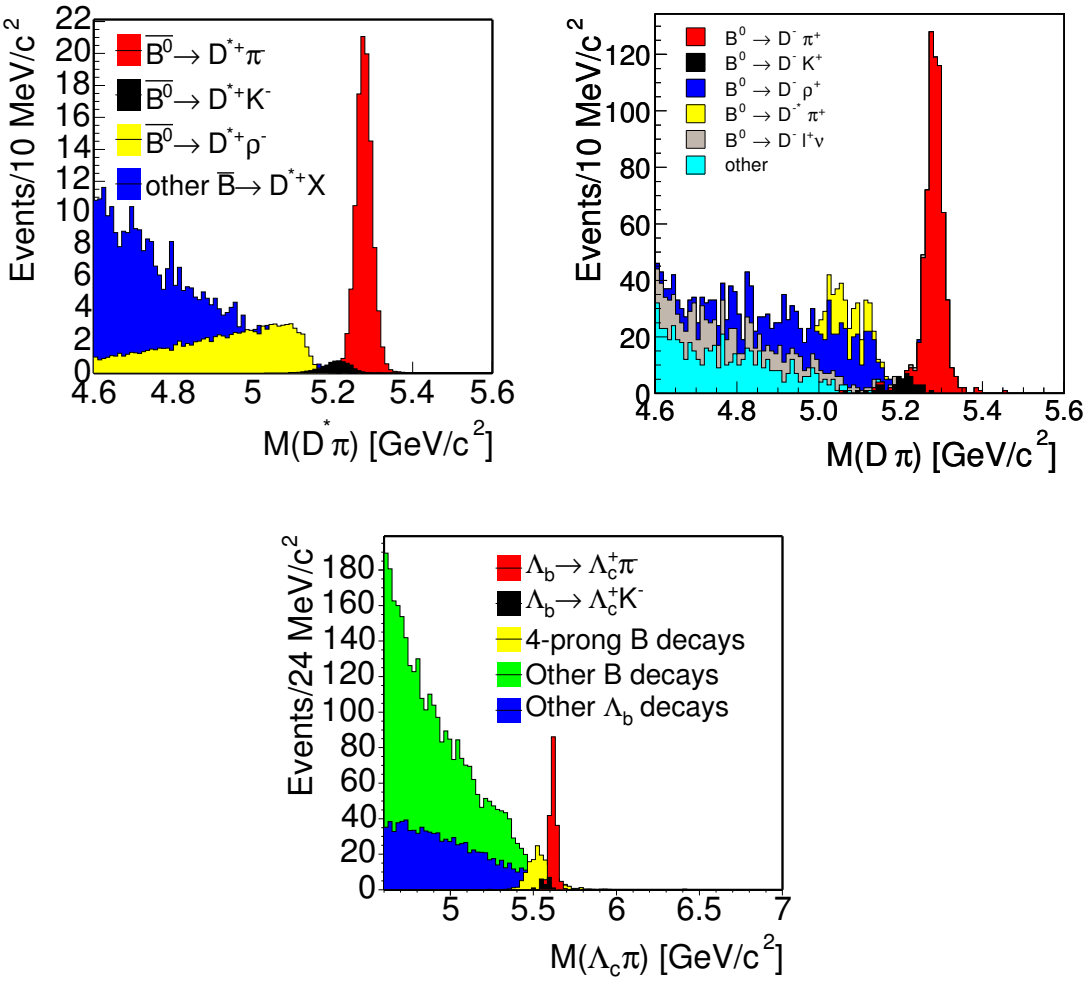


Figure 25: Background composition for $\overline{B}^0 \to D^{*+}\pi^-$ (top), $\overline{B}^0 \to D^+\pi^-$ (middle) from Furic [3] and $\Lambda_b \to \Lambda_c^+\pi^-$ (bottom) from Martin and Maksimović [4]. Note that the mis-reconstructed $\overline{B}_s \to D_s^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\pi^-$ in the $D\pi$ mass is not shown.

See Figure 26 (middle) for the fit.

3. Continuum: remaining $\overline{B} \to D^{*+}X$ decays partially reconstructed. These backgrounds have similar mass spectrum and are group together. The shape is modeled by a first-order polynomial with a negative slope and a turn-off at $M_{\rm off}^{\rm otherB}$; when $m < M_{\rm off}^{\rm otherB}$:

$$\mathcal{H}(m) = \frac{2}{(M_{\rm off}^{\rm otherB} - M_{\rm min})^2} \cdot (M_{\rm off}^{\rm otherB} - m), \tag{21}$$

and when $m > M_{\rm off}^{\rm otherB}$:

$$\mathcal{H}(m) = 0. \tag{22}$$

The lowest boundary of the $D^*\pi$ mass window, $M_{\rm min}$, is 4.6 GeV/c^2. See Figure 26 (bottom) for the fit to these MC samples.

In the unbinned fit, the extended log likelihood function is expressed by the sum of five likelihoods: one describing the data, and the other four describing the MC samples from each type of background and the signal:

$$\log \mathcal{L} = \log \mathcal{L}^{\rm data} + \log \mathcal{L}^{\rm MC}_{D^*\pi} + \log \mathcal{L}^{\rm MC}_{D^*K} + \log \mathcal{L}^{\rm MC}_{D*\rho} + \log \mathcal{L}^{\rm MC}_{\rm otherB}, \tag{23}$$

The likelihood function $\log \mathcal{L}^{\rm data}$ is a sum of a signal Gaussian, a constant combinatorial background, the functions for D^*K ($\mathcal{E}$), $D^*\rho$ ($\mathcal{T}$), and the continuum ($\mathcal{H}$). In addition, there is a constraint on each of the following parameters: the signal width, relative amount of D^*K to the signal (f_{D^*K}), and the fraction of $D^*\rho$ in $D^*\rho$ + remaining $\overline{B} \to D^{*+}X$ ($f_{D^*\rho}$) The reason for the last constraint is because $\overline{B}^0 \to D^{*+}\rho^-$ and the remaining $\overline{B} \to D^{*+}X$ decays occupy the same mass region. Therefore, the likelihood fit converges faster if we constrain $f_{D^*\rho}$.

$$\begin{aligned}
\log \mathcal{L}^{\rm data} &= \sum_i \log\{N_{\rm sig} \cdot (\mathcal{G}(m_i, \mu, \sigma) + f_{D^*K} \cdot \mathcal{E}(m_i)) \\
&\quad + N_{\rm bg} \cdot [f_{\rm combg} \cdot \frac{1}{M_{\rm max} - M_{\rm min}} \\
&\quad + (1 - f_{\rm combg}) \cdot [f_{D^*\rho} \cdot \mathcal{T}(m_i) + (1 - f_{D^*\rho}) \cdot \mathcal{H}(m_i)]\} \\
&- N_{\rm sig} - N_{D^*K} - N_{\rm bg} \\
&+ \log \mathcal{C}_1 + \log \mathcal{C}_2 + \log \mathcal{C}_\sigma,
\end{aligned} \tag{24}$$

where $\mathcal{E}(m_i)$, $\mathcal{T}(m_i)$ and $\mathcal{H}(m_i)$ are expressed in Equations 19–21. The $M_{\rm max}$ and $M_{\rm min}$ specify the mass window: $4.6 < M_{D^*\pi} < 5.6$ GeV/c^2. The parameters f_{D^*K}, $N_{\rm bg}$, $f_{\rm combg}$ and $f_{D^*\rho}$ are defined as follow:

$$\begin{aligned}
f_{D^*K} &\equiv \frac{N_{D^*K}}{N_{\rm sig}}, \\
N_{\rm bg} &\equiv N_{\rm combg} + N_{\rm otherB} + N_{D^*\rho}, \\
f_{\rm combg} &\equiv \frac{N_{\rm combg}}{N_{\rm bg}}, \\
f_{D^*\rho} &\equiv \frac{N_{D^*\rho}}{N_{\rm otherB} + N_{D^*\rho}}.
\end{aligned}$$

The constraints are expressed as:

$$\begin{aligned}
\mathcal{C}_1 &= \mathcal{G}(f_{D^*K}, \mu_1, \sigma_1), \\
\mathcal{C}_2 &= \mathcal{G}(f_{D^*\rho}, \mu_2, \sigma_2), \\
\mathcal{C}_\sigma &= \mathcal{G}(\sigma, \mu_p, \sigma_p),
\end{aligned}$$

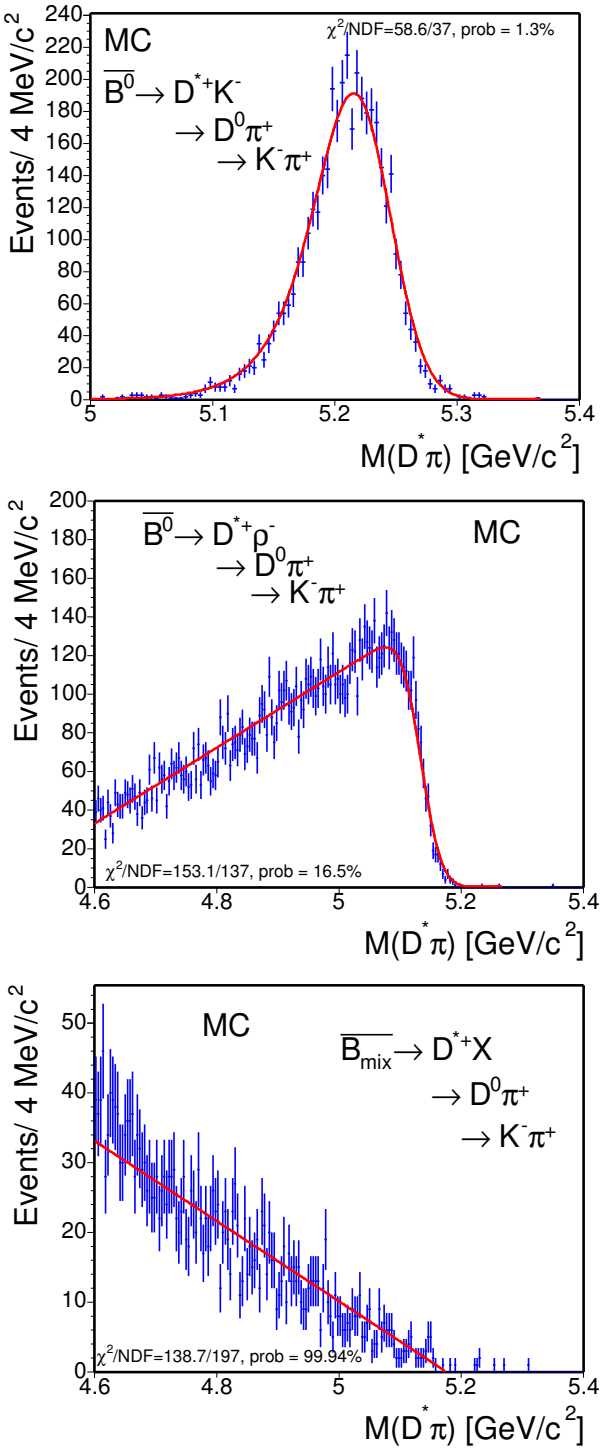


Figure 26: Various MC samples reconstructed as $\overline{B}^0 \to D^{*+}\pi^-$. From the top to the bottom are $\overline{B}^0 \to D^{*+}K^-$, $\overline{B}^0 \to D^{*+}\rho^-$, and the remaining $\overline{B} \to D^{*+}X$.

Table 10: Branching ratios and relative efficiencies for $\overline{B}^0 \to D^{*+}\pi^-$ background

	$\overline{B}^0 \to D^{*+}K^-$	$\overline{B}^0 \to D^{*+}\pi^-$
B(%)	0.276 ± 0.021	0.020 ± 0.005
ϵ ratio	1	1.02 ± 0.02
f_{D^*K}		0.071 ± 0.019
	$\overline{B}^0 \to D^{*+}\rho^-$	remaining $\overline{B} \to D^{*+}X$
N_{MC}	758	2371
$f_{D^*\rho}$		0.242 ± 0.008

where $\mu_1 = 0.071$, $\sigma_1 = 0.019$, $\mu_2 = 0.242$, $\sigma_2 = 0.008$, $\mu_p = 0.0259$ GeV/c^2, and $\sigma_p = 0.0012$ GeV/c^2.

Here, μ_p and σ_p are determined using the $\overline{B}^0 \to D^+\pi^-$ signal in the data, $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ MC as described earlier. The μ_1 and σ_1 are determined using the world averaged branching ratios, and the efficiencies from the MC listed in Table 10:

$$f_{D^*K} = \frac{\mathcal{B}(\overline{B}^0 \to D^{*+}K^-)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)} \cdot \frac{\epsilon_{\overline{B}^0 \to D^{*+}K^-}}{\epsilon_{\overline{B}^0 \to D^{*+}\pi^-}}. \tag{25}$$

The μ_2 and σ_2 are determined by counting the number of reconstructed $D^*\rho$ and the remaining $\overline{B} \to D^{*+}X$ events in the MC after all the analysis cuts.

The three likelihoods for the background MC are used to obtain the parameterization of $\mathcal{E}(m)$, $\mathcal{T}(m)$, and $\mathcal{H}(m)$. The normalizations do not matter here.

$$\log \mathcal{L}^{\rm MC}_{D^*K} = \sum_i \log \mathcal{E}(m_i), \tag{26}$$

$$\log \mathcal{L}^{\rm MC}_{D*\rho} = \sum_i \log \mathcal{T}(m_i), \tag{27}$$

$$\log \mathcal{L}^{\rm MC}_{\rm otherB} = \sum_i \log \mathcal{H}(m_i). \tag{28}$$

In addition, $\log \mathcal{L}^{\rm MC}_{D^*\pi}$ is used to obtain the reconstructed mass difference between MC and data, $m_{\rm diff}$. In the $\log \mathcal{L}^{\rm data}$, all the parameters except the normalization and the resolution parameters (σ) for the signal Gaussian and the background functions, differ by $m_{\rm diff}$ from those in the $\log \mathcal{L}^{\rm MC}$. The resolutions for all the backgrounds are kept the same between MC and data, while the resolution of the signal Gaussian in the data is a separate free parameter from that in the MC.

We use the total likelihood to fit the data and MC simultaneously. Table 11 lists the pull means and widths of toy MC test and the unbinned likelihood fit result to the data. Figures 27–28 give the pull of each fit parameter. The pull mean of $M^{\rm otherB}_{\rm off}$ in Equation 21 is -0.220 ± 0.031, but this value corresponds to a ~ 0.02 % shift in the central value. The pull widths of $M_0^{D^*\rho}$ and $f_{\rm combg}$ are about 3 σ away from one. However, the number of signal events is not affected. All the other pull means and widths are consistent with zero and one. Table 12 gives the correlation coefficients returned from the likelihood fit to the data. Figure 29 shows the fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\overline{B}^0 \to D^{*+}\pi^-} = 106 \ \pm 11.$$

If we remove the constraint on the signal width, we find $N_{\overline{B}^0 \to D^{*+}\pi^-} = 110 \pm 11$ and $\sigma_{\rm data} = 0.0295 \pm 0.0033$GeV/c^2. Removing the constraint on f_{D^*K} gives us $N_{\overline{B}^0 \to D^{*+}\pi^-} = 107 \pm 11$ and $f_{D^*K} = 0.053 \pm 0.053$. Removing the constraint on $f_{D^*\rho}$ gives us $N_{\overline{B}^0 \to D^{*+}\pi^-} = 107 \pm 11$ and $f_{D^*\rho} = 0.38 \pm 0.07$.

Table 11: $\overline{B}^0 \to D^{*+}\pi^-$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	$N_{\rm sig}$		-0.019 ± 0.031	0.964 ± 0.022	106 ± 11
2	μ	[GeV/c^2]	0.013 ± 0.033	1.032 ± 0.024	5.2772 ± 0.0002
3	$\sigma_{\rm MC}$	[GeV/c^2]	-0.043 ± 0.033	1.036 ± 0.024	0.0262 ± 0.0002
4	$f_{D^*\rho}$		-0.009 ± 0.032	1.006 ± 0.023	0.244 ± 0.008
5	$M_0^{D^*\rho}$	[GeV/c^2]	0.027 ± 0.033	0.943 ± 0.020	4.43 ± 0.01
6	$M_{off}^{D^*\rho}$	[GeV/c^2]	0.045 ± 0.032	0.993 ± 0.023	5.134 ± 0.001
7	$\sigma_{D^*\rho}$	[GeV/c^2]	-0.077 ± 0.033	1.001 ± 0.024	0.026 ± 0.001
8	$f_{\rm combg}$		-0.045 ± 0.032	0.940 ± 0.023	0.09 ± 0.03
9	$N_{\rm bg}$		-0.010 ± 0.033	1.019 ± 0.023	428 ± 21
10	$M_{\rm off}^{\rm otherB}$	[GeV/c^2]	-0.220 ± 0.031	0.972 ± 0.022	5.174 ± 0.004
11	f_{D^*K}		-0.047 ± 0.033	1.016 ± 0.023	0.069 ± 0.018
12	μ_{D^*K}	[GeV/c^2]	-0.032 ± 0.033	1.024 ± 0.024	5.2345 ± 0.0009
13	τ_{D^*K}	[GeV/c^2]$^{-1}$	-0.017 ± 0.032	0.986 ± 0.023	0.0287 ± 0.0009
14	σ_{D^*K}	[GeV/c^2]	0.029 ± 0.033	1.029 ± 0.024	0.0254 ± 0.0006
15	$m_{\rm diff}$	[GeV/c^2]	-0.034 ± 0.032	0.992 ± 0.023	0.005 ± 0.003
16	$\sigma_{\rm data}$	[GeV/c^2]	-0.050 ± 0.031	0.971 ± 0.022	0.026 ± 0.001

In conclusion, the un-constrained fits return a value consistent with the constrained fit, but with larger uncertainties. The fit χ^2/NDF are 20.0/12, 20.8/12, 16.5/12 and the fit probabilities are 6.7%, 5.4 %, 16.9% for the three different unconstrained fits.

Table 12: Correlation coefficients returned from the fit ($\overline{B}^0 \to D^{*+}\pi^-$ data)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1.000															
2	0.000	1.000														
3	0.000	0.001	1.000													
4	0.006	0.004	0.000	1.000												
5	0.001	0.000	0.000	-0.010	1.000											
6	-0.001	0.001	0.000	-0.005	-0.308	1.000										
7	0.000	0.000	0.000	0.003	0.164	-0.531	1.000									
8	-0.221	0.001	0.000	-0.037	-0.003	0.000	-0.001	1.000								
9	-0.062	0.000	0.000	-0.003	0.000	0.001	0.000	0.125	1.000							
10	-0.001	0.001	0.000	-0.008	-0.001	0.000	0.000	-0.004	0.001	1.000						
11	-0.124	0.001	0.000	-0.007	0.000	-0.003	-0.001	-0.057	-0.024	-0.007	1.000					
12	-0.001	0.000	0.000	-0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.006	1.000				
13	0.001	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	-0.001	0.000	0.003	0.765	1.000			
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	-0.002	-0.547	-0.532	1.000		
15	-0.071	0.065	0.000	0.050	-0.001	0.016	-0.007	0.215	0.050	0.008	-0.126	-0.004	-0.001	0.005	1.000	
16	0.128	0.000	0.000	0.005	0.001	-0.001	0.000	-0.242	-0.065	-0.001	-0.048	0.000	0.000	0.002	-0.123	1.000

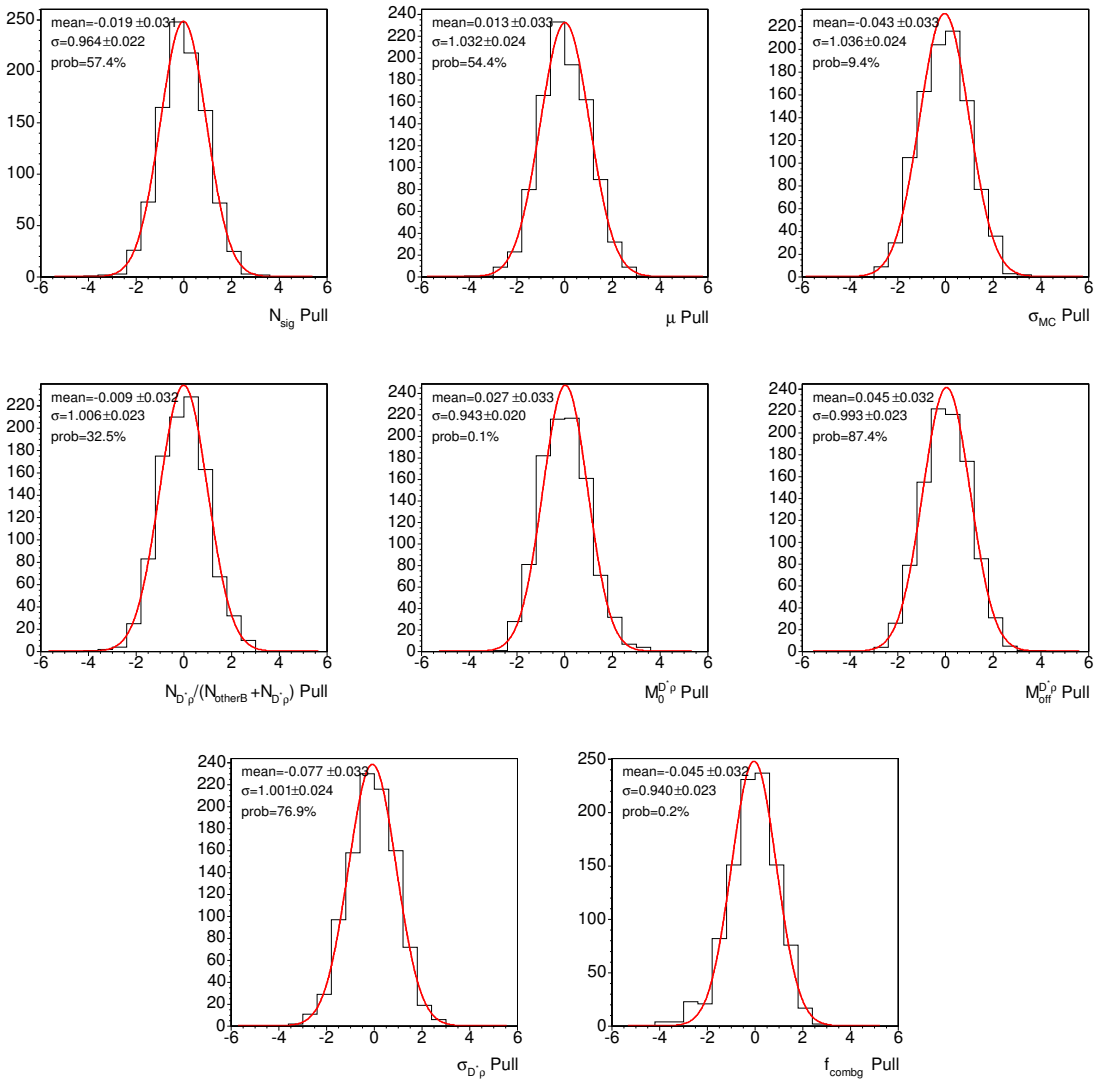


Figure 27: Pull of each fit parameter in the unbinned likelihood fit I ($D^*\pi$)

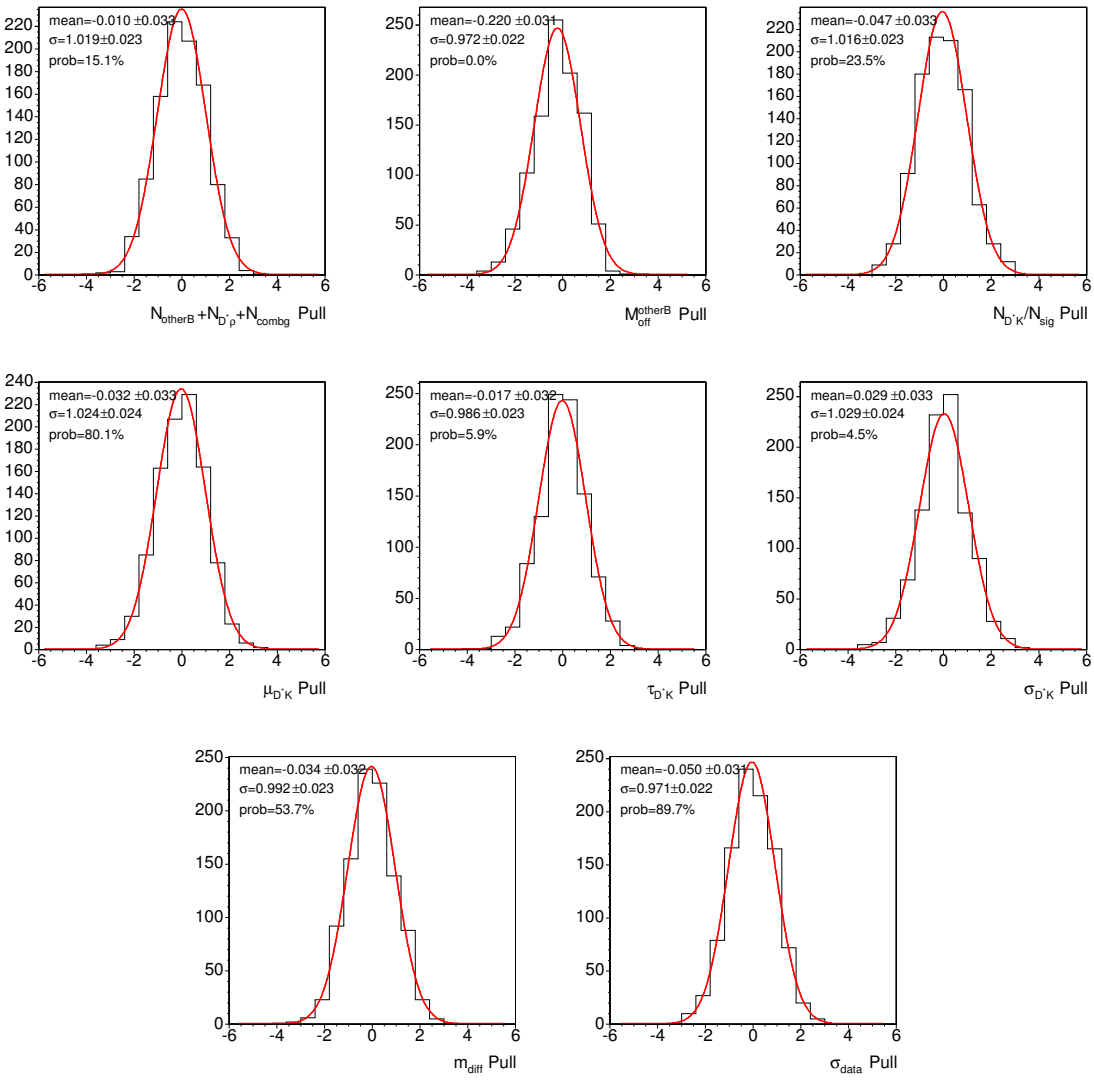


Figure 28: Pull of each fit parameter in the unbinned likelihood fit II ($D^*\pi$)

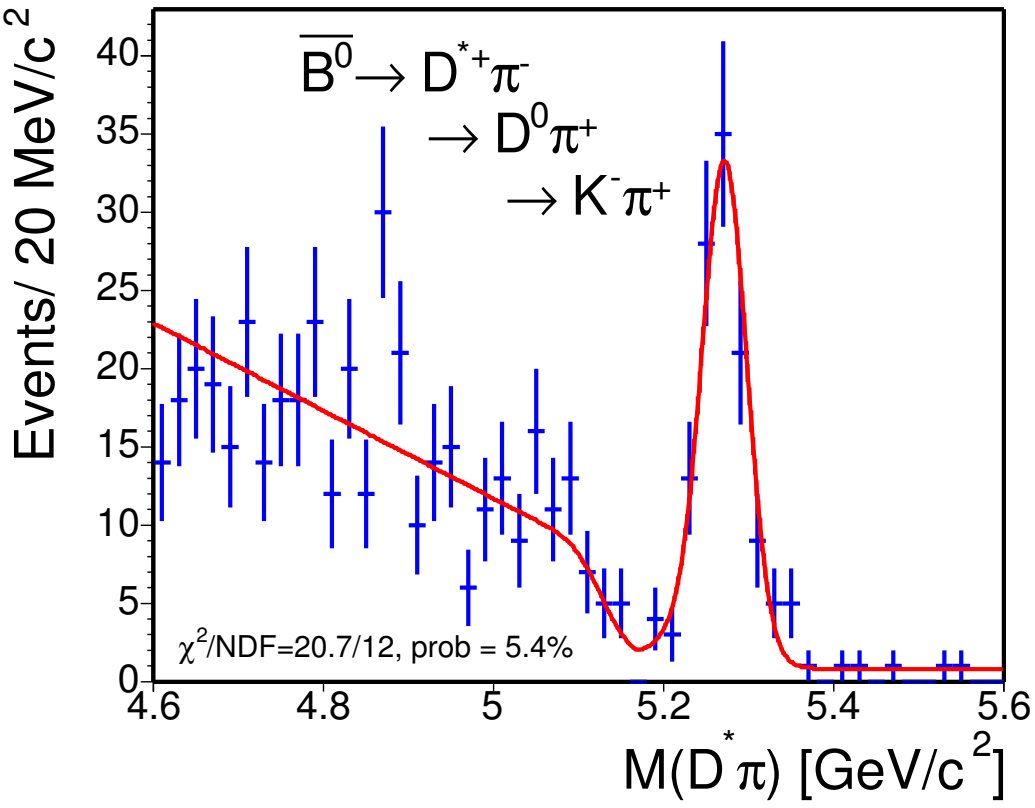


Figure 29: $M_{D^*\pi}$ from the $\overline{B}^0 \to D^{*+}\pi^-$ events is fit to a Gaussian (signal), a constant (combinatorial), and the background functions for the lower mass spectrum as described in the text. The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated. Note that the bins with less than 20 entries are combined.

4.2.2 $\overline{B}^0 \to D^+\pi^-$ Yield

As noted earlier, we make use of the mass function derived in Furic's analysis for the $\overline{B}^0 \to D^+\pi^-$ mode. The parameters which are kept constant in Furic's mass function remain constant in our analysis. The following backgrounds contribute to the mass spectrum of B^0 from Furic's study: Cabibbo suppressed decay $\overline{B}^0 \to D^+K^-$, $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B}^0 \to D^+\rho^-$, remaining $\overline{B} \to D^+X$ and the combinatorial background. Recent study by Belloni, Martin and Piedra *et al.* [4] [5] shows that the mis-reconstructed $\overline{B}_s \to D_s^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\pi^-$ produce small contamination in the $\overline{B}^0 \to D^+\pi^-$ signal. See the text below for the detailed descriptions.

1. $\overline{B}^0 \to D^+K^-$: fully reconstructed Cabibbo suppressed decays. The mass spectrum is a peak about 60 MeV/c^2 below the $\overline{B}^0 \to D^+\pi^-$ signal. The shape is modeled by a single Gaussian;

$$\mathcal{DK}(m) = \mathcal{G}(m, \mu - \Delta M_{DK}, \sigma_{DK}), \tag{29}$$

   where the shift of Gaussian mean from the $\overline{B}^0 \to D^+\pi^-$ signal, ΔM_{DK}, and the width, σ_{DK}, are extracted from the MC.

2. $\overline{B}_s \to D_s^+\pi^-$, where $D_s^+ \to \phi\pi^+$ and $\phi \to K^+K^-$: this decay produces a peak at around 5.31 GeV/c^2 when the pion mass is assigned to one of the kaons. The spectrum is modeled by double Gaussians with the same mean;

$$\mathcal{B}_\mathcal{S}(m) = f_1 \cdot \mathcal{G}(m, \mu_{B_s}, \sigma_1) + (1 - f_1) \cdot \mathcal{G}(m, \mu_{B_s}, \sigma_2), \tag{30}$$

   where the fraction f_1, μ_{B_s}, σ_1 and σ_2 of each Gaussian are obtained from the fit to the MC as shown in Figure 30 (top left).

3. $\Lambda_b \to \Lambda_c^+\pi^-$, where $\Lambda_c^+ \to pK^-\pi^+$: this background produces a broad peak around 5.4 GeV/c^2, the region where the pion mass is mis-assigned to the proton. The spectrum is modeled by a lifetime function;

$$\mathcal{L}_\mathcal{B}(m) = \exp(m, \tau_{\Lambda_b}) \otimes \mathcal{G}(m, \mu_{\Lambda_b}, \sigma_{\Lambda_b}), \tag{31}$$

   where μ_{Λ_b} and σ_{Λ_b} are the zero point and the resolution of the lifetime function. See Figure 30 (top right) for the fit to the $\Lambda_b \to \Lambda_c^+\pi^-$ MC when reconstructed as $D^+\pi^-$.

4. $\overline{B}^0 \to D^+\rho^-$, where $\rho^- \to \pi^0\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ where $D^{*+} \to D^+\pi^0$: These two backgrounds are combined. The spectrum of $\overline{B}^0 \to D^+\rho^-$ looks like $\overline{B}^0 \to D^{*+}\rho^-$ in Figure 26 (middle) and is modeled by a lifetime function. The spectrum of $\overline{B}^0 \to D^{*+}\pi^-$ is composed of two horns and is modeled by two Gaussians with different means.

   The structure of double horns arises for the following reasons: When $\overline{B}^0 \to D^{*+}\pi^-$, $D^{*+} \to D^+\pi^0$, is reconstructed as $D^+\pi^-$, the mass is lower than the world averaged B^0 mass due to the missing π^0. The amount of the negative mass shift, ΔM, is determined by the angle between the π^0 and the D^{*+} flight direction, $d\phi$. Because both B^0 and π^- are scalars (spin=0), to conserve the total angular momentum in the decay, the vector particle (spin=1), D^{*+}, is transversely polarized. The angle $d\phi$ from a transversely polarized D^{*+} is $\cos^2\theta$ distributed and the most probable $d\phi$ is either 0 or 180 degrees. Therefore, ΔM is quantized and this forms a double-horns spectrum.

   After combing $\overline{B}^0 \to D^+\rho^-$ and $\overline{B}^0 \to D^{*+}\pi^-$, we have:

$$\begin{aligned} \mathcal{R}(m) &= (1 - f_H) \cdot \exp(m, \tau_{\rm ref}) \otimes \mathcal{G}(m, \mu_{\rm ref}, \sigma_{\rm ref}) \\ &+ f_H \cdot (0.5 \cdot \mathcal{G}(m, \mu_{\rm ref} - \nu_{\rm ref} - \delta_{\rm ref}, \sigma_H) \\ &\quad +0.5 \cdot \mathcal{G}(m, \mu_{\rm ref} - \nu_{\rm ref} + \delta_{\rm ref}, \sigma_H)). \end{aligned} \tag{32}$$

The exact form of the lifetime function is found in Appendix B.2. The zero point of the lifetime function is $\mu_{\rm ref}$ and $\nu_{\rm ref}$ is the offset of the mid point between two horns from the lifetime function. The $\mu_{\rm ref}$ and $\nu_{\rm ref}$ are left free in the likelihood fit to the data. The values of the following parameters are extracted from the fit to the MC, as shown in Figure 30 (bottom left), and kept constant in the fit to the data: the lifetime ($\tau_{\rm ref}$), the fraction of horns (f_H), the half distance between the peak of two horns ($\delta_{\rm ref}$), the resolution of the lifetime function ($\sigma_{\rm ref}$) and the width of both horns (σ_H).

5. Continuum: remaining $\overline{B} \to D^+X$ decays and partially reconstructed. These backgrounds have similar mass spectrum and are group together. The shape is modeled by a first-order polynomial with a negative slope and a turn-off at $M_{\rm off}$; when $m < M_{\rm off}$:

$$\mathcal{H}(m) = \frac{2}{(M_{\rm off} - M_{\rm min})^2} \cdot (M_{\rm off} - m), \tag{33}$$

and when $m > M_{\rm off}$:

$$\mathcal{H}(m) = 0. \tag{34}$$

The lowest boundary of the $D\pi$ mass window, $M_{\rm min}$, is 4.6 GeV/c^2. See Figure 30 (bottom right) for the fit to these MC samples from Furic's analysis [3].

6. combinatorial: modeled by an exponential function. When the slope of the exponential, p_0, is not zero,

$$\mathcal{E}_{\mathcal{XP}}(m) = p_0 \cdot \frac{e^{-p_0 \cdot M_{\rm mid}}}{e^{-p_0 \cdot M_{\rm min}} - e^{-p_0 \cdot M_{\rm max}}} \cdot e^{-p_0 \cdot (x - M_{\rm mid})}, \tag{35}$$

and when p_0 is zero,

$$\mathcal{E}_{\mathcal{XP}}(m) = \frac{1}{M_{\rm max} - M_{\rm min}}, \tag{36}$$

where $M_{\rm max}$ and $M_{\rm min}$ specify the mass window: $4.6 < M_{D\pi} < 5.6$ GeV/c^2 and $M_{\rm mid}$ is the average of $M_{\rm max}$ and $M_{\rm min}$.

In the unbinned fit, the extended log likelihood function is expressed as a sum of a signal Gaussian, the functions for the DK mode ($\mathcal{DK}$), $D_s\pi$ ($\mathcal{B_S}$), $\Lambda_c\pi$ ($\mathcal{L_B}$), $D^*\pi$ plus $D\rho$ ($\mathcal{R}$), the remaining $\overline{B} \to D^+X$ decays ($\mathcal{H}$), and the combinatorial background ($\mathcal{E_{XP}}$):

$$\begin{aligned}
\log \mathcal{L} \;=\; & \sum_i \log\{N_{\rm sig} \cdot [\mathcal{G}(m_i, \mu, \sigma) + f_{DK} \cdot \mathcal{DK}(m_i) \\
& + f_{B_s} \cdot \mathcal{B_S}(m_i) + f_{\Lambda_b} \cdot \mathcal{L_B}(m_i)] \\
& + N_{\rm bg} \cdot [(1 - f_{\rm combg}) \cdot [(1 - f_{\rm otherB}) \cdot \mathcal{R}(m_i) + f_{\rm otherB} \cdot \mathcal{H}(m_i)] \\
& + f_{\rm combg} \cdot \mathcal{E_{XP}}(m_i)]\} \\
- \;\; & N_{\rm sig} \cdot (1 + f_{DK} + f_{B_s} + f_{\Lambda_b}) - N_{\rm bg},
\end{aligned} \tag{37}$$

where $\mathcal{DK}(m_i)$, $\mathcal{B_S}(m_i)$, $\mathcal{L_B}(m_i)$, $\mathcal{R}(m_i)$, $\mathcal{H}(m_i)$ and $\mathcal{E_{XP}}(m_i)$ are expressed in Equations 29–36. The fractions f_{DK}, f_{B_s} and f_{Λ_b} are the ratios of N_{DK}, N_{B_s} and N_{Λ_b} to the signal, $N_{\rm sig}$. The total amount of combinatorial background, the backgrounds from the $D\rho$, $D^*\pi$, and the remaining B decays is denoted as $N_{\rm bg}$. The parameters $f_{\rm combg}$ and $f_{\rm otherB}$ are defined as follows:

$$\begin{aligned}
f_{\rm combg} &\equiv \frac{N_{\rm combg}}{N_{\rm bg}}, \\
f_{\rm otherB} &\equiv \frac{N_{\rm otherB}}{N_{\rm otherB} + N_{D^*\pi} + N_{D\rho}}.
\end{aligned}$$

All the fractions and ratios here except $f_{\rm combg}$ are kept constant in the likelihood fit. The $\overline{B}^0 \to D^+K^-$ fraction, f_{DK}, is determined from the world average branching ratios;

$$f_{DK} = \frac{\mathcal{B}(\overline{B}^0 \to D^+K^-)}{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)}. \tag{38}$$

Table 13 lists the values of the branching ratios in Equation 38. We have $f_{DK} = 0.073 \pm 0.023$.

The B_s fraction, f_{B_s}, is obtained using the formula:

$$\begin{aligned} f_{B_s} &= \frac{f_s}{f_d} \cdot \frac{\mathcal{B}(\overline{B}_s \to D_s^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)} \cdot \frac{\mathcal{B}(D_s^+ \to \phi\pi^+)\mathcal{B}(\phi \to K^+K^-)}{\mathcal{B}(D^+ \to K^-\pi^+\pi^+)} \\ &\cdot \frac{\Gamma(D_s^+ \to K^+K^-\pi^-)}{\Gamma(D_s^+ \to \phi(K^+K^-)\pi^-)} \cdot \frac{\epsilon^{MC}_{\overline{B}_s \to D_s^+\pi^-}}{\epsilon^{MC}_{\overline{B}^0 \to D^+\pi^-}}, \end{aligned} \tag{39}$$

where the branching ratios are from the 2004 PDG and the CDF II measurement $\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(\overline{B}_s \to D_s^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)}$ by Furic [3]. The efficiency ratio is obtained by applying our $D\pi$ analysis cuts on the B_s MC. Inserting the numbers listed in Table 13 into Equation 39, we obtain $f_{Bs} = 0.006 \pm 0.001$. Note that the uncertainties from the branching ratios of ϕ, D_s, and D decays vanish after multiplying Furic's result with the ratio: $\frac{\mathcal{B}(D_s^+ \to \phi\pi^+)\mathcal{B}(\phi \to K^+K^-)}{\mathcal{B}(D^+ \to K^-\pi^+\pi^+)}$.

The Λ_b fraction, f_{Λ_b}, is obtained using a similar formula;

$$f_{\Lambda_b} = \frac{\sigma_{\Lambda_b}(P_T > 6.0)\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T > 6.0)\mathcal{B}(\overline{B}^0 \to D^+\pi^-)} \times \frac{\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)}{\mathcal{B}(D^+ \to K^-\pi^+\pi^+)} \times \frac{\epsilon^{MC}_{\Lambda_b \to \Lambda_c^+\pi^-}}{\epsilon^{MC}_{\overline{B}^0 \to D^+\pi^-}}, \tag{40}$$

where the product of the first and the second terms come from 2004 PDG and CDF II measurements by Le, *et al.* [4]. The uncertainties from the branching ratios of Λ_c and D decays vanish in Equation 40. The efficiency ratio is obtained using the $\Lambda_b \to \Lambda_c^+\pi^-$ MC. The value of f_{Λ_b} is then 0.031 ± 0.005. Table 13 lists the numerical values of Le's result and the MC efficiency. Finally, $f_{\rm otherB}$ is obtained using Furic's $\overline{B} \to D^+X$ MC. We apply our analysis cuts and count the number of $D^*\pi + D\rho$ and the remaining $\overline{B} \to D^+X$ events. We find $f_{\rm otherB} = 0.569 \pm 0.011$.

Table 14 lists the constant parameters with their values and uncertainties obtained from the fit to the MC. Table 15 lists the mean, width of the pulls from the toy MC test and the value of each fit parameter from the fit to the data. Figure 31 gives the pull of each fit parameter. All the pull means and widths are consistent with zero and one, except the pull mean of $f_{\rm combg}$, which is -0.145±0.034. However, this only corresponds to a 1.1 % shift on the central value. The fit for the number of signal events, $N_{\rm sig}$, is not affected. Table 16 gives the correlation coefficients returned from the likelihood fit to the data. Figure 32 shows the fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\overline{B}^0 \to D^+\pi^-} = 579 \ \pm 30.$$

Table 13: Parameter values used to determine f_{DK}, f_{B_s} and f_{Λ_b}

$\mathcal{B}(\overline{B}^0 \to D^+K^-)$	$(2.0 \pm 0.6)\times10^{-4}$
$\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$	$(2.76 \pm 0.25)\times10^{-3}$
f_{DK}	0.073 ± 0.023
$\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(\overline{B}_s \to D_s^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)}$	$0.35 \pm 0.05(stat) \pm 0.02\ (syst) \pm 0.09\ (BR)$
$\mathcal{B}(D_s^+ \to \phi\pi^+)$	$(3.6 \pm 0.9)\%$
$\mathcal{B}(\phi \to K^+K^-)$	$(49.1 \pm 0.6)\%$
$\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$	$(9.2 \pm 0.6)\%$
$\frac{\Gamma(D_s^+ \to K^+K^-\pi^-)}{\Gamma(D_s^+ \to \phi(K^+K^-)\pi^-)}$	0.81 ± 0.08
$\epsilon^{MC}_{\overline{B}_s \to D_s^+\pi^-}/\epsilon^{MC}_{\overline{B}^0 \to D^+\pi^-}$	0.071 ± 0.004
f_{B_s}	0.006 ± 0.001
$\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0 \to D^+\pi^-)}$	$0.82 \pm 0.08\ (stat) \pm 0.11\ (syst) \pm 0.22\ (BR)$
$\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$	$(5.0 \pm 1.3)\%$
$\epsilon^{MC}_{\Lambda_b \to \Lambda_c^+\pi^-}/\epsilon^{MC}_{\overline{B}^0 \to D^+\pi^-}$	0.069 ± 0.002
f_{Λ_b}	0.031 ± 0.005

Table 14: Fixed parameters in the $\overline{B}^0 \to D^+\pi^-$ unbinned likelihood fit.

Parameter	Meaning	Value
f_{DK}	$N_{\overline{B}^0 \to D^+K^-}/N_{\overline{B}^0 \to D^+\pi^-}$	0.073 $\pm$ 0.023
ΔM_{DK}	mass shift of $\overline{B}^0 \to D^+K^-$ [GeV/c^2]	0.067 $\pm$ 0.006
σ_{DK}	width of $\overline{B}^0 \to D^+K^-$[GeV/c^2]	0.032 $\pm$ 0.009
f_{B_s}	$N_{\overline{B}_s \to D_s^+\pi^-}/N_{\overline{B}^0 \to D^+\pi^-}$	0.006 $\pm$ 0.001
μ_{B_s}	mean of B_s background [GeV/c^2]	5.307 $\pm$ 0.001
f_1	fraction of the narrow B_s Gaussian	0.773 $\pm$ 0.002
σ_1	width of the narrow B_s Gaussian [GeV/c^2]	0.021 $\pm$ 0.002
σ_2/σ_1	width ratio of the B_s Gaussians	1.8 $\pm$ 0.3
f_{Λ_b}	$N_{\Lambda_b \to \Lambda_c^+\pi^-}/N_{\overline{B}^0 \to D^+\pi^-}$	0.031 $\pm$ 0.005
μ_{Λ_b}	mean of Λ_b [GeV/c^2]	5.416 $\pm$ 0.002
σ_{Λ_b}	width of Λ_b background [GeV/c^2]	0.024 $\pm$ 0.002
τ_{Λ_b}	lifetime of Λ_b background [GeV/c$^{2-1}$]	0.052 $\pm$ 0.002
$\tau_{\rm ref}$	lifetime of $D\rho$ background [GeV/c$^{2-1}$]	0.36 $\pm$ 0.06
$\sigma_{\rm ref}$	width of $D\rho$ background [GeV/c^2]	0.039 $\pm$ 0.008
f_H	fraction of $D^*\pi$ horns	0.20 $\pm$ 0.06
$\delta_{\rm ref}$	distance between two horns [GeV/c^2]	0.039 $\pm$ 0.003
σ_H	width of the horns [GeV/c^2]	0.019 $\pm$ 0.003
$f_{\rm otherB}$	fraction of the remaining $\overline{B} \to D^+X$	0.569 $\pm$ 0.011
$M_{\rm off}$	cut off for $\overline{B} \to D^+X$ mass [GeV/c^2]	5.112 $\pm$ 0.007

Table 15: $\overline{B}^0 \to D^+\pi^-$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	$N_{\rm sig}$		0.012 ± 0.035	1.021 ± 0.026	579 ± 30
2	μ	[GeV/c^2]	-0.026 ± 0.034	0.989 ± 0.025	5.278 ± 0.001
3	σ	[GeV/c^2]	-0.040 ± 0.035	1.015 ± 0.026	0.0235 ± 0.0012
4	$N_{\rm bg}$		0.017 ± 0.034	0.990 ± 0.025	4049 ± 67
5	$\mu_{\rm ref}$	[GeV/c^2]	-0.036 ± 0.036	1.037 ± 0.026	5.145 ± 0.015
6	$\nu_{\rm ref}$	[GeV/c^2]	-0.037 ± 0.038	1.085 ± 0.028	0.068 ± 0.020
7	$f_{\rm combg}$		-0.145 ± 0.034	0.988 ± 0.025	0.583 ± 0.044
8	p_0		-0.051 ± 0.034	0.976 ± 0.024	1.75 ± 0.15

Table 16: Correlation coefficients returned from the fit ($\overline{B}^0 \to D^+\pi^-$ data)

	1	2	3	4	5	6	7	8
1	1.000							
2	-0.021	1.000						
3	0.308	-0.013	1.000					
4	-0.201	0.011	-0.151	1.000				
5	-0.034	0.105	-0.037	0.017	1.000			
6	-0.042	0.095	-0.043	0.020	0.946	1.000		
7	-0.312	0.023	-0.233	0.154	-0.005	0.021	1.000	
8	-0.076	0.053	0.056	0.037	0.190	0.188	0.731	1.000

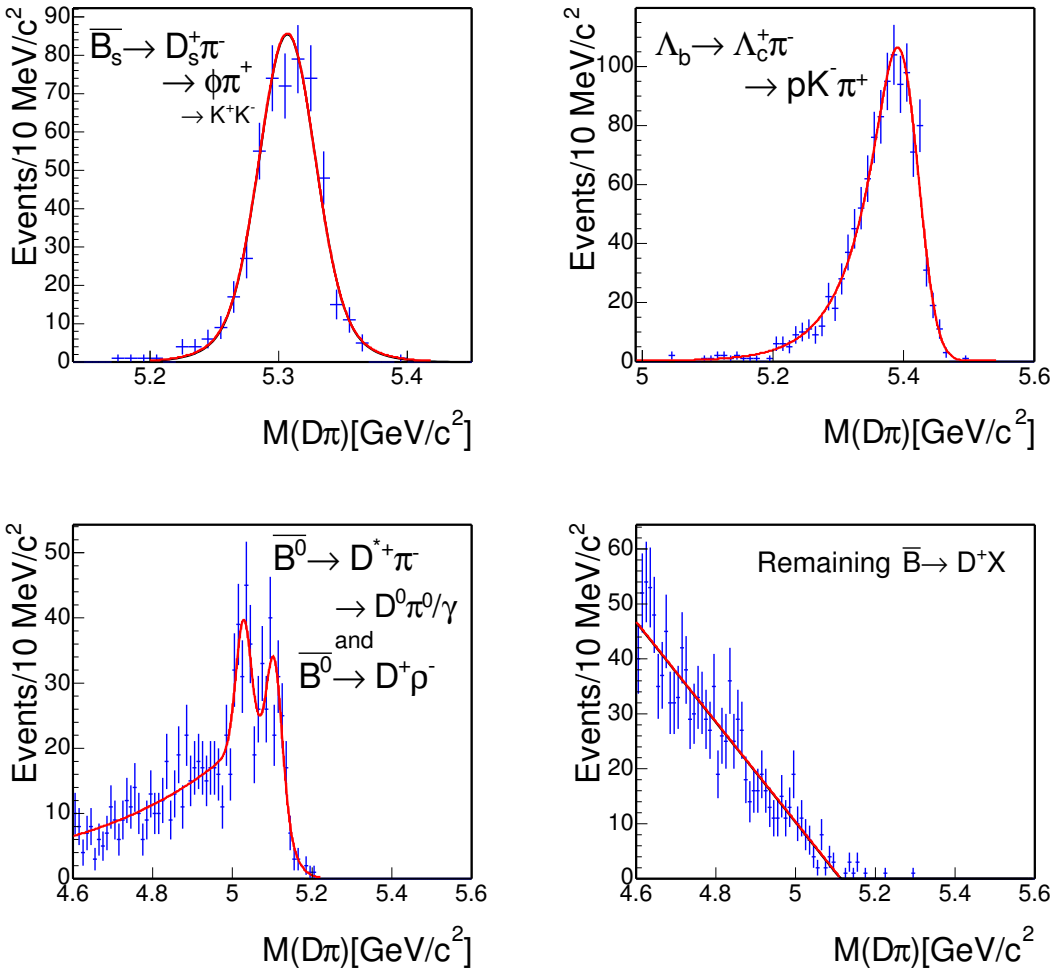


Figure 30: Various MC samples reconstructed as $\overline{B}^0 \to D^+\pi^-$. From the top left to the bottom right are $\overline{B}_s \to D_s^+\pi^-$, $\Lambda_b \to \Lambda_c^+\pi^-$, $\overline{B}^0 \to D^{*+}\pi^-$ + $\overline{B}^0 \to D^+\rho^-$, and the remaining $\overline{B} \to D^+X$. The fit probabilities are 32.5%, 66.8%, 16.5% and 32.9 %

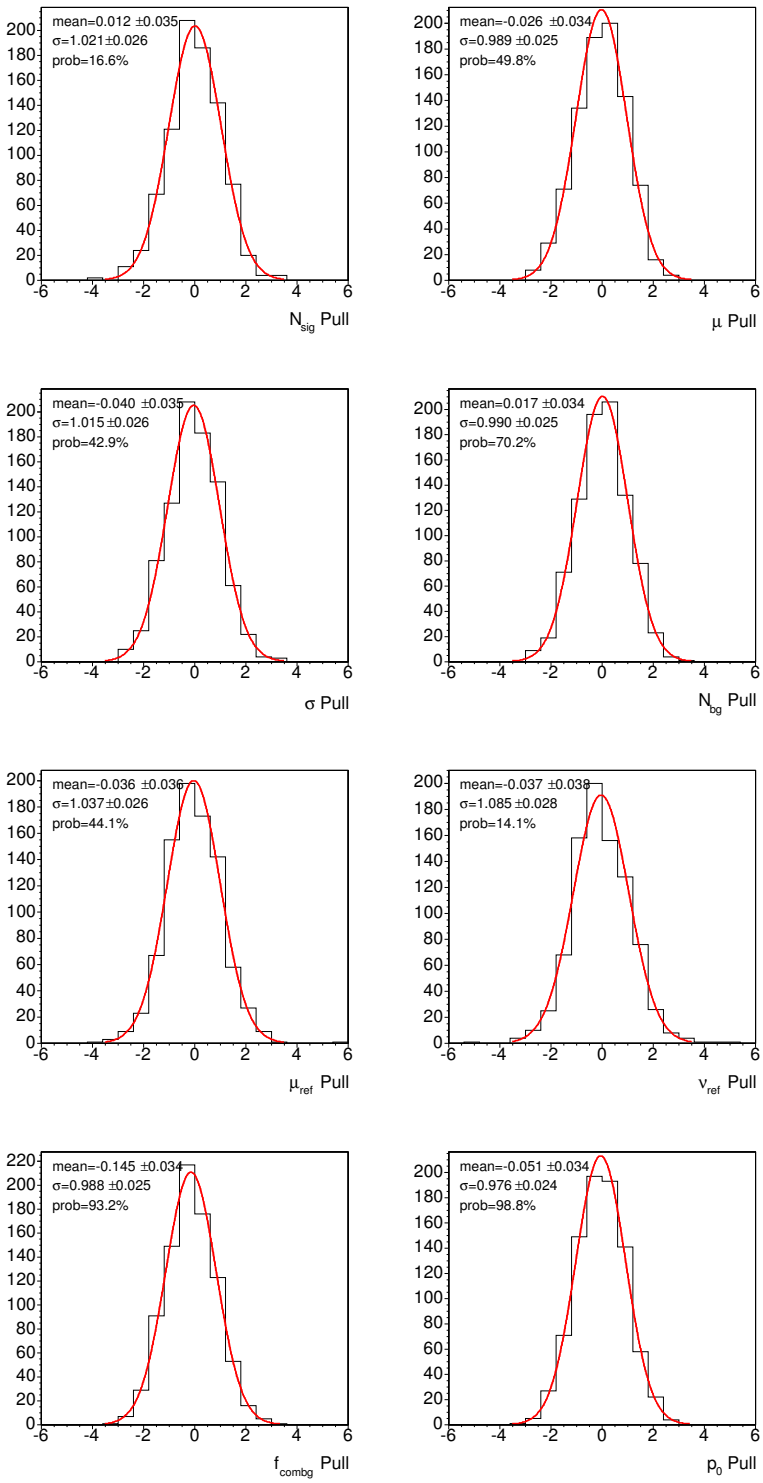


Figure 31: Pull of each fit parameter in the unbinned likelihood fit ($D\pi$)

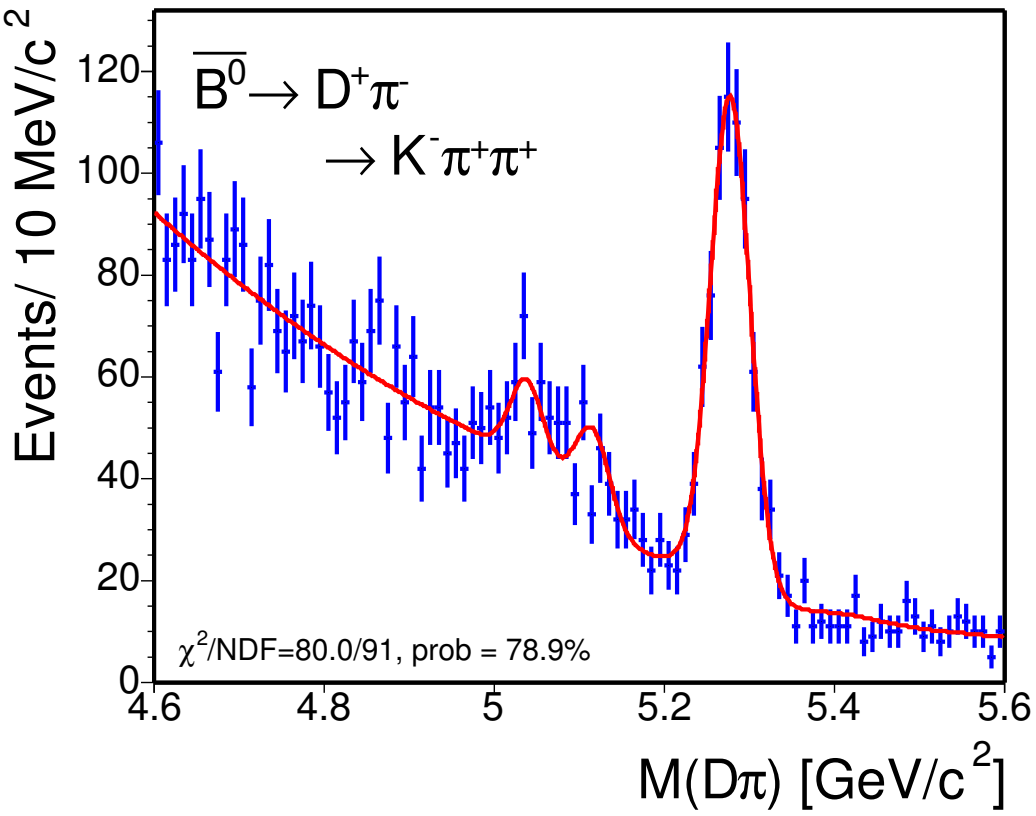


Figure 32: $M_{D\pi}$ from the $\overline{B}^0 \to D^+\pi^-$ events is fit to a Gaussian (signal), an exponential (combinatorial), and the background functions for the lower mass spectrum as described in the text. The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated.

4.2.3 $\Lambda_b \to \Lambda_c^+\pi^-$ Yield

We use the mass function derived in the analysis of Martin and Maksimović for the $\Lambda_b \to \Lambda_c^+\pi^-$ mode [4]. The parameters which were kept constant remain constant in our analysis. We cross-check the values of the parameters by applying our analysis cuts on the MC used in the Martin and Maksimović analysis and find the same numbers can be used for this analysis. The following backgrounds contribute to the mass spectrum of Λ_b from their study: Cabibbo decay $\Lambda_b \to \Lambda_c^+K^-$, four-prong mis-identified B meson, the remaining B meson decays, the remaining Λ_b decays and the combinatorial background. Detailed descriptions of each background are find below.

1. $\Lambda_b \to \Lambda_c^+K^-$: fully reconstructed Cabibbo suppressed decays. The mass spectrum is a peak about 50 MeV/c^2 below the $\Lambda_b \to \Lambda_c^+\pi^-$ signal. The shape is modeled by two Gaussians of different mean and width;

$$\mathcal{L_C K}(m) = f_1 \cdot \mathcal{G}(m, \mu^1_{\Lambda_c K}, \sigma_1) + (1 - f_1) \cdot \mathcal{G}(m, \mu^2_{\Lambda_c K}, \sigma_2), \tag{41}$$

   where f_1, $\mu^1_{\Lambda_c K}$, σ_1, $\mu^2_{\Lambda_c K}$ and σ_2 are from the fit to the MC.

2. mis-identified four-prong B mesons: all the B mesons with four tracks in the final states and fully reconstructed. $\overline{B}^0 \to D^+\pi^-$ contributes about 50% of this type of background. Since these decays have similar final state as our $\Lambda_b \to \Lambda_c^+\pi^-$ signal, they produce a distinguished peak to the left of the signal Gaussian. This background ($\mathcal{B}_{4\mathcal{PRONG}}$) is modeled by the sum of a Landau ($\mathcal{L}_{\mathcal{AND}}$) and a Gaussian function:

$$\mathcal{B}_{4\mathcal{PRONG}}(m) = f_{\rm L} \cdot \mathcal{L}_{\mathcal{AND}}(m, \mu_{\rm BPL}, \sigma_L) + (1 - f_{\rm L}) \cdot \mathcal{G}(m, \mu_{\rm BPG}, \sigma_G), \tag{42}$$

   where $f_{\rm L}$, $\mu_{\rm BPL}$ and σ_L are the fraction, mean and the width of Landau distribution. The mean and the width of the Gaussian are denoted as $\mu_{\rm BPG}$ and σ_G. These parameters are extracted from fit to the MC as shown in Figure 33 (bottom).

3. remaining B meson decays: this background ($\mathcal{O}_{\mathcal{B}}$) spectrum is modeled by the sum of an exponential function and a product of a bifurcated Gaussian ($\mathcal{B}_{\mathcal{F}}$) with a step-down function:

$$\mathcal{O}_{\mathcal{B}}(m) = \mathcal{E}_{\mathcal{XP}}(m) + f_{\rm bifg} \cdot \mathcal{B}_{\mathcal{F}}(m, \mu_{\rm ob}, \sigma^L_{\rm ob}, \sigma^R_{\rm ob}) \cdot (1 - \frac{1}{1 + e^{(\mu_{obst} - m)/a^{\rm ob}_0}}) \tag{43}$$

   where $\mathcal{E}_{\mathcal{XP}}(m)$ is expressed in Equations 35–36. The parameters $f_{\rm bifg}$, $\mu_{\rm ob}$, $\sigma^L_{\rm ob}$, and $\sigma^R_{\rm ob}$ are the fraction, mean, left sigma, right sigma of bifurcated Gaussian. The step-down function parameters, μ_{obst} and $a^{\rm ob}_0$, together with the parameters for the bifurcated Gaussian, are extracted from the MC as shown in Figure 33 (top left). The exact form of the bifurcated Gaussian is found in Appendix B.3.

4. remaining Λ_b decays: this background ($\mathcal{O}_{\mathcal{L}}$) spectrum is modeled by the sum of two Gaussians and the product of a bifurcated Gaussian and a step-down function

$$\begin{aligned} \mathcal{O}_{\mathcal{L}}(m) \quad = \quad & f^{\rm ol}_1 \cdot \mathcal{G}(m, \mu^{olg}_1, \sigma^{\rm ol}_1) + f^{\rm ol}_2 \cdot \mathcal{G}(m, \mu^{olg}_2, \sigma^{\rm ol}_2) \\ & + \cdot \mathcal{B}_{\mathcal{F}}(m, \mu_{\rm ol}, \sigma^L_{\rm ol}, \sigma^R_{\rm ol}) \cdot (1 - \frac{1}{1 + e^{(\mu_{\rm olst} - m)/a^{\rm ol}_0}}) \end{aligned} \tag{44}$$

   where the parameters in the function are from the fit to the MC as shown in Figure 33 (top right).

5. combinatorial background: described by an exponential function

In the unbinned fit, the extended log likelihood function is expressed as a sum of a signal Gaussian and the functions for $\Lambda_c K$ ($\mathcal{L_C K}$), four-prong B meson ($\mathcal{B}_{4\mathcal{PRONG}}$), remaining B meson decays ($\mathcal{O_B}$), remaining Λ_b decays ($\mathcal{O_L}$) and the combinatorial background ($\mathcal{E_{XP}}$). In addition, there is a constraint on the width of the signal Gaussian determined using the $\overline{B}^0 \to D^+\pi^-$ data, $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\pi^-$ MC as described earlier.

$$\begin{aligned}
\log \mathcal{L} &= \sum_i \log\{N_{\rm sig} \cdot [\mathcal{G}(m_i,\mu,\sigma) + f_{\Lambda_c K} \cdot \mathcal{L_C K}(m_i)] + N_{\rm B4prong} \cdot \mathcal{B}_{4\mathcal{PRONG}}(m_i) \\
&\quad + N_{\rm OB} \cdot \mathcal{O_B}(m_i) + N_{\rm OL} \cdot \mathcal{O_L}(m_i) + N_{\rm combg} \cdot \mathcal{E_{XP}}(m_i)\} \\
&- N_{\rm sig} \cdot (1 + f_{\Lambda_c K}) - N_{\rm B4prong} - N_{\rm OB} - N_{\rm OL} - N_{\rm combg} \\
&+ \log \mathcal{C}_\sigma, \qquad (45)
\end{aligned}$$

where $\mathcal{L_C K}(m_i)$, $\mathcal{B}_{4\mathcal{PRONG}}(m_i)$, $\mathcal{O_B}(m_i)$, $\mathcal{O_L}(m_i)$ and $\mathcal{E_{XP}}(m_i)$ are expressed in Equations 41–44 and Equations 35 –36. The fraction $f_{\Lambda_c K}$ is defined as:

$$f_{\Lambda_c K} = \frac{N_{\Lambda_b \to \Lambda_c^+ K^-}}{N_{\Lambda_b \to \Lambda_c^+ \pi^-}}, \qquad (46)$$

and is fixed to 0.08; the number is suggested by the branching ratio of the Cabibbo suppressed relative to the Cabibbo favored decay in the B meson system. The Gaussian constraint of the signal width, $\mathcal{C}_\sigma$ is expressed as:

$$\mathcal{C}_\sigma = \mathcal{G}(\sigma, \mu_p, \sigma_p), \qquad (47)$$

where $\mu_p = 0.0231$ GeV/c^2, and $\sigma_p = 0.0012$ GeV/c^2.

Table 17 lists the values of the constant parameters imported from the analysis of Martin and Maksimović. Table 18 lists the mean, width of the pulls from the toy MC test and the unbinned likelihood fit result to the data. Figure 34 gives the pull of each fit parameter. Each pull mean is consistent with zero and the pull width is consistent with one. Table 19 gives the correlation coefficients returned from the likelihood fit to the data. Figure 35 shows the fit result superimposed on the data histogram. We have obtained from the fit:

$$N_{\Lambda_b \to \Lambda_c^+ \pi^-} = 179 \ \pm 19.$$

We also cross-check by removing the constraint on signal width and obtain $N_{\Lambda_b \to \Lambda_c^+ \pi^-} = 177 \pm 22$, and $\sigma = 0.022 \pm 0.004$, which are consistent with the fit result in Table 18. The fit without constraint has a χ^2/NDF of 123.2/111 and fit probability of 20.2 %.

Table 17: Fixed parameters in the $\Lambda_b \to \Lambda_c^+ \pi^-$ unbinned likelihood fit.

Parameter	Meaning	Value
$f_{\Lambda_c K}$	$N_{\Lambda_b \to \Lambda_c^+ K^-}/N_{\Lambda_b \to \Lambda_c^+ \pi^-}$	0.080
f_1	fraction of the narrow $\Lambda_C K$ Gaussian	0.902
$\mu^1_{\Lambda_c K}$	mean of the narrow $\Lambda_C K$ Gaussian [GeV/c^2]	5.573
σ_1	width of the narrow $\Lambda_C K$ Gaussian [GeV/c^2]	0.029
$\mu^2_{\Lambda_c K}$	mean of the wide $\Lambda_C K$ Gaussian [GeV/c^2]	5.529
σ_2	width of the wide $\Lambda_C K$ Gaussian [GeV/c^2]	0.075
$f_{\rm L}$	fraction of the Landau, 4-prong	0.413
$\mu_{\rm BPL}$	mean of the Landau, 4-prong [GeV/c^2]	5.486
σ_L	width of the Landau, 4-prong [GeV/c^2]	0.025
$\mu_{\rm BPG}$	mean of the Gaussian, 4-prong [GeV/c^2]	5.526
σ_G	width of the Gaussian, 4-prong [GeV/c^2]	0.078
s_0	slope of the exponential, other B	2.180
$f_{\rm bifg}$	fraction of the bifurcated Gaus, other B	0.106
$\mu_{\rm ob}$	mean of the bifurcated Gaus, other B [GeV/c^2]	5.598
$\sigma^L_{\rm ob}$	left σ of the bifurcated Gaus, other B [GeV/c^2]	10.0
$\sigma^R_{\rm ob}$	right σ of the bifurcated Gaus, other B [GeV/c^2]	4.800
μ_{obst}	mean of "step-down", other B [GeV/c^2]	5.436
$a_0^{\rm ob}$	slope of the "step-down", other B	0.079
$\mu_{\rm ol}$	mean of the bifurcated Gaus, other Λ_b [GeV/c^2]	3.469
$\sigma^L_{\rm ol}$	left σ of the bifurcated Gaus, other Λ_b [GeV/c^2]	10.0
$\sigma^R_{\rm ol}$	right σ of the bifurcated Gaus, other Λ_b [GeV/c^2]	1.236
$\mu_{\rm olst}$	mean of "step-down", other Λ_b [GeV/c^2]	5.451
$a_0^{\rm ol}$	slope of "step-down", other Λ_b [GeV/c^2]	0.091
$f_1^{\rm ol}$	fraction of first Gaus, other Λ_b	0.0005
$\mu_1^{\rm ol}$	mean of first Gaus, other Λ_b	5.644
$\sigma_1^{\rm ol}$	width of first Gaus, other Λ_b	0.019
$f_2^{\rm ol}$	fraction of second Gaus, other Λ_b	0.0034
$\mu_2^{\rm ol}$	mean of second Gaus, other Λ_b	5.459
$\sigma_2^{\rm ol}$	width of second Gaus, other Λ_b	0.030

Table 18: $\Lambda_b \to \Lambda_c^+ \pi^-$ results from the unbinned likelihood fit.

Index	Parameter		1000 toy MC pull mean	1000 toy MC pull width	Data fit value
1	$N_{\rm sig}$		0.007 ± 0.032	0.995 ± 0.023	179 ± 19
2	μ	[GeV/c^2]	0.021 ± 0.033	1.031 ± 0.024	5.621 ± 0.003
3	σ	[GeV/c^2]	0.026 ± 0.031	0.976 ± 0.022	0.023 ± 0.001
4	$N_{\rm B4prong}$		0.002 ± 0.032	1.018 ± 0.023	150 ± 32
5	$N_{\rm OB}$		0.038 ± 0.033	1.046 ± 0.024	3170 ± 291
6	$N_{\rm OL}$		-0.048 ± 0.033	1.030 ± 0.023	962 ± 324
7	$N_{\rm combg}$		-0.023 ± 0.032	1.013 ± 0.023	1971 ± 171
8	p_0		-0.027 ± 0.032	1.010 ± 0.023	0.63 ± 0.10

Table 19: Correlation coefficients returned from the fit ($\Lambda_b \to \Lambda_c^+ \pi^-$ data)

	1	2	3	4	5	6	7	8
1	1.000							
2	-0.077	1.000						
3	0.212	-0.100	1.000					
4	-0.244	0.163	-0.104	1.000				
5	-0.011	0.023	-0.005	0.278	1.000			
6	0.075	0.014	0.030	-0.109	-0.862	1.000		
7	-0.138	-0.088	-0.053	-0.394	-0.060	-0.393	1.000	
8	-0.121	-0.072	-0.046	-0.333	-0.098	-0.348	0.898	1.000

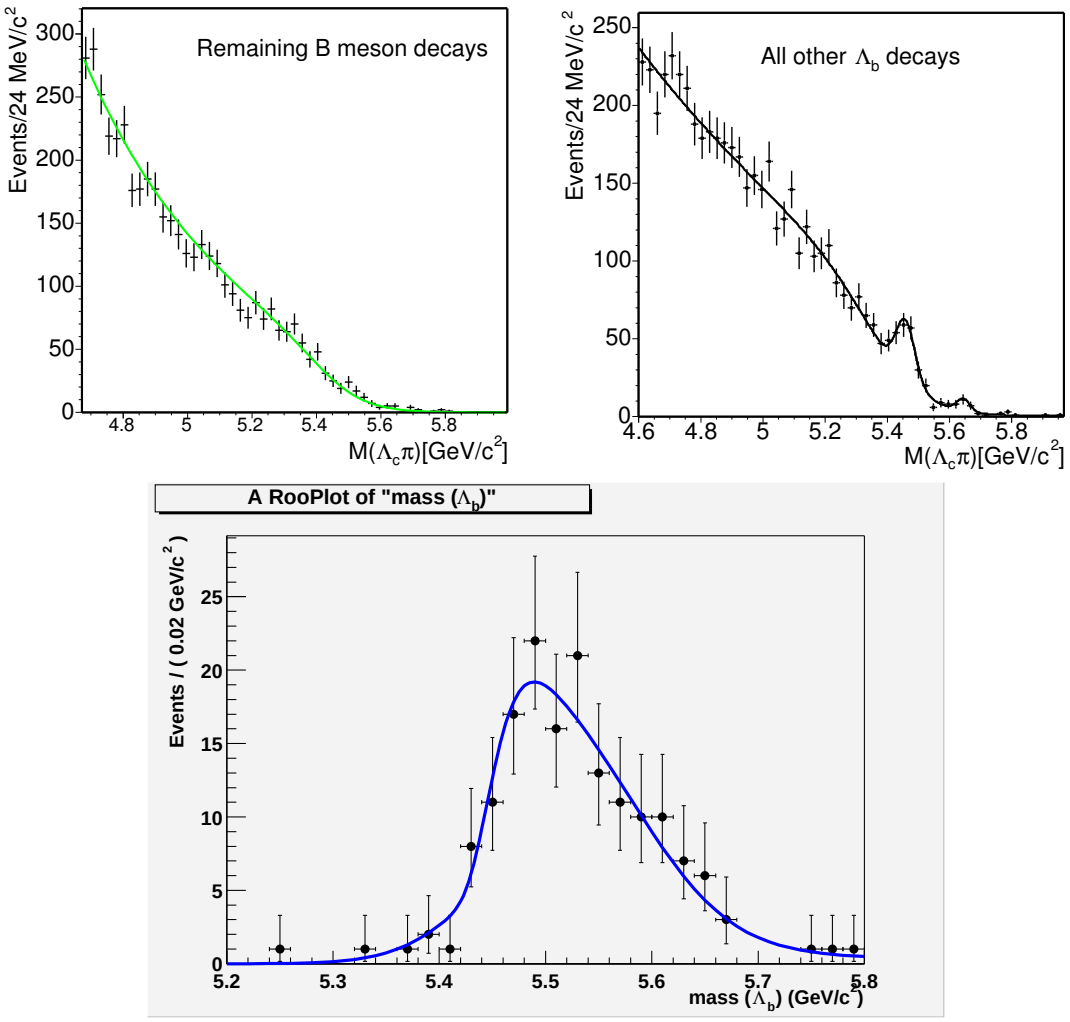


Figure 33: Various MC samples reconstructed as $\Lambda_b \to \Lambda_c^+\pi^-$. From the top left to the bottom are the remaining B meson decays, the remaining Λ_b decays and four-prong B meson decays.

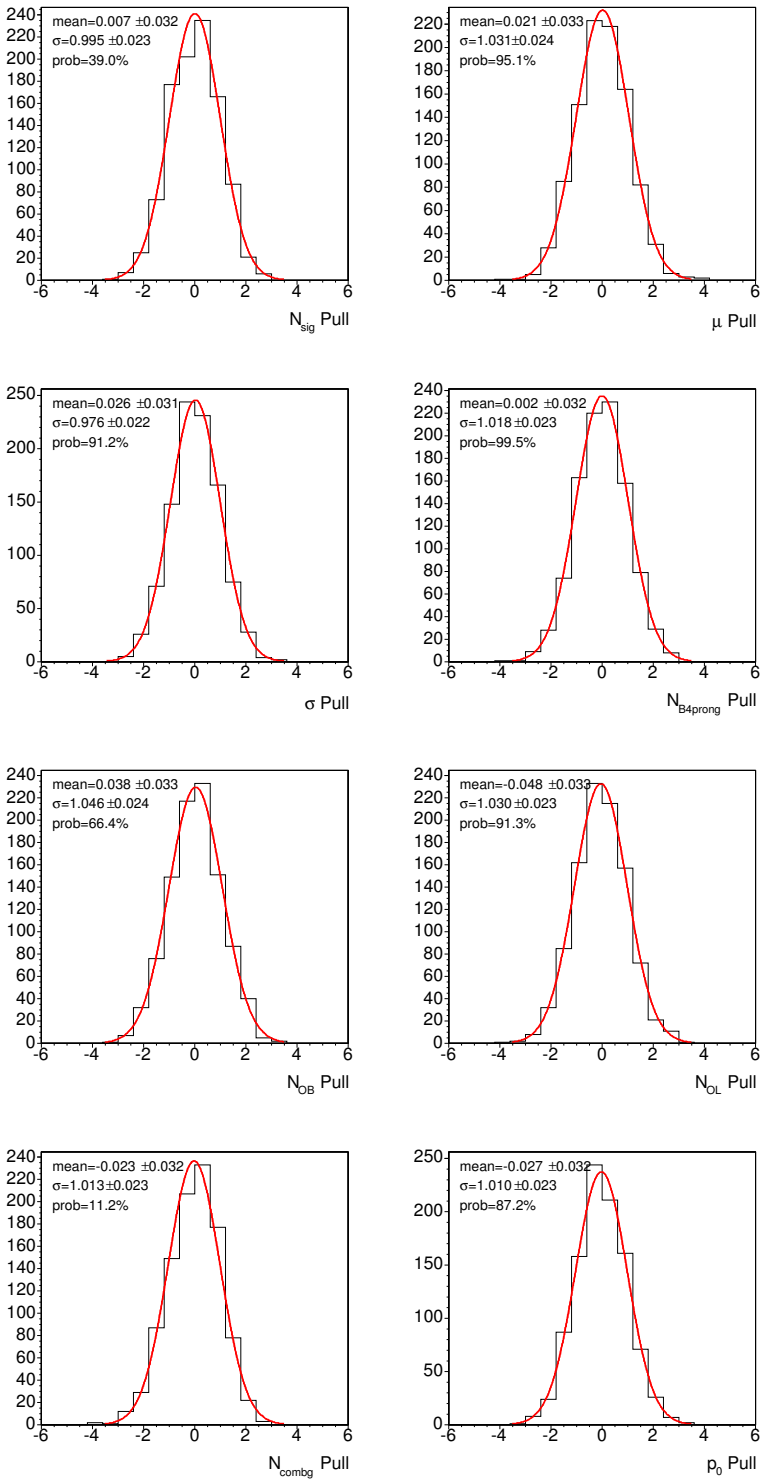


Figure 34: Pull of each fit parameter in the unbinned likelihood fit ($\Lambda_c\pi$)

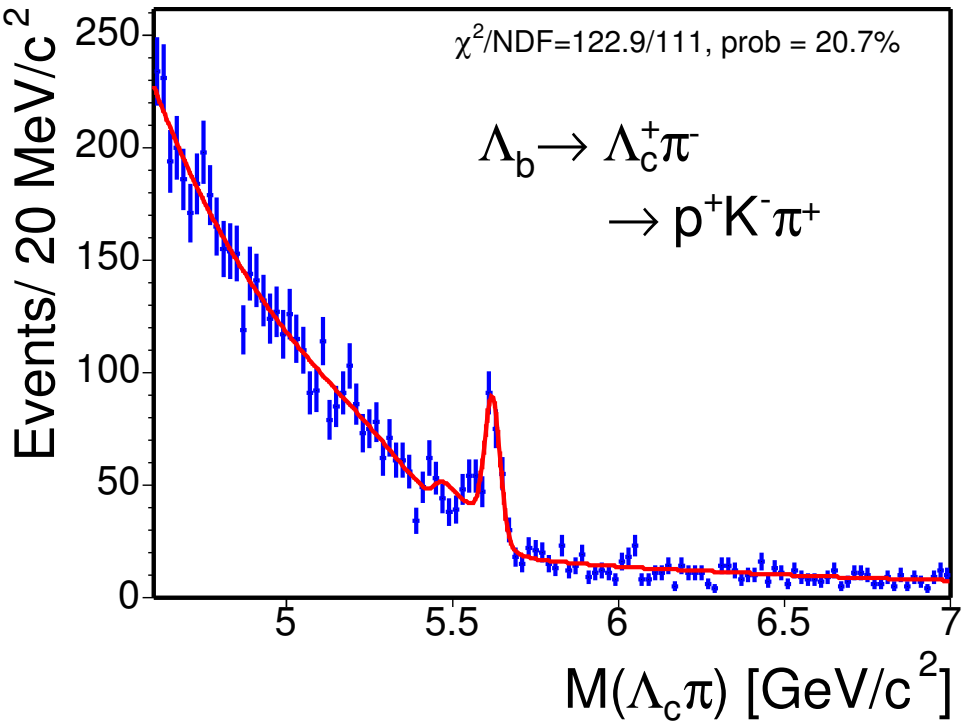


Figure 35: $M_{\Lambda_c\pi}$ from the $\Lambda_b \rightarrow \Lambda_c^+\pi^-$ events is fit to a Gaussian (signal), an exponential (combinatorial), and the background functions for the lower mass spectrum as described in the text. The result of the unbinned likelihood fit is projected on the histogram and a χ^2 probability is calculated.

4.3 Summary

Using the unbinned, extended log likelihood technique, we fit the charm and B hadron mass spectra to obtain the number of events. The yield for each mode is listed below. The performance of the fitter is validated using 1000 toy MC test for each mode. In general, the mean of each pull distribution from the toy MC test is consistent with zero and the width is consistent with one. For the fit parameter with a pull mean deviated from zero and a width deviated from unity, the fitter only indicates a less than 1% bias on the central value. Besides, these fit parameters are not correlated with the number of signal events and do not affect the yield we obtain. The fit result to the data is also superimposed on the data histograms and a χ^2 is computed. We have obtained good χ^2 for each mode.

Mode	Yield
$\overline{B}^0 \to D^{*+}\pi^-$	106 ± 11
$\overline{B} \to D^{*+}\mu^- X$	1059 ± 33
$\overline{B}^0 \to D^{+}\pi^-$	579 ± 30
$\overline{B} \to D^{+}\mu^- X$	4721 ± 104
$\Lambda_b \to \Lambda_c^+\pi^-$	179 ± 19
$\overline{B} \to \Lambda_c^+\mu^- X$	1237 ± 97

5 Monte Carlo Samples, Acceptance and Efficiencies

With the raw yield in hand, we now turn to the correction which must be applied to obtain the value of the ratio of branching fractions, that is the acceptance, trigger and reconstruction efficiency which may only be calculated using a Monte Carlo program. The Monte Carlo (MC) simulation plays a crucial role in this analysis. In addition to the acceptance and efficiencies for our signals and backgrounds, as described in Section 5.3 and Section 6, the MC is used for the optimization of signals in Section 3. MC is also used to find out the function form that describes the mass spectrum of the background due to partial- or mis-reconstruction in Section 4.2. In this section, we first explain the components of Monte Carlo samples and show that, in general, the MC reproduces the data. Then we present the acceptance, trigger and reconstruction efficiencies obtained from the MC.

5.1 Monte Carlo Simulation Components

There are several components in the MC simulation:

- production and decay of the B hadrons
- detector simulation
- trigger simulation

We use two types of event generators: `Bgenerator`and `PYTHIA`. `Bgenerator` is the primary generator used in this analysis for calculating the acceptance and efficiencies of our signals and most backgrounds. The `PYTHIA` provides more realistic simulation of an event than `Bgenerator`, and produces multi-particle final states like the hadron collider data. However the generation using the `PYTHIA` is also more time consuming than the `Bgenerator`. This makes `PYTHIA` inefficient to understand the acceptance and efficiency of a single decay mode. Therefore, `PYTHIA` has been used in this analysis only to study the background from $b\bar{b}$ and $c\bar{c}$ decays.

`Bgenerator` generates a single b-quark according to the $P_T(b)$ spectrum which follows the NLO calculation by Nason, Dawson, and Ellis (NDE) [6]. For the B meson MC sample, the b-quarks are generated with a P_T threshold of 4.0 GeV/c over the range in rapidity $|y| < 2.5$, and then fragmented into B mesons with the CDF default Peterson fragmentation parameter [7], ϵ_B, set to 0.006. Figure 36 shows a small discrepancy in the reconstructed $P_T(B^0)$ between data and MC. The slope of the data to MC ratio is about 2 σ away from zero. The MC events which survive the trigger simulation, reconstruction and the analysis cuts, will be re-weighted according to the ratio numerically, i.e. we multiply each event with the ratio, w. We then calculate the efficiencies using the re-weighted MC events;

$$R_{\rm pass} = \sum_{i}^{N_{\rm pass}} w_i$$

$$\epsilon = \frac{R_{\rm pass}}{N_{\rm gen}}$$

Figure 36 also shows a discrepancy in the reconstructed $P_T(\Lambda_b)$ between data and MC from the `Bgenerator`. In order to correctly assess the acceptance and efficiency of the Λ_b, we need to skip the fragmentation process inside `Bgenerator`. The Λ_b needs to be generated directly with a P_T spectrum which reproduces the data. This spectrum is obtained in the following way: We first obtain the default generated Λ_b P_T spectrum from the `Bgenerator`. Then, the default generated $P_T(\Lambda_b)$ is re-weighted with the exponential slope of the ratio data/MC shown in Figure 36, using the "acceptance-rejection (Von Neumann)" method [8]. See Figure 37 for the Λ_b P_T spectra before and after our reweighting. We also confirm that the reconstructed $P_T(\Lambda_b)$ from the MC using the re-weighted spectrum reproduces the data, see Figure 49.

After the event generation, the hadrons are allowed to decay using the `EvtGen`. This package is maintained by `BABAR` and mainly tuned by the results from the experiments at the $\Upsilon(4S)$ resonance. The decay model and branching ratios for B^0 and B^+ are well described but not necessarily those of the B_s and the B baryons. As a proper decay model for the semileptonic Λ_b decays is not implemented in the `EvtGen` yet, we use a phase space to decay the Λ_b first. Then, we will apply a scaling factor on the acceptance after taking into account the effect of the semileptonic form factors (see Section 5.3). The particles from the output of `Bgenerator` and `EvtGen` are then run through a full ("realistic") simulation of the CDF detector and trigger. The software version for the simulation is 4.11.2 with patches which implement the most up-to-date configuration of SVT. Most of the detector subsystems, like COT and CMU, are assumed to be in a time-independent and perfect condition, which means there are no dead channels and the high voltages are constantly at full value. Selecting the data when these systems are in good condition helps to ensure that MC reproduces the data. Because the SVX active coverage and the configuration for the XFT and SVT systems change on various occasions, we divide the data taking period into eight sub-periods, where the detector and trigger performance is constant. We generate our MC samples for these eight sub-periods by choosing the runs with maximum number of L3 triggered events as the representative runs. Each run has its own parameters for the performance of the detector and triggers. For the sub-periods with large integrated luminosity, we choose more representative runs so that each run corresponds to a period with integrated luminosity around 3–6 pb^{-1}. See Table 20 for the representative runs in the MC. The number of generated events is proportional to the integrated luminosity of the sub-period each run represents. The positions of the beamline for each run is taken directly from the database and simulated in the MC.

After the detector and trigger simulation, the MC events are run through a trigger decision program, `svtfilter`. `svtfilter` takes the information from the simulated SVT data and makes the `B_CHARM Scenario A` requirements. The events which pass `svtfilter` are processed with the same `Production` executable (version 4.9.1hpt3) as that which is run on the data. The `Production` executable reconstructs higher level objects, such as electrons, muons, tracks and missing energy, from the simulated detector and trigger data. The resulting MC events have the same structure and format as the data and are then run through the same analysis program described in Section 3.

5.2 Monte Carlo and Data Comparison

To confirm that the simulation accurately reproduces the data, we compare various reconstructed distributions from the MC with the same distribution from the data. To ensure a fair comparison, the combinatorial background present in the signal region of data has to be removed . We perform a sideband subtraction for the $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B} \to D^{*+}\mu^- X$, $\overline{B}^0 \to D^+\pi^-$ and $\overline{B} \to D^+\mu^- X$ decays. For the $\Lambda_b \to \Lambda_c^+\pi^-$ and $\overline{B} \to \Lambda_c^+\mu^- X$ decays, a sideband subtraction can not remove all the backgrounds in the signal region as explained later in the text and in Section 4.2. Instead, a signal distribution of variable "X" is obtained by fitting $M_{\Lambda_c\pi}$ and $M_{pK\pi}$ to get the number of signal events in bins of variable "X". For all the semileptonic modes, we include the MC samples of the physics backgrounds described in Section 6.1. The distribution from each physics background is scaled according to the assumed or measured branching ratio for that background. In addition, the distribution of each compared variable from the fake muons is subtracted from the data. The distribution from the fake muons is obtained by reconstructing the "fake muon-charm" final state as described in Section 6.2. The combinatorial background in the "fake muon-charm" is removed using the same method as described above for the real muon. See Figure 38 for the $M_{D\mu}$ from the muon fakes.

For the B meson semileptonic channels, the mass difference between D^{*+} and D^0 in the $\overline{B} \to D^{*+}\mu^- X$ mode ($M_{K\pi\pi}$ - $M_{K\pi}$), and mass of D^+ in the $\overline{B} \to D^+\mu^- X$ mode ($M_{K\pi\pi}$), are used as the variables to perform the sideband subtraction. The signal region for both these modes is defined as:

$$|M - M_{\rm PDG}| < 2\sigma, \tag{48}$$

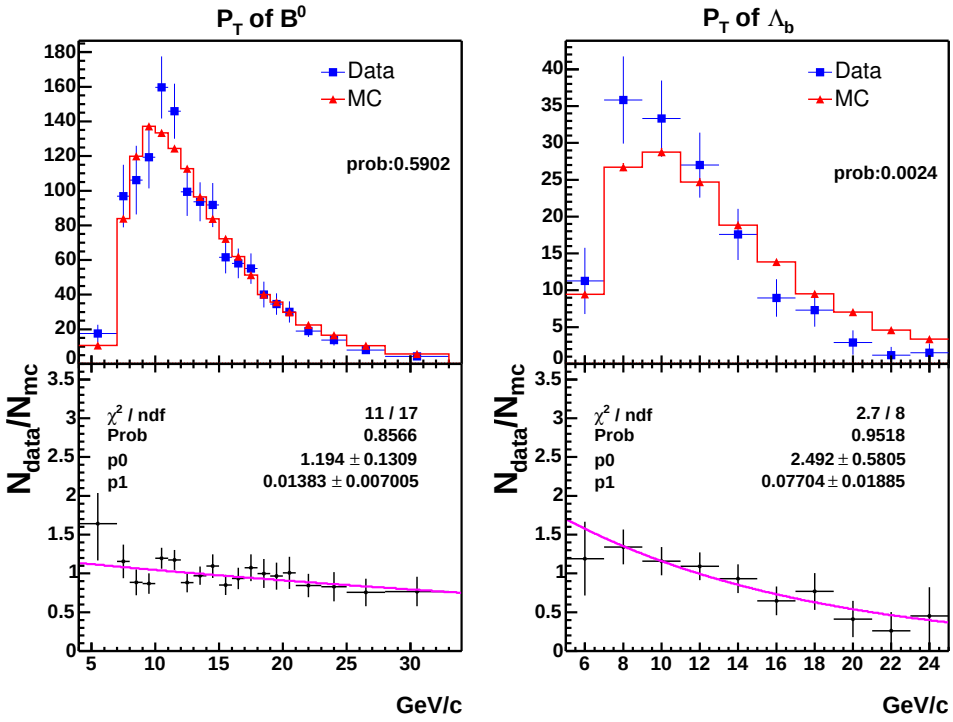


Figure 36: Comparison of reconstructed B^0 (left) and Λ_b (right) P_T spectrum between data and `Bgenerator` (MC). The top figures show the P_T distribution while the bottom figures show the ratio data/MC. The curves in the bottom figures are the result of an exponential fit to the ratio. It is evident that the MC P_T spectrum is harder than that of the data.

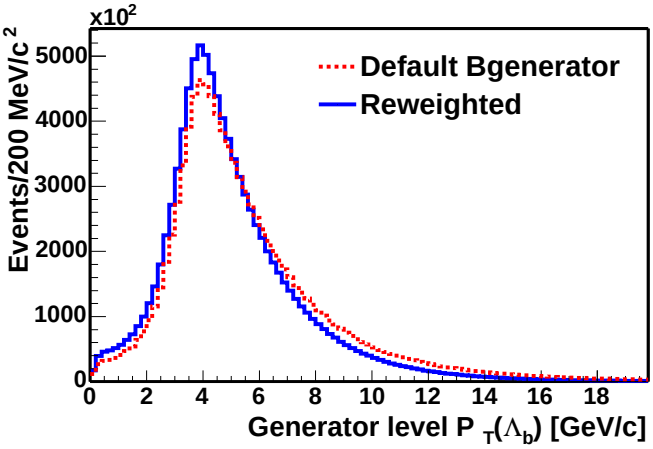


Figure 37: Generator level Λ_b P_T spectra before and after the reweighting.

Table 20: Simulated Runs in the MC sample.

Range	Run	$\int \mathcal{L}\, dt$ (pb^{-1})	Comment
138809–143000	140129	3.4	Scenario A implemented
143001–146000	145005	4.0	Tevatron incident
146001–149659	148824	4.2	SVX coverage improved
	149387	2.9	
149660–150009	149663	0.6	SVT optimization (coverage+patterns)
150010–152668	150820	4.1	$L_{xy} > 200$ μm cut added
	151844	3.7	
	152520	3.5	
152669–156487	152967	3.6	XFT from 2-miss to 1-miss
	153327	3.7	
	153447	3.7	
	153694	2.4	
	154452	4.2	
	154654	4.9	
	155364	4.3	
	155795	2.5	
	155895	3.6	
	156116	3.7	
	156484	2.6	
159603–164302	160230	3.7	data taken after the shutdown
	160441	3.4	
	160823	3.7	
	161029	3.8	
	161379	3.3	
	161678	3.9	
	162130	3.6	
	162393	3.6	
	162498	5.6	
	162631	5.7	
	162857	4.4	
	163064	3.7	
	163431	4.3	
164303–167715	164451	4.6	SVT change from 4/4 to 4/5
	164844	3.5	
	165121	2.9	
	165271	3.9	
	165412	3.6	
	166008	6.0	
	166063	2.9	
	166567	5.2	
	166662	5.3	
	167053	5.9	
	167186	2.2	
	167506	4.0	
	167551	2.7	
Total		170.9	

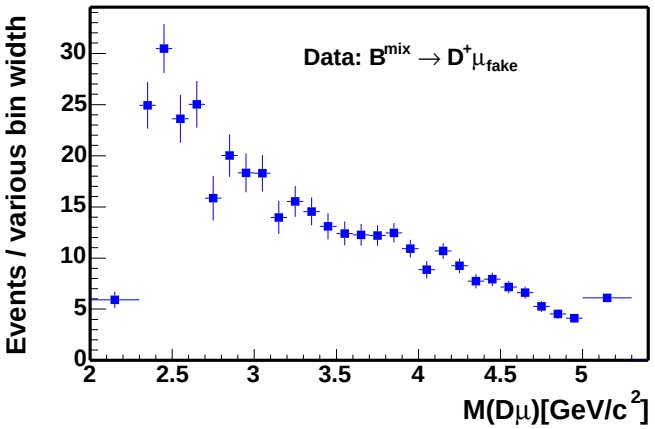


Figure 38: $M_{D\mu}$ from the $B \to D^+\mu_{fake}$ data. The distribution is sideband subtracted using $M_{K\pi\pi}$.

and the sideband region is defined as:

$$4\sigma < |M - M_{\rm PDG}| < 6\sigma. \tag{49}$$

The background function is assumed to be a straight line. Therefore, the amount of background in our signal region is the same as that in our sideband regions. We obtain a clean signal distribution by subtracting the histogram in the sideband region from the histogram in the signal region. Figure 39 displays the signal and sideband regions of $M_{K\pi\pi}$ - $M_{K\pi}$ and $M_{K\pi\pi}$.

For the B meson hadronic modes, we use the upper mass sideband above the signal peak to perform the sideband subtraction. The lower mass region below the signal peak consists of both combinatorial background and partially reconstructed B decays. However, the background in the signal region and in the upper mass region above the peak is mainly combinatorial as shown in Figures 29 – 32. We have learned in Section 4.2 that the combinatorial background is adequately described by an exponential function. Therefore, we fit the upper mass region to an exponential function. We further extrapolate the exponential to the signal region and obtain the ratio of the background in our signal region to that in our upper mass sideband, $R_{\rm bg}$. The histogram of the compared variable extracted from the upper mass sideband is scaled by $R_{\rm bg}$ and subtracted from the histogram in the signal region. See Figure 40 for the B^0 mass signal region we define and the upper mass region we fit to an exponential.

For the $\overline{B} \to \Lambda_c^+\mu^- X$ and $\Lambda_b \to \Lambda_c^+\pi^-$ modes, there are non-negligible backgrounds under the signal peak from the reflections due to a mis-assignment of the mass for one of the particles, see Section 4 for more details. This type of background has a different behavior from the combinatorial background in the sideband region. Since a background-free sideband subtraction is difficult to perform, we choose to fit the number of signal events in each bin of the variables which we want to compare. For the number of $\overline{B} \to \Lambda_c^+\mu^- X$ candidates, the $M_{pK\pi}$ distribution is fitted to a signal Gaussian and a second-order polynomial background as shown in Figure 41. For the number of $\Lambda_b \to \Lambda_c^+\pi^-$candidates, the $M_{pK\pi\pi}$ distribution is fitted to a simplified model: a Gaussian signal and an exponential background, as shown in Figure 42. Note that although the Λ_b fit model is simplified, the systematic uncertainty due to the naive model is no more than 3% of the number of signal events in each bin compared with the 15% statistical uncertainty. The widths of $M_{pK\pi}$ and $M_{K\pi}$ are fixed to the values obtained from the full statistics when doing the fit. Figure 43 shows the data and MC comparison using the fit values obtained from Figures 41– 42.

When comparing the MC and data distributions, if the number of data signal events in one bin is less than 20, that bin is combined with the next bin until the sum of the events is over 20. Then a χ^2

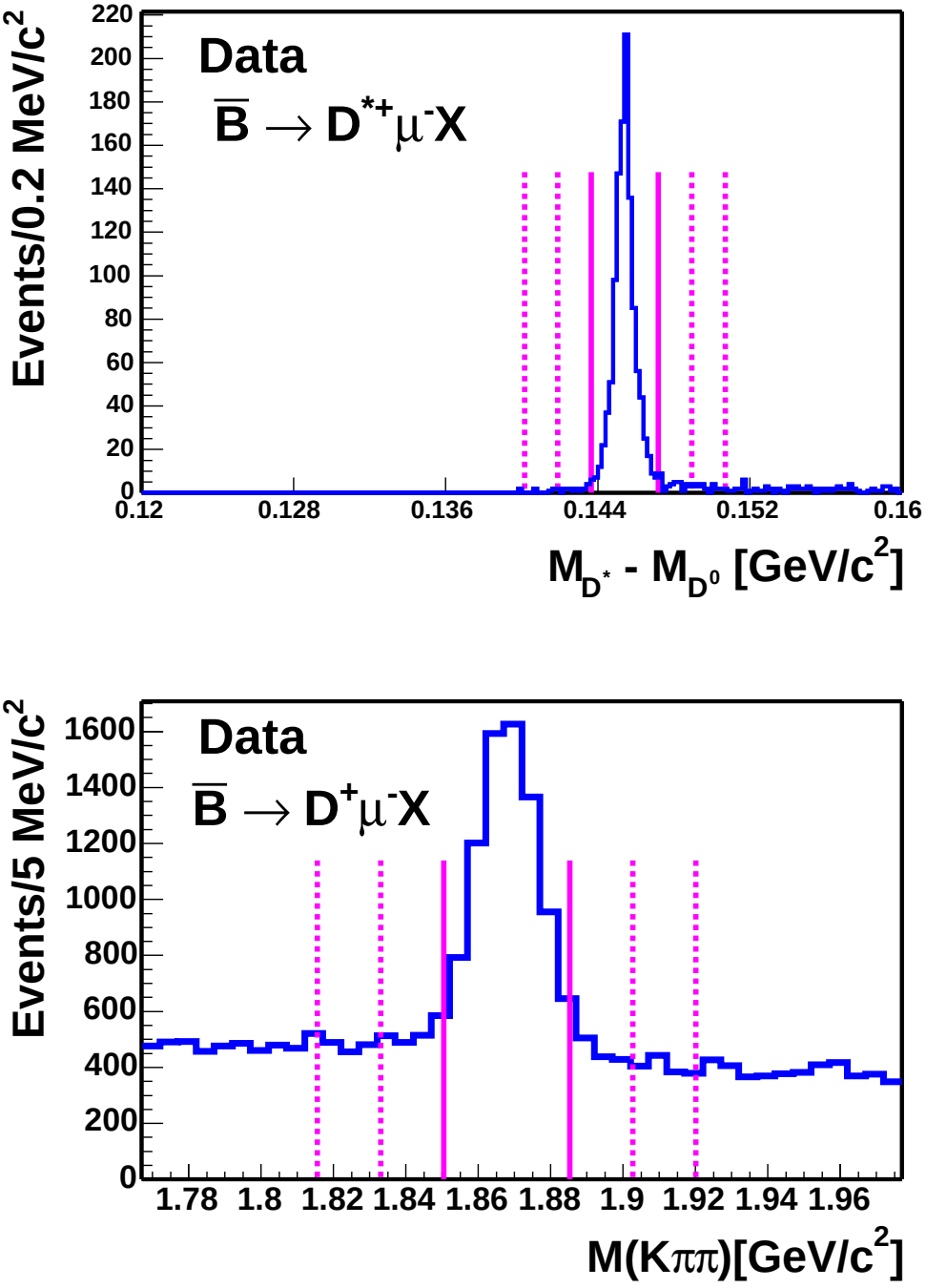
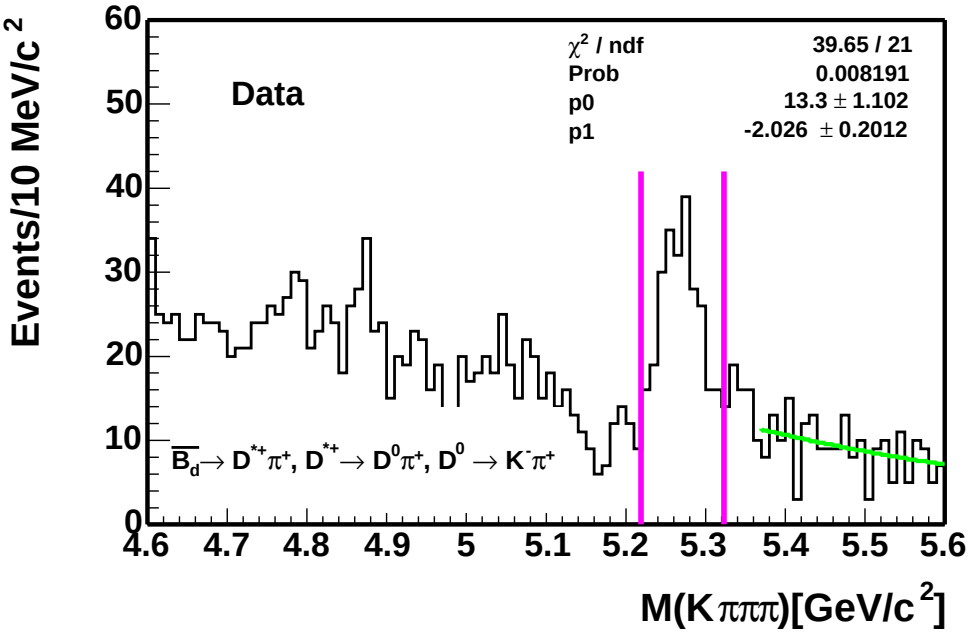


Figure 39: Invariant mass difference $M_{K\pi\pi} - M_{K\pi}$ (top) and invariant mass $M_{K\pi\pi}$ (bottom) showing our reconstructed $\overline{B} \to D^{*+}\mu^- X$ and $\overline{B} \to D^{+}\mu^- X$ signals. The vertical solid (dashed) lines indicate the signal (sideband) regions.



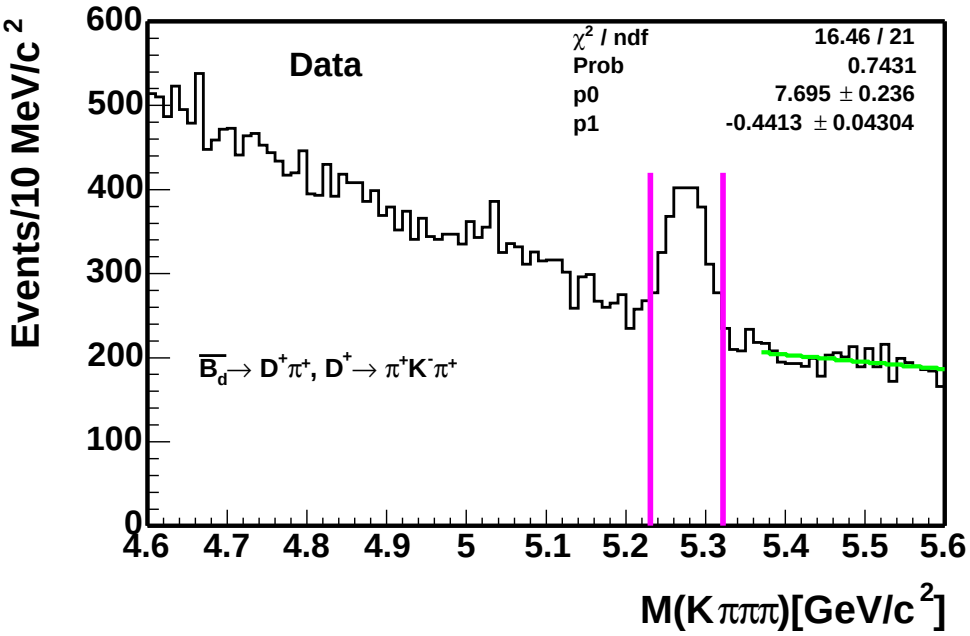


Figure 40: Invariant mass $M_{K\pi\pi\pi}$ for $\overline{B}^0 \to D^{*+}\pi^-$ (top) and $\overline{B}^0 \to D^+\pi^-$ (bottom) signals. The vertical solid lines indicate the signal region. The upper mass regions on the right are fitted to exponential functions as indicated by the green solid line.

is computed,

$$\chi^2 = \sum_i^n \frac{(N_{\rm MC}(i) - N_{data}(i))^2}{\sigma_{\rm MC}(i)^2 + \sigma_{data}(i)^2} \tag{50}$$

where i stands for i^{th} bin and total number of bins in a histogram is n. The number of degree freedom is $n-1$. For the $\overline{B} \to \Lambda_c^+ \mu^- X$ and $\Lambda_b \to \Lambda_c^+ \pi^-$ modes, a χ^2 is also calculated except that the bin width of each variable is fixed in this case. Besides the χ^2 test, we also plot the ratio data/MC. We fit the ratio to a first-order polynomial and check if the slope, $\mathcal{M}$, is consistent with zero. In the first pass, we find discrepancies in the P_T spectra of B^0 and Λ_b between MC and data (see Figure 36). As the semileptonic modes are three-body decays and two-body decays for the hadronic modes, the efficiency of the trigger and analysis P_T cut depends strongly on the P_T of B hadron (see Figure 44). We decide to reweight the P_T spectra of B^0 and Λ_b as described in Section 5.1. Figures 45– 50 show the comparison between MC and data for the analysis cut variables. Figures 77– 94 in Appendix C shows other distributions important for this analysis. Figure 51 shows the comparison for the $M_{D^*\mu}$, $M_{D\mu}$, and the $M_{\Lambda_c\mu}$ from the phase space MC before and after multiplying each bin entry with a scaling factor. The scaling factor is obtained by dividing the $M_{\Lambda_c\mu}$ distribution from the form factor weighted (see Section 5.3) by that from the phase space generator-level MC. In general, the MC describes the data well except for the pseudo $c\tau$ of $\Lambda_c\mu$, and the $\chi^2_{r-\phi}$ of the B and charm vertex fits. For the disagreement in the vertex fit $\chi^2_{r-\phi}$, as it is beyond the scope of this analysis to scale the measurement errors in the MC, we choose to make a loose cuts on the data $\chi^2_{r-\phi}$. In Section 7.4, we perform a cross-check of the relative branching ratio variation by dividing the data into two subsets, according to the cuts on the pseudo $c\tau$ of $\Lambda_c\mu$, $\chi^2_{r-\phi}$ and other variables. We do not see any significant inconsistency. Therefore, we do not assign any systematic uncertainties.

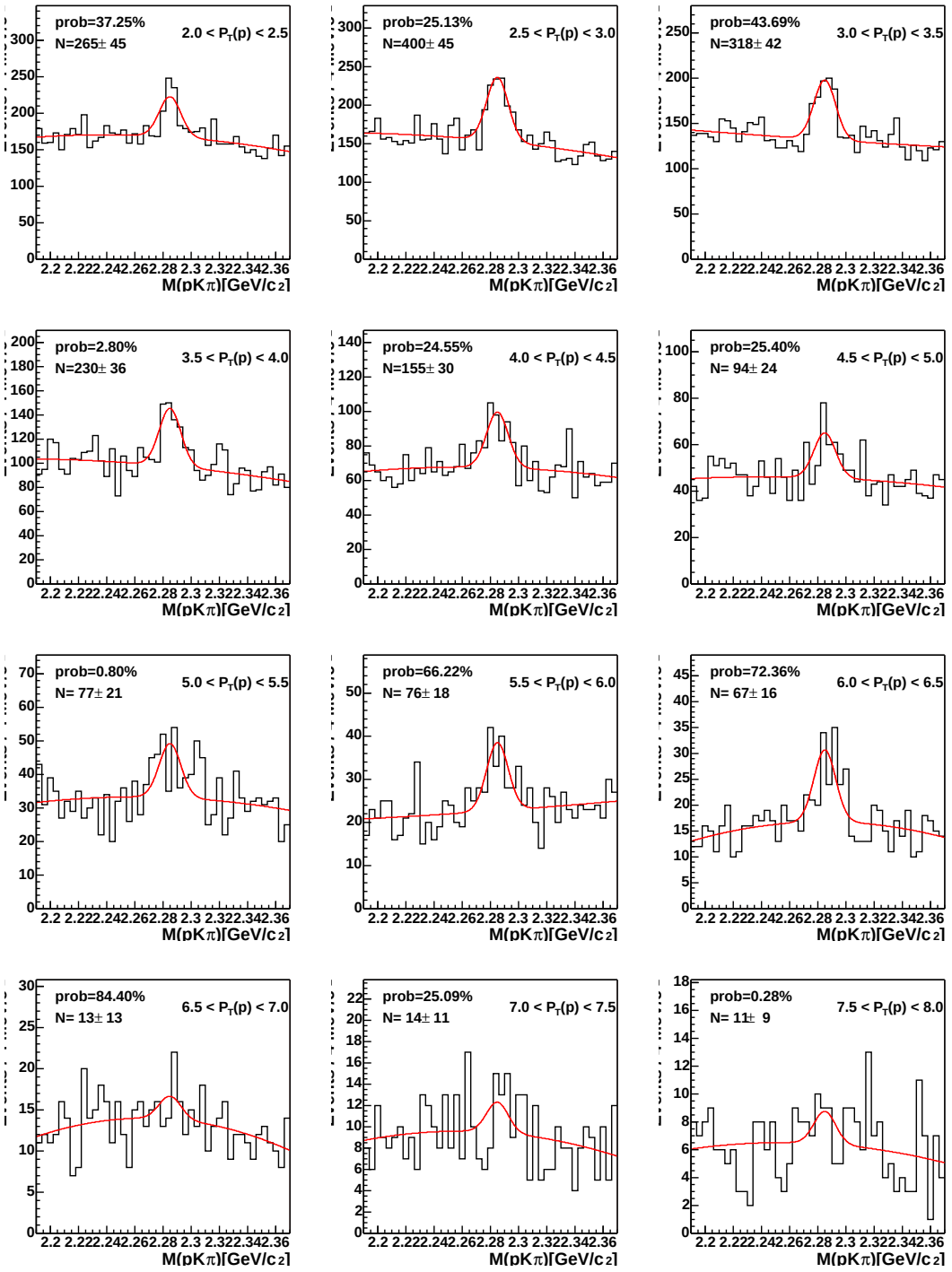


Figure 41: Example of Λ_c^+ mass fit for the MC and data comparison. The variable to compare is the P_T of proton, from 2 to 8 GeV/c, in bins of 0.5 GeV/c. $M_{pK\pi}$ is fitted to a signal Gaussian and a second-order polynomial.

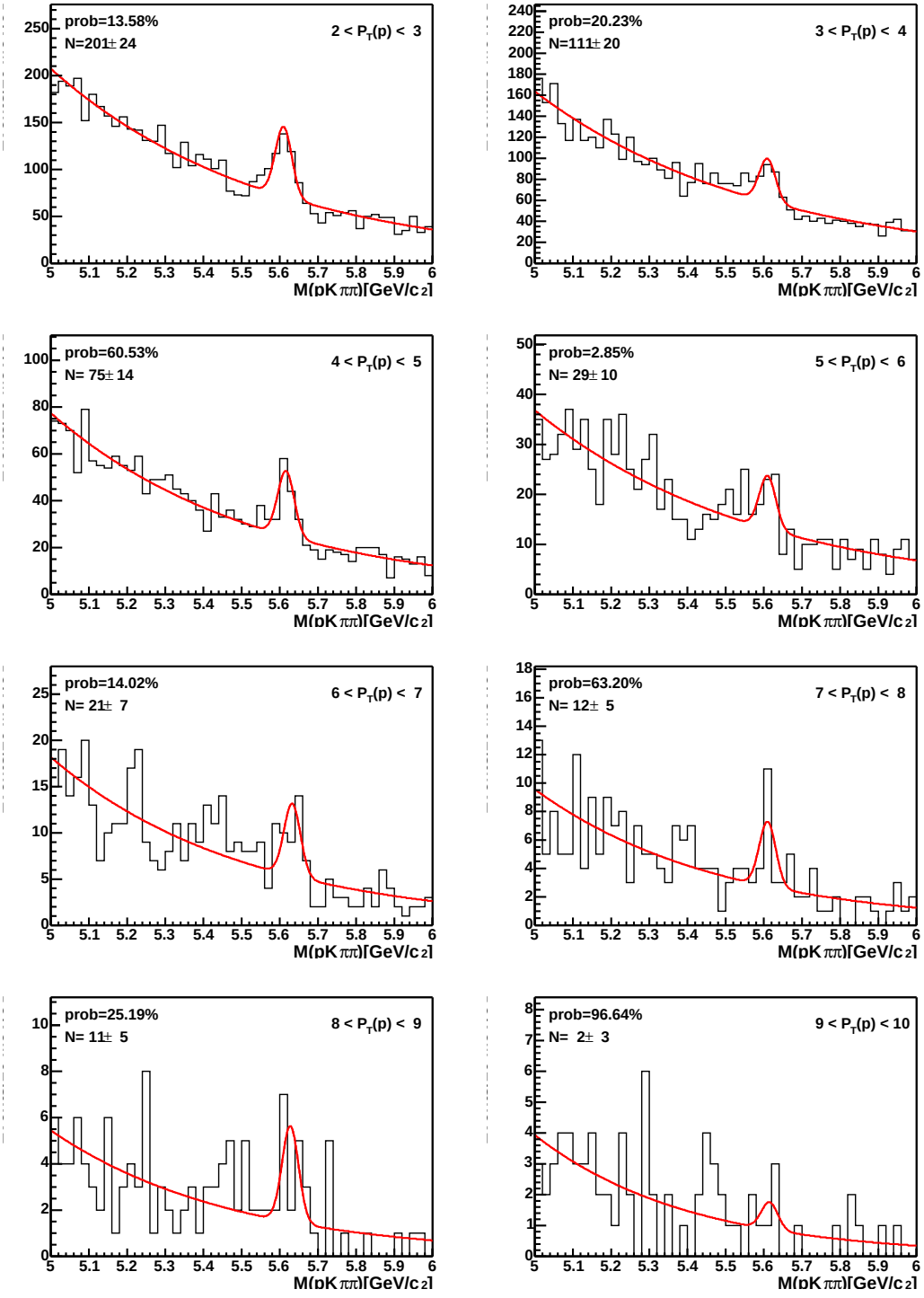


Figure 42: Example of Λ_b mass fit for the MC and data comparison. The variable to compare is the P_T of proton, from 2 to 10 GeV/c, in bins of 1 GeV/c. $M_{pK\pi\pi}$ is fitted to a signal Gaussian and an exponential background.

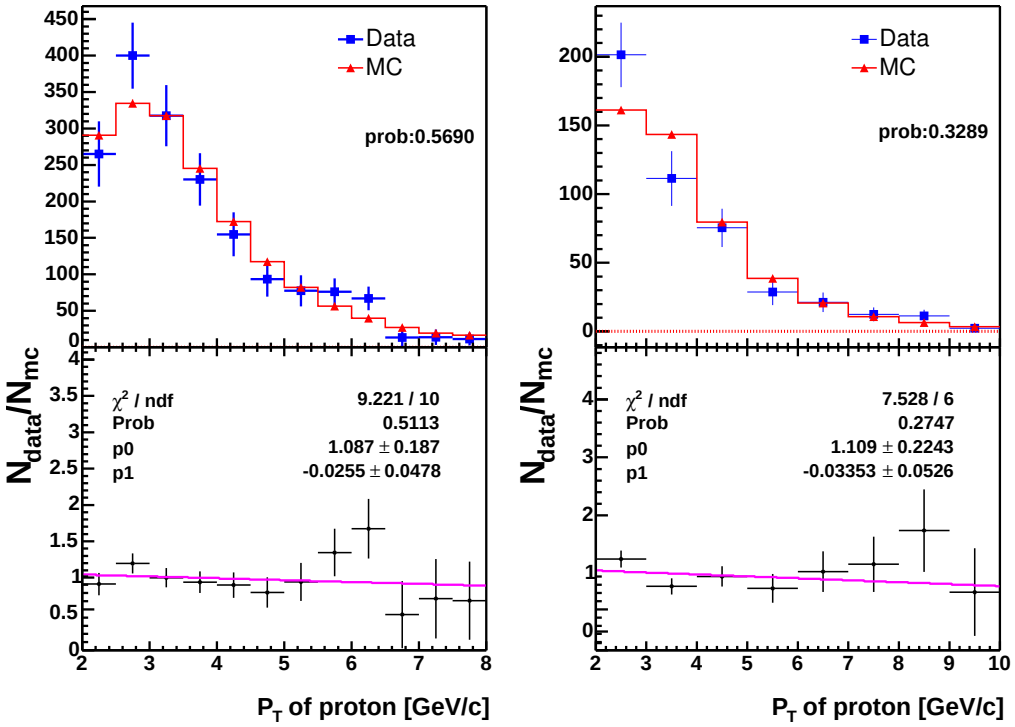


Figure 43: MC and data comparison of P_T(proton). The data points come from the fit to $M_{pK\pi}$ in Figure 41 (left) and the fit to $M_{pK\pi\pi}$ in Figure 42 (right).

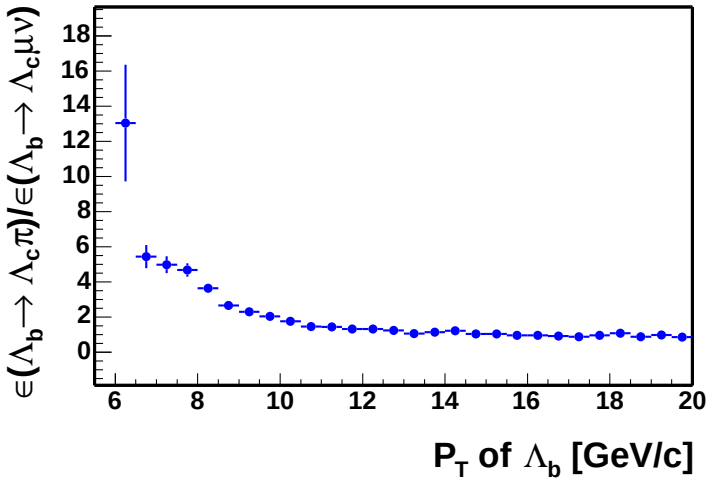


Figure 44: MC Efficiency ratio of $\Lambda_b \to \Lambda_c^+\pi^-$ to $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ as a function of the transverse momentum of Λ_b.

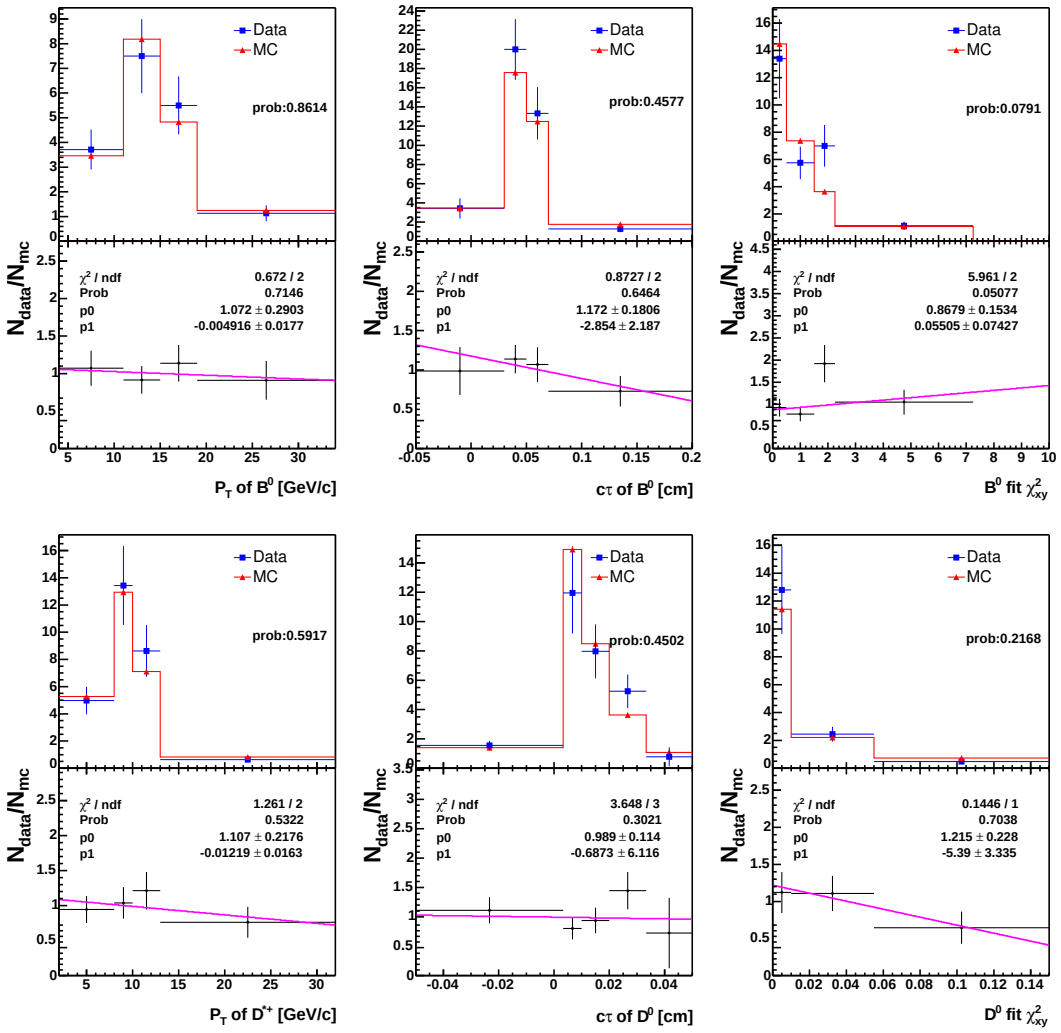


Figure 45: $\overline{B}^0 \to D^{*+}\pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(B^0)$, $c\tau(B^0)$, vertex fit $\chi^2_{r-\phi}$ for the B^0 vertex, $P_T(D^{*+})$, $c\tau(D^0)$, and vertex fit $\chi^2_{r-\phi}$ for the D^0 vertex.

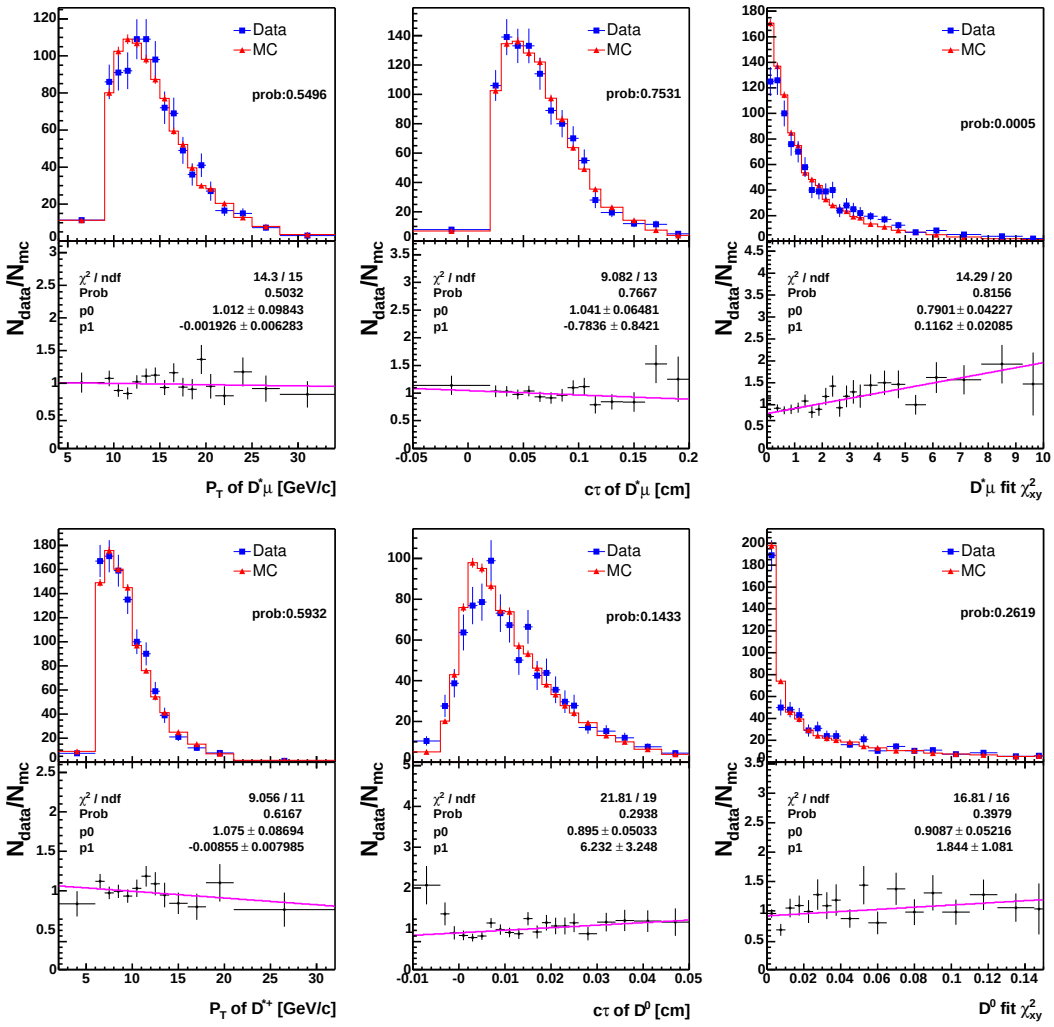


Figure 46: $\overline{B} \to D^{*+}\mu^- X$ MC and data comparison: from the top left to the bottom right are: $P_T(D^*\mu)$, $c\tau(D^*\mu)$, vertex fit $\chi^2_{r-\phi}$ for the $D^*\mu$ vertex, $P_T(D^{*+})$, $c\tau(D^0)$, and vertex fit $\chi^2_{r-\phi}$ for the D^0 vertex.

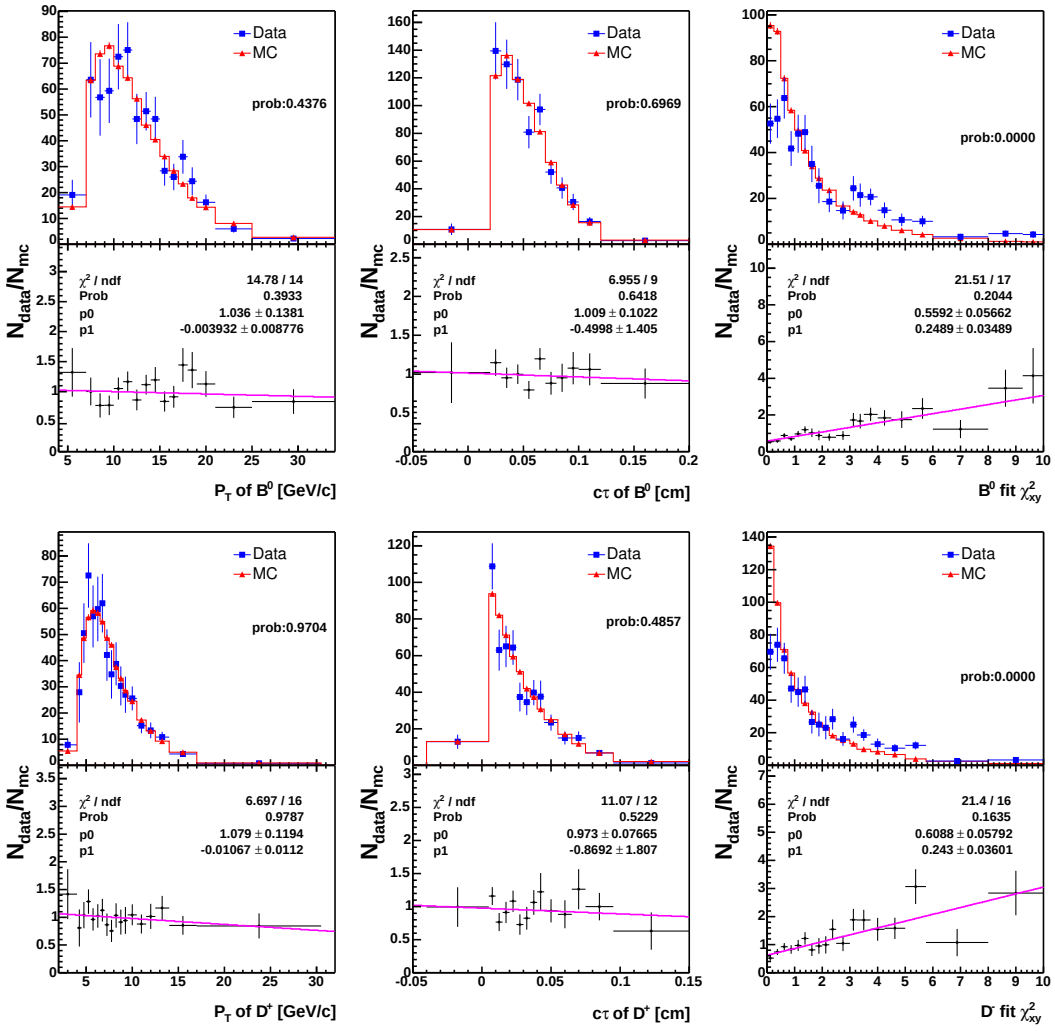


Figure 47: $\overline{B}^0 \to D^+\pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(B^0)$, $c\tau(B^0)$, vertex fit $\chi^2_{r-\phi}$ for the B^0 vertex, $P_T(D^+)$, $c\tau(D^+)$, and vertex fit $\chi^2_{r-\phi}$ for the D^+ vertex.

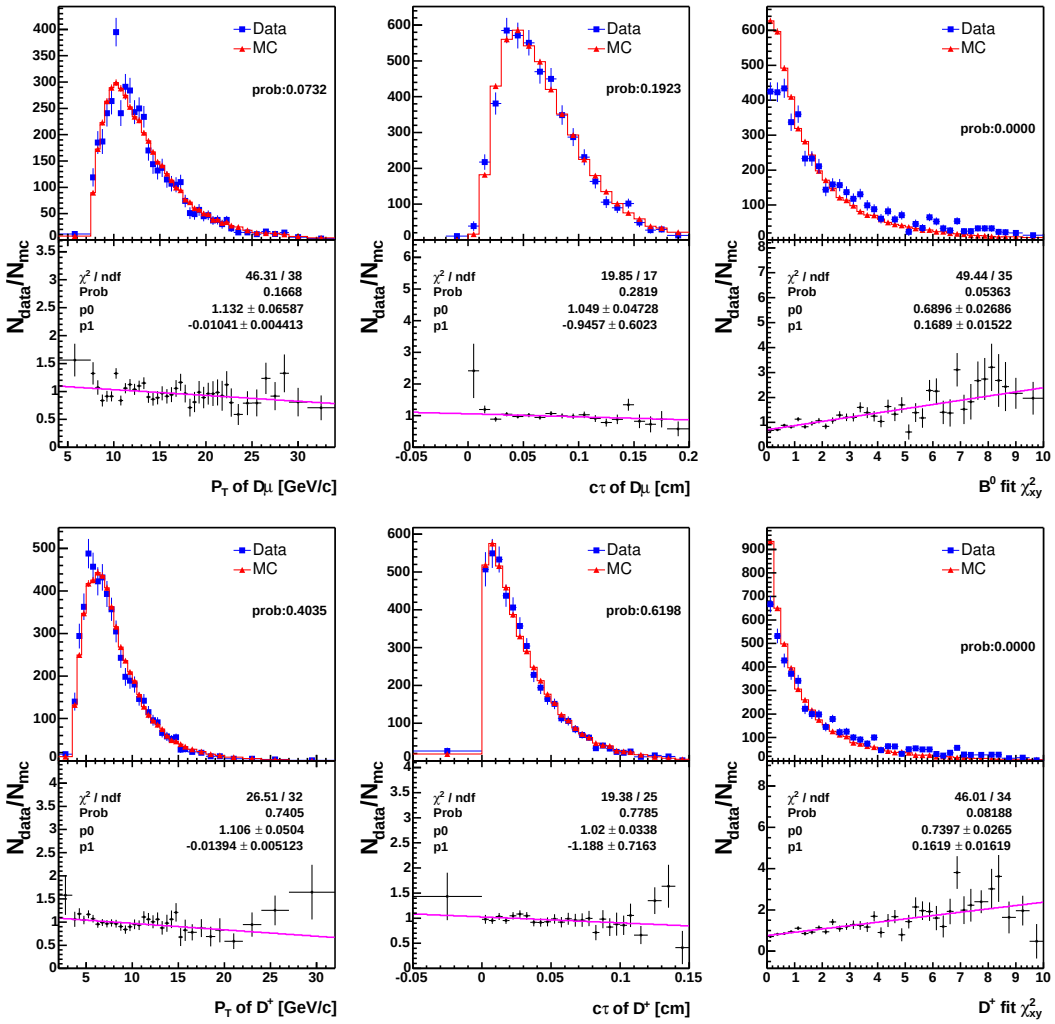


Figure 48: $\overline{B} \to D^+\mu^-X$ MC and data comparison: from the top left to the bottom right are: $P_T(D\mu)$, $c\tau(D\mu)$, vertex fit $\chi^2_{r-\phi}$ for the $D\mu$ vertex, $P_T(D^+)$, $c\tau(D^+)$, and vertex fit $\chi^2_{r-\phi}$ for the D^+ vertex.

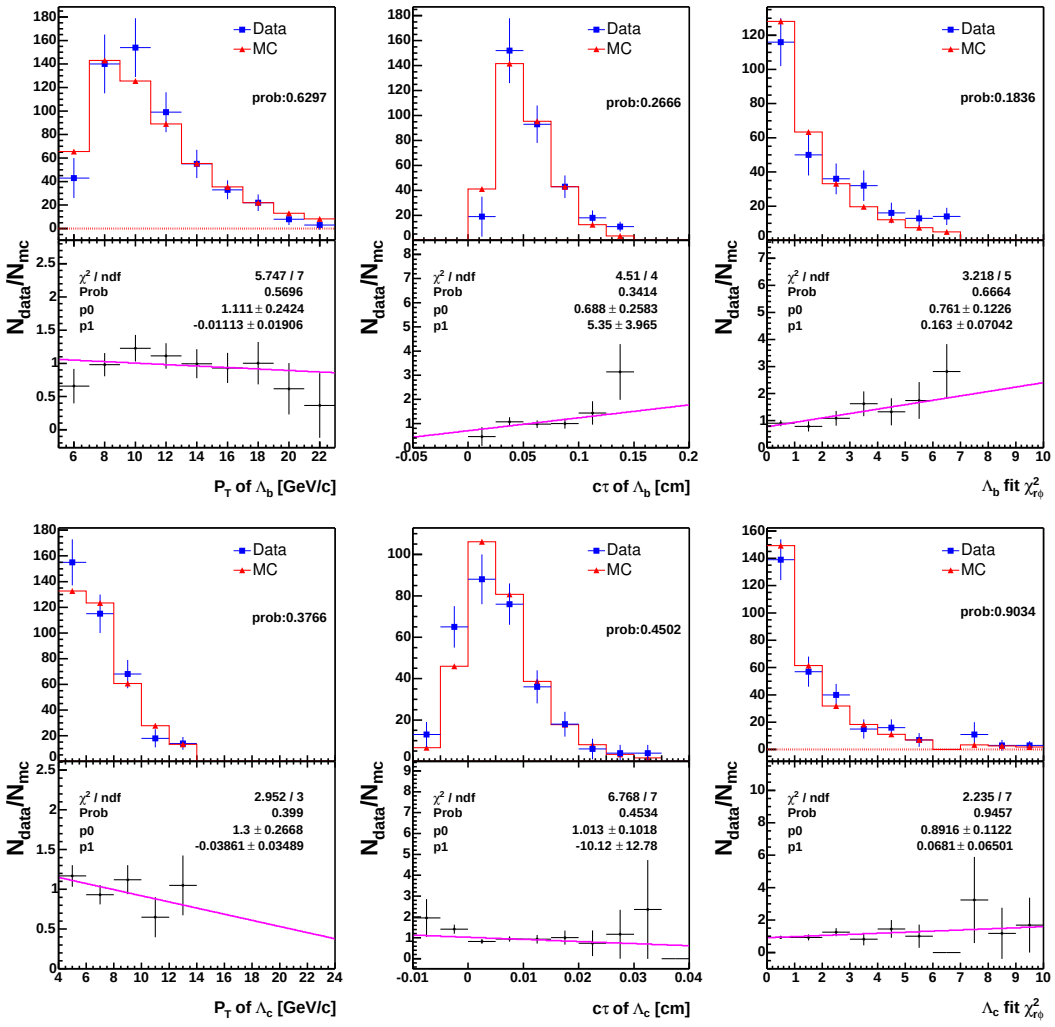


Figure 49: $\Lambda_b \to \Lambda_c^+ \pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(\Lambda_b)$, $c\tau(\Lambda_b)$, vertex fit $\chi^2_{r-\phi}$ for the Λ_b vertex, $P_T(\Lambda_c^+)$, $c\tau(\Lambda_c^+)$, and vertex fit $\chi^2_{r-\phi}$ for the Λ_c^+ vertex.

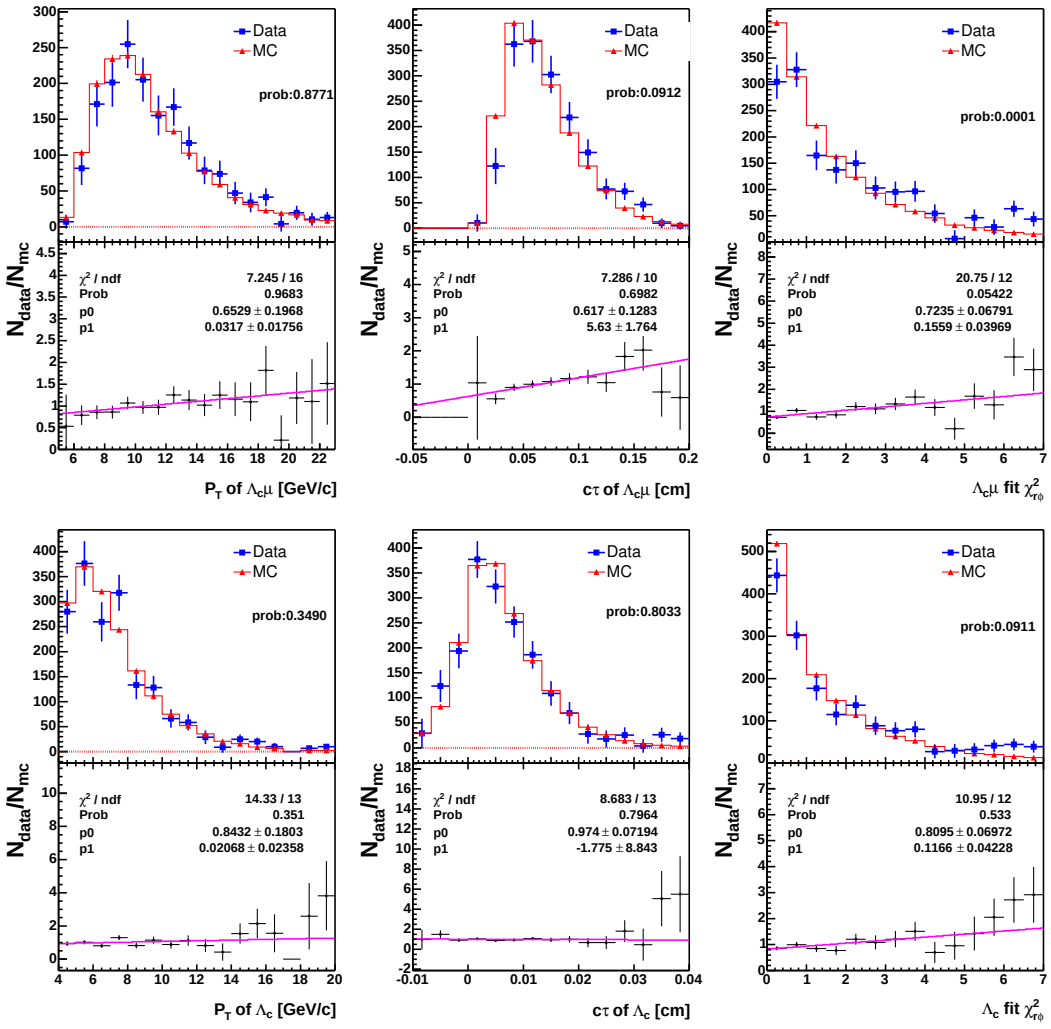


Figure 50: $\overline{B} \to \Lambda_c^+ \mu^- X$ MC and data comparison: from the top left to the bottom right are: $P_T(\Lambda_c \mu)$, $c\tau(\Lambda_c \mu)$, vertex fit $\chi^2_{r-\phi}$ for the $\Lambda_c \mu$ vertex, $P_T(\Lambda_c^+)$, $c\tau(\Lambda_c^+)$, and vertex fit $\chi^2_{r-\phi}$ for the Λ_c^+ vertex.

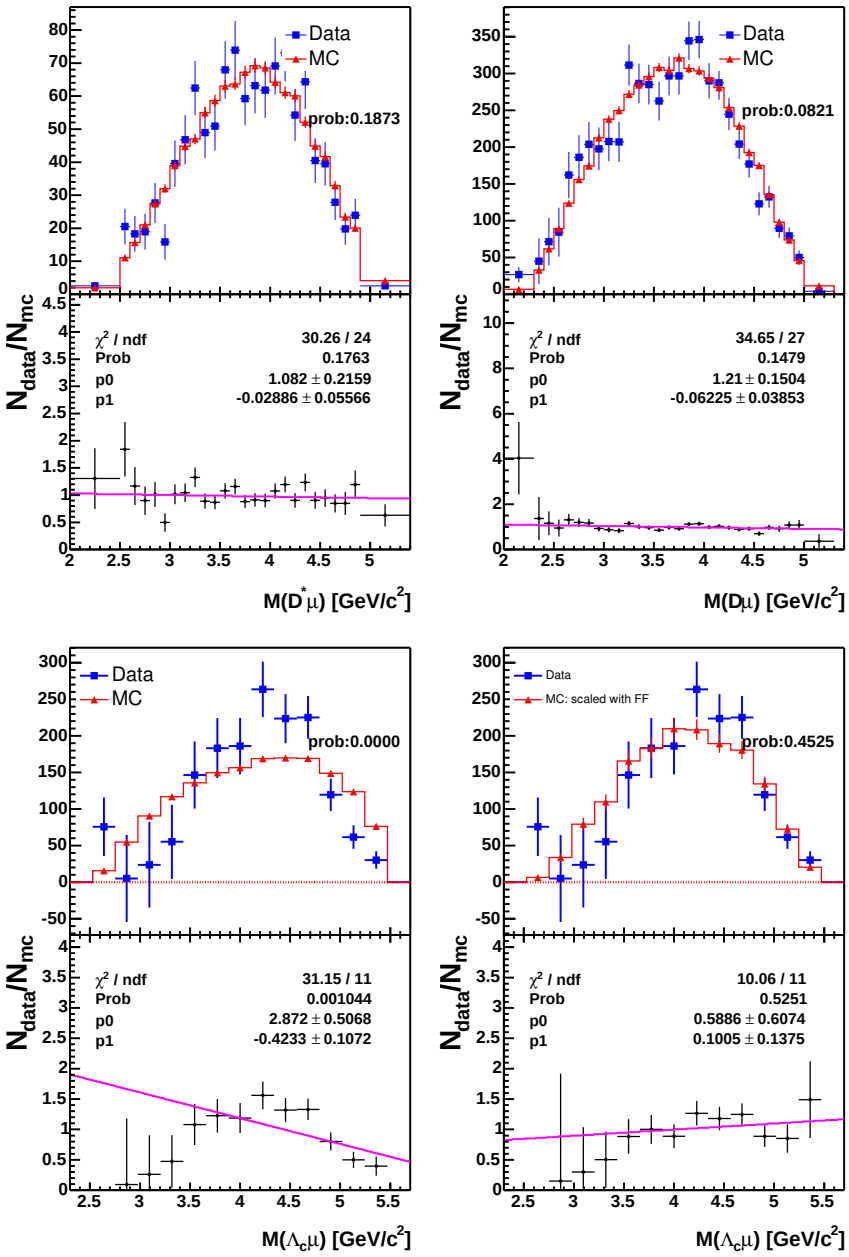


Figure 51: four track invariant mass MC and data comparison: from the top left to the bottom right are: $M_{D^*\mu}$, $M_{D\mu}$, $M_{\Lambda_c\mu}$ (phase space MC without scaling), $M_{\Lambda_c\mu}$ (phase space MC after scaling).

5.3 Acceptance, Trigger and Reconstruction Efficiencies of Signal

We obtain the product of the acceptance, trigger and reconstruction efficiencies using the MC described in Section 5.1. The total efficiency is defined as: the number of reconstructed events surviving the trigger simulation and analysis cuts, divided by the total number of events generated. Efficiencies for the backgrounds are found in Section 6. Our data could be divided into eight sub-periods under different trigger and hardware configurations, see Table 20. In this analysis, because the final states are nearly alike, we expect that the ratio of the efficiencies to be independent of the detector, trigger and calibration effects. To confirm this, we divide our signal MC samples into eight sub-periods and calculate their efficiencies and the ratio of hadronic to semileptonic modes. Tables 21– 23 list the efficiencies for our signals. It is shown as an example in Figure 53 that the efficiency of $\overline{B}^0 \to D^+\pi^-$ varies dramatically in each period. However the efficiency ratio of the hadronic to semileptonic mode is quite stable, as shown in Figures 54– 56.

Note that because Bloom and Dagenhart [9] find a difference in the CMU muon reconstruction efficiency between MC and data, we apply a scaling factor on the efficiencies of the semileptonic signals and backgrounds: $R = 0.986 \pm 0.003$. The uncertainty on R is used to estimate the systematic uncertainty on the relative branching fractions. In addition, Giagu, Herndon and Rescigno [10] [11] notice that there are differences in the XFT efficiencies for the charged kaons, pions, and protons, when the XFT configuration is switched to the "1-miss" mode, i.e. when the tracking algorithm in the XFT requires at least 11 hits from each COT axial superlayer. The COT frond-end electronics requires a minimum input charge from the ionization of the incident particle. At a fixed momentum, protons and kaons deposit less charge than the pions, have more hits below the electronics threshold, and fail the stringent XFT "1-miss" requirement. Therefore, in general, the proton and kaon XFT efficiencies are lower than that of the pion.

Figure 57(top) shows that kaon and pion XFT efficiencies are identical in the MC and need to be corrected. Giagu, Herndon and Rescigno measure the ratio, data/MC for the XFT efficiencies of pions, kaons and protons, as shown in Figure 57. We reweight the MC events using the "acceptance-rejection (Von Neumann)" method [8] according to the ratio:

$$\begin{aligned} C_\pi &= 1.002 - \frac{0.067}{P_T}, \\ C_K &= 0.969 - \frac{0.094}{P_T}, \\ C_p &= 1.06 - \frac{1.3}{P_T} + \frac{3.2}{P_T^2} - \frac{2.2}{P_T^3}, \end{aligned}$$

where P_T is the transverse momentum of the track that passes the trigger cuts in our reconstruction program.

Finally, one additional scaling factor has to be applied on all the Λ_b decays with $\Lambda_c\mu$ in the final state. We have mentioned in Section 5.1 that a phase space decay model was used for these decays. In a phase space, the event density in the w-$\cos\theta$ plane is a constant within the kinematic boundary. The w is the scalar product of the Λ_b and Λ_c four-velocities, and θ is the angle between the muon and the neutrino momentum vectors in the Λ_b rest frame. The form factors that describe the strong interaction in the Λ_b semileptonic decay modify the event distribution in the phase space and change the fraction of events accepted. We obtain the scaling factor in the following way: Using the "acceptance-rejection" method, we reweight the generator-level $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ MC according to:

$$f_c = \frac{d\Gamma(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}{dw} \cdot \frac{T(\cos\theta, w)}{P(w)}, \tag{51}$$

where the differential semileptonic decay rate, $\frac{d\Gamma(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}{dw}$, is obtained from Huang [12]. The

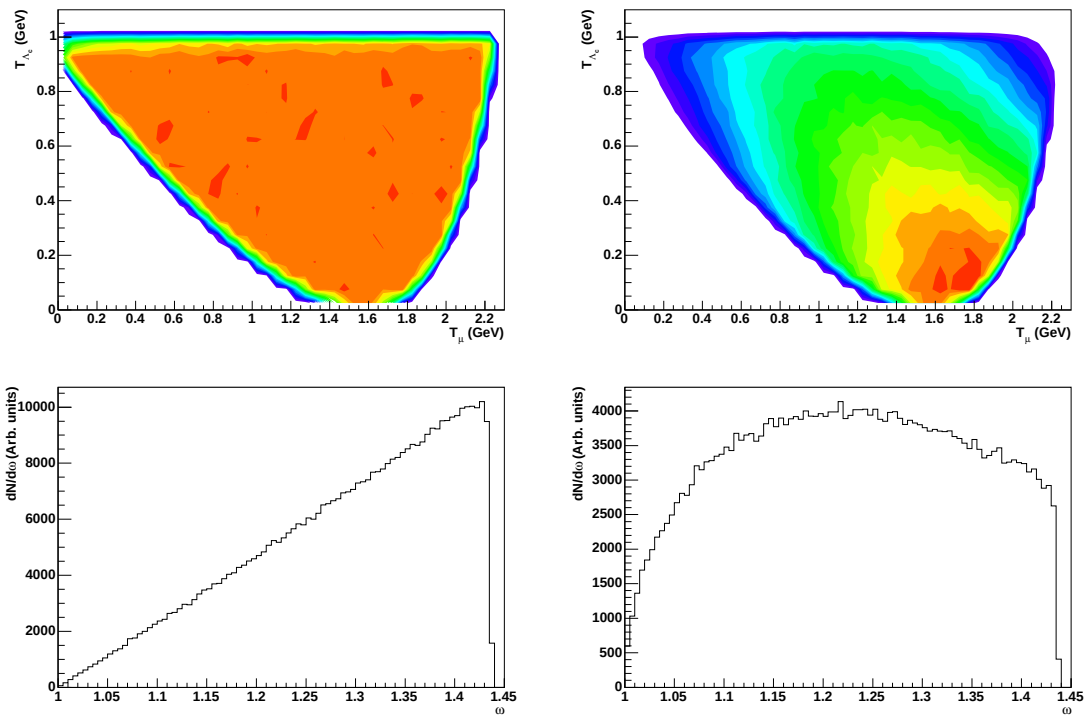


Figure 52: The phase space (top) and the w (bottom) distribution obtained from the $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$ phase space MC, before (left) and after (right) reweighting the events according to the semileptonic form factors from Huang [12].

$T(\cos\theta, w)$ includes the W spin effect and describes the correlation between the μ and $\overline{\nu}_\mu$, and

$$P(w) = \int_{\cos\theta_{\rm min}(w)}^{\cos\theta_{\rm max}(w)} T(\cos\theta, w) d\cos\theta. \tag{52}$$

Here, $\cos\theta_{\rm max}$ and $\cos\theta_{\rm min}$ specify the kinematic range and are functions of w. Figure 52 shows the phase space and the w distribution from the phase space and the form factor reweighted MC. Then, we apply generator-level analysis cuts to obtain the acceptance. We further divide this acceptance by that from the phase space MC and obtain a scaling factor of 0.994±0.025, where the uncertainty is dominated by the size of the MC sample.

5.4 Summary

We have described the procedure of generating MC samples and compared the MC distributions with those in the data. In general, the MC and data are in good agreement. For the variables which MC does not reproduce the data, we will study the systematic uncertainties from the disagreement. We also obtain the signal efficiencies from the MC. It is confirmed that the efficiency ratios from both the Λ_b and B^0 modes are insensitive to the time variation of beam lines and SVT trigger configurations.

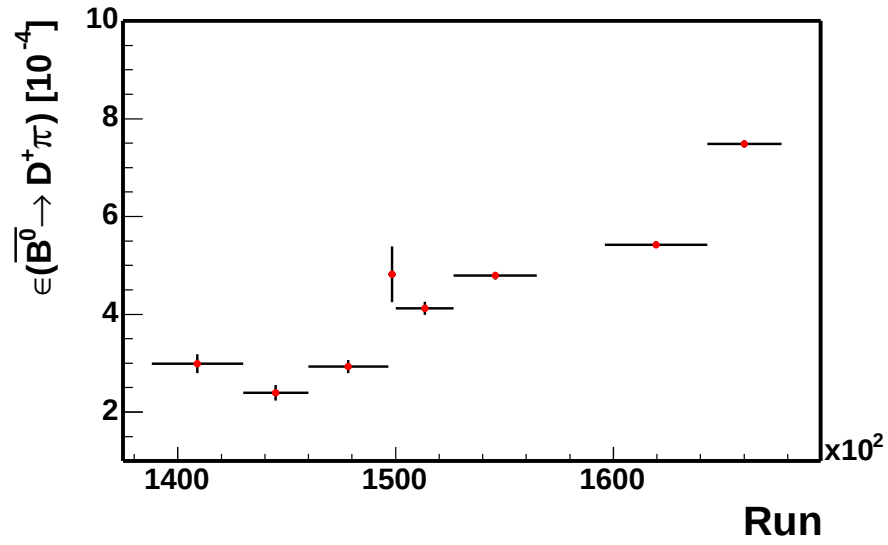


Figure 53: $\overline{B}^0 \to D^+\pi^-$ MC total efficiency as a function of run number

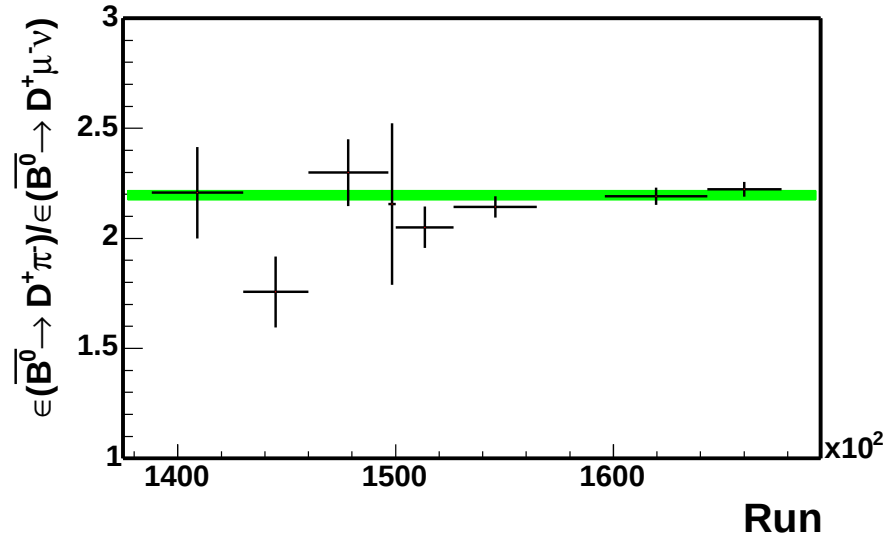


Figure 54: Total efficiency ratio of $\overline{B}^0 \to D^+\pi^-$ to $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ MC in eight different hardware configurations. The shaded area represents the total average efficiency ratio including the uncertainty.

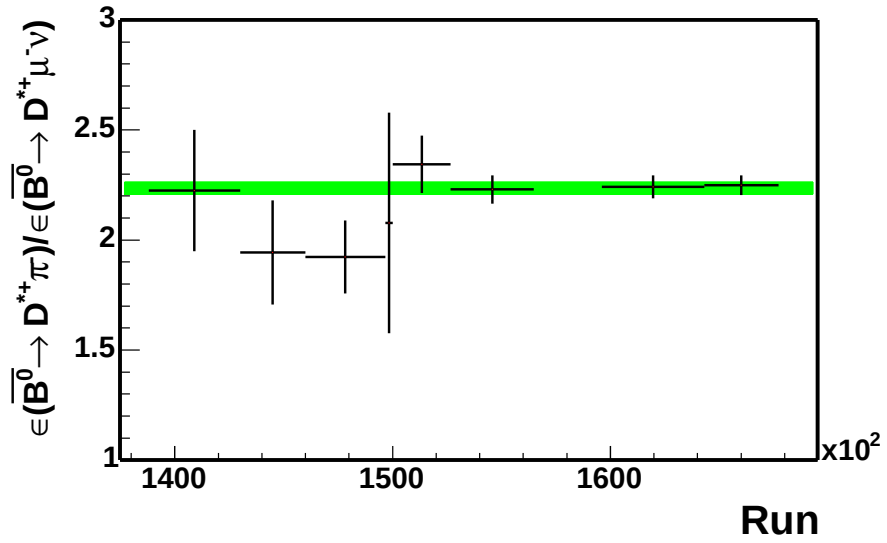


Figure 55: Total efficiency ratio of $\overline{B}^0 \to D^{*+}\pi^-$ to $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ MC in eight different hardware configurations. The shaded area represents the total average efficiency ratio including the uncertainty.

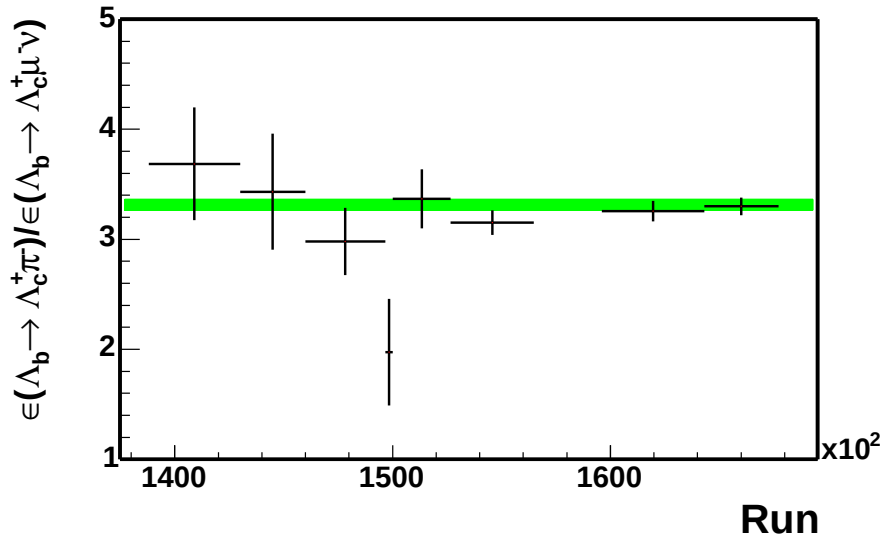


Figure 56: Total efficiency ratio of $\Lambda_b \to \Lambda_c^+\pi^-$ to $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ MC in eight different hardware configurations. The shaded area represents the total average efficiency ratio including the uncertainty.

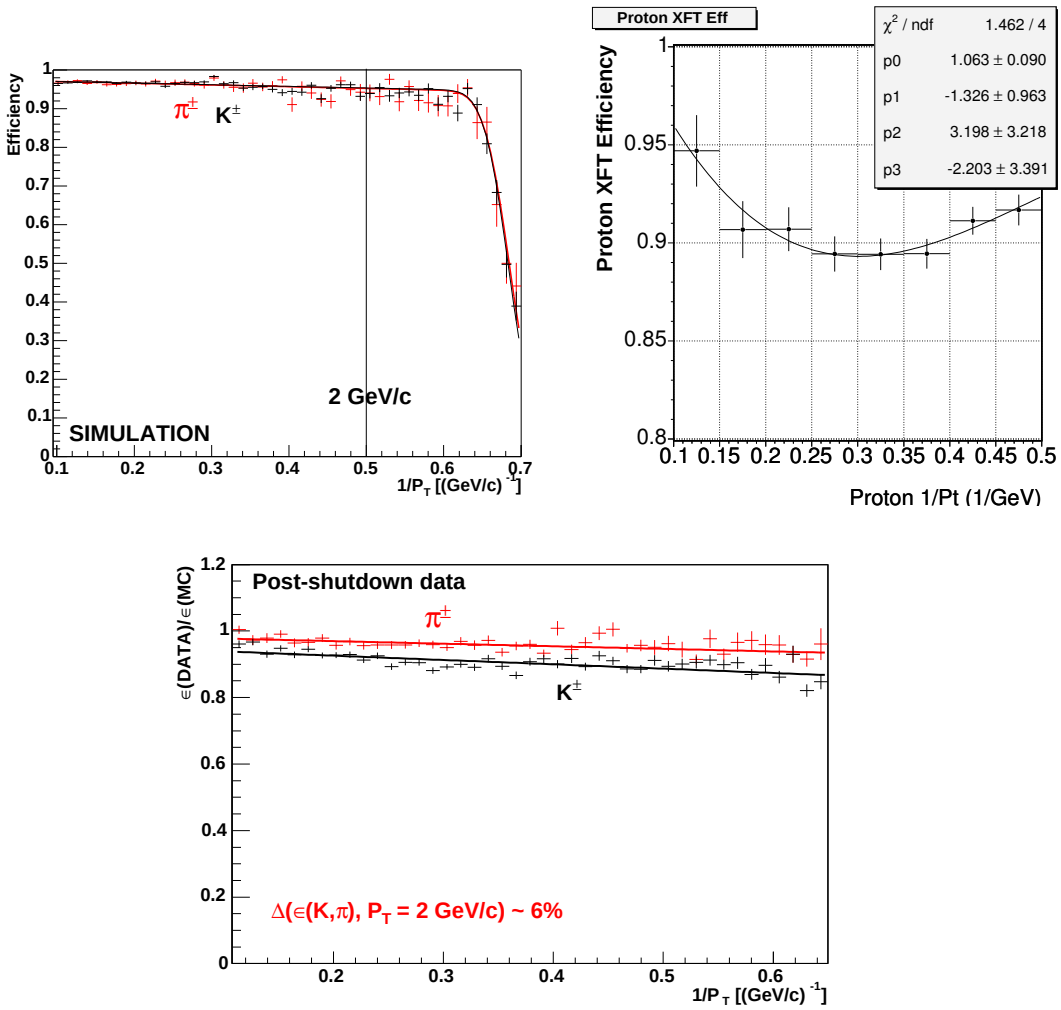


Figure 57: Giagu and Rescigno [10] find that K and π XFT efficiencies in the MC (top left) are identical. The relative proton (top right) and K, π (bottom) XFT efficiencies between MC and data, in bins of $1/P_T$, are fit to a third order polynomial by Herndon [11], and to a first order polynomial by Giagu and Rescigno [10], respectively.

Table 21: Total efficiency and ratios for $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$.

Run Range	$\int \mathcal{L}\, dt$ (pb^{-1})	$\epsilon_{\overline{B}^0 \to D^{*+}\pi^-}$ (10^{-4})	$\epsilon_{\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu}$ (10^{-4})	ϵ Ratio
138809–143000	3.4	1.72 ± 0.15	0.77 ± 0.07	2.22 ± 0.28
143001–146000	4.0	1.42 ± 0.12	0.73 ± 0.06	1.94 ± 0.24
146001–149659	7.1	1.57 ± 0.10	0.82 ± 0.05	1.92 ± 0.17
149660–150009	0.6	2.35 ± 0.40	1.13 ± 0.19	2.08 ± 0.50
150010–152668	11.3	2.65 ± 0.10	1.13 ± 0.05	2.34 ± 0.13
152669–156487	39.2	2.74 ± 0.06	1.23 ± 0.03	2.23 ± 0.07
159603–164302	52.7	3.10 ± 0.05	1.38 ± 0.02	2.24 ± 0.05
164303–167715	52.7	4.27 ± 0.06	1.90 ± 0.03	2.25 ± 0.05
Total average	171.0	3.22 ± 0.03	1.44 ± 0.01	2.24 ± 0.03

Table 22: Total efficiency and ratios for $\overline{B}^0 \to D^{+}\pi^-$ and $\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu$.

Run Range	$\int \mathcal{L}\, dt$ (pb^{-1})	$\epsilon_{\overline{B}^0 \to D^{+}\pi^-}$ (10^{-4})	$\epsilon_{\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu}$ (10^{-4})	ϵ Ratio
138809–143000	3.4	2.99 ± 0.19	1.35 ± 0.09	2.21 ± 0.21
143001–146000	4.0	2.39 ± 0.16	1.36 ± 0.09	1.76 ± 0.16
146001–149659	7.1	2.93 ± 0.13	1.28 ± 0.06	2.30 ± 0.15
149660–150009	0.6	4.82 ± 0.57	2.23 ± 0.27	2.16 ± 0.37
150010–152668	11.3	4.12 ± 0.14	2.01 ± 0.06	2.05 ± 0.09
152669–156487	39.2	4.79 ± 0.08	2.24 ± 0.04	2.14 ± 0.05
159603–164302	52.7	5.43 ± 0.07	2.48 ± 0.03	2.19 ± 0.04
164303–167715	52.7	7.49 ± 0.08	3.37 ± 0.04	2.22 ± 0.03
Total average	171.0	5.67 ± 0.04	2.58 ± 0.02	2.20 ± 0.02

Table 23: Total efficiency and ratios for $\Lambda_b \to \Lambda_c^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$.

Run Range	$\int \mathcal{L}\, dt$ (pb^{-1})	$\epsilon_{\Lambda_b \to \Lambda_c^+\pi^-}$ (10^{-4})	$\epsilon_{\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu}$ (10^{-4})	ϵ Ratio
138809–143000	3.4	1.84 ± 0.15	0.50 ± 0.06	3.69 ± 0.51
143001–146000	4.0	1.23 ± 0.12	0.36 ± 0.04	3.43 ± 0.53
146001–149659	7.1	1.42 ± 0.09	0.48 ± 0.04	2.98 ± 0.30
149660–150009	0.6	2.23 ± 0.39	1.13 ± 0.19	1.98 ± 0.48
150010–152668	11.3	2.13 ± 0.12	0.63 ± 0.04	3.37 ± 0.27
152669–156487	39.2	2.37 ± 0.05	0.75 ± 0.02	3.15 ± 0.11
159603–164302	52.7	2.67 ± 0.05	0.82 ± 0.02	3.26 ± 0.09
164303–167715	52.7	3.76 ± 0.06	1.14 ± 0.02	3.30 ± 0.08
Total average	171.0	2.86 ± 0.03	0.86 ± 0.01	3.31 ± 0.05

6 Backgrounds of the Semileptonic Modes

The $D^*\mu$, $D\mu$ and $\Lambda_c\mu$ combinations we observe in the data consist of the exclusive semileptonic signals, $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$, $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$, $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$, in the presence of other backgrounds. These backgrounds arise from three sources:

- physics backgrounds: B hadron decays into similar final states, a charm hadron, a real muon and other tracks.
- muon fakes: a charm hadron and a track which fakes a muon.
- $b\bar{b}$ and $c\bar{c}$: two B or charm hadrons from the $b\bar{b}$ and $c\bar{c}$ pairs decay into a D^{*+} (D^+, Λ_c^+) and a muon, respectively.

The goal is to measure the relative partial decay widths of the exclusive semileptonic decays to hadronic decays. The backgrounds listed above should be subtracted from the observed inclusive semileptonic signal in the data. The ratio of branching fractions is then calculated as follows:

$$\frac{\mathcal{B}_{\rm semi}}{\mathcal{B}_{\rm had}} = \left(\frac{N_{\rm inclusive\ semi} - N_{\rm physics} - N_{{\rm fake}\mu} - N_{c\bar{c},\ b\bar{b}}}{N_{\rm had}}\right) \cdot \frac{\epsilon_{\rm had}}{\epsilon_{\rm semi}}, \tag{53}$$

where $\mathcal{B}$ stands for the branching ratio, ϵ is the efficiency from the MC. We estimate the contribution from the physics and $b\bar{b}$, $c\bar{c}$ backgrounds, using the efficiencies from the MC and the branching ratios from the PDG. We normalize the backgrounds to the observed number of events in the fully reconstructed hadronic signal in the data,

$$N_{{\rm physics}\ (b\bar{b},c\bar{c})} = N_{\rm had} \cdot \Sigma \frac{\mathcal{B}_i \cdot \epsilon_i}{\mathcal{B}_{\rm had} \cdot \epsilon_{\rm had}}. \tag{54}$$

The branching ratios of B^0 hadronic modes in Equation 54 come from the world average in the PDG. For the Λ_b mode, we extract $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ from the recent CDF result, $\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b\to\Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0\to D^+\pi^-)}$ by Le, *et al.* [4] in Yu [1]. Sections 6.1– 6.3 estimate the amount of backgrounds in the semileptonic signal. We will show that the dominant signal contamination is from the physics background. The second largest background arises from muon fakes. The smallest background source is from $b\bar{b}$ and $c\bar{c}$.

6.1 Physics Backgrounds

Physics backgrounds come from the decays of B hadrons into similar final state as our semileptonic signal: a D^{*+} (D^+, Λ_c^+), a μ^- and missing particles. Branching ratios and efficiencies of these physics decays are needed to normalize the background contribution to the observed number of hadronic signal events in the data;

$$\frac{N_{\rm physics}}{N_{\rm had}} = \frac{\Sigma\ \mathcal{B}_i \cdot \epsilon_i}{\mathcal{B}_{\rm had} \cdot \epsilon_{had}}. \tag{55}$$

For the backgrounds to the $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ decays, we find the modes which give similar final states as our semileptonic signals in the decays listed in the PDG summary and the default decay table inside `EvtGen` package. Many decays of B and D mesons have been measured by other experiments, such as `CLEO`, `BELLE` and `BABAR`. These measurements serve as inputs to the `EvtGen` decay package. Since `BELLE` and `BABAR` also use the `EvtGen` package, they have included decay modes into `EvtGen` which have not yet been measured and estimate the branching ratios. For the backgrounds to the $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ decay, none of the B hadrons decays to $\Lambda_c\mu$ final states have been measured in the CDF-I and other experiments, or estimated inside `EvtGen`. We use the results from the preliminary measurements by Litvintsev, *et al.* [13] and the PDG upper limit to obtain the background branching ratios.

After we have a list of decays which share similar final states as our signals, we use the generator level simulation to estimate the composition of the inclusive semileptonic signal from each physics background. Details of this procedure maybe be found in the study by Tesarek, *et al.* [14] [15]. The decays which contribute $\geq 1\%$ to the semileptonic signal after trigger-like and the four track invariant mass $M_{D^*(D,\Lambda_c)\mu}$ cuts are selected for further consideration. We generate each selected decay separately and run through the full CDF detector simulation as described in Section 5.1. Then we run the same signal reconstruction program used for the data on the MC and divide the number of reconstructed events by the number of generated events to obtain the efficiency.

6.1.1 Physics backgrounds of $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$

A detailed description about the estimate of the branching fractions of the $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ physics backgrounds can be found in Tesarek [14]. Tables 24– 25 summarize the physics background in $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ which contribute $\geq 1\%$. The second column in the table lists the measured or estimated branching ratios. The third column lists their efficiencies relative to the hadronic signals with statistical errors. The fourth column lists the normalization of each background relative to the exclusive semileptonic signal. The last column lists the number of events from each background after multiplying the relative branching ratio and efficiencies with the number of hadronic signal in the data, as expressed in Equation 55. The uncertainty in the last column only includes the statistical uncertainty of the hadronic yield. The dominant background of $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ is $B^- \to D_1^0\mu^-\overline{\nu}_\mu$ where $D_1^0 \to D^{*+}\pi^-$. The total physics background in Table 24 is about 15% of $\overline{B} \to D^{*+}\mu^- X$ events in the data after all the cuts. The dominant background of $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ is $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$ where $D^{*+} \to D^+\pi^0$. The total physics background in Table 25 contributes about 40% of $\overline{B} \to D^+\mu^- X$ events in the data after all the cuts. As shown in Figure 58, a cut on the invariant mass of $D^{*+}(D^+)\mu^-$ can reduce or eliminate the background from B^0, B^+ decaying semileptonically to more particles or a higher mass charm state.

6.1.2 Physics backgrounds of $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$

A more detailed description about the estimate of the branching fractions of the $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ physics backgrounds can be found in Tesarek [15]. Table 26 summarizes the physics background from the other Λ_b semileptonic decays discussed in Tesarek [15] and their relative efficiencies to the hadronic signal. The dominant backgrounds are $\Lambda_b \to \Lambda_c(2593)^+\mu^-\overline{\nu}_\mu$ and $\Lambda_b \to \Lambda_c(2625)^+\mu^-\overline{\nu}_\mu$. Total physics background in Table 26 is about 9.2% of the $\overline{B} \to \Lambda_c^+\mu^- X$ events in the data.

We also consider the background from the baryonic semileptonic decays of B mesons. While there are branching ratio measurements of the B baryonic hadronic decay, eg: $\overline{B}^0 \to \Lambda_c^+\overline{p}\pi^+\pi^-$, there is only an upper limit for the semileptonic decay of B_u/B_d mixture. We assume the upper limit for the muon-neutron or muon-proton final state should be the same as those decays with a proton and electron in the final state:

$$\mathcal{B}(B^0/B^+ \to \Lambda_c^- pe\nu_e) < 0.15\%.$$

and use the limit for the branching ratios of the following modes:

$$\mathcal{B}(B^- \to \Lambda_c^+\overline{p}\mu^-\overline{\nu}_\mu) = 0.15\%,$$
$$\mathcal{B}(\overline{B}^0 \to \Lambda_c^+\overline{n}\mu^-\overline{\nu}_\mu) = 0.15\%.$$

We then generate Monte Carlo to obtain the efficiencies for these two decays. Since we find the P_T spectra of B mesons and Λ_b are quite different, it is least ambiguous to calculate the quantity:

$$\frac{N_{B^-\to\Lambda_c^+\overline{p}\mu^-\overline{\nu}_\mu}}{N_{\Lambda_b\to\Lambda_c^+\pi^-}} = \frac{\sigma_{\Lambda_b}(P_T > 4.0)}{\sigma_{B^0}(P_T > 4.0)} \cdot \frac{\mathcal{B}(B^- \to \Lambda_c^+\overline{p}\mu^-\overline{\nu}_\mu)\cdot\epsilon_{B^-\to\Lambda_c^+\overline{p}\mu^-\overline{\nu}_\mu}}{\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)\cdot\epsilon_{\Lambda_b\to\Lambda_c^+\pi^-}}. \quad (56)$$

Table 24: Physics backgrounds in $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$.

Mode	BR (%)	ϵ ratio	Norm	N_{event}
$\overline{B}^0 \to D^{*+}\pi^-$	0.276± 0.021	1	–	106 ± 11
$\overline{B} \to D^{*+}\mu^- X$	–	–	–	1059 ± 33
$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	5.44 ± 0.23	0.447 ± 0.006	1.000	–
$B^- \to D_1^0\mu^-\overline{\nu}_\mu$	0.56 ± 0.16	0.356 ± 0.008	0.055	51 ± 5
$\hookrightarrow D^{*+}\pi^-$	66.67 ± ?			
$\overline{B}^0 \to D_1^+\mu^-\overline{\nu}_\mu$	0.56 ± ?	0.349 ± 0.008	0.027	25 ± 3
$\hookrightarrow D^{*+}\pi^0$	33.33 ± ?			
$B^- \to D_1'^0\mu^-\overline{\nu}_\mu$	0.37 ± ?	0.351 ± 0.008	0.036	33 ± 4
$\hookrightarrow D^{*+}\pi^-$	66.67 ± ?			
$\overline{B}^0 \to D_1'^+\mu^-\overline{\nu}_\mu$	0.37 ± ?	0.336 ± 0.008	0.017	16 ± 2
$\hookrightarrow D^{*+}\pi^0$	33.33 ± ?			
$B^- \to D^{*+}\pi^-\mu^-\overline{\nu}_\mu$	0.20 ± ?	0.242 ± 0.007	0.020	19 ± 2
$\overline{B}^0 \to D^{*+}\pi^0\mu^-\overline{\nu}_\mu$	0.100± ?	0.239 ± 0.006	0.010	9 ± 1
$\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau$	1.60 ± ?	0.136 ± 0.005	0.016	15 ± 2
$\hookrightarrow \mu^-\overline{\nu}_\mu$	17.36 ± 0.06			

Table 25: Physics backgrounds in $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$.

Mode	BR (%)	ϵ ratio	Norm	N_{event}
$\overline{B}^0 \to D^+\pi^-$	0.276± 0.025	1.000	–	579 ± 30
$\overline{B} \to D^+\mu^- X$	–	–	–	4721 ± 104
$\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$	2.14 ± 0.20	0.455 ± 0.004	1.000	–
$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	5.44 ± 0.23	0.372 ± 0.005	0.671	1373 ± 71
$\hookrightarrow D^+\pi^0/\gamma$	32.30 ± 0.64			
$\overline{B}^0 \to D^+\pi^0\mu^-\overline{\nu}_\mu$	0.30 ± ?	0.165 ± 0.004	0.051	104 ± 5
$B^- \to D^+\pi^-\mu^-\overline{\nu}_\mu$	0.60 ± ?	0.165 ± 0.004	0.102	208 ± 11
$B^- \to D_1^0\mu^-\overline{\nu}_\mu$	0.56 ± 0.16	0.278 ± 0.005	0.034	70 ± 4
$\hookrightarrow D^{*+}\pi^-$	66.67 ± ?			
$\hookrightarrow D^+\pi^0/\gamma$	32.30 ± 0.64			
$B^- \to D_1^0\mu^-\overline{\nu}_\mu$	0.37 ± ?	0.273 ± 0.005	0.022	46 ± 3
$\hookrightarrow D^{*+}\pi^-$	66.67 ± ?			
$\hookrightarrow D^+\pi^0/\gamma$	32.30 ± 0.64			
$\overline{B}^0 \to D^+\tau^-\overline{\nu}_\tau$	0.70 ± ?	0.100 ± 0.004	0.013	26 ± 1
$\hookrightarrow \mu^-\overline{\nu}_\mu$	17.36 ± 0.06			
$\overline{B}_s \to D^+K^0\mu^-\overline{\nu}_\mu$	0.30 ± ?	0.137 ± 0.005	0.011	23 ± 1

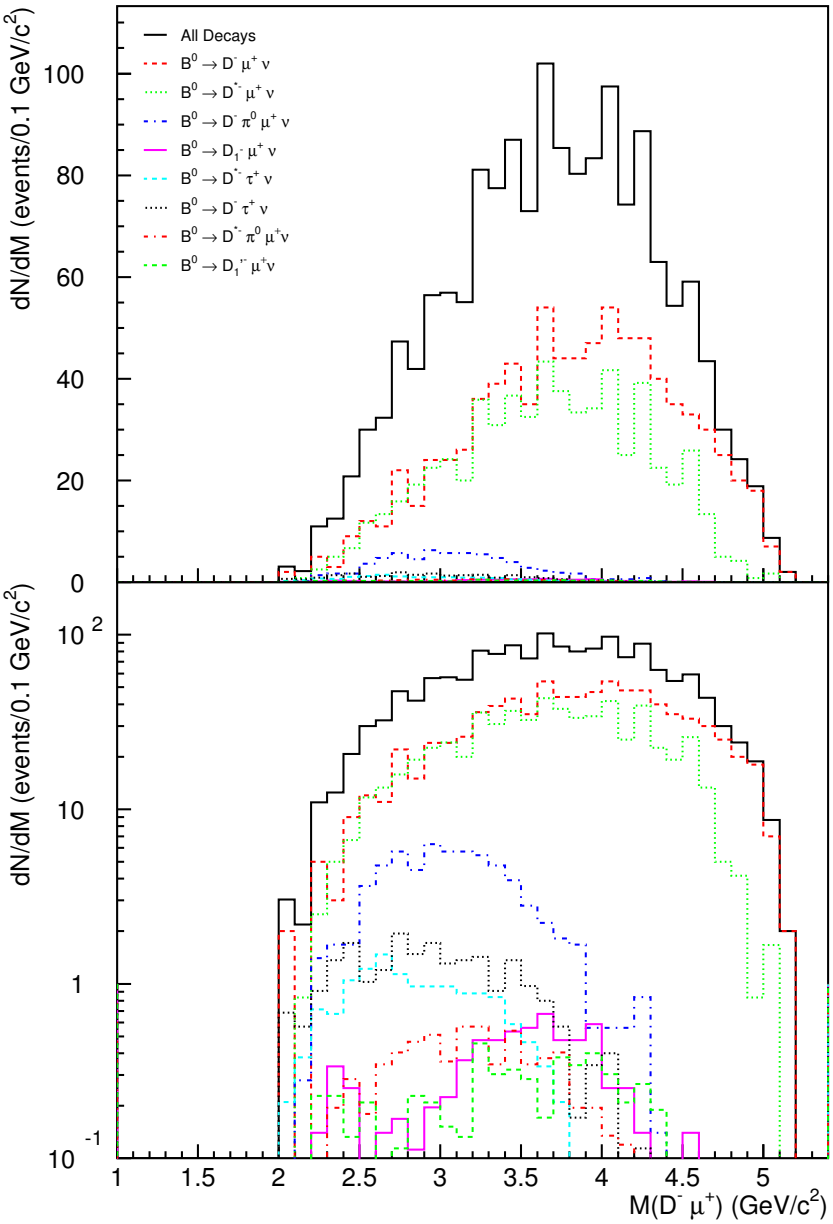


Figure 58: Invariant mass of $D^+\mu^-$ for the signal and physics backgrounds from semileptonic B^0 decays [14]. The top and bottom histograms are the same plot on a linear and log scale, respectively. Note that the backgrounds are concentrated in the low mass region. The signal to background ratio is larger at the higher mass region.

Table 26: Physics backgrounds in $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$ from other Λ_b semileptonic decays.

Mode	BR (%)	ϵ ratio	Norm	N_{event}
$\Lambda_b \to \Lambda_c^+ \pi^-$	0.41 ± 0.21	1	–	179 ± 19
$\overline{B} \to \Lambda_c^+ \mu^- X$	–	–	–	1237 ± 97
$\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$	6.26 ± 0.21	0.300 ± 0.004	1	–
$\Lambda_b \to \Lambda_c(2593)^+ \mu^- \overline{\nu}_\mu$	0.295 ± 0.098	0.196 ± 0.003	0.031	26 ± 3
$\hookrightarrow \Sigma_c^{++} \pi^-$	24 ± 7			
$\hookrightarrow \Lambda_c^+ \pi^+$	100			
$\hookrightarrow \Sigma_c^0 \pi^+$	24 ± 7			
$\hookrightarrow \Lambda_c^+ \pi^-$	100			
$\hookrightarrow \Sigma_c^+ \pi^0$	$24 \pm ?$			
$\hookrightarrow \Lambda_c^+ \pi^0$	100			
$\hookrightarrow \Lambda_c^+ \pi^+ \pi^-$	18 ± 10			
$\hookrightarrow \Lambda_c^+ \pi^0 \pi^0$	$9 \pm ?$			
$\hookrightarrow \Lambda_c^+ \gamma$	$1 \pm ?$			
$\Lambda_b \to \Lambda_c(2625)^+ \mu^- \overline{\nu}_\mu$	0.492 ± 0.095	0.191 ± 0.003	0.050	42 ± 4
$\hookrightarrow \Lambda_c^+ \pi^+ \pi^-$	$66 \pm ?.$			
$\hookrightarrow \Lambda_c^+ \pi^0 \pi^0$	$33 \pm ?$			
$\hookrightarrow \Lambda_c^+ \gamma$	$1 \pm ?$			
$\Lambda_b \to \Lambda_c^+ f^0 \mu^- \overline{\nu}_\mu$	$0.25 \pm ?$	0.023 ± 0.002	0.003	2.6 ± 0.3
$\Lambda_b \to \Lambda_c^+ \pi^+ \pi^- \mu^- \overline{\nu}_\mu$	$0.50 \pm ?$	0.032 ± 0.002	0.009	7 ± 1
$\Lambda_b \to \Lambda_c^+ \pi^0 \pi^0 \mu^- \overline{\nu}_\mu$	$0.25 \pm ?$	0.033 ± 0.002	0.004	3.6 ± 0.4
$\Lambda_b \to \Sigma_c^0 \pi^+ \mu^- \overline{\nu}_\mu$	0.264 ± 0.103	0.081 ± 0.004	0.011	10 ± 1
$\hookrightarrow \Lambda_c^+ \pi^-$	100			
$\Lambda_b \to \Sigma_c^+ \pi^0 \mu^- \overline{\nu}_\mu$	0.264 ± 0.103	0.072 ± 0.004	0.010	8 ± 1
$\hookrightarrow \Lambda_c^+ \pi^0$	100			
$\Lambda_b \to \Sigma_c^{++} \pi^- \mu^- \overline{\nu}_\mu$	0.264 ± 0.103	0.077 ± 0.004	0.011	9 ± 1
$\hookrightarrow \Lambda_c^+ \pi^+$	100			
$\Lambda_b \to \Lambda_c^+ \tau^- \overline{\nu}_\tau$	$1.74 \pm ?$	0.040 ± 0.003	0.007	5 ± 1
$\hookrightarrow \mu^- \overline{\nu}_\mu$	17.36 ± 0.06			

Table 27: Physics backgrounds in $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$ from B mesons.

Mode	BR (%)	Relative ϵ	N_{event}
$\Lambda_b \to \Lambda_c^+ \pi^-$	0.41 ± 0.21	1	179 ± 19
$\overline{B} \to \Lambda_c^+ \mu^- X$			1237 ± 97
$\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$	6.26 ± 0.21	0.265 ± 0.004	–
$B^- \to \Lambda_c^+ \overline{p} \mu^- \overline{\nu}_\mu$	$0.15 \pm$?	0.045 ± 0.002	4.7 ± 0.5
$\overline{B}^0 \to \Lambda_c^+ \overline{n} \mu^- \overline{\nu}_\mu$	$0.15 \pm$?	0.044 ± 0.002	4.6 ± 0.5

We use a low P_T threshold ($P_T > 4$ GeV/c) because we wish to accurately assess the acceptance of the B hadron after applying the reconstruction requirements. Specifically we are concerned about the case where the neutrino is emitted in the direction opposite to the direction that B hadron is traveling. This case increases the P_T of the remaining daughters and may make their total P_T greater than the P_T of the B hadron. Therefore, the denominator of the efficiency is the number of events with B mesons or Λ_b $P_T > 4$ GeV/c and rapidity < 2.0. The numerator is the number of events which pass all the trigger and analysis cuts. Using the method presented in Yu [1], we obtain a corrected number for the ratio of B meson to Λ_b cross-sections:

$$\frac{\sigma_{\Lambda_b}(P_T > 4.0)}{\sigma_{B^0}(P_T > 4.0)} = 0.63 \pm 0.23(stat \oplus other\ syst) \begin{matrix} +0.24 \\ -0.14 \end{matrix} (P_T). \tag{57}$$

The uncertainty due to the Λ_b P_T spectrum is separated from the other systematic and statistical uncertainties for the systematics study on the relative branching fractions in Section 7.1. Table 27 summarizes the B meson to $\Lambda_c^+ \mu^-$ backgrounds. The contribution of B^0 and B^+ in the $\overline{B} \to \Lambda_c^+ \mu^- X$ events is about 0.4% each.

6.2 Fake Muons

Another source of background originates from a charm hadron together with a hadron track (π, K, proton) misidentified as a muon. Physics processes that generate these hadrons are direct production, inelastic collisions with the detector material, fragmentation or the decays of charm and B hadrons. Fragmentation is the process by which a b or c quark combines with additional quarks and gluons to form a $q\overline{q}$ or qqq bound state. Fake muons from the first three categories tend to have a softer P_T spectrum than the real muons from B decays. A tighter P_T cut on the muon candidate largely removes these backgrounds. Fake muons from the charm hadrons which are produced at the primary vertex are also suppressed. For the reason that we require the muon candidate should be matched to an SVT track with a d_0 greater than 120 μm, while fake muons from the promptly produced charm tend to have smaller impact parameter. Also for the reason that we require the charm hadron and the muon candidate to form a vertex displaced from the beam line and make a strict requirement on the pseudo $c\tau$.

$$\text{pseudo } c\tau \ = \frac{M_B}{P_T(\text{charm} + \mu)} \cdot L_{xy}. \tag{58}$$

Here $P_T(\text{charm} + \mu)$ is the total transverse momentum of charm hadron and the muon.

Therefore, our principle source of fake muons comes from two types of B hadron decays:

- $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+) X_{\text{had}}$ anything: hadronic decays of any B hadrons, where X_{had} is π, K or proton which fakes the muon.
- $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+) X_{\text{had}}\ l\ \overline{\nu}_l$ anything: semileptonic B decay into a charm, a hadron track X_{had}, and any leptons (e, μ, τ). The muon is not reconstructed but X_{had} fakes the muon.

In this section, we estimate the fake muon contamination for our three signals: $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$, $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ and $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$. Note the charge conjugates of the modes listed are also included.

6.2.1 Background Estimate

We use two methods to estimate the amount of contamination from fake muons in our semileptonic signal. Each method uses a different way to obtain the number of (hadron track, charm hadron) candidates in our data. Both methods apply the previous CDF measurements of the probabilities for a real pion, kaon and proton to be misidentified as a muon. These measurements are performed by Ashmanskas and Harr [16] using a pion and kaon sample from the $D^{*+} \to D^0\pi^+$decays, where $D^0 \to K^-\pi^+$, and by Litvintsev [17] using a proton sample from the $\Lambda \to p^+\pi^-$ decays. The fake probability $\mathcal{P}_\pi$ ($\mathcal{P}_K$, $\mathcal{P}_p$) is defined as the number of pions (kaons or protons) that pass the following muon identification cuts divided by the total number of pions (kaons or protons) inside the fiducial volume of CDF Central Muon Detector (CMU) and matched to an SVT track.

- The track is fiducial to the CMU and matched to an SVT track
- The track is associated with hits in the CMU
- The matching χ^2 between the track and the hits in the CMU is less than 9.

Figure 59 shows the $\mathcal{P}_\pi$, $\mathcal{P}_K$ measured in sixteen and $\mathcal{P}_p$ measured in twelve transverse momentum bins for positive and negative charged tracks, separately.

Method I The first method uses data to obtain the number of (hadron track,charm hadron) candidates, then Monte Carlo to determine the ratio of pions, kaons and protons in the hadron tracks. We run the same signal reconstruction program on the secondary datasets `hbot0h` and `hbot1i`. We do not use the skimmed tertiary datasets (see Yu [2]) as the samples are biased by requiring at least one track in the event matched to a muon stub in the muon detector. We look for a charged track which fails the muon identification cuts (TRK^{fail}). TRK^{fail} and a charm hadron should form a displaced vertex and pass the same analysis cuts we apply to the signal. Each event is weighted with the fake probability ($\mathcal{P}_{\rm avg}$) according to the momentum and the charge of TRK^{fail}. We then fit the weighted charm hadron mass distribution, i.e. $M_{K\pi\pi} - M_{K\pi}$, $M_{K\pi\pi}$ and $M_{pK\pi}$, using the same functions as described in Section 4.1, to obtain the signal contamination from the fake muons. Since an event-weighted likelihood fit will not give a proper error for the yield, a binned χ^2 fit is performed. $\mathcal{P}_{\rm avg}$ is a weighted average of pion, kaon and proton fake probability ($\mathcal{P}_\pi$, $\mathcal{P}_K$, $\mathcal{P}_p$). The weight R_i is determined by the fraction of pions, kaons and protons in the $\overline{B} \to D^{*+}$ $(D^+, \Lambda_c^+)X_{\rm had}$ l $\overline{\nu}_l$ anything and $\overline{B} \to D^{*+}$ $(D^+, \Lambda_c^+)X_{\rm had}$ anything MC after analysis cuts:

$$\mathcal{P}_{\rm avg} = R_\pi\mathcal{P}_\pi + R_K\mathcal{P}_K + R_p\mathcal{P}_p, \tag{59}$$

where

$$R_i = \frac{N_i}{N_\pi + N_K + N_p},$$

and i is π, K or proton. The Monte Carlo is generated as described in Section 5.1. Decays of B^0, B^+, B_s and Λ_b are generated separately and decay tables include all the possible decays. Each kind of B hadron gives different R_i and is weighted with the product of the production fractions, total branching ratios and the number of generated events. Table 28 summarizes the pion to kaon ratio and the number of fake muon candidates before and after weighting the events in our three different signals. See Figures 60– 62 for the weighted mass distribution for each mode.

The uncertainties of the number of fake muon candidates come from three sources: 1. the uncertainty from the binned χ^2 fit, 2. the uncertainties on the pion, kaon and proton fractions due to the finite Monte Carlo sample size, the uncertainties of the branching ratios and production fractions, and 3. the uncertainty on the measured fake probability. For the last source, we vary the fake rate in

each momentum bin ± 1 sigma, independently. We then add the systematic shifts in quadrature to get the accumulative uncertainty. The number of fake muons using this method is about 3–5% of the inclusive semileptonic signals in the data.

Method II The second method relies on the MC and the external input of the branching ratios from the PDG. MC is run through the same reconstruction program for the data. Then we apply the same cuts as signal reconstruction and obtain the trigger and reconstruction efficiencies of $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}$ anything and $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}\ l\ \overline{\nu}_l$ anything decays. We weight the MC events which pass the analysis cuts with the fake probability according to the momentum, the charge and the particle type of the track, $X_{\rm had}$. The particle identification of the track, $X_{\rm had}$, is obtained by matching the hits on the reconstructed track with those on the input simulated track. Together with the efficiency of hadronic mode, branching ratios of our hadronic signals, $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}$ anything and $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}\ l\ \overline{\nu}_l$ anything from the PDG, we normalize the background to the observed number of hadronic signals in the data,

$$\frac{N_{\rm fake\ \mu}}{N_{\rm had}} = \frac{\mathcal{B}(\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X) \cdot \epsilon_{\rm fake\ \mu}}{f_{d,\rm baryon} \cdot \mathcal{B}_{\rm had}\epsilon_{\rm had}}. \tag{60}$$

Equations 61– 62 use D^+ as an example to show how we derive the $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}$ anything and $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}\ l\ \overline{\nu}_l$ anything branching ratios from the existing information in the PDG.

$$\begin{aligned} \mathcal{B}(\overline{B} \to D^+ X_{\rm had}\ l\nu_l \text{ anything}) &= \frac{7}{3} \cdot \mathcal{B}(\overline{B} \to D^+\pi^+\mu \text{ anything}) \\ &+ \frac{7}{3} \cdot \mathcal{B}(\overline{B} \to D^+\pi^-\mu \text{ anything}), \qquad (61) \\ \mathcal{B}(\overline{B} \to D^+ X_{\rm had} \text{ anything}) &= \mathcal{B}(\overline{B} \to D^+ \text{anything}) \\ &- \frac{7}{3} \cdot \mathcal{B}(\overline{B} \to D^+\mu \text{ anything}), \qquad (62) \end{aligned}$$

where the factor, $\frac{7}{3}$, comes from the fact that the branching ratios of muon and electron channels are equal and the branching ratio of the tau channel is scaled down by the ratio of the phase space, $\sim \frac{1}{3}$. Therefore, we have to scale up the branching ratio of the muon channel by $1 + 1 + \frac{1}{3} = \frac{7}{3}$ to get the total branching ratio of all the lepton channels.

Table 29 summarizes the parameters used to calculate the number of fake muon events, where the decay $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}$ anything is denoted as mode "1" and $\overline{B} \to D^{*+}\ (D^+, \Lambda_c^+)X_{\rm had}\ l\ \overline{\nu}_l$ anything is denoted as mode "2" in the table. The uncertainties on the number of fake muons originate from: the uncertainty on the hadronic yield, the relative efficiency, the uncertainty on the fake rate and the relative branching ratios. The dominant uncertainty is from the relative branching ratios. The number of fake muon backgrounds from method I is consistent with the result using method II. We use the results of method I in the calculation of our final result of the relative branching ratios. In general, the fraction of fake muons is about 5% of the total semileptonic yield in the data.

6.2.2 Like-sign Combination

Note that we do not use the like sign combination (i.e. the charm hadron and the muon have the same sign of charges) to estimate the fake muon background for two reasons: First, two different B hadrons from the $b\bar{b}$ in the event can produce a real muon and a real charm of the same charge sign when the B hadrons in the event have opposite flavors and one B hadron decays semileptonically. Second, the two track trigger, used for this analysis, requires a pair of tracks with opposite charges. The trigger requirement greatly reduces the number of like-sign (wrong-sign) candidates and introduces large statistical errors for the number of fake muons.

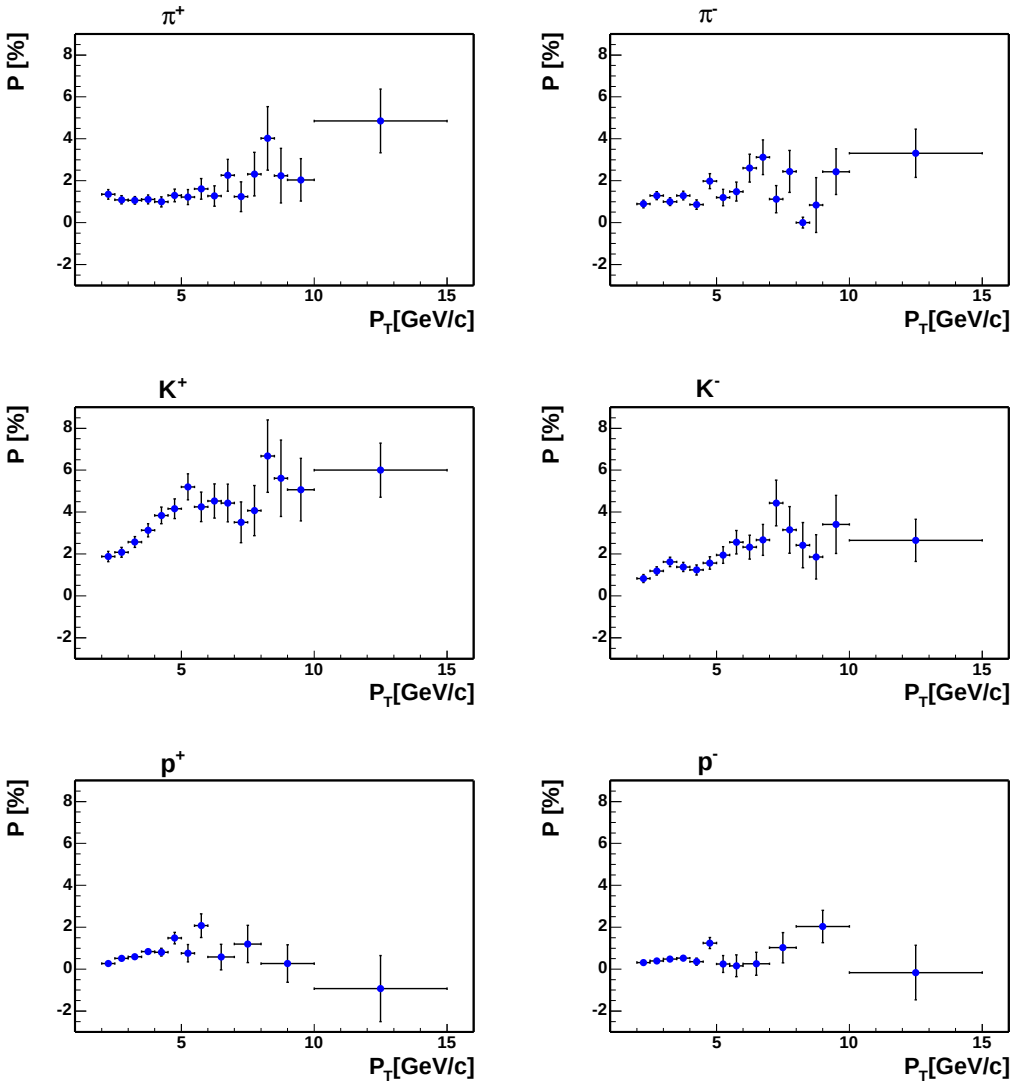


Figure 59: The probability for a pion, kaon or proton being misidentified as a muon in bins of transverse momentum (P_T) from the measurements by Ashmanskas, Harr [16] and Litvintsev [17]. From the top left to the bottom right are π^+, π^-, K^+, K^-, p and $\overline{p}$ fake probabilities.

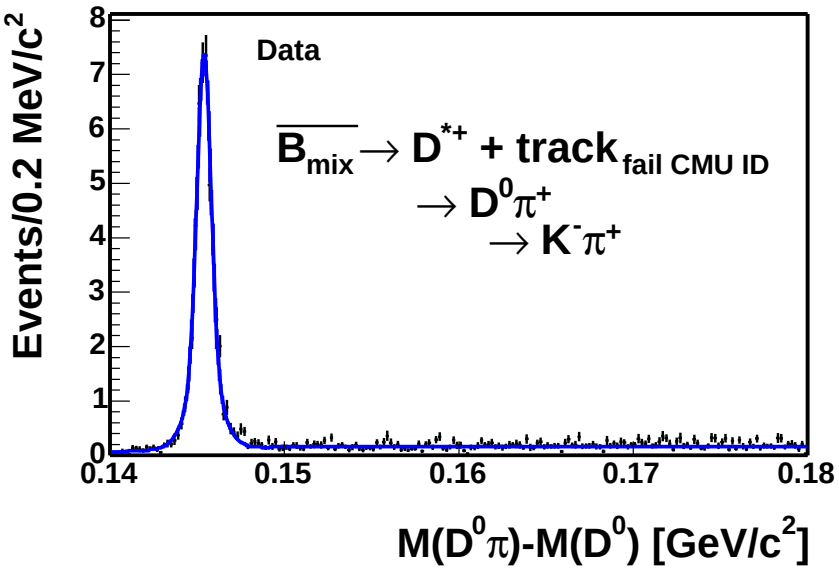


Figure 60: Fit of $\overline{B} \rightarrow D^{*+}\mu_{\rm fake}$ yield after weighting the charged track which fails the muon ID cut with an averaged muon fake probability. There are 44 ± 1 events in the peak. Fit χ^2/NDF = 134/92, probability = 0.3%. A sideband subtraction yields 46 ± 7 events in the signal peak. Note that the uncertainty from the sideband subtraction does not take into account the event weighting with the fake rate and is over-estimated.

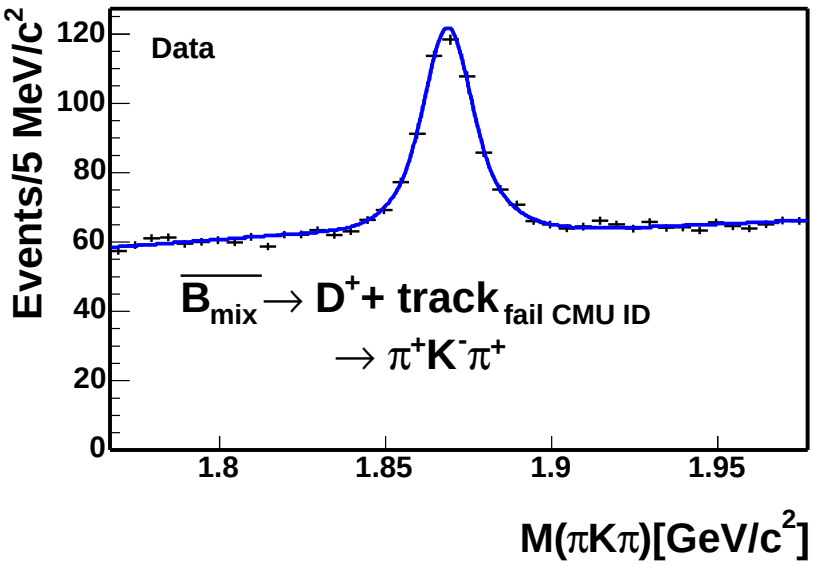


Figure 61: Fit of $\overline{B} \rightarrow D^+\mu_{\rm fake}$ yield after weighting the charged track which fails the muon ID cut with an averaged muon fake probability. There are 230 $\pm$ 5 events in the peak. Fit χ^2/NDF = 36.3/35, probability = 19.7%

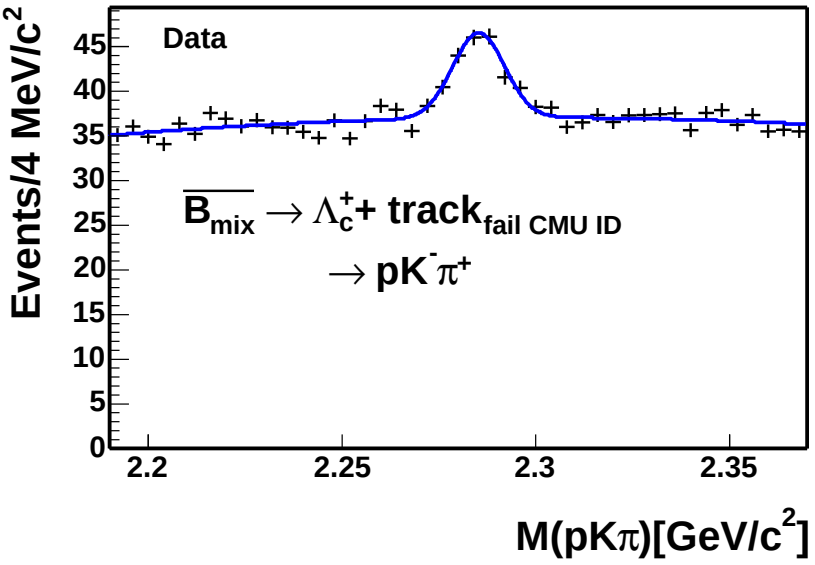


Figure 62: Fit of $\overline{B} \rightarrow \Lambda_c^+\mu_{\rm fake}$ yield after weighting the charged track which fails the muon ID cut with an averaged muon fake probability. There 40 $\pm$ 6 events in the peak. Fit χ^2/NDF = 55.6/39, probability = 4.1%. A sideband subtraction yields 44 $\pm$ 25 events in the signal peak.

Table 28: Parameters for the number of fake muons: Method I.

	$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	$\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$	$\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$
N before weighting	2953 ± 57	15343 ± 303	3560 ± 198
R_π	0.937 ± 0.009	0.909 ± 0.005	0.71 ± 0.16
R_K	0.063 ± 0.009	0.091 ± 0.005	0.05 ± 0.08
R_p	–	–	0.24 ± 0.16
$N_{\text{fake }\mu}$	44 ± 3	230 ± 19	40 ± 9

Table 29: Parameters for the number of fake muons: Method II.

	$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	$\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$	$\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$
$\mathcal{B}_{\text{had}}$ %	0.276 ± 0.021	0.276 ± 0.025	0.41 ± 0.19
$\mathcal{B}_1$ %	10.9 ± 2.1	17.7 ± 2.4	4.8 ± 3.0
$\frac{\epsilon_1}{\epsilon_{\text{had}}}$	0.0038 ± 0.0004	0.0022 ± 0.0002	0.0029 ± 0.0003
$\mathcal{B}_2$ %	1.3 ± 0.3	1.8 ± 0.6	< 1.23
$\frac{\epsilon_2}{\epsilon_{\text{had}}}$	0.0005 ± 0.0002	0.0010 ± 0.0002	0.0002 ± 0.0001
N_{had}	106 ± 11	579 ± 30	179 ± 19
$N_{\text{fake }\mu}$	45 ± 11	220 ± 41	28 ± 34

Table 30: Fake muons from $b\bar{b}$ and $c\bar{c}$.

	$b\bar{b}$	$c\bar{c}$
N_{gen}	43454949	89718181
Real muon N_{pass}	15	35
Fake muon N_{pass}	1.8	0.4

6.2.3 Fake muons from $b\bar{b}$ and $c\bar{c}$

One type of fake muons is not included in the previous subsections. These fake muons stem from $b\bar{b}$, $c\bar{c}$ to two B or charm hadrons then decay into a charm signal, a hadron track misidentified as a muon and other missing particles. A study at the generator level for the $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ mode is done using the $b\bar{b}$ and $c\bar{c}$ `PYTHIA` Monte Carlo datasets as described in Section 6.3.2. We apply analysis-like cuts on the Monte Carlo. We weight the events that pass the cuts with the muon fake probability according to the "muon" candidate momentum, charge and the true particle identification: a kaon, a pion or a proton. Then we compare the number of weighted events with the number of charm hadron and real muon combinations, i.e, a $b\bar{b}$ and $c\bar{c}$ background as described in Section 6.3. We find that fake muons from $b\bar{b}$ and $c\bar{c}$ is about 10% of the $b\bar{b}$ and $c\bar{c}$ background with real muons. See Table 30. From Section 6.3, we show that the $b\bar{b}$ and $c\bar{c}$ background with real muons is at the 1% level. Therefore, we conclude that $b\bar{b}$ and $c\bar{c}$ background with fake muons is about or less than 0.1% and can be ignored.

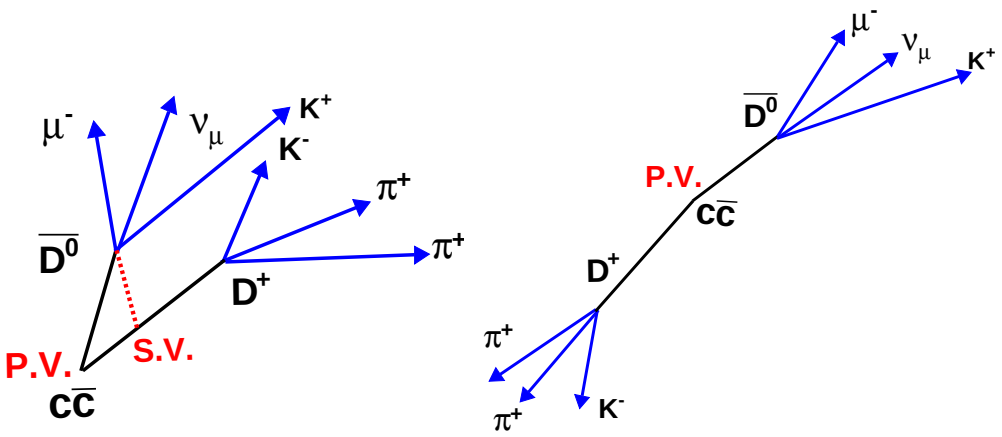


Figure 63: Charm hadrons from $c\bar{c}$ with small (left) and big (right) $\Delta\phi$. In the left figure, the muon from the semileptonic decay of $\overline{D^0}$ and the D^+ forms a secondary vertex and fake our $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ signal. In the right figure, $\Delta\phi$ between two charm hadrons is too big and the daughters can not form a secondary vertex.

6.3 $b\bar{b}$ and $c\bar{c}$ Backgrounds

When the azimuthal angle ($\Delta\phi$) between $b\bar{b}$ or $c\bar{c}$ quark pair is small, daughters of two heavy flavor hadrons from the fragmentation of $b\bar{b}$ or $c\bar{c}$ appear to come from the same decay vertex, see Figure 63 . Here $\Delta\phi$ is defined as the opening angle in the plane perpendicular to the proton and antiproton beam axis. If one hadron decays semileptonically, and the other hadron decays into a charm final state, such as $D^{*+} \to D^0\pi^+$, $D^+ \to K^-\pi^+\pi^+$, and $\Lambda_c^+ \to pK^-\pi^+$, the muon from the semileptonic decay, together with the charm, may fake our semileptonic signal. Production mechanisms and an estimate of the amount of $b\bar{b}$ and $c\bar{c}$ backgrounds are discussed below.

6.3.1 $b\bar{b}$ and $c\bar{c}$ Production Mechanism

In $p\overline{p}$ collisions, the b or c quarks may be single or pair produced by the electroweak and the strong (QCD) processes. The b or c quark production cross-section for the electroweak process $\sigma \cdot \mathcal{B}(p\overline{p} \to W \to bc)$ is around 0.01 μb and is derived from the CDF measurement of the inclusive W cross-section by Halkiadakis, *et al.* [18]. The $b\bar{b}$ and $c\bar{c}$ production cross-sections for the QCD process are around 50 and 200μb respectively from the `PYTHIA` Monte Carlo, when the total transverse momenta of the hard scattering, i.e. the part of the interaction with the largest momentum scale, is greater than 5 GeV/c and at least one b or c quark has $P_T > 4.0$ GeV/c, pseudo-rapidity $\eta < 1.5$. The $b\bar{b}$ and $c\bar{c}$ production rates from the electroweak process are about five thousand times smaller than the QCD processes. Therefore, only the QCD processes are discussed here. Figure 64 shows the leading and next-to-leading order Feynman diagrams for $b\bar{b}(c\bar{c})$ production by the QCD processes from Lannon [19]. The QCD process that contributes the production at leading order is flavor creation. The next-to-leading order (NLO) processes are flavor excitation and gluon splitting, contribute at the same level as the flavor creation [19].

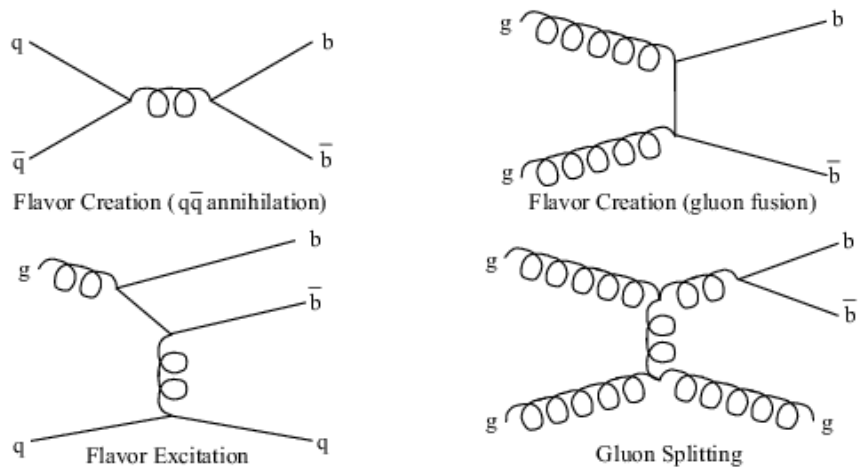


Figure 64: Representative lowest order Feynman diagrams (without loops or radiative corrections) of quark annihilation, gluon fusion, flavor excitation and gluon splitting. Details of these processes may be found in Lannon [19].

6.3.2 Background Estimate

The amount of $b\bar{b}$ and $c\bar{c}$ background is normalized to the number of events observed in the hadronic modes in the the data,

$$\frac{N_{b\bar{b},c\bar{c}}}{N_{\rm had}} = \frac{\sigma_{b\bar{b},c\bar{c}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j}{\sigma_{B^0,\Lambda_b} \cdot \mathcal{B}_{\rm had} \epsilon_{\rm had}}. \tag{63}$$

Here, i represents the species of b(c) hadrons and j represents the decay modes which could contribute to $b\bar{b}$ and $c\bar{c}$ backgrounds. f^i stands for the production fraction ratio for species i. $\mathcal{B}_j$ and ϵ_j are the branching ratio and the efficiency of j^{th} decay mode. Here, "had" refers to $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B}^0 \to D^+\pi^-$, or $\Lambda_b \to \Lambda_c^+\pi^-$, which are the normalization modes in our relative branching ratio measurements. The following subsections detail the methods to estimate $\sigma_{b\bar{b},c\bar{c}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ and $\sigma_{B^0,\Lambda_b} \cdot \mathcal{B}_{\rm had}\epsilon_{\rm had}$ in Equation 63. We do not use detector and trigger simulations to obtain the efficiencies for the following reasons: First, detector and trigger simulations are time and CPU intensive. Second, we will find the contribution of this background is quite small compared with the other backgrounds. Third, we care about the efficiency ratio of the background to the signal, not the absolute efficiency. Our studies show that generator level Monte Carlo gives a good approximation. For instance, the relative efficiency $\frac{\epsilon(\Lambda_b \to \Lambda_c^+ \pi^-)}{\epsilon(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}$ is 3.31 ± 0.05 from the full detector simulation and 3.23 ± 0.01 from the generator level simulation. The difference is only about 2.5%. Similar results are obtained from the relative efficiencies of our other signals.

Background: $\sigma_{b\bar{b},c\bar{c}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ Our estimate of $\sigma_{b\bar{b},c\bar{c}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ relies heavily on the Monte Carlo. We use `PYTHIA` version 6.2 and to generate $b\bar{b}$ and $c\bar{c}$ events, we include the QCD processes mentioned in Section 6.3.1: flavor creation, flavor excitation and gluon splitting (MSEL=1). We further require the P_T of the hard scattering be greater than 5 GeV/c. Events with b quarks P_T greater than 4.0 GeV/c and pseudo-rapidity less than 1.5 are collected into the `nbot90` sample. Events with c quarks which satisfy the same kinematic cuts are collected into the `nbota0` sample. Note that `nbot90` and `nbota0` have small overlap when both b and c quarks are produced and are above the P_T and pseudo-rapidity thresholds. Details of `nbot90` and `nbota0` datasets could be found in [20]. `PYTHIA` $b\bar{b}$ and $c\bar{c}$ cross-sections are used for $\sigma_{b\bar{b},c\bar{c}}$.

The product of the efficiency, branching ratio and production fraction from all the modes, $\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$, is obtained using the following steps: First, we identify the $b\bar{b}$ ($c\bar{c}$) production mechanism to which we are most sensitive. In this pass, we only study the background that forms a $D^+\mu^-$ signature since D^+ has longer lifetime than D^{*+}, Λ_c^+ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ suffers larger $c\bar{c}$ background contamination compared to the other two modes. We need to achieve higher accuracy for the estimate of $\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$. Consequently, for the $b\bar{b}$ background, we re-decay `nbot90` sample ten times, i.e. we re-use the same kinematic distribution of the parent hadrons from `nbot90` ten times but decay the hadrons with independent random numbers and force the decay $D^+ \to K^-\pi^+\pi^+$. For the $c\bar{c}$ background, we force the decay of $D^+ \to K^-\pi^+\pi^+$ and require that all the negative charged charm hadron decay semileptonically.

Then, a generator level two track trigger filter (`SvtFilter`) is applied. We further identify any combination of D^+ and a muon which passes the generator-level, analysis-like cuts found in Table 32. In order to avoid double counting (count $c\bar{c}$ as $b\bar{b}$ background in `nbot90` and $b\bar{b}$ as $c\bar{c}$ background in `nbota0`) due to the overlap of `nbot90` and `nbota0` samples, the ancestors of the muon and charm hadrons are retrieved by tracing the true information from the generator. If both muon and charm hadron come from the same B hadron, the combination is rejected. If the muon and charm hadron come from different B hadrons, the combination is categorized into $b\bar{b}$ background, otherwise, the combination is categorized as a $c\bar{c}$ background. We find that for both $b\bar{b}$ and $c\bar{c}$, more than 90% of the events that pass the cuts are from gluon splitting. Therefore, we are most sensitive to the "gluon splitting" mechanism. Table 31 summarizes the background contributions from different production processes.

Table 31: Summary of $b\bar{b}$ and $c\bar{c}$ production mechanisms and our relative sensitivity for reconstructing the event in our semileptonic sample.

	$b\bar{b}$ background	$c\bar{c}$ background
N_{gen}	219093011	21996889
N_{pass}	75	62
N_{gluon}	70	57
$N_{excitation}$	5	5
$N_{creation}$	0	0
$f_{gluon}(\%)$	93 ± 3	92 ± 3
$f_{excitation}(\%)$	7 ± 3	8 ± 4
$f_{creation}(\%)$	0	0

Second, we filter the gluon splitting events and re-decay the b(c) hadrons in `nbot90` and `nbota0` ten times with the procedure described above. For the $b\bar{b}$ background estimate, we let all the b hadrons and negative charged charm hadrons decay freely, but force the decays of the positive charged charm hadrons in two ways:

- $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ for the background of $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$
- $D^+ \to K^-\pi^+\pi^+$ and $\Lambda_c^+ \to pK^-\pi^+$ for the background of $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ and $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$

Then the `SvtFilter` and the cuts listed in Table 32 are applied. We divide the number of reconstructed events by the number of generated events and get $\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$. Table 33 lists the parameters for the $b\bar{b}$ background.

For the $c\bar{c}$ background estimate, we force the decays of both positive and negative charged charm hadrons. The positive charged charm hadrons are forced to decay into the modes listed above. The negative charged charm hadrons are forced to decay into semileptonic modes individually for D^-,

$\overline{D}^0$, D_s^- and Λ_c^-. As the semileptonic decay modes of these four charm hadrons are all different, we separate the events into four classes denoted by the parent charm particles. After applying `SvtFilter` and the cuts listed in Table 32, we obtain $\sum_j f^i \mathcal{B}_j \epsilon_j$ for each class. Then we multiply the semileptonic branching ratios for each kind of charm hadron with its $\sum_j f^i \mathcal{B}_j \epsilon_j$ and sum them up to get the total amount of $\sum_j f^i \mathcal{B}_j \epsilon_j$ for the $c\bar{c}$ background,

$$\begin{aligned}\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j(total) &= \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j(D^-)\mathcal{B}(D^- \to \mu X) \\ &+ \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j(\overline{D}^0)\mathcal{B}(\overline{D}^0 \to \mu X) \\ &+ \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j(D_s^-)\mathcal{B}(D_s^- \to \mu X) \\ &+ \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j(\Lambda_c^-)\mathcal{B}(\Lambda_c^- \to \mu X). \end{aligned} \quad (64)$$

Table 34 lists the parameters for the $c\bar{c}$ background.

In both $b\bar{b}$ and $c\bar{c}$ background estimates, since we force the positive charged charm hadron to decay into the same final state as our charm signals, we have to multiply the final result by two to include the contribution from both charge states. Table 36 lists $N_{b\bar{b}}$ and $N_{c\bar{c}}$ in our three different signals after multiplying the ratio in Equation 63 with the observed number of events in the hadronic signals.

Table 32: Generator-level analysis-like cuts for $b\bar{b}$ and $c\bar{c}$ background study.

Parameter	Cut Value
P_T of all tracks	> 0.5 GeV/c
P_T of μ (π_B)	> 2.0 GeV/c
η of all tracks	< 1.2
η of μ (π_B)	< 0.6
P_T of four tracks	> 6.0 GeV/c
P_T of charm hadron	> 5.0 GeV/c
$c\tau$ of four tracks	> 200 μm (B), > 250 μm (Λ_b)
$c\tau$ of charm hadron	$>$ -70 μm (D^*, Λ_c), $>$ -30 μm (D^+)
$3.0 < M_{D^{(*)}\mu} < 5.5$ GeV/c^2	
$3.7 < M_{\Lambda_c \mu} < 5.7$ GeV/c^2	
μ (π_B) match to a SVT track	
charm hadron and μ (π_B) have opposite charge signs	
2 out of 4 tracks of B candidate pass two track trigger cuts	

Hadronic signal: $\sigma_{B^0,\Lambda_b} \cdot \mathcal{B}_{\rm had}\epsilon_{\rm had}$ In order to normalize the background to the observed number of events in the hadronic mode, the B^0 or Λ_b production cross-section, the efficiency and the branching ratio of the hadronic signal, have to come from external input or must be calculated using MC (see Equation 63). We obtain σ_{B^0} by multiplying the previous CDF σ_{B^+} measurement by Keaffaber, *et al.* [21] with the production fraction ratios, f_d/f_u, from the 2004 PDG. B^0 decay branching ratios are also obtained from the PDG. The product of σ_{Λ_b} and $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ is obtained by multiplying the CDF measurement of $\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0 \to D^+ \pi^-)}$ by Le, *et al.* [4] with the σ_{B^0} we derive and

Table 33: Parameters used for $b\bar{b}$ background estimate.

PYTHIA $\sigma_{b\bar{b}}$ (μb)	49.6		
	$b\bar{b} \rightarrow D^*\mu$	$b\bar{b} \rightarrow D\mu$	$b\bar{b} \rightarrow \Lambda_c\mu$
$N_{\rm gen}$	221606748	221619610	221619610
$N_{\rm pass}$	43	80	9
$\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ (10^{-7})	1.9 $\pm$ 0.3	3.6 $\pm$ 0.4	0.41 $\pm$0.14
$2 \cdot \sigma_{b\bar{b}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ (10^{-5} μ b)	1.9 $\pm$ 0.3	3.6 $\pm$ 0.4	0.41 $\pm$ 0.14

Table 34: Parameters used for $c\bar{c}$ background estimate.

PYTHIA $\sigma_{c\bar{c}}$ (μb)	198.4		
	$c\bar{c} \rightarrow D^*\mu$	$c\bar{c} \rightarrow D\mu$	$c\bar{c} \rightarrow \Lambda_c\mu$
$N_{\rm gen}$	720741510	698988700	698988700
$N_{\rm pass}$	214	396	7
D^- : $\mathcal{B}(D^- \rightarrow \mu X) = 14.22$ (%)	117	205	4
$\overline{D}^0$: $\mathcal{B}(\overline{D}^0 \rightarrow \mu X) = 6.15$ (%)	76	157	2
D_s^- : $\mathcal{B}(D_s^- \rightarrow \mu X) = 13.32$ (%)	19	34	1
Λ_c^- : $\mathcal{B}(\Lambda_c^- \rightarrow \mu X) = 4.5$ (%)	2	0	0
$\sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ (10^{-8})	3.3 $\pm$ 0.6	6.2 $\pm$ 1.0	0.12 $\pm$ 0.05
$2 \cdot \sigma_{c\bar{c}} \cdot \sum_i \sum_j f^i \mathcal{B}_j \epsilon_j$ (10^{-5} μ b)	1.3 $\pm$ 0.2	2.5 $\pm$ 0.4	0.047 $\pm$ 0.020

the PDG $\mathcal{B}(\overline{B}^0 \rightarrow D^+\pi^-)$.

Since we reconstruct both b and anti-b hadrons in the data, we should multiply the measured cross-section by two. For the efficiencies, we use the MC to generate and decay B hadrons into our signals as described in Section 5.1. The CDF σ_{B^+} measurement is restricted to the B^+ with P_T greater than 6 GeV/c and rapidity (y) less than 1.0. Therefore, the denominator of the efficiency is the number of events in which the B hadrons have $P_T > 6$ GeV/c and $|y| < 1.0$. The numerator of the efficiency is the number of events which pass the cuts in Table 32 except the cut on four track invariant mass. Table 35 lists the parameters that are used to calculate $\sigma_{B^0,\Lambda_b} \cdot \mathcal{B}_{\rm had}\epsilon_{\rm had}$.

Semileptonic signal Table 35 also lists the efficiency and branching ratio of the semileptonic signal mode for a comparison with the background $N_{b\bar{b},c\bar{c}}$. In the case of the Λ_b, we lack the external input for the branching ratio of $\Lambda_b \rightarrow \Lambda_c^+\mu^-\overline{\nu}_\mu$. Therefore, instead of the exclusive mode, we list the branching ratio of the inclusive mode: $\overline{B} \rightarrow \Lambda_c^+\mu^- X$ from the 2004 PDG as an upper bound. We multiply the ratio $\frac{\sigma_{\Lambda_b}}{\sigma_{B^0}}(P_T > 6.0)$ from Yu [1] with the Keaffaber σ_{B^+} result to get σ_{Λ_b} for P_T greater than 6.0 GeV/c. The efficiency of $\Lambda_b \rightarrow \Lambda_c^+\mu^-\overline{\nu}_\mu$ is listed for a comparison with $\epsilon_{b\bar{b},c\bar{c}}$. Note that the amount of $b\bar{b}$ and $c\bar{c}$ background relative to the signal is around 1%. The numbers from three different modes should not be compared directly without multiplying the branching ratios of the charm decays.

6.3.3 Comparison of Data and MC Cross Section

While there are precise measurements of the single charm hadron, B^+ and inclusive b hadron cross-sections, there are no accurate measurements of the total $b\bar{b}$ and $c\bar{c}$ cross-section ($\sigma_{b\bar{b}}$, $\sigma_{c\bar{c}}$) at the

Table 35: Parameters used to calculate σ_{B^0,Λ_b}, $\mathcal{B}_{had}\epsilon_{had}$ and $\mathcal{B}_{semi}\epsilon_{semi}$.

σ_{B^+} (μb)	3.6 ± 0.6		
	$\overline{B} \to D^{*+}X$	$\overline{B} \to D^{+}X$	$\Lambda_b \to \Lambda_c^+ X$
f_x/f_u	1.00 ± 0.04	1.00 ± 0.04	0.25 ± 0.04
σ_{B^0,Λ_b} (μb)	3.6 ± 0.6	3.6 ± 0.6	2.2 ± 0.5
	$\overline{B}^0 \to D^{*+}\pi^-$	$\overline{B}^0 \to D^{+}\pi^-$	$\Lambda_b \to \Lambda_c^+\pi^-$
$\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b\to\Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0\to D^+\pi^-)}$	–	–	0.82 ± 0.26
$\mathcal{B}_{\rm had}$ (%)	0.276 ± 0.021	0.276 ± 0.025	–
$\sigma(\Lambda_b)\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ (μb)	–	–	0.008 ± 0.003
$N_{\rm gen}$	4242100	4242100	39999996
$N_{\rm pass}$	70147	130433	843693
$\epsilon_{\rm had}$ (10^{-2})	1.654 ± 0.006	3.075 ± 0.008	2.109 ± 0.002
$2\sigma_{B^0,\Lambda_b}\mathcal{B}_{had}\epsilon_{had}$ (10^{-5} μb)	33 ± 6	61 ± 12	34 ± 13
	$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	$\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu$	$\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$
$\mathcal{B}_{\rm semi}$ (%)	5.44 ± 0.23	2.14 ± 0.20	9.2 ± 2.1 (%)
$N_{\rm gen}$	4242100	4242100	39999996
$N_{\rm pass}$	32620	66854	264484
$\epsilon_{\rm semi}$ (10^{-2})	0.769 ± 0.004	1.576 ± 0.006	0.661 ± 0.001
$2\sigma_{B^0,\Lambda_b}\mathcal{B}_{\rm semi}\epsilon_{\rm semi}$ (10^{-5} μb)	300 ± 50	240 ± 50	270 ± 90

Tevatron, yet. To understand how well PYTHIA predicts $\sigma_{b\bar{b}}$ and $\sigma_{c\bar{c}}$, we cross-check indirectly by comparing the "differential cross-section" of D^0, B^+ and inclusive b hadrons in PYTHIA with CDF Run I and II measurements by Chen [22], Keaffaber [21], and Bishai [23] *et al.*. We count the number of D^0, B^+ or b hadrons from nbot90 and nbota0 in bins of $P_T(D^0)$, $P_T(B^+)$ and $P_T(J/\psi)$. The bin width and the P_T ranges are the same as Chen, Keaffaber and Bishai analyses. We divide the number of hadrons in each P_T bin by the total number of generated events. Then we multiply Pythia assumed $\sigma_{b\bar{b}}$ and $\sigma_{c\bar{c}}$ (see Tables 33– 34) to get the cross section of hadrons in each P_T bin. We further divide the number by the bin width to obtain the "differential cross-section". The agreements between Monte Carlo and data cross-sections are generally within 10% for charm hadrons and 40% for B hadrons (see Figure 65).

Besides the total cross-section of $b\bar{b}$ and $c\bar{c}$, the ratio of gluon splitting relative to the other two processes, flavor creation and flavor excitation, also affects the amount of $b\bar{b}$ and $c\bar{c}$ backgrounds. Previous CDF Run I measurement of $b\bar{b}$ azimuthal production correlations by Lannon [24] concludes that Pythia gives reasonable prediction of the relative $b\bar{b}$ production rates from the three processes. However, due to the lack of measurements of the $c\bar{c}$ relative production rates, we do not yet have a comparison of the fraction of $c\bar{c}$ gluon splitting between Monte Carlo and data. Therefore, we assign 100% uncertainty when calculating the systematic errors for the estimate of $b\bar{b}$ and $c\bar{c}$ backgrounds. As the contribution of $b\bar{b}$ and $c\bar{c}$ background is at the 1% level, the systematic errors from 100% uncertainty is also about 1%.

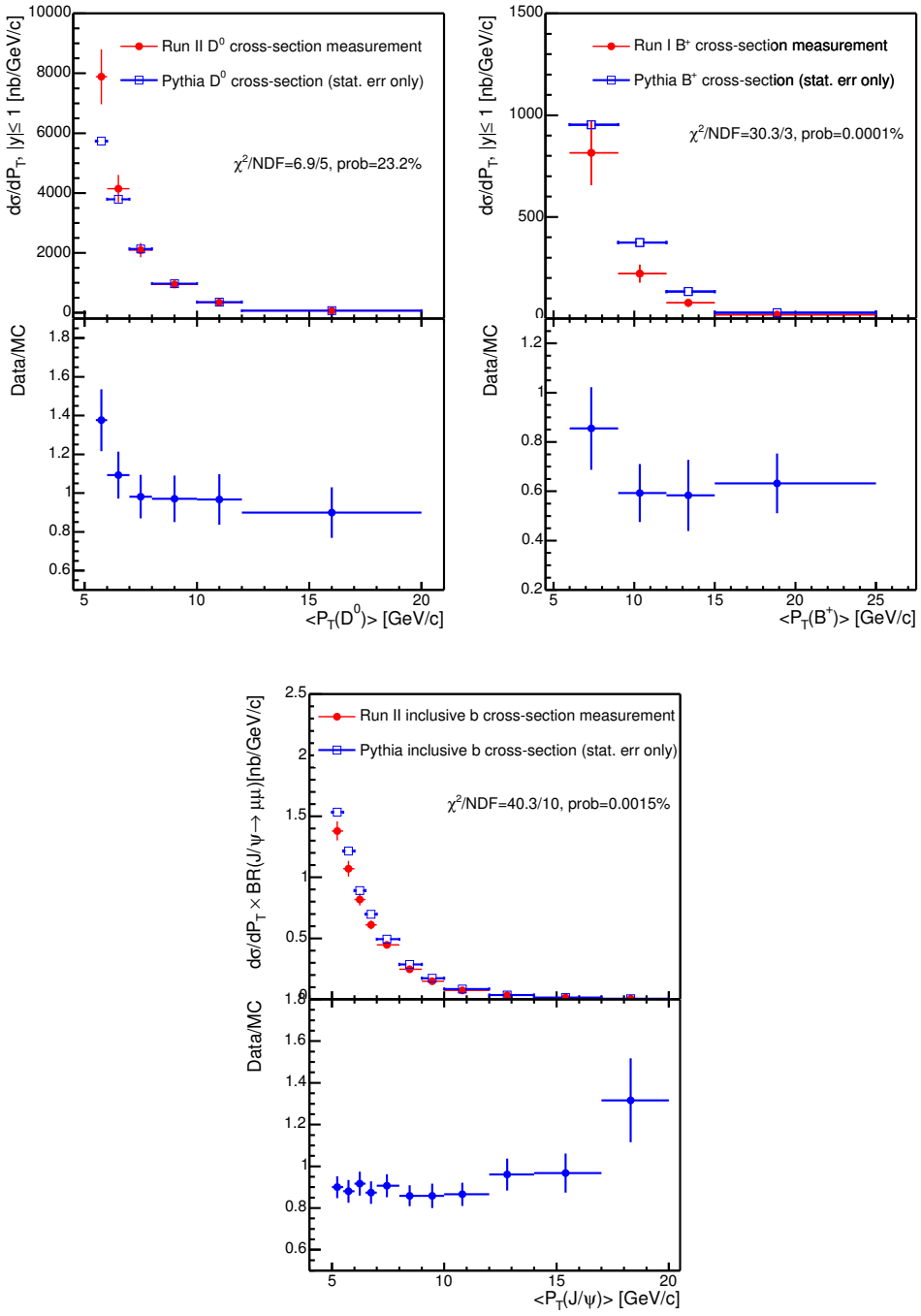


Figure 65: D^0 (top left), B^+ (top right), and inclusive b (bottom) differential cross-sections. The upper plot in each figure shows the differential cross-section for data (closed circles) by Chen [22], Keaffaber [21], and Bishai [23] and MC (open squares). The lower plot in each figure shows the data to MC ratio.

Table 36: The amount of $b\bar{b}$ and $c\bar{c}$ background.

	$\overline{B} \to D^{*+}X$	$\overline{B} \to D^{+}X$	$\Lambda_b \to \Lambda_c^+ X$
$N_{\rm had}$	106 ± 11	579 ± 30	179 ± 19
$N_{\rm semi}$	1059 ± 33	4721 ± 104	1237 ± 97
$N_{b\bar{b}}$	6 ± 0.6	34 ± 2	2.1 ± 0.2
$N_{c\bar{c}}$	4 ± 0.4	23 ± 1	0.2 ± 0.03

6.3.4 Comparison of Data and MC Impact Parameter

We compare the distribution of the impact parameter of charm hadrons with respect to the beam spot in MC and data. The impact parameter, d_0, is the distance from the closest approach of the charm hadron trajectory to the beam spot when projected onto the r-ϕ plane. Figure 66 shows a good agreement of the MC with the data. No excess of charm hadrons with small d_0 is found in the data. This indicates that the promptly produced charm from $c\bar{c}$ has negligible contribution to the background in the semileptonic B decays, which is consistent with our estimate using `PYTHIA`.

6.4 Background Summary

The fraction of each type of background in our semileptonic signal is summarized below. The dominant signal contamination is from the physics background. The second largest background arises from muon fakes. The smallest background source is from $b\bar{b}$ and $c\bar{c}$.

Table 37: Summary of the backgrounds of the semileptonic modes.

	Fraction (%)		
Background Type	$D^*\mu$	$D\mu$	$\Lambda_c\mu$
Physics	15	40	9.8
Muon fakes	4.3	4.9	3.2
$b\bar{b}$ and $c\bar{c}$	0.9	1.2	0.2
Total	20.2	46.1	13.8

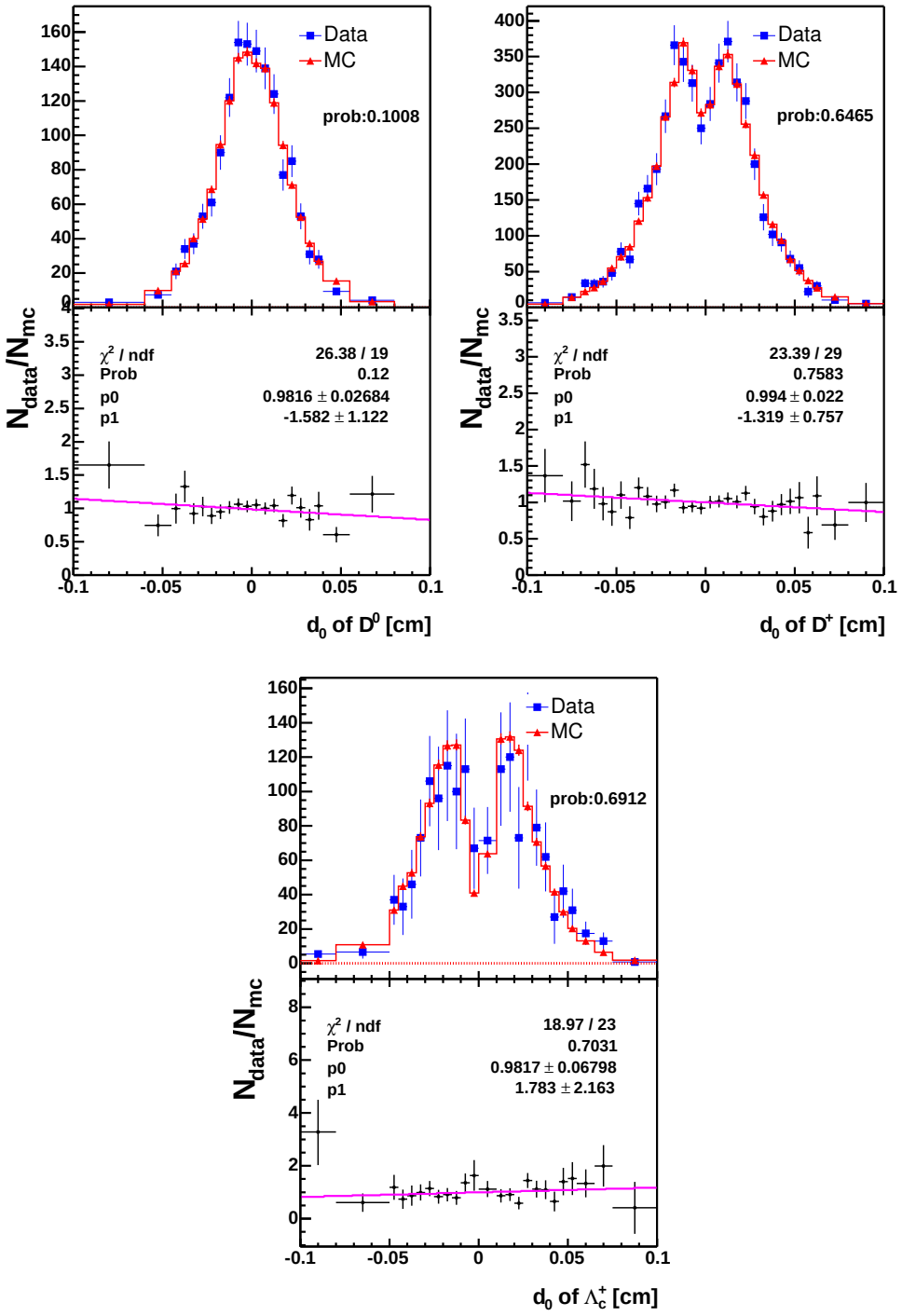


Figure 66: Impact parameters of the charm hadrons, from the top left to the bottom are D^0, D^+ and Λ_c^+: MC and data comparison. The good agreement of the MC with the data indicates that background from the promptly produced charm ($c\bar{c}$) is negligible.

7 Relative Branching Fraction Results and Systematics

We reconstruct hadronic and inclusive semileptonic signals in the data and fit the mass spectra to obtain the yields, $N_{\rm had}$ and $N_{\rm inclusive\ semi}$, as described in Section 4. We estimate the backgrounds of our semileptonic signals, $N_{\rm physics}$, $N_{{\rm fake}\mu}$ and $N_{c\bar{c},\ b\bar{b}}$, in Section 6. The ratio of branching fractions is the yield ratio with each mode corrected for acceptance and reconstruction efficiency.

$$\frac{\mathcal{B}_{\rm semi}}{\mathcal{B}_{\rm had}} = \left(\frac{N_{\rm inclusive\ semi} - N_{\rm physics} - N_{{\rm fake}\mu} - N_{c\bar{c},\ b\bar{b}}}{N_{\rm had}}\right) \times \frac{\epsilon_{\rm had}}{\epsilon_{\rm semi}}, \tag{65}$$

where $\mathcal{B}$ stands for the branching ratio, ϵ is the efficiency from the MC as described in Section 5.3. With every component on the right hand side of Equation 65 in hand, we are ready to calculate the relative branching fractions. The relative branching fractions with the statistical uncertainties only are:

$$\begin{aligned}
\frac{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)} &= 17.7\ \pm 2.3, \\
\frac{\mathcal{B}(\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{+}\pi^-)} &= 9.8\ \pm 1.0, \\
\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)} &= 20.0\ \pm 3.0.
\end{aligned}$$

In this section, we first discuss and estimate the systematic uncertainties. Then, we show the result of the relative branching fractions for each mode. Finally, a conclusion is given at the end of the section.

7.1 Systematic Uncertainties

7.2 Sources of Systematics

Systematic uncertainties in our measurements may arise from the difference in the semileptonic and hadronic decays, from the lack of knowledge of certain backgrounds, and from the uncertainties on the external information. Most of the sources are common to all the decay modes. Systematic uncertainties which affect only one mode are discussed separately. To simplify the notation, we define our measurement of the relative branching fractions as R:

$$R \equiv \frac{\mathcal{B}_{\rm semi}}{\mathcal{B}_{\rm had}},$$

and σ_R is the systematic uncertainty. We also denote the term, "branching ratio" in the text, as $\mathcal{BR}$, while the branching ratio of one specific mode is denoted as $\mathcal{B}$(mode), eg: $\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$.

Mass Fitting

- $\overline{B} \to D^{*+}\mu^- X$ and $\overline{B} \to \Lambda_c^+\mu^- X$: The mass functions are general and cover all the possible backgrounds. Because the functions do not involve any external $\mathcal{BR}$ or MC efficiencies, we do not assign systematic uncertainty for the mass fitting of these two modes.

- $\overline{B} \to D^+\mu^- X$: The uncertainties on the D_s decay $\mathcal{BR}$ can modify the misreconstructed D_s mass spectrum in $\overline{B} \to D^+\mu^- X$. In addition, the mean of the Gaussian constraint for the amount of D_s background in Equation 16, μ_p, also changes accordingly. We study this effect by varying the following numbers in Table 5 $\pm$ 1 σ independently: the $\mathcal{BR}$ of $D_s^+ \to \phi\pi^+$, and the $\mathcal{BR}$ of each selected D_s decay relative to $\mathcal{B}(D_s^+ \to \phi\pi^+)$, because these D_s $\mathcal{BR}$ were measured relative to the $\phi\pi$ mode [25]. The corresponding D_s background shape and the μ_p are re-evaluated for each

change of D_s $\mathcal{BR}$. The changes in the yield are added in quadrature to get the accumulative variation. Table 38 summarizes the yield variations. Modes which are not listed give identical results to the central value. We find total $\Delta(N_{\overline{B}\to D^+\mu^- X}) = 33$ events and $\sigma_R = 0.13$. Note that a few D_s decays in Table 5 only have an upper limit in the PDG and the estimated values in the `EvtGen` decay table are used. We assign 100% uncertainty for these modes.

Table 38: $\overline{B} \to D^+\mu^- X$ yield change due to the variation of D_s $\mathcal{BR}$.

	$\Delta(N_{\overline{B}\to D^+\mu^- X})$
$D_s^+ \to K^+K^-\pi^+$	± 23
$D_s^+ \to \phi K^+$	$\pm$ 2
$D_s^+ \to \eta\pi^+$	$\pm$ 2
$D_s^+ \to \eta'\pi^+$	$\pm$ 3
$D_s^+ \to \omega\pi^+$	$\pm$ 1
$D_s^+ \to \rho^0 K^+$	$\pm$ 1
$D_s^+ \to f_2\pi^+$	$\pm$ 1
$D_s^+ \to \rho^+\eta$	$\pm$ 3
$D_s^+ \to \rho^+\eta'$	$\pm$ 2
$D_s^+ \to K^0K^+$	$\pm$ 2
$D_s^+ \to K^{*0}K^+$	± 22
$D_s^+ \to \overline{K}^{*0}\pi^+$	$\pm$ 2
Total	± 33

- $\overline{B}^0 \to D^{*+}\pi^-$: the composition of the remaining $\overline{B} \to D^{*+}X$ background can affect the shape of its mass spectrum, and its ratio to the $D^*\rho$ background. The latter changes the mean of the Gaussian constraint, μ_2, in Equation 24. We study the systematics by varying the $\mathcal{BR}$ of $\overline{B}^0 \to D^{*+}\rho^-$ and the dominant modes in the remaining $\overline{B} \to D^{*+}X$ background. The change of signal yield from each variation of $\mathcal{BR}$ is listed in Table 39. The accumulative yield change is only $^{+0.1}_{-0.2}$ events, which is insignificant. Therefore, we do not assign systematic uncertainty for the mass fitting of this decay mode.
- $\overline{B}^0 \to D^+\pi^-$: the systematic uncertainties come from three sources: the normalizations of

Table 39: $\overline{B}^0 \to D^{*+}\pi^-$ yield change due to the variation of the background $\mathcal{BR}$.

Mode	$\mathcal{BR}$ (%)	ΔN
$\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$	5.44 ± 0.23	< 0.1
$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	5.44 ± 0.23	< 0.1
$\overline{B}^0 \to D^{*+}\pi^-\pi^0$	$0.7 \pm$?	< 0.1
$\overline{B}^0 \to D^{*+}a_1^-$	1.30 ± 0.27	< 0.1
$\overline{B}^0 \to D^{*+}\rho^-$	0.68 ± 0.09	$^{+0.1}_{-0.2}$
Total		$^{+0.1}_{-0.2}$

the Cabibbo suppressed decay, B_s and Λ_b backgrounds, the uncertainties of the background function fit to the MC, and the $\mathcal{BR}$ of the backgrounds. We study the effect of the first two sources in the following way: We vary each constant parameter including the normalizations and the shape parameters in Table 14 $\pm$ 1 σ, independently. The changes of yield (ΔN) are listed in Table 40. The normalizations of the backgrounds from the Cabibbo suppressed decay, B_s and Λ_b decays are independent from the shape parameters. But, several shape parameters for the same background are correlated, as shown in Tables 43–45. In order to take into account the correlation properly, the correlated shape parameters are grouped together. We calculate the product of the correlation coefficient matrix ($\mathbf{M}$), with the row and column vectors of $\mathbf{\Delta N}$, to obtain a total systematic uncertainty. For instance, the systematic uncertainty from the B_s background shape parameters is:

$$\sigma_N^2 = \left(\begin{array}{cccc} \Delta N_\mu & \Delta N_{\sigma_1} & \Delta N_{f_1} & \Delta N_{\frac{\sigma_2}{\sigma_1}} \end{array} \right) \mathbf{M} \left(\begin{array}{c} \Delta N_\mu \\ \Delta N_{\sigma_1} \\ \Delta N_{f_1} \\ \Delta N_{\frac{\sigma_2}{\sigma_1}} \end{array} \right), \tag{66}$$

where $\mathbf{M}$ is a 4×4 correlation coefficient matrix returned from the fit to the B_s MC (see Table 43). The value of ΔN for each parameter is listed in Table 40.

For the systematics associated with the $\mathcal{BR}$, we vary the $\mathcal{BR}$ of $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B}^0 \to D^+\rho^-$ and the dominant modes in the remaining $\overline{B} \to D^+X$ backgrounds, $\pm$ 1 σ independently. We re-fit the background shapes using the MC, fix the shape parameters and re-fit the data. Table 41 lists the signal yield change due to the variation of the $\mathcal{BR}$. Table 42 summarizes the signal yield change from the variation of the shape parameters and the $\mathcal{BR}$. These changes are added in quadrature to get the accumulative difference. The total change in the yield is 13 events, which modifies R by 0.38.

- $\Lambda_b \to \Lambda_c^+\pi^-$: we follow the same scheme applied by Martin [26]. Using a generic B-decay MC, we first extract the top twenty largest contributing modes in the mass region $5.3 < M_{\Lambda_c\pi} < 6.0$ GeV/c^2, from each type of background: four-prong B meson, the remaining B meson decays, and the remaining Λ_b decays. Each dominant decay contributes $N^i_{\rm base}$ events. Then, we generate a new distribution for each dominant mode, according to the shape determined from a large single-decay MC. The normalization of the new distribution is first Gaussian fluctuated with a mean $N^i_{\rm base}$, a sigma of $\Delta(\mathcal{BR})/(\mathcal{BR})$ and then Poisson fluctuated. For the measured decays, $\Delta(\mathcal{BR})$ is the uncertainty reported in the PDG. For the unmeasured B meson decays, $\Delta(\mathcal{BR})$ is assumed to be three times the uncertainty of the closest equivalent mode in the measured B meson decays. For the unmeasured Λ_b decays, $\Delta(\mathcal{BR})$ is hypothesized to be $^{+100}_{-50}$% of the $\mathcal{BR}$. These Gaussian and Poisson fluctuated distributions are then re-combined with the other non-dominant modes. The combined background mass spectrum is refitted and the newly derived shape parameters are fixed in the fit to the data. The whole procedure is repeated 1000 times with different random seeds for the Gaussian and Poisson fluctuations. We plot the distribution of the $\Lambda_b \to \Lambda_c^+\pi^-$ yield and record the RMS as the change in the yield due to the variation of the $\mathcal{BR}$. Figure 67 shows an example of the $\Lambda_b \to \Lambda_c^+\pi^-$ yield distribution from the $\mathcal{BR}$ variation of the four-prong B meson background. In addition, we vary the fraction of the Cabibbo suppressed mode, $f_{\Lambda_c K}$, $^{+100}_{-50}$% and record the yield change. Table 46 summarizes the change of $\Lambda_b \to \Lambda_c^+\pi^-$ yield. The accumulative σ_R is 0.63.

Measured Branching Fractions

We use the $\mathcal{BR}$ from the world average in the PDG to estimate the physics backgrounds in our semileptonic signals as described in Section 6.1. We vary the $\mathcal{BR}$ of these measured physics backgrounds by $\pm$ 1 σ. We then calculate σ_R. Note that here the variation of the $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ does not include

Table 40: $\overline{B}^0 \to D^+\pi^-$ yield change due to an independent variation of the fixed parameter value.

Parameter		ΔN
f_{DK}	$N_{\overline{B}^0 \to D^+K^-}/N_{\overline{B}^0 \to D^+\pi^-}$	± 4.97
ΔM_{DK}	mass shift of $\overline{B}^0 \to D^+K^-$	± 3.82
σ_{DK}	width of $\overline{B}^0 \to D^+K^-$	± 0.82
f_{B_s}	$N_{\overline{B}_s \to D_s^+\pi^-}/N_{\overline{B}^0 \to D^+\pi^-}$	± 0.51
μ_{B_s}	mean of B_s background	± 0.02
f_1	fraction of the narrow B_s Gaussian	± 0.00
σ_1	width of the narrow B_s Gaussian	± 0.10
σ_2/σ_1	width ratio of the B_s Gaussians	± 0.04
f_{Λ_b}	$N_{\Lambda_b \to \Lambda_c^+\pi^-}/N_{\overline{B}^0 \to D^+\pi^-}$	± 0.22
μ_{Λ_b}	mean of Λ_b	± 0.22
σ_{Λ_b}	width of Λ_b background	± 0.02
τ_{Λ_b}	lifetime of Λ_b background	± 0.21
$\tau_{\rm ref}$	lifetime of $D\rho$ background	± 2.62
$\sigma_{\rm ref}$	width of $D\rho$ background	± 0.27
f_H	fraction of $D^*\pi$ horns	± 6.38
$\delta_{\rm ref}$	distance between two horns	± 1.04
σ_H	width of the horns	± 2.70
$f_{\rm otherB}$	fraction of the remaining $\overline{B} \to D^+X$	± 1.90
$M_{\rm off}$	cut off for $\overline{B} \to D^+X$ mass	± 1.02

Table 41: $\overline{B}^0 \to D^+\pi^-$ yield change due to the variation of the background $\mathcal{BR}$.

Mode	$\mathcal{BR}$ (%)	ΔN
$\overline{B}^0 \to D^+\rho^-$	0.77 ± 0.13	$^{+0.5}_{-3.4}$
$\overline{B}^0 \to D^{*+}\pi^-$	0.276 ± 0.021	$^{+2.2}_{-1.3}$
$\overline{B}^0 \to D^+e^-\overline{\nu}_e$	2.14 ± 0.20	$^{+0.6}_{-0.2}$
$\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$	2.14 ± 0.20	$^{+1.0}_{-0.4}$
$\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$	5.44 ± 0.23	$^{+0.1}_{-0.2}$
$\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$	5.44 ± 0.23	± 0.2
$\overline{B}^0 \to D^+\pi^-\pi^0$	$0.1 \pm ?$	± 0.5
$\overline{B}^0 \to D^{*+}\pi^-\pi^0$	$0.7 \pm ?$	± 1.0
$\overline{B}^0 \to D^+a_1^-$	0.60 ± 0.33	± 3.0
$\overline{B}^0 \to D^{*+}a_1^-$	1.30 ± 0.27	± 0.1
$\overline{B}^0 \to D^{*+}\rho^-$	0.68 ± 0.09	± 0.6
Total		± 4.5

Table 42: Systematic uncertainty on the $\overline{B}^0 \to D^+\pi^-$ yield from each independent parameter group.

	ΔN
f_{DK}	± 5.0
DK shape	± 3.9
f_{B_s}	± 0.5
B_s shape	± 0.1
f_{Λ_b}	± 0.2
Λ_b shape	± 0.4
$D\rho + D^*\pi$ shape	± 9.9
f_{otherB}	± 1.9
M_{off}	± 1.0
$\mathcal{BR}$	± 4.5
Total	± 12.8

Table 43: Correlation coefficients returned from the fit to B_s MC.

	N	μ_{B_s}	σ_1	f_1	σ_2/σ_1
N	1.000				
μ_{B_s}	0.024	1.000			
σ_1	-0.003	-0.084	1.000		
f_1	0.007	-0.004	0.847	1.000	
σ_2/σ_1	0.023	0.133	0.268	0.647	1.000

Table 44: Correlation coefficients returned from the fit to Λ_b MC.

	N	μ_{Λ_b}	σ_{Λ_b}	τ_{Λ_b}
N	1.000			
μ_{Λ_b}	0.000	1.000		
σ_{Λ_b}	0.000	-0.624	1.000	
τ_{Λ_b}	0.000	0.699	-0.508	1.000

Table 45: Correlation coefficients returned from the fit to $D^*\pi$ and $D\rho$ MC.

	N	τ_{ref}	μ_{ref}	σ_{ref}	f_H	δ_{ref}	σ_H	ν_{ref}
N	1.000							
$\tau_{\rm ref}$	0.000	1.000						
$\mu_{\rm ref}$	0.000	0.013	1.000					
$\sigma_{\rm ref}$	0.000	0.169	-0.841	1.000				
f_H	0.000	0.549	-0.688	0.720	1.000			
$\delta_{\rm ref}$	0.000	0.294	0.284	-0.429	-0.435	1.000		
σ_H	0.000	0.407	-0.507	0.507	0.699	-0.381	1.000	
$\nu_{\rm ref}$	0.000	0.029	0.971	-0.786	-0.624	0.261	-0.455	1.000

Table 46: $\Lambda_b \to \Lambda_c^+\pi^-$ yield change due to the variation of the background $\mathcal{BR}$.

	ΔN
$\Lambda_b \to \Lambda_c^+ K^-$	$^{+1.1}_{-2.8}$
four-prong B meson decays	$\pm$ 2.9
remaining B meson decays	$\pm$ 0.9
all the other Λ_b decays	$\pm$ 2.8
Total	$^{+4.3}_{-5.0}$

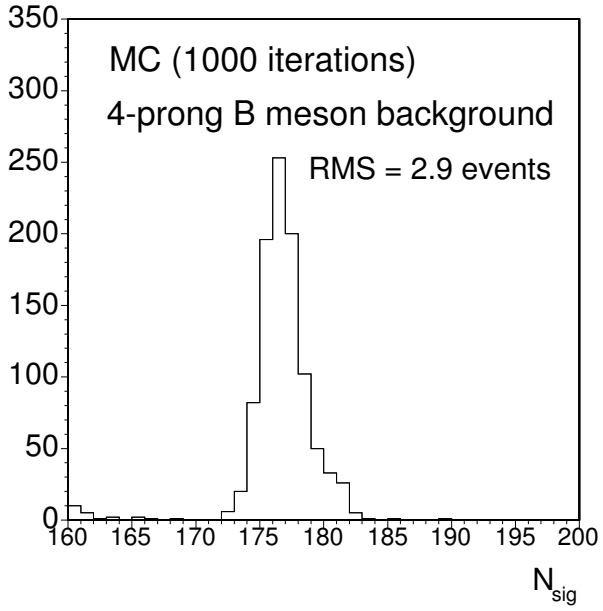


Figure 67: $\Lambda_b \to \Lambda_c^+ \pi^-$ yield from 1000 variations of the 4-prong B meson background $\mathcal{BR}$. The RMS is recorded as the yield change.

the uncertainty due to the measured Λ_b P_T spectrum. The σ_R due to the measured $\mathcal{BR}$ is 0.43, 0.75 and $^{+0.73}_{-2.07}$ for $\overline{B}^0 \to D^{*+}$, $\overline{B}^0 \to D^+$ and $\Lambda_b \to \Lambda_c^+$ modes, respectively. For the remainder of this section, we quote the systematic uncertainties in the same order.

Unmeasured Branching Fractions

The $\mathcal{BR}$ of several physics backgrounds in Section 6.1 have not yet been measured, *e.g.*: $\Lambda_b \to \Lambda_c^+ f^0 \mu^- \overline{\nu}_\mu$, or have just been measured by us for this analysis, *e.g.*:$\Lambda_b \to \Lambda_c(2593)^+ \mu^- \overline{\nu}_\mu$. For the first case, we use the estimated $\mathcal{BR}$ from the decay file of `EvtGen`, and our own derivation based on HQET. As we have no uncertainty input from the estimated $\mathcal{BR}$, we assign a 5% uncertainty to the $\mathcal{BR}$ of the excited charm meson decays and a 100% uncertainty to the $\mathcal{BR}$ of the B hadron decays. Because the excited charm hadrons decay via strong interaction and conserve isospin symmetry, their $\mathcal{BR}$ could be inferred from Clebsch-Gordan Coefficients. While for the weak decays of B hadrons, allowable decay spectrum is wider. For the second case, we add $(20 \oplus 20)\%$ uncertainty in quadrature with the uncertainty from the preliminary measurement (see Table 26). The first 20% arises from the unresolved disagreement of measured τ_{Λ_b} with that from the HQET prediction. The second 20% is due to the difference of the soft pion reconstruction efficiency between MC and data. We vary the $\mathcal{BR}$ by the uncertainties we assigned and calculate the shift of our measurement. The shift due to the unmeasured $\mathcal{BR}$ is 1.09, 0.91 and 0.50.

Fake μ estimate

As noted in Section 6.2.1, the systematic uncertainties from the fake μ estimate originate from:

1. The uncertainty from the fit to the charm mass spectra.

Table 47: Summary of fake muon contamination.

$\overline{B} \to D^{*+}\mu_{fake}$	$\overline{B} \to D^{+}\mu_{fake}$	$\overline{B} \to \Lambda_c^{+}\mu_{fake}$
45 ± 3	230 ± 19	40 ± 9

2. The uncertainty on the probabilities for the pions, kaons and protons to fake muons.

3. The uncertainty on the fraction of pion, kaon and proton in the hadron tracks.

For each category, we vary the central value ± 1 σ, independently. More detailed description can be found in Section 6.2.1. The resulting uncertainty on the amount of fake muons together with the central value are summarized in Table 47. We then vary the number of fake muons ± 1 σ and insert the new number into Equation 65. The total variation on R due to the fake μ estimate is 0.07, 0.07, 0.17.

$b\bar{b}$ and $c\bar{c}$ background

In Section 6.3, we notice a 10–40% discrepancy of the D^0, B^+, and inclusive b cross-section from `PYTHIA` with those from the data. In addition, we do not possess information about the relative $b\bar{b}$ and $c\bar{c}$ production rates between flavor creation, flavor excitation, and gluon splitting. Therefore, we assign a 100% uncertainty to the amount of $b\bar{b}$ and $c\bar{c}$ backgrounds. This changes R by 0.22, 0.22, 0.04.

MC sample size

We have generated large MC samples for calculating the efficiencies of our signals and backgrounds, but there is a small statistical uncertainty due to the finite MC sample size. We use the uncertainties on the efficiencies to calculate σ_R. σ_R is 0.28, 0.18, and 0.32.

MC $P_T(B)$ Spectrum

We find discrepancies between data and MC in the P_T spectrum of B^0 and Λ_b, as described in Section 5.1. After reweighting the P_T spectrum of B^0 and Λ_b, we have observed good agreement of MC with the data as seen in Section 5.2. However, there is an uncertainty on the exponential slope of data/MC, p_1 in Figure 36, which is limited by the amount of data used for comparison with the MC. We vary p_1 ± 1 σ and re-weight the MC events after the analysis cuts numerically to calculate the efficiency change. In addition, we vary the variables which depend on the Λ_b P_T spectrum accordingly, eg: cross-section correction factors and the result of $\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b\to\Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0\to D^+\pi^-)}$ by Le, *et al.* [4]. The total variation on R due to the MC P_T spectrum of B hadrons is 0.38, 0.32, and $^{+0.28}_{-0.50}$.

Pion Interaction with the Material

One difference between our semileptonic and hadronic final states is the muon and the pion. The muon does not interact with the material via the hadronic (strong) interaction while the pion does. In order to model the track reconstruction efficiency correctly, two things have to be right:

1. The type and the amount of material in the detector.

2. The model that describes the hadronic interaction cross-section, the final state multiplicities and kinematics.

We generate MC for the signals as described in Section 5.1 except that we switch off the hadronic interaction in the detector simulation. We compare the difference in the hadronic to semileptonic signal efficiency ratio, between the normal MC and the MC with the hadronic interaction off. This difference gives us an idea for the extreme situation, when the material is 100% wrong. For both the B^0 and Λ_b modes, the efficiency ratio changes by 4%. From the study of Korn [27], we know that the available CDF detector simulation underestimate the amount of material by 15%. In addition, a comparison between two programs which model the hadronic interactions, `GHEISHA` and `FLUKA` [28], has been done by Michael [29]. The `FLUKA` package is known to better reproduce the experimental data but currently it is not available in the CDF detector simulation. The effect of the hadronic interaction model estimated by Michael is 20%. Adding 15% and 20% in quadrature, we get 25%. We multiply the 100% efficiency ratio difference described earlier, with 0.25, and get 1%. We apply a scaling factor, 1.01, to all the relative efficiencies, including the semileptonic background to hadronic efficiency ratios. We then re-calculate R and find σ_R is 0.22, 0.17 and 0.22.

CMU reconstruction efficiency scaling factor

The scaling factor to correct the difference of CMU muon reconstruction efficiency between MC and data has an uncertainty as described in Section 5.3. We vary the scaling factor $\pm$ 1 σ and calculate σ_R= 0.07, 0.05, and 0.07.

XFT efficiency scaling factor

We apply the XFT efficiency scaling factors data/MC in bins of inverse P_T from Herndon [11] to correct the signal and background efficiencies. The uncertainty on the kaon and pion scaling factors are varied $\pm$ 1 σ to evaluate σ_R. For the proton scaling factor, due to the limited statistics, we evaluate the systematic uncertainty following the suggestion in Herndon's analysis: we compare the difference by applying a constant efficiency scaling factor as shown in Figure 68, instead of the one based on the third order polynomial in Figure 57. For all the three modes, the systematic uncertainties are negligible as expected, since the final states of our semileptonic and hadronic modes are almost identical and the difference in the ionization of the pion and muon is insignificant. σ_R is less than 0.01.

Λ_b and Λ_c polarizations

There is not yet a precise measurement of the production polarizations of Λ_b and Λ_c, while the Standard Model predicts both particles are produced polarized. The angular distribution of of the Λ_b daughters is parameterized by

$$\frac{dN}{d\cos\Theta} \propto 1 + \mathcal{P}_{\mathcal{B}} \cos\Theta, \tag{67}$$

where $\mathcal{P}_{\mathcal{B}}$ is the product of the Λ_b polarization and the asymmetry parameter of the weak decay. Θ is defined as the angle between the Λ_c^+ momentum in the Λ_b rest frame and the axis normal to the beam proton-Λ_b production plane, $\hat{\mathbf{n}}$. Therefore,

$$\cos\Theta = \hat{\mathbf{P}}(\Lambda_c) \cdot \hat{\mathbf{n}}, \tag{68}$$

where

$$\hat{\mathbf{n}} \equiv \frac{\hat{\mathbf{P}}(p) \times \hat{\mathbf{P}}(\Lambda_b)}{|\hat{\mathbf{P}}(p) \times \hat{\mathbf{P}}(\Lambda_b)|}, \tag{69}$$

Here "$\times$" ("$\cdot$") means vector (scalar) product of two vectors. See Figure 69 for the definition of Θ and $\hat{\mathbf{n}}$. The angular distribution of Λ_c daughters is parameterized in a similar way;

$$\frac{dN}{d\cos\theta} \propto 1 + \mathcal{P}_{\mathcal{C}} \cos\theta, \tag{70}$$

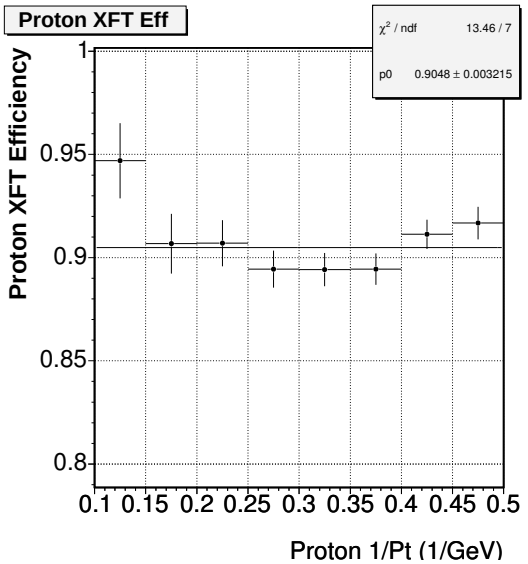


Figure 68: The relative proton XFT efficiency between MC and data in bins of $1/P_T$ fit to a constant by Herndon [11].

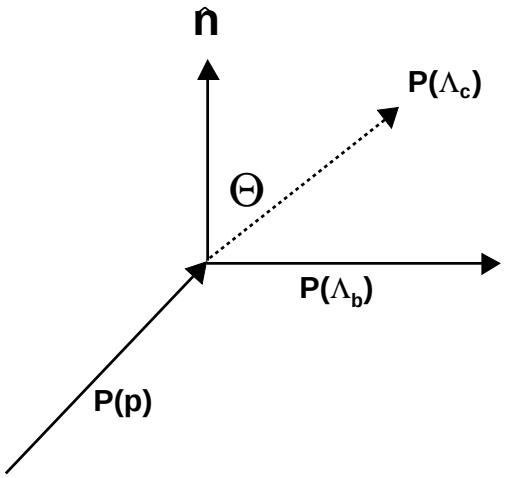


Figure 69: Angle definition for the Λ_b production polarization, where the dashed line indicates the momentum of Λ_c in the rest frame of Λ_b, and $\hat{\mathbf{n}}$ is the polarization axis normal to the beam proton-Λ_b production plane.

Table 48: $\epsilon(\Lambda_b \to \Lambda_c^+\pi^-)/\epsilon(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)$ from each combination of $\mathcal{P}_\mathcal{B}$ and $\mathcal{P}_\mathcal{C}$.

$\mathcal{P}_\mathcal{B}$	$\mathcal{P}_\mathcal{C}$	$\frac{\epsilon(\Lambda_b\to\Lambda_c^+\pi^-)}{\epsilon(\Lambda_b\to\Lambda_c^+\mu^-\overline{\nu}_\mu)}$	Scaling factor
0	0	3.225 ± 0.010	1
0	-1	3.193 ± 0.007	0.990 ± 0.004
0	1	3.281 ± 0.006	1.017 ± 0.004
1	0	3.232 ± 0.006	1.002 ± 0.004
1	-1	3.193 ± 0.009	0.990 ± 0.004
1	1	3.275 ± 0.009	1.016 ± 0.004
-1	0	3.235 ± 0.006	1.003 ± 0.004
-1	1	3.274 ± 0.009	1.015 ± 0.004
-1	-1	3.175 ± 0.011	0.985 ± 0.005

where θ is defined as the angle between the proton momentum in the rest frame of Λ_c and the Λ_c momentum in the lab frame, i.e.

$$\cos\Theta = \hat{\mathbf{P}}(\Lambda_c) \cdot \hat{\mathbf{p}}, \tag{71}$$

Note that here $\hat{\mathbf{P}}(\Lambda_c)$ is in a different frame from that in Equation 68. The value of $\mathcal{P}_\mathcal{B}$ ($\mathcal{P}_\mathcal{C}$) is $\pm$ 1 for the polarized state and 0 for the unpolarized state.

The `Bgenerator` and `EvtGen` do not include the polarization of Λ_b and Λ_c. We use the default settings for the central value of R. We study the systematics due to a non-null polarization using the generator-level signal MC without the detector and trigger simulation. We use the "acceptance-rejection (Von Neumann)" method [8] and reweight the MC according to:

$$(1 + \mathcal{P}_\mathcal{B}\cos\Theta) \cdot (1 + \mathcal{P}_\mathcal{C}\cos\theta),$$

where all combinations of $\mathcal{P}_\mathcal{B}$ and $\mathcal{P}_\mathcal{C}$ for values at -1, 0, 1 are used. Each MC starts from a different random seed. We apply generator-level analysis-like cuts and obtain the efficiency ratio for each combination. We compare these efficiency ratios with that from the MC generated with zero $\mathcal{P}_\mathcal{B}$ and $\mathcal{P}_\mathcal{C}$. We find the efficiency ratios are mainly determined by the $\mathcal{P}_\mathcal{C}$ as seen in Table 48, i.e. the efficiency ratios with the same $\mathcal{P}_\mathcal{C}$, but different $\mathcal{P}_\mathcal{B}$ are consistent with each other. Therefore, we apply scaling factors from the two $\mathcal{P}_\mathcal{C}$ values: -1 and 1 ($\pm$ 1.017%) on all the relative efficiencies and re-calculate R. We find $\sigma_R = 0.37$.

Λ_c Dalitz structure

The Λ_c from our $\Lambda_b \to \Lambda_c^+\pi^-$ and $\overline{B} \to \Lambda_c^+\mu^- X$ signal, decays into p, K, and π in the final state. However, any two Λ_c daughters could form an intermediate resonant state, see Table 49. The resonant structure is called the "Dalitz" structure in the literature, and is usually displayed with a Dalitz plot [30], where the invariant mass square of one pair of daughters is plotted versus another pair in the two-dimension. Figure 70 (left) shows the Dalitz plot from the $\overline{B} \to \Lambda_c^+\mu^- X$ data after sideband subtraction. If a resonance exists, a concentrated area near the mass of the resonant particle will be visible. The momenta of p, K and π are affected by the Dalitz structure and Λ_b decays have different efficiencies for various structures. However, `EvtGen` does not take into account the interference of each resonant state. Each state is considered as an independent decay with a $\mathcal{BR}$ measured by E791 [31] and listed in Table 49. See Figure 70 (right) for the Λ_c Dalitz structure in the MC.

Without a better model to describe the Λ_c Dalitz structure, we study the change in the efficiency ratio of $\Lambda_b \to \Lambda_c^+\pi^-$ to $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ by varying the $\mathcal{BR}$ in Table 49 $\pm$ 1σ. We generate four sets

Table 49: Λ_c decays with p, K, π in the final state.

Decay Mode	$\mathcal{BR}(\%)$	$\frac{\epsilon(\Lambda_b\to\Lambda_c^+\pi^-)}{\epsilon(\Lambda_b\to\Lambda_c^+\mu^-\overline{\nu}_\mu)}$	R_j
$pK^*(890)^0$	1.6 $\pm$ 0.5	3.17$\pm$ 0.01	3.22
$\Delta(1232)^{++}K^-$	0.86$\pm$ 0.30	3.24$\pm$ 0.01	3.23
$\Lambda(1520)\pi^+$	0.59$\pm$ 0.21	3.30$\pm$ 0.01	3.23
non-resonant	2.8 $\pm$ 0.8	3.24$\pm$ 0.01	3.23
R_c			3.23 $\pm$ 0.01

of $\Lambda_b \to \Lambda_c^+\pi^-$ and $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ MC samples without detector and trigger simulation, where Λ_c decay is forced to one single mode. The efficiency of Λ_b decay with $\Lambda_c^+ \to pK^-\pi^+$(total), is then the sum of the $\mathcal{BR}$ weighted efficiency of each individual Λ_c mode.

$$\mathcal{E}_c = \frac{\sum_i^4 \mathcal{BR}_i \cdot \epsilon_i}{\sum_i^4 \mathcal{BR}_i}, \tag{72}$$

$$R_c = \frac{\mathcal{E}_c(\Lambda_b \to \Lambda_c^+\pi^-)}{\mathcal{E}_c(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}, \tag{73}$$

where $\mathcal{E}_c$ (R_c) is the total weighted efficiency (ratio) using the central value of each Λ_c $\mathcal{BR}$. We re-calculate the absolute and relative efficiency, by varying $\mathcal{BR}$ of each Λ_c decay $\pm$ 1 σ;

$$\mathcal{E}_j = \frac{\sum_i^3 \mathcal{BR}_i \cdot \epsilon_i + (\mathcal{BR}_j + \sigma_j) \cdot \epsilon_j}{\sum_i^4 \mathcal{BR}_i + \sigma_j}, \tag{74}$$

$$R_j = \frac{\mathcal{E}_j(\Lambda_b \to \Lambda_c^+\pi^-)}{\mathcal{E}_j(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}, \tag{75}$$

where $\mathcal{E}_j$ (R_j) is the total weighted efficiency (ratio) with $\mathcal{BR}$ of j^{th} mode varied by $\pm$ 1 σ and other $\mathcal{BR}$ fixed.

We find a fractional change of 0.3% after adding the difference of each R_j from R_c in quadrature. We apply this fractional change to the relative efficiencies of all the semileptonic backgrounds to the hadronic signal and calculate σ_R= 0.07.

Λ_b Lifetime

The world average Λ_b lifetime is lower than the theoretical prediction. A smaller Λ_b lifetime gives a smaller efficiency for reconstructing Λ_b decays. While we cut on the $c\tau$ of $\Lambda_b \to \Lambda_c^+\pi^-$, the $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ is not fully reconstructed and we actually cut on the pseudo-$c\tau$ of the inclusive semileptonic decays. Therefore, systematics due to the uncertainty on the Λ_b lifetime may not cancel in our measurement. We study this effect by generating $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ and $\Lambda_b \to \Lambda_c^+\pi^-$ MC without detector and trigger simulation. We vary the lifetime of Λ_b $\pm$ 15% around the central value: 1.229 ps. We compare the difference of the relative efficiency ratio from the central value. We then apply a scaling factor from the signals, on the efficiency ratios of the semileptonic backgrounds to the hadronic mode, and calculate σ_R=0.22.

Semileptonic Λ_b decay model

In Section 5.3, we introduce a scaling factor, f_c, which accommodates the acceptance difference between the flat phase space MC and the form factor weighted MC. We vary the f_c $\pm$ 1 σ according to its statistical uncertainty and obtain $\sigma_R = \pm 0.57$.

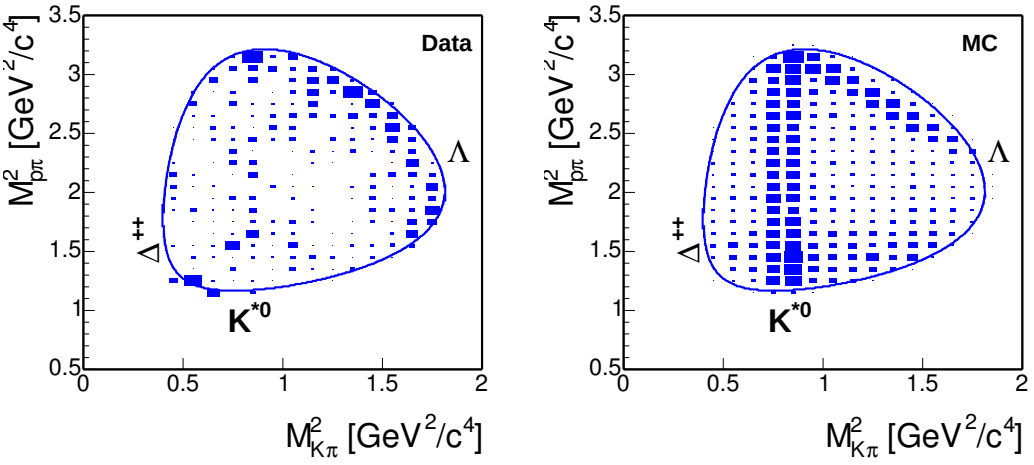


Figure 70: Λ_c Dalitz structure in the sideband subtracted $\overline{B} \to \Lambda_c^+ \mu^- X$ data (left) and MC (right). The concentrated areas in the top figure indicate the existence of $K^*(892)^0$ and $\Lambda(1520)$. Clearly, the destructive interference between the resonant states are not simulated in the MC.

7.3 Systematic Uncertainty for Each Mode

Tables 50–52 list the result of systematic uncertainties as discussed above. The systematics from the external information are separated from the ones from the CDF MC and measurements. Table 53 summarizes the uncertainties from each category. The statistical uncertainties on the relative branching fractions are also listed for comparison.

7.4 Consistency Check of R

In order to detect any unexpected systematics in R, we separate the data and MC into several groups of independent subsets according to the run number, vertex position, $c\tau$ and P_T of the charm and B hadrons, and etc. We cross-check the consistency of the R within each group. Figure 71 displays the result of the cross-check, where the uncertainties in the figure are statistical only. The R from all the subsets are consistent with the other subsets in the same group.

7.5 Measurement Result

For the control modes, we measure the relative branching fractions to be:

$$\frac{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)} = 17.7 \ \pm 2.3\ (stat) \pm 0.6\ (syst) \pm 0.4\ (BR) \pm 1.1\ (UBR),$$

which is consistent with the ratio obtained by the PDG, 19.7 ± 1.7, at the 0.7 σ level. And

$$\frac{\mathcal{B}(\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{+}\pi^-)} = 9.8 \ \pm 1.0\ (stat) \pm 0.6\ (syst) \pm 0.8\ (BR) \pm 0.9\ (UBR),$$

which is consistent with the PDG ratio, 7.8 ± 1.0, at the 1.1 σ level. Finally, we measure the relative Λ_b branching fraction to be:

$$\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)} = 20.0 \ \pm 3.0\ (stat) \pm 1.2\ (syst) \ {}^{+0.7}_{-2.1}\ (BR) \pm 0.5\ (UBR).$$

Table 50: Statistical and systematic uncertainties of $\frac{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)}$.

Source	σ_R
Statistical	$\pm$ 2.3
Measured $\mathcal{BR}$	
$\overline{B}^0 \to D^{*+}\pi^-$	$\pm$ 0.29
$B^- \to D_1^0\mu^-\overline{\nu}_\mu$	$\pm$ 0.31
$\tau \to \mu\overline{\nu}_\mu\nu_\tau$	$<$ 0.01
	$\pm$ 0.43
Unmeasured $\mathcal{BR}$	
$D_1^0 \to D^{*+}\pi^-$	$\pm$ 0.05
$D_1'^0 \to D^{*+}\pi^-$	$\pm$ 0.04
$D_1^+ \to D^{*+}\pi^0$	$\pm$ 0.03
$D_1'^+ \to D^{*+}\pi^0$	$\pm$ 0.02
$B^- \to D_1'^0\mu^-\overline{\nu}_\mu$	$\pm$ 0.70
$B^- \to D^{*+}\pi^-\mu^-\overline{\nu}_\mu$	$\pm$ 0.39
$\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau$	$\pm$ 0.31
$\overline{B}^0 \to D_1^+\mu^-\overline{\nu}_\mu$	$\pm$ 0.53
$\overline{B}^0 \to D_1'^+\mu^-\overline{\nu}_\mu$	$\pm$ 0.34
$\overline{B}^0 \to D^{*+}\pi^0\mu^-\overline{\nu}_\mu$	$\pm$ 0.19
	$\pm$ 1.09
CDF Internal Systematics	
Fitting of $\overline{B}^0 \to D^{*+}\pi^-$	$<$ 0.01
Fake μ estimate	$\pm$ 0.07
$b\bar{b}$ and $c\bar{c}$ background	$\pm$ 0.22
MC sample size	$\pm$ 0.28
MC $P_T(B^0)$	$\pm$ 0.38
π interaction with the material	$\pm$ 0.22
CMU reconstruction efficiency scaling factor	$\pm$ 0.07
XFT efficiency scaling factor	$<$ 0.01
	$\pm$ 0.58

Table 51: Statistical and systematic uncertainties of $\frac{\mathcal{B}(\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^+ \pi^-)}$.

Source	σ_R
Statistical	$\pm$ 1.0
Measured $\mathcal{BR}$	
$\overline{B}^0 \to D^+ \pi^-$	$\pm$ 0.70
$\overline{B}^0 \to D^{*+} \mu^- \overline{\nu}_\mu$	$\pm$ 0.22
$B^- \to D_1^0 \mu^- \overline{\nu}_\mu$	$\pm$ 0.08
$D^{*+} \to D^+ \pi^0$	$\pm$ 0.11
$\tau \to \mu \overline{\nu}_\mu \nu_\tau$	< 0.01
f_s / f_d	$\pm$ 0.01
	$\pm$ 0.75
Unmeasured $\mathcal{BR}$	
$D_1^0 \to D^{*+} \pi^-$	$\pm$ 0.01
$D_1^{\prime 0} \to D^{*+} \pi^-$	$\pm$ 0.01
$B^- \to D_1^{\prime 0} \mu^- \overline{\nu}_\mu$	$\pm$ 0.17
$B^- \to D^+ \pi^- \mu^- \overline{\nu}_\mu$	$\pm$ 0.79
$\overline{B}^0 \to D^+ \pi^0 \mu^- \overline{\nu}_\mu$	$\pm$ 0.39
$\overline{B}^0 \to D^+ \tau^- \overline{\nu}_\tau$	$\pm$ 0.10
$\overline{B}_s \to D^+ K^0 \mu^- \overline{\nu}_\mu$	$\pm$ 0.09
	$\pm$ 0.91
CDF Internal Systematics	
Fitting of $\overline{B}^0 \to D^+ \pi^-$	$\pm$ 0.38
Fitting of $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$	$\pm$ 0.13
Fake μ estimate	$\pm$ 0.07
$b\bar{b}$ and $c\bar{c}$ background	$\pm$ 0.22
MC sample size	$\pm$ 0.18
MC $P_T(B^0)$	$\pm$ 0.32
π interaction with the material	$\pm$ 0.17
CMU reconstruction efficiency scaling factor	$\pm$ 0.05
XFT efficiency scaling factor	< 0.01
	$\pm$ 0.62

Table 52: Statistical and systematic uncertainties of $\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)}$.

Source	σ_R
Statistical	$\pm$ 3.0
Measured $\mathcal{BR}$	
$\Lambda_b \to \Lambda_c^+ \pi^-$	$^{+0.73}_{-2.07}$
$\tau \to \mu \overline{\nu}_\mu \nu_\tau$	< 0.01
Total	$^{+0.73}_{-2.07}$
Unmeasured $\mathcal{BR}$	
$\Lambda_b \to \Lambda_c(2593)^+ \mu^- \overline{\nu}_\mu$	$\pm$ 0.21
$\Lambda_b \to \Lambda_c(2625)^+ \mu^- \overline{\nu}_\mu$	$\pm$ 0.27
$\Lambda_b \to \Sigma_c^0 \pi^+ \mu^- \overline{\nu}_\mu$, $\Lambda_b \to \Sigma_c^+ \pi^0 \mu^- \overline{\nu}_\mu$, $\Lambda_b \to \Sigma_c^{++} \pi^- \mu^- \overline{\nu}_\mu$	$\pm$ 0.24
$\Lambda_b \to \Lambda_c^+ f^0 \mu^- \overline{\nu}_\mu$	$\pm$ 0.05
$\Lambda_b \to \Lambda_c^+ \pi^0 \pi^0 \mu^- \overline{\nu}_\mu$, $\Lambda_b \to \Lambda_c^+ \pi^+ \pi^- \mu^- \overline{\nu}_\mu$	$\pm$ 0.20
$\Lambda_b \to \Lambda_c^+ \tau^- \overline{\nu}_\tau$	$\pm$ 0.10
$B^- \to \Lambda_c^+ \overline{p} \mu^- \overline{\nu}_\mu$	$\pm$ 0.11
$\overline{B}^0 \to \Lambda_c^+ \overline{n} \mu^- \overline{\nu}_\mu$	$\pm$ 0.11
Total	$\pm$ 0.50
CDF Internal Systematics	
Fitting of $\Lambda_b \to \Lambda_c^+ \pi^-$	$\pm$ 0.63
Fake μ estimate	$\pm$ 0.17
$b\bar{b}$ and $c\bar{c}$ background	$\pm$ 0.04
MC sample size	$\pm$ 0.32
MC $P_T(\Lambda_b)$	$^{+0.28}_{-0.50}$
π interaction with the material	$\pm$ 0.22
CMU reconstruction efficiency scaling factor	$\pm$ 0.07
XFT efficiency scaling factor	< 0.01
Λ_b and Λ_c polarizations	$\pm$ 0.37
Λ_c^+ Dalitz structure	$\pm$ 0.07
Λ_b lifetime	$\pm$ 0.22
Semileptonic Λ_b decay model	$\pm$ 0.57
	$^{+1.09}_{-1.16}$

Table 53: Summary of statistical and systematic uncertainties.

	$\frac{\sigma_R}{R}$ (%)		
Source	$\frac{\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)}$	$\frac{\mathcal{B}(\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\overline{B}^0 \to D^{+}\pi^-)}$	$\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)}$
Measured $\mathcal{BR}$ (%)	2.4	7.7	$^{+3.5}_{-10.5}$
Unmeasured $\mathcal{BR}$ (%)	6.2	9.3	2.5
CDF internal (%)	3.3	6.4	6.0
Statistical (%)	13.1	10.2	15

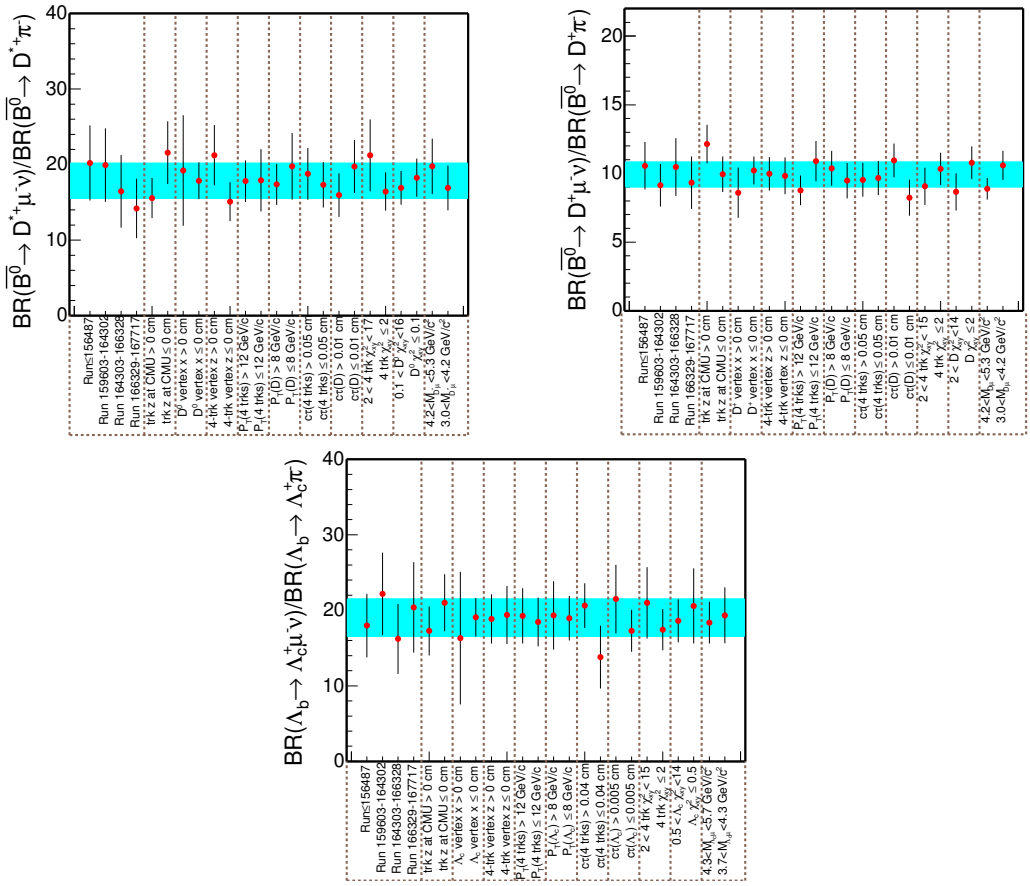


Figure 71: Consistency check of $\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)/\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)$ (top left), $\mathcal{B}(\overline{B}^0 \to D^{+}\mu^-\overline{\nu}_\mu)/\mathcal{B}(\overline{B}^0 \to D^{+}\pi^-)$ (top right) and $\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)/\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ (bottom). The uncertainty on each point is statistical only. Each independent group is separated by a vertical dashed line.

The first uncertainty of the relative branching fraction is due to the statistics of the data, the second is the systematic uncertainty due to the MC and CDF internal measurements used in this analysis, the third is due to the uncertainty of the world averaged branching ratios and production fractions, the last uncertainty is due to the estimate of the unmeasured branching ratios. If we use the branching ratio prediction by Leibovich [32] for the $\Lambda_b \to \Lambda_c(2593)^+\mu^-\overline{\nu}_\mu$ and $\Lambda_b \to \Lambda_c(2625)^+\mu^-\overline{\nu}_\mu$ decays, the central value of the relative branching ratio drops to 17.7; the uncertainties from the measured and unmeasured branching fractions increase to $^{+1.5}_{-4.1}$ and ± 3.0, respectively. If we use the prediction from Huang [33], the central value goes up to 20.4; the uncertainties from the measured and unmeasured branching fractions change to $^{+0.6}_{-1.7}$ and ± 0.8, respectively. Our result above and the numbers using the theoretical predictions are consistent within the uncertainty from the unmeasured branching fractions.

8 Estimate of the $\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)$

We have just presented the first measurement of the ratio of Λ_b exclusive semileptonic to hadronic branching fractions. The ratio provides important input for the absolute branching fraction of $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ or $\Lambda_b \to \Lambda_c^+\pi^-$. Leibovich *et al.* [34] predict $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-) = 0.45\%$ and $\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu) = 6.6\%$, which gives a relative branching fraction of 14.7. However, the largest theoretical uncertainty from the functional form of the Isgur-Wise function is 30%, due to the assumption of the large N_c limit. Our measurement of the ratio has a 19% uncertainty and may stimulate additional theoretical work. Multiplying our Λ_b relative branching fraction, with our derivation of $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ from the CDF measurement of $\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b\to\Lambda_c^+\pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0\to D^+\pi^-)}$ [4] in Yu [1]:

$$\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-) = \left(0.41 \pm 0.19\ (stat \oplus syst)\ ^{+0.06}_{-0.08}\ (P_T)\right)\%,$$

we obtain

$$\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu) = \left(8.1\ \pm 1.2\ (stat)\ ^{+1.1}_{-1.6}\ (syst) \pm 4.3\ (\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-))\right)\%.$$

which is also consistent with a recent DELPHI result derived from the $\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu$ form factor measurement [35],

$$\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)^{\rm DELPHI} = \left(5.0\ ^{+1.1}_{-0.9}(stat)\ ^{+1.6}_{-1.2}(syst)\right)\%$$

Combining our and DELPHI's numbers, we obtain $(5.5 \pm 1.8(stat \oplus syst))\,\%$. Our relative branching ratios and the derived $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$, $\mathcal{B}(\Lambda_b \to \Lambda_c^+\mu^-\overline{\nu}_\mu)$ are all in agreement with the predictions by Leibovich *et al.*, within large uncertainties. Note that the dominant uncertainties of $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ arise from $\frac{\sigma_{\Lambda_b}}{\sigma_{B^0}}$ and $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$. New CDF-II measurements of $\frac{\sigma_{\Lambda_b}}{\sigma_{B^0}}$ are anticipated with reduced uncertainties. A better measurement of $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$ has been proposed by Dunietz [36] and Migliozzi [37]. Improvements in the $\mathcal{B}(\Lambda_b \to \Lambda_c^+\pi^-)$ will reduce the uncertainties in our determination of the exclusive semileptonic branching ratio.

9 Conclusion

We analyze 171.5 pb^{-1} of data collected with the CDF-II detector in the $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Using a novel secondary vertex track trigger, we reconstruct $1237\ \pm 97\ \overline{B} \to \Lambda_c^+\mu^-X$ decays and $179\ \pm 19\ \Lambda_b \to \Lambda_c^+\pi^-$ decays. This is the largest Λ_b sample in the world, which enables us to measure the relative Λ_b branching fractions and examine Heavy Quark Effective Theory. We have also observed several Λ_b semileptonic decays which have never been seen in the other experiments: $\Lambda_b \to \Lambda_c(2593)^+\mu^-X$, $\Lambda_b \to \Lambda_c(2625)^+\mu^-X$, $\Lambda_b \to \Sigma_c^0\pi^+\mu^-X$, and $\Lambda_b \to \Sigma_c^{++}\pi^-\mu^-X$. In addition,

we reconstruct the $\overline{B}^0 \to D^{*+}$ and $\overline{B}^0 \to D^+$ decays similar to our Λ_b decays and use them as the control samples to understand the issues associated with the Λ_b measurement. After the estimate and the subtraction of the background in the inclusive semileptonic signal, we correct the yield observed in the data with the trigger and reconstruction efficiencies obtained from the Monte Carlo. We find the relative branching fraction of the control modes in good agreement with the values obtained by the PDG. We measure the ratio of Λ_b branching fraction to be:

$$\frac{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)} = 20.0 \pm 3.0(stat) \pm 1.2(syst) \begin{matrix} +0.7 \\ -2.1 \end{matrix} (BR) \pm 0.5(UBR).$$

The uncertainty is dominated by the size of the data sample and the branching ratio of $\Lambda_b \to \Lambda_c^+ \pi^-$. More data and a more precise measurement of $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ in the future will immediately improve our relative branching fraction measurement and our determination of $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)$.

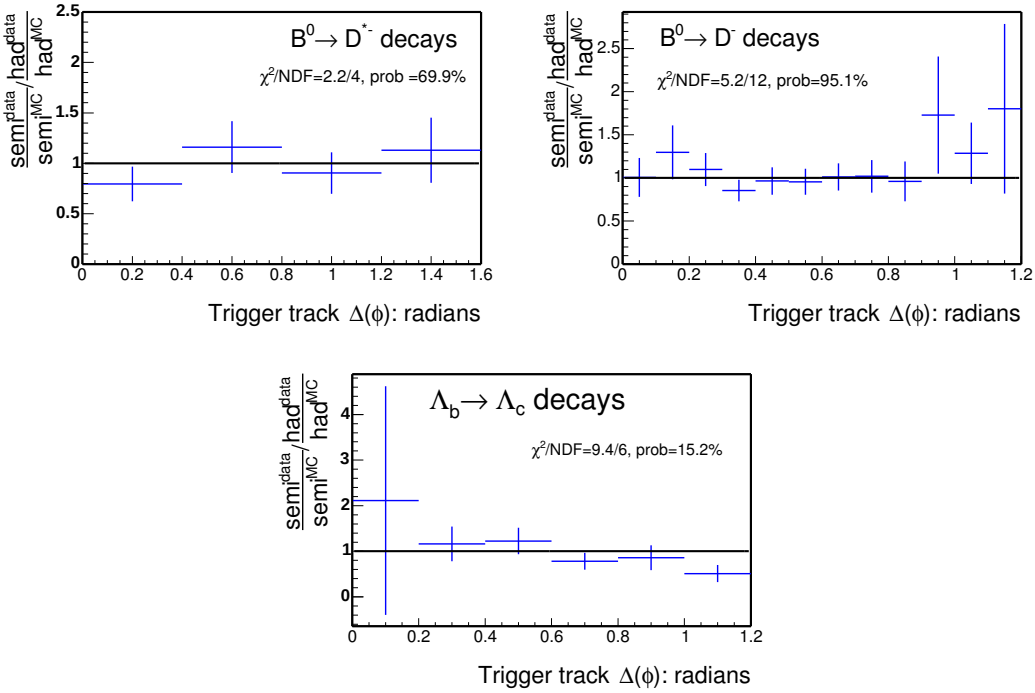


Figure 72: Ratio of $\Delta\phi$ between two trigger tracks of the semileptonic relative to the hadronic decays, from the $\overline{B}^0 \to D^{*+}$ (top left), $\overline{B}^0 \to D^+$ (top right), and $\Lambda_b \to \Lambda_c^+$ (bottom) modes. The ratio is compared to unity and a χ^2 probability is calculated.

A Questions and Answers for the Blessing

In these sections, we list the questions raised by the colleagues in the B group and the answers to them.

Q.1: Does the MC reproduce the angle between the two SVT tracks in the numerator and denominator? This could affect the efficiency. Calculate a systematic error if there is a difference.

Figure 72 shows the double ratio {semi/had(data)}{semi/had(MC)} for the $\Delta\phi$ between the two SVT trigger tracks in our three data samples. The ratio is compared to unity and a χ^2 probability is calculated. All the ratios are consistent with unity. Therefore, we do not assign any systematic uncertainty.

Q.2: Consider whether only using Wendy Taylor's analysis and the corrections you apply to determine f_{Λ_b} makes more sense than correcting the PDG average number.

Since we do not have access to the details of the LEP analyses and it is still questionable whether the f_{Λ_b} should be the same for the LEP and the Tevatron experiments, we use Wendy Taylor's result and apply corrections following the procedure documented in CDF note 7558 [1]. The change of the $\frac{\sigma_{\Lambda_b}}{\sigma_{B^0}}$ is reflected in CDF note 7558.

Q.3: From the optimization plots many of the cuts appear to be non optimal. Specifically the P_T cuts on the Λ_b and the Λ_c and equivalent B^0 and D cuts in the light B states. Could you explain concisely which of these choices were forced by what considerations? (MC, preselection ...) Also can you make a back of the envelope(or better) estimation of how much better the analysis could be statistically after performing a more rigorous optimization. Though I don't think we will chose to change the cuts now this may be useful the next time around.

The optimization procedure is standard and described in Section 3. As noted in Section 3, we make tighter cuts on the P_T of charm and B hadrons, instead of cutting on the optimization points, for the following reasons: First, the existing `PYTHIA` MC for the $b\bar{b}$ and $c\bar{c}$ background study, `nbot90` and `nbota0`, has preselection cuts on the P_T of the b and c quarks at 4 GeV/c. This P_T threshold makes the reconstruction of charm hadrons below 4 GeV/c inefficient. It took one month [38] to produce `nbot90` and `nbota0` and will take more time to produce another sample with a lower P_T threshold. Therefore, we choose to make a requirement of P_T greater than 5 GeV/c for all the charm hadrons. Second, to normalize the $b\bar{b}$ and $c\bar{c}$ backgrounds to the hadronic signals, we need an input of σ_{B^+} and $\sigma_{\Lambda_b} \cdot \mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)$ (see Section 6.3.2). As there are no measurements of an exclusive B hadron production cross-section at Run II, yet, we use the results of CDF Run I σ_{B^+} [21] and CDF Run II $\frac{\sigma_{\Lambda_b}(P_T>6.0)\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)}{\sigma_{B^0}(P_T>6.0)\mathcal{B}(\overline{B}^0 \to D^+ \pi^-)}$ measurements [4]. Both measurements were made with a P_T threshold of B hadrons greater than 6 GeV/c, thus, we make the same requirement in our analysis.

If we lower the P_T cuts to the data skimming requirements, i.e. P_T of charm hadrons greater than 2 GeV/c, and P_T of B hadrons greater than 4 GeV/c, the amount of the $\overline{B} \to D^{*+}X$ signal events is not affected. For the $\overline{B} \to D^+X$ and $\Lambda_b \to \Lambda_c^+ X$ decays, the amount of signal events is increased by 15% and 25%, respectively. The statistical uncertainties on the relative branching ratios only reduce by 10%. The significance is increased by 6% for both decays and the signal to background ratios drop by 7 and 23 %, respectively. The gains made by reducing the P_T thresholds are marginal.

Q.4: Would it help to cut explicitly on the D_s mass hypothesis with the proper particle type assignments to reduce background?

Figure 73 shows the $M_{K\pi\pi}$ distribution when the M_{xyz} is within (left) and outside (right) of the 2 σ D_s mass window, where xyz can be any K, π combinations except the $K\pi\pi$. The peak at 1.868 GeV/c^2 in the left plot indicates large contribution from the real D^+ inside the D_s mass window. Therefore, if we cut explicitly on the D_s mass hypothesis with the proper particle type assignments, the amount of signal events will reduce by a factor of 2 as shown from the fit in the right plot. Besides, about 10% of the D_s decays, *e.g.*: $D_s^+ \to K^{*0}K^+, K^{*0} \to K^0\pi(\gamma)$, are not fully reconstructed and lie underneath our signal. Consequently, they will not be removed by cutting on the D_s mass window. At the same time, we lose the ability to estimate these backgrounds by normalizing them to the $D_s \to \phi\pi$ mode.

Q.5: I notice that in the $D\mu$ MC pull tests, $N_{\rm combg}$ is mis-estimated. Nominally one would expect that if $N_{\rm combg}$ or N_{D_s} are systematically mis-estimated this could lead to a mis-estimation of $N_{\rm sig}$. Though in fact $N_{\rm combg}$ and N_{D_s} are very highly correlated but $N_{\rm combg}$ is not correlated to $N_{\rm sig}$. Perhaps an equal size or greater size set of additional toy MC would demonstrate that this is just a fluctuation(which is quite probable) and that we don't have anything to worry about.

We have performed test on 1000 toy MC samples. Each sample has a factor of 5 more statistics than the data. The results of the pull mean and width for $N_{\rm sig}$, $N_{\rm combg}$ and N_{D_s} are listed in Table 54. Figure 74 shows the pull distribution. We conclude that the previous 2 σ shift in the pull mean of $N_{\rm combg}$ (see Table 6) is only a fluctuation.

Table 54: Pull mean and width of $N_{\rm sig}$, $N_{\rm combg}$ and N_{D_s} from 1000 toy MC tests with 5 times more statistics than the data.

	pull mean	pull width
$N_{\rm sig}$	-0.015 $\pm$ 0.031	1.019 $\pm$ 0.022
$N_{\rm combg}$	0.011 $\pm$ 0.032	1.066 $\pm$ 0.023
N_{D_s}	0.025 $\pm$ 0.031	1.033 $\pm$ 0.022

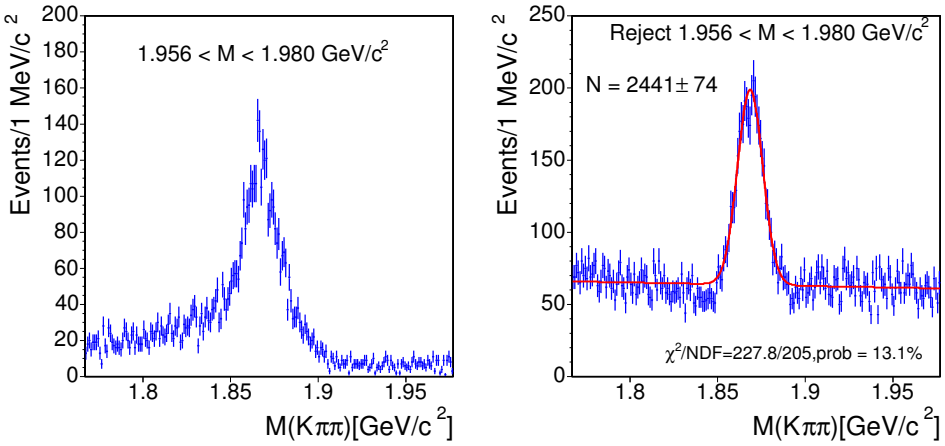


Figure 73: Left: $M_{K\pi\pi}$ distribution by requiring M_{xyz} within the D_s mass window, where x, y and z can be a K or a π. Right: $M_{K\pi\pi}$ distribution after removing the events with M_{xyz} in the D_s mass window. The data points are fit to a signal Gaussian and a first-order polynomial background.

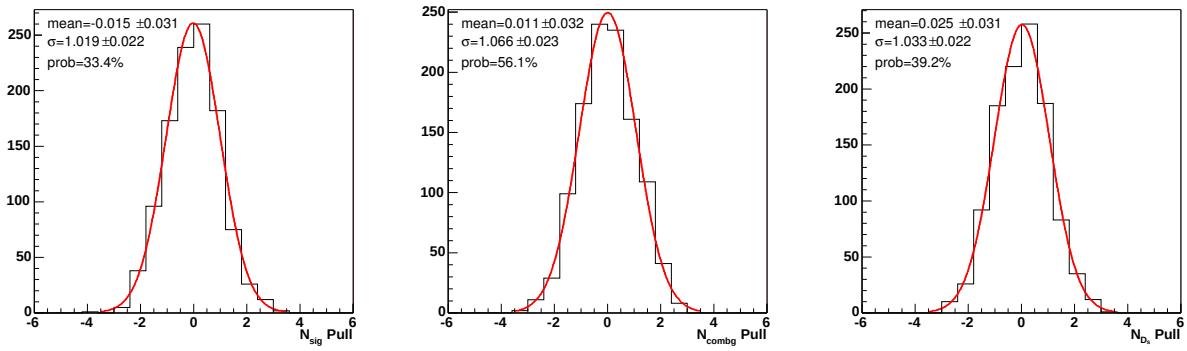


Figure 74: $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$: Pull mean and width of $N_{\rm sig}$, $N_{\rm combg}$ and N_{D_s} from the fit to 1000 toy MC samples with 5 times more statistics than the data.

Table 55: Pull mean and width of $N_{\rm sig}$ and σ from 1000 toy MC tests with 10 times more statistics than the data.

	pull mean	pull width
$N_{\rm sig}$	0.004 ± 0.031	0.970 ± 0.022
σ	0.008 ± 0.032	1.020 ± 0.023

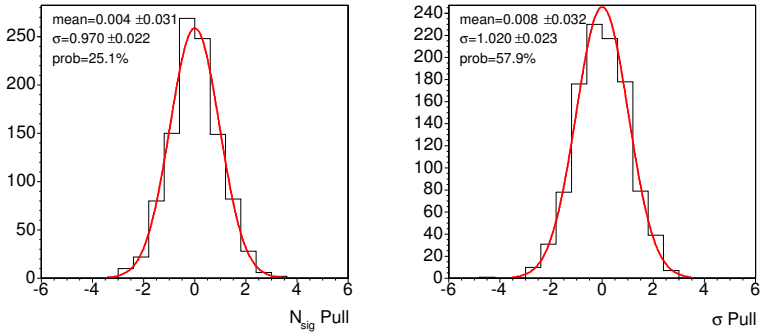


Figure 75: $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$: Pull mean and width of $N_{\rm sig}$ and σ from the fit to 1000 toy MC samples with 10 times more statistics than the data.

Q.6: for the $\Lambda_c\mu$ results on page 35. You let the width float for the Λ_c and the pull for the Λ_c width is slightly off and highly correlated to the number of signal events. Here you could also run more toy MC and check that the width that is extracted in the data is consistent with that expected from MC(with appropriate scaling).

Using the $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$, $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$ MC and $\overline{B} \to D^+ \mu^- X$ data, we predict the width of the $\overline{B} \to \Lambda_c^+ \mu^- X$ signal Gaussian to be 6.9 ± 0.2 MeV/c^2. The fit to the data yields 7.4 ± 0.6 MeV/c^2 (see Table 8), which is consistent with the MC prediction. We also perform test on 1000 toy MC samples. Each sample has a factor of 10 more statistics than the data. The results of the pull mean and width for $N_{\rm sig}$ and σ are listed in Table 55. Figure 75 shows the pull distribution. We conclude that the previous 2 σ shift in the pull mean of σ (see Table 8) is only a fluctuation.

Q.7: for the $\overline{B}^0 \to D^+\pi^-$ results. $f_{\rm comb}$ is off by a large number of sigma - though I understand this is small on the absolute scale. $f_{\rm comb}$ is also highly correlated to $N_{\rm sig}$. Is the effect on $N_{\rm sig}$ also small when this parameter is varied?

We study the change on $N_{\rm sig}$ due to the variation of $f_{\rm comb}$ by fixing all the fit parameters, except $N_{\rm sig}$ and $f_{\rm comb}$, to the fit result in Table 15. In Table 15, $f_{\rm comb} = 0.583 \pm 0.044$. Therefore, $f_{\rm comb}$ is now fixed to $0.583 + 0.044 \cdot 0.145$, since the pull mean from the toy MC test is -0.145 ± 0.034. The number of signal events is changed by 2 events out of 579 events, which is only a 0.35% shift.

Q.8: For the Cabibbo suppressed mode $\Lambda_b \to \Lambda_c^+ K^-$ you vary the possible branching fraction by -50% + 100%. This lies right under the peak so is potentially a large systematic error. 1st, why in general the asymmetric errors? Second is there any expectation for how well we can predict this from the light B modes that would allow us to chose the range for the systematic variation in a more justifiable way. Do we expect this % to carry over from the light B mesons?

We assign the uncertainty on the $\mathcal{B}(\Lambda_b \to \Lambda_c^+ K^-)$ in the same way as Le, Martin and Maksimović [4, 26]. We expect the ratio of the Cabibbo suppressed to Cabibbo favored modes to carry over from the B mesons. However, if we use the fractional uncertainty from the ratio $\mathcal{B}(\overline{B}^0 \to D^+ K^-)/\mathcal{B}(\overline{B}^0 \to D^+ \pi^-)$ in Table 14 ($\sim 30\%$) instead, the uncertainty from the fit to $\Lambda_b \to \Lambda_c^+ \pi^-$ is changed from 0.63 to 0.53, and the total CDF systematic uncertainty is changed from 1.2 to 1.15, which is only a marginal gain.

Q.9: You compute a correction to the MC pt spectrum based on the $\Lambda_b \to \Lambda_c^+ \pi^-$. Do we observe the same discrepancy in the semileptonic decay P_T distribution after accounting for the contributions from other sources?

The $P_T(\Lambda_c \mu)$ from the MC produced with the corrected Λ_b P_T spectrum agree with that from the data (see Figure 50). Besides, compared to the fully reconstructed modes, we do not believe the partially reconstructed modes can provide an adequate P_T spectrum for correcting the default MC P_T spectrum.

Q.10: In the MC the Λ_b semileptonic modes are decayed via phase space. I notice that the data spectra actually looks more like that for the B mesons. Is there an expectation for a more realistic model for the Λ_b and is the data qualitatively following that expectation? Can we update the decay table? Also could you made a histogram of the Lb pt spectra for `Bgenerator` available so that I can put it in the official MC directories so anyone can easily use it.

In Section 5.3, we have developed a decay model to provide a better description of the decay dynamics. From this better decay model, we find the acceptance unchanged, but the $\Lambda_c \mu$ invariant mass from MC agrees with the data better. This work will be fed back to the authors of `EvtGen`. The Λ_b P_T spectrum is also made public by the B MC group.

Q.11: Figure 33. Maybe you say that in the note, but do you know which kind of events have a mass above 5.5 GeV/c^2? I think they are not statistical fluctuations and they are included in your templates.

I assume you are asking if we have included the background template for the mass region above 5.5 GeV/c^2. Yes, we have. In the unbinned likelihood fit, we use the functional forms in Figure 33 for the backgrounds, which are derived by Maksimović [4]. The backgrounds that peak at 5.6 GeV/c^2 near the signal are mostly $\Lambda_b \to \Sigma_c \pi$ and $\Lambda_b \to \Lambda_c^+ \pi$, where the Λ_c^+ either decay semileptonically or to other hadronic final states instead of $pK\pi$. A more detailed list of the decays can be found in CDF note 6953 [26].

Q.12: Figure 39. Do you really include both sidebands in the D*-D, do you have any events on the left side with respect to the D*-D PDG value?. I think the common thing is to increase the right sideband.

Figure 76 shows a fit to the $M_{D^*} - M_{D^0}$ distribution after removing the events within 4 σ around 0.1454 GeV/c^2. The good χ^2 probability indicates that it is adequate to describe the background with a constant. Therefore, we can include both sidebands when performing a sideband subtraction to compare the data with the MC distribution. In addition, the signal to background ratio in the

signal region is about 63, as noted in Section 3. If we overestimate or underestimate the amount of the background by a factor of 2, we are only changing the data distributions by 1/63=2%. The agreement of the MC distribution with the data is still valid.

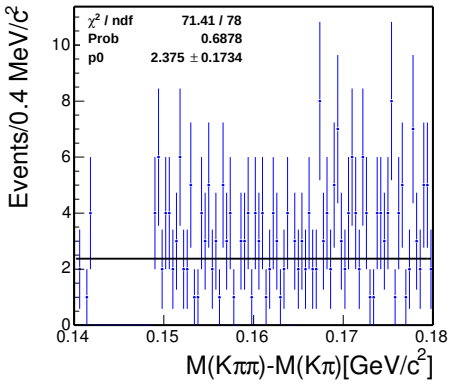


Figure 76: $M_{D^*} - M_{D^0}$ fit to a constant after removing the events 4 σ around 0.1454 GeV/c^2.

Q.13: Figure 44. Do you consider the first bin of the efficiency ratios as a statistical fluctuation? The ratio is much higher in that bin than any other.

The first bin represents a 2 σ deviation from the adjacent higher P_T bins. There are 28 bins in Figure 44 and we expect 1.5 bins with a 2 σ fluctuation. We believe this to be a fluctuation.

Q.14: Section 4.2. You are comparing data and MC, but on the other hand you know that there is a 10% of physics backgrounds which (I think!) are not included in the MC to make the plot. So, for several variables I would expect some disagreement due to that fact. For example l-D mass, vertex fits and Pt, you should have some disagreement due to that effect. Did you say that in the note, do you agree on it?

See the first paragraph of Section 5.2. We did include the physics backgrounds and the fake muon contribution in the MC.

Q.15: Section 5.3.3. When we compare data and PYTHIA MC cross-sections and I expect to have a disagreement of about 20-30%. Pythia is a LO generator, so once you include the NLO corrections you expect a factor 1.1-1.3, right?. I know it doesn't matter once you include a 100% syst. error. But I think it's well overestimated that error, but on the other hand " no-effect" or "no-effect/4" is pretty much the same.

The 100% uncertainty is a combination of two sources: 1. the relative contribution of each production mechanism: flavor creation, flavor excitation, and gluon splitting. 2. the difference of the B^+, D^0, and inclusive b cross-section between PYTHIA and the CDF Run-I, Run-II measurements (see Section 6.3.3). Even if this is an overestimate, the systematic uncertainty is still small compared to the other uncertainties.

Table 56: Background contribution from the $\overline{B}_s \to D^+ K^0 \mu^- \overline{\nu}_\mu$ and $\overline{B}^0 \to D^+ D_s^-$ decays to the $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$ signal.

Mode	BR (%)	Norm
$\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$	2.14 ± 0.20	1.000
$\overline{B}_s \to D^+ K^0 \mu^- \overline{\nu}_\mu$	0.30 ± ?	0.013
$\overline{B}^0 \to D^+ D_s^-$	0.80 ± 0.30	0.006
$\hookrightarrow (\phi, \eta, \eta') \mu^- \overline{\nu}_\mu$	5.51 ± 0.92	

Q.16:I think there are some missing backgrounds in the $B \to lD$ modes. (a) $B_s \to l\nu D_s^{**}, D_s^{**} \to D^{(*)} K^0$, (b) $B \to D_s^{(*)} DX, D_s \to lX$, (c) $B \to D1^{(*)} D2^{(*)} KX$.

The decays which contribute $\geq 1\%$ each to the semileptonic signal are included in the final calculation of the relative branching fractions (see Section 6.1). CDF note 6599 [14] details the list of decays which have $D\mu$ in the final state. For a comparison with the backgrounds you mentioned, we list two decays in Table 56 which have similar decay chains as (a), (b) and (c), extracted from Table 1 and 2 of CDF note 6599. The normalization has included the factor, $f_s/f_d = 0.27 \pm 0.03$.

Decays (a) have not been observed, yet. However, the decay table in the `EvtGen` provides an estimate of the branching fractions. The $\overline{B}_s \to l D_s^{**}$ decays that can fake our signals are $\overline{B}_s \to D_{s1}^+ \mu \overline{\nu}_\mu$ and $\overline{B}_s \to D_{s2}^{*+} \mu \overline{\nu}_\mu$, where $D_{s1}^+ \to D^{*+}, D^{*+} \to D^+$ and $D_{s2}^{*+} \to (D^{*+}, D^+) K^0, D^{*+} \to D^+$. After taking into account the branching ratios of the $D_{s1,s2}$ decays, the total branching ratios of the D_{s1} and D_{s2} modes are about 0.06% and 0.3%, respectively. The efficiencies of these two decays are supposed to be only as much as that of $\overline{B}_s \to D^+ K^0 \mu^- \overline{\nu}_\mu$, since the final states are identical or with additional π^0, γ. Using the information in Table 56, D_{s1} and D_{s2}^* are estimated to contribute about 0.2%, and 1% to the signal. The D_{s1} contribution can be neglected and the existence of D_{s2}^* is yet to be confirmed. Therefore, we do not include these two decays in the relative branching ratio calculation.

Decays (b) and (c) can be compared to the decay $\overline{B}^0 \to D^+ D_s^-$, where $D_s^- \to (\phi, \eta, \eta') \mu \overline{\nu}_\mu$. Decays (b) and (c) are expected to have smaller branching ratios than the $\overline{B}^0 \to D^+ D_s^-$ mode, since the decays of the excited charm hadrons into D^+ have branching fractions at most $\sim 32\%$, and additional K^0 in the final state of decays (c) means creating extra quark pairs from the vacuum. Decays (b) and (c) are also expected to have smaller efficiencies, since there are more particles in the final state and they will be more suppressed after the cut on the $M_{D\mu}$. Because the contribution of $\overline{B}^0 \to D^+ D_s^-$ is only 0.6%, decays (b) and (c) are expected to contribute less than 1% each and can be ignored. Table 2 in CDF note 6599 gives a partial list of decays which fall into the same category as (b) and (c), and confirms that these decays contribute not more than 0.6% to the signal.

B Derivation of the Function for the Mass Fit

B.1 Convolution of Gaussian with Triangular Distribution

The derivation comes from J. Heinrich.

Gaussian:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$$

Triangular (also normalized to 1):

$$h(x) = \frac{2}{a^2}x \qquad (0 \le x \le a)$$

Convolution:

$$f(x) = (g * h)(x) = \int_{-\infty}^{\infty} g(\xi)h(x-\xi)d\xi = \frac{1}{a^2\sigma}\sqrt{\frac{2}{\pi}}\int_{x-a}^{x}(x-\xi)e^{-\frac{1}{2}(\frac{\xi}{\sigma})^2}d\xi$$

Substitute $\xi = \sqrt{2}\sigma u$:

$$f(x) = \frac{2}{a^2\sqrt{\pi}}\int_{\frac{x-a}{\sqrt{2}\sigma}}^{\frac{x}{\sqrt{2}\sigma}}(x-\sqrt{2}\sigma u)e^{-u^2}du$$

Error functions:

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt \qquad \mathrm{erf}(\infty) = 1$$

$$\mathrm{erf}(-z) = -\mathrm{erf}(z) \qquad \mathrm{erfc}(z) = 1 - \mathrm{erf}(z)$$

$$f(x) = \frac{x}{a^2}\left[\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) + \mathrm{erf}\left(\frac{a-x}{\sqrt{2}\sigma}\right)\right] + \frac{\sigma}{a^2}\sqrt{\frac{2}{\pi}}\left[e^{-\frac{1}{2}(\frac{x}{\sigma})^2} - e^{-\frac{1}{2}(\frac{x-a}{\sigma})^2}\right]$$

The form above is well suited for $0 \le x \le a$. When $x \gg a$, a better numerical form is

$$f(x) = \frac{x}{a^2}\left[\mathrm{erfc}\left(\frac{x-a}{\sqrt{2}\sigma}\right) - \mathrm{erfc}\left(\frac{x}{\sqrt{2}\sigma}\right)\right] + \frac{\sigma}{a^2}\sqrt{\frac{2}{\pi}}\left[e^{-\frac{1}{2}(\frac{x}{\sigma})^2} - e^{-\frac{1}{2}(\frac{x-a}{\sigma})^2}\right]$$

When $x \ll 0$,

$$f(x) = \frac{x}{a^2}\left[\mathrm{erfc}\left(\frac{-x}{\sqrt{2}\sigma}\right) - \mathrm{erfc}\left(\frac{a-x}{\sqrt{2}\sigma}\right)\right] + \frac{\sigma}{a^2}\sqrt{\frac{2}{\pi}}\left[e^{-\frac{1}{2}(\frac{x}{\sigma})^2} - e^{-\frac{1}{2}(\frac{x-a}{\sigma})^2}\right]$$

is better numerically.

Next we want the integral $F(x) = \int_x^\infty f(t)dt$. We have

$$F(x) = \int_x^\infty\int_{-\infty}^\infty g(t-\xi)h(\xi)d\xi\,dt = \int_{-\infty}^\infty h(\xi)\int_x^\infty g(t-\xi)dt\,d\xi$$

$$= \frac{1}{2}\int_{-\infty}^\infty h(\xi)\mathrm{erfc}\left(\frac{x-\xi}{\sqrt{2}\sigma}\right)d\xi = \frac{1}{a^2}\int_0^a \xi\,\mathrm{erfc}\left(\frac{x-\xi}{\sqrt{2}\sigma}\right)d\xi$$

Substitute $\xi = x - \sqrt{2}\sigma v$:

$$F(x) = \frac{\sqrt{2}\sigma}{a^2}\int_{\frac{x-a}{\sqrt{2}\sigma}}^{\frac{x}{\sqrt{2}\sigma}}(x-\sqrt{2}\sigma v)\mathrm{erfc}(v)dv$$

Generic integrals:

$$\int \mathrm{erfc}(z)dz = z\,\mathrm{erfc}(z) - \frac{1}{\sqrt{\pi}}e^{-z^2} + \mathrm{const.}$$

$$\int z\,\mathrm{erfc}(z)dz = \left(\frac{z^2}{2} - \frac{1}{4}\right)\mathrm{erfc}(z) - \frac{z}{2\sqrt{\pi}}e^{-z^2} + \text{const.}$$

Result:

$$F(x) = \frac{x^2+\sigma^2}{2a^2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}\sigma}\right) - \frac{x^2-a^2+\sigma^2}{2a^2}\mathrm{erfc}\left(\frac{x-a}{\sqrt{2}\sigma}\right) + \frac{(x+a)\sigma}{\sqrt{2\pi}a^2}e^{-\frac{1}{2}(\frac{x-a}{\sigma})^2} - \frac{x\sigma}{\sqrt{2\pi}a^2}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$$

or equivalently

$$F(x) = \frac{1}{2} + \frac{x^2-a^2+\sigma^2}{2a^2}\mathrm{erf}\left(\frac{x-a}{\sqrt{2}\sigma}\right) - \frac{x^2+\sigma^2}{2a^2}\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) + \frac{(x+a)\sigma}{\sqrt{2\pi}a^2}e^{-\frac{1}{2}(\frac{x-a}{\sigma})^2} - \frac{x\sigma}{\sqrt{2\pi}a^2}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$$

B.2 Convolution of Gaussian with Exponential

Gaussian:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$$

Exponential:

$$h(x) = e^{-\frac{x}{\tau}}$$

Unnormalized lifetime distribution after convolution:

$$f(x) = (g*h)(x) = \frac{1}{2\tau}\mathrm{erfc}(\frac{1}{\sqrt{2}}(\frac{\sigma}{\tau} - \frac{x}{\sigma}))e^{\frac{\sigma^2}{2\tau^2}-\frac{x}{\tau}} \tag{76}$$

Integral:

$$F(x) = \frac{1}{2}\{\mathrm{erfc}(-\frac{x}{\sqrt{2}\sigma}) - \mathrm{erfc}(\frac{1}{\sqrt{2}}(\frac{\sigma}{\tau} - \frac{x}{\sigma}))e^{\frac{\sigma^2}{2\tau^2}-\frac{x}{\tau}}\} \tag{77}$$

The normalization could be obtained by dividing Equation 76 by the integral in the mass window of concern using Equation 77.

B.3 Bifurcated Gaussian

Notation:

$$\mu \equiv \text{mean},$$
$$\sigma_L \equiv \text{left sigma},$$
$$\sigma_R \equiv \text{right sigma}.$$

Function:
when $x \geq \mu$,

$$f(x) = e^{-\frac{1}{2}(\frac{x}{\sigma_R})^2}$$

when $x < \mu$,

$$f(x) = e^{-\frac{1}{2}(\frac{x}{\sigma_L})^2}$$

Integral from a to b where $a < b$
if $b < \mu$:

$$F(x) = \sqrt{\frac{\pi}{2}}\sigma_L\,[\mathrm{erf}(\frac{b-\mu}{\sqrt{2}\sigma_L}) - \mathrm{erf}(\frac{a-\mu}{\sqrt{2}\sigma_L})]$$

if $a > \mu$:

$$F(x) = \sqrt{\frac{\pi}{2}}\sigma_R \left[\operatorname{erf}(\frac{b-\mu}{\sqrt{2}\sigma_R}) - \operatorname{erf}(\frac{a-\mu}{\sqrt{2}\sigma_R})\right]$$

otherwise:

$$F(x) = \sqrt{\frac{\pi}{2}}[\sigma_R \operatorname{erf}(\frac{b-\mu}{\sqrt{2}\sigma_R}) - \sigma_L \operatorname{erf}(\frac{a-\mu}{\sqrt{2}\sigma_L})]$$

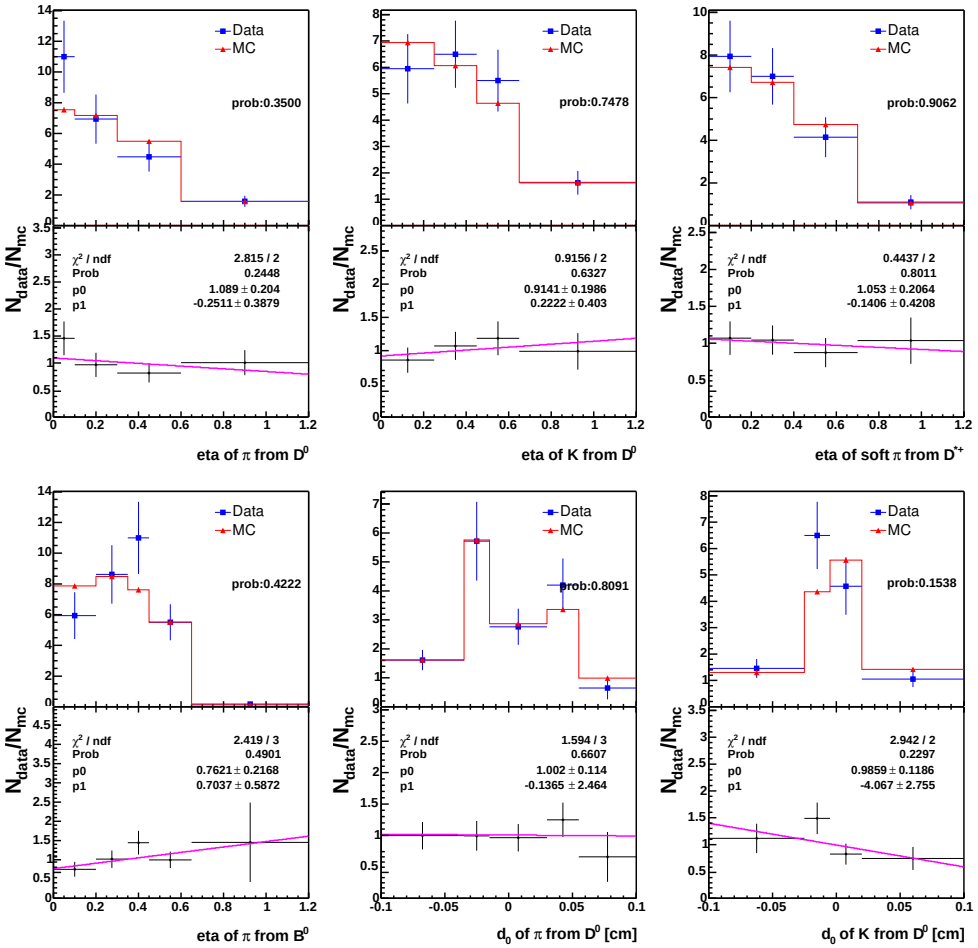


Figure 77: $\overline{B}^0 \to D^{*+}\pi^-$ MC and data comparison: from the top left to the bottom right are: $\eta(\pi_{D^0})$, $\eta(K_{D^0})$, $\eta(\pi_{D^*})$, $\eta(\pi_B)$, $d_0(\pi_{D^0})$, $d_0(K_{D^0})$.

C Comparison of Data and MC

This section includes various distributions important for this analysis. The MC is generated as described in Section 5.1. The background in the data distribution is removed using the method described in Section 5.2. Comparison between MC and data for the following variables: P_T, $c\tau$, fit $\chi^2_{r-\phi}$ of the B and charm hadrons, P_T of the muon and pion from the B and invariant mass of the four tracks, are found in Section 5.2.

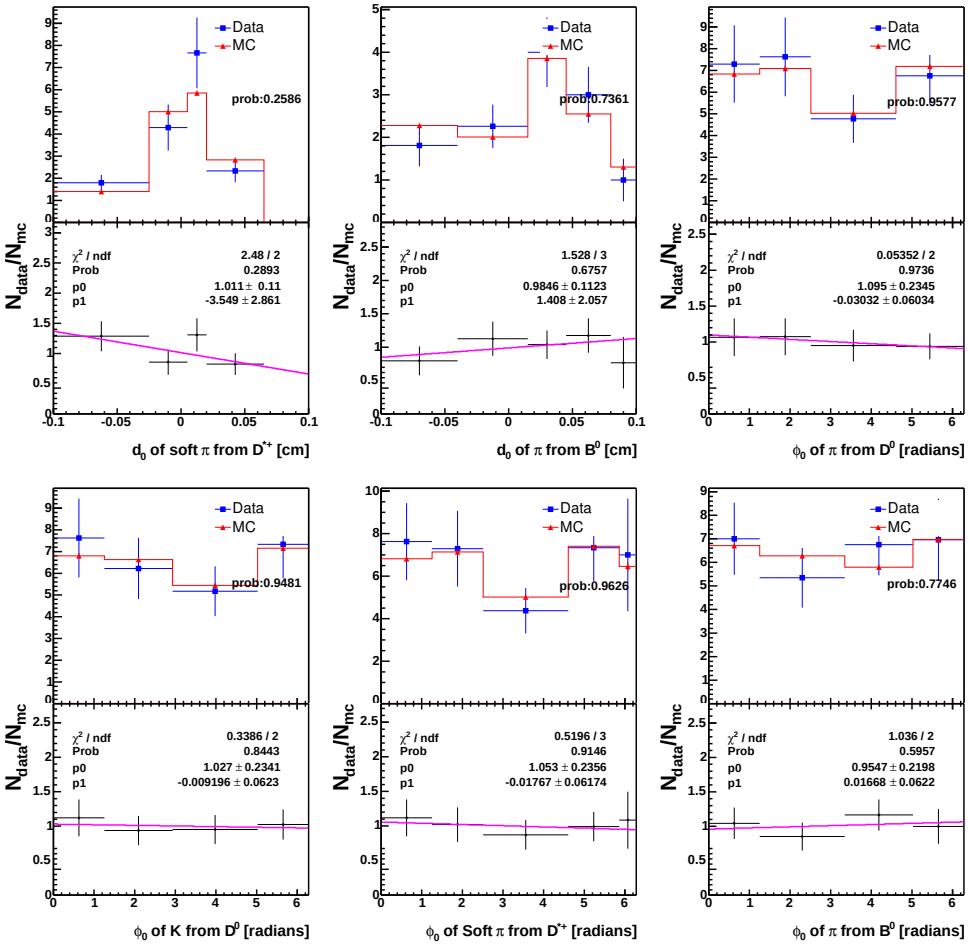


Figure 78: $\overline{B}^0 \to D^{*+}\pi^-$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi_{D^*})$, $d_0(\pi_B)$, $\phi_0(\pi_{D^0})$, $\phi_0(K_{D^0})$, $\phi_0(\pi_{D^*})$, $\phi_0(\pi_B)$

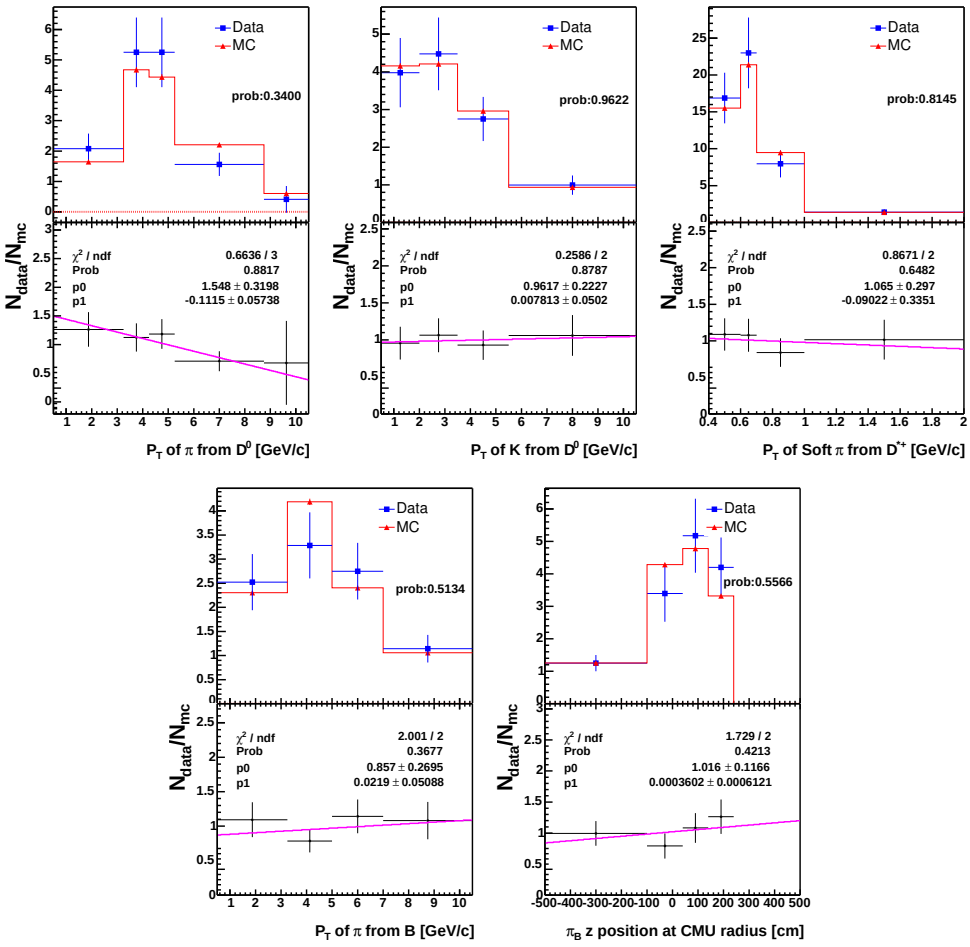


Figure 79: $\overline{B}^0 \rightarrow D^{*+}\pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(\pi_{D^0})$, $P_T(K_{D^0})$, $P_T(\pi_{D^*})$, $P_T(\pi_B)$ and extrapolated z position at CMU of π_B.

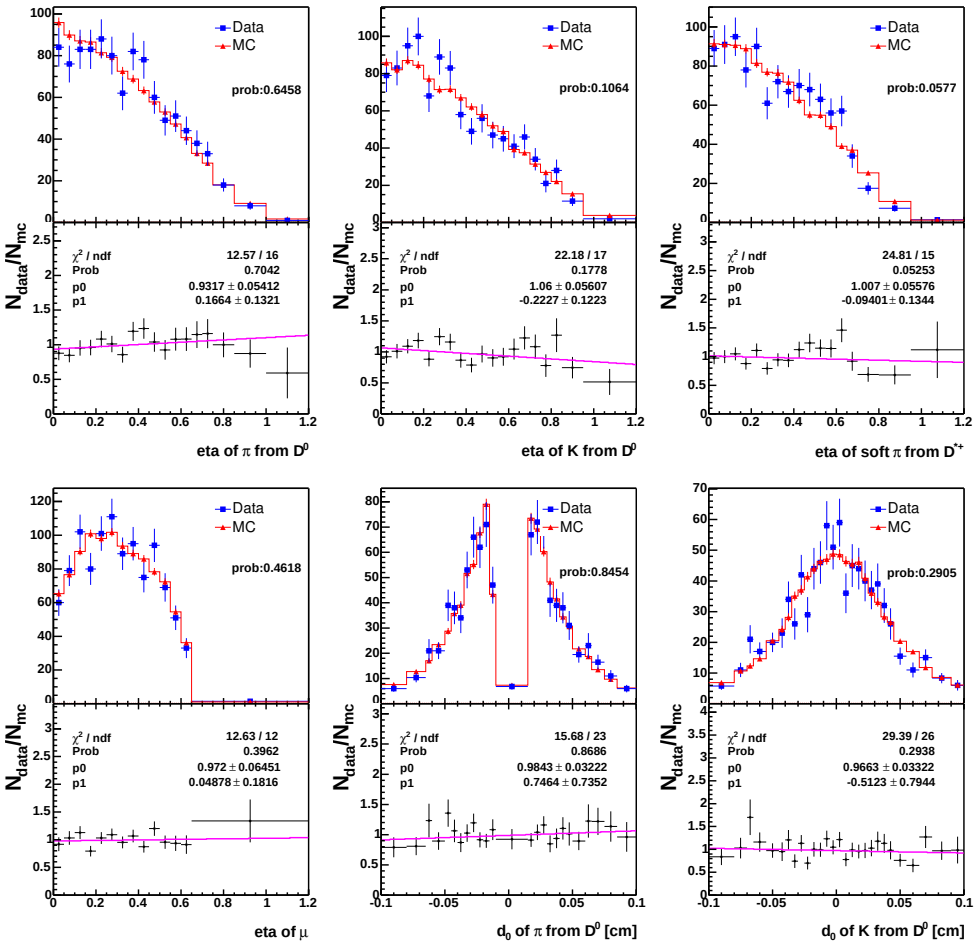


Figure 80: $\overline{B} \to D^{*+}\mu^- X$ MC and data comparison: from the top left to the bottom right are: $\eta(\pi_{D^0})$, $\eta(K_{D^0})$, $\eta(\pi_{D^*})$, $\eta(\mu)$, $d_0(\pi_{D^0})$, $d_0(K_{D^0})$.

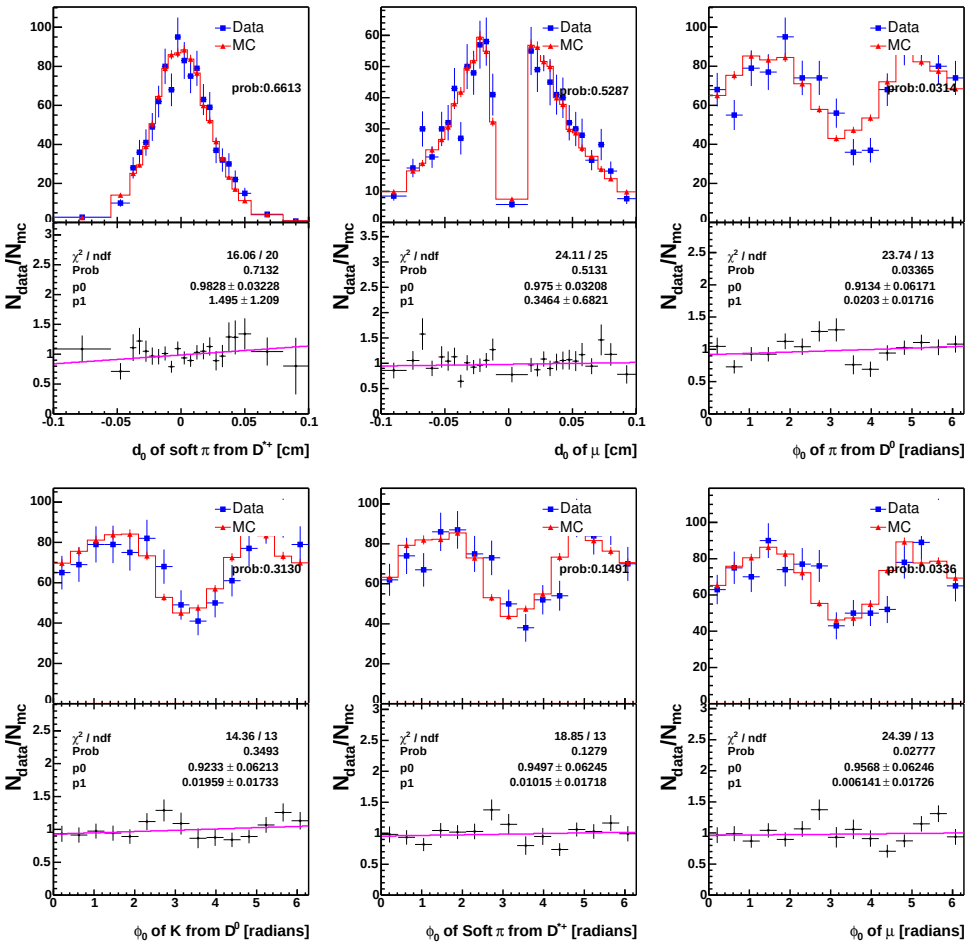


Figure 81: $\overline{B} \to D^{*+}\mu^- X$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi_{D^*})$, $d_0(\mu)$, $\phi_0(\pi_{D^0})$, $\phi_0(K_{D^0})$, $\phi_0(\pi_{D^*})$, $\phi_0(\mu)$

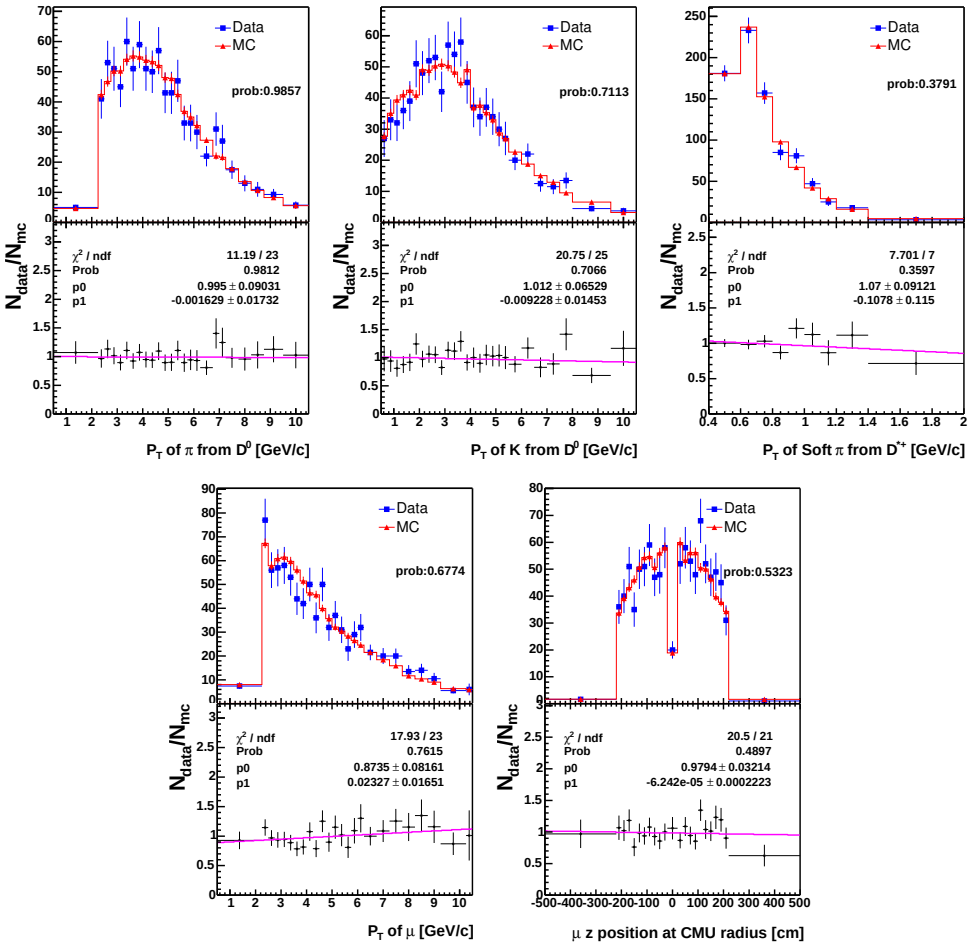


Figure 82: $\overline{B} \to D^{*+}\mu^- X$ MC and data comparison: from the top left to the bottom right are: $P_T(\pi_{D^0})$, $P_T(K_{D^0})$, $P_T(\pi_{D^*})$, $P_T(\mu)$, extrapolated z position at CMU of μ

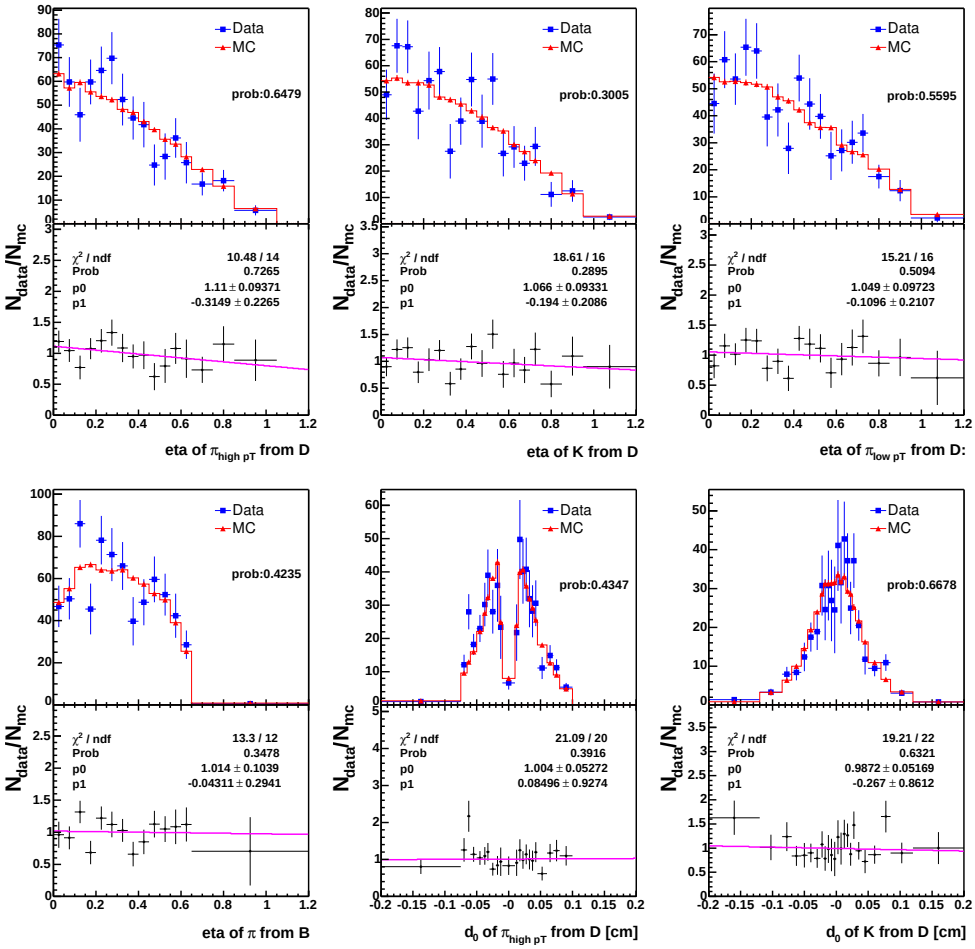


Figure 83: $\overline{B}^0 \to D^+\pi^-$ MC and data comparison: from the top left to the bottom right are: $\eta(\pi_D^{high})$, $\eta(K_D)$, $\eta(\pi_D^{low})$, $\eta(\pi_B)$, $d_0(\pi_D^{high})$, $d_0(K_D)$.

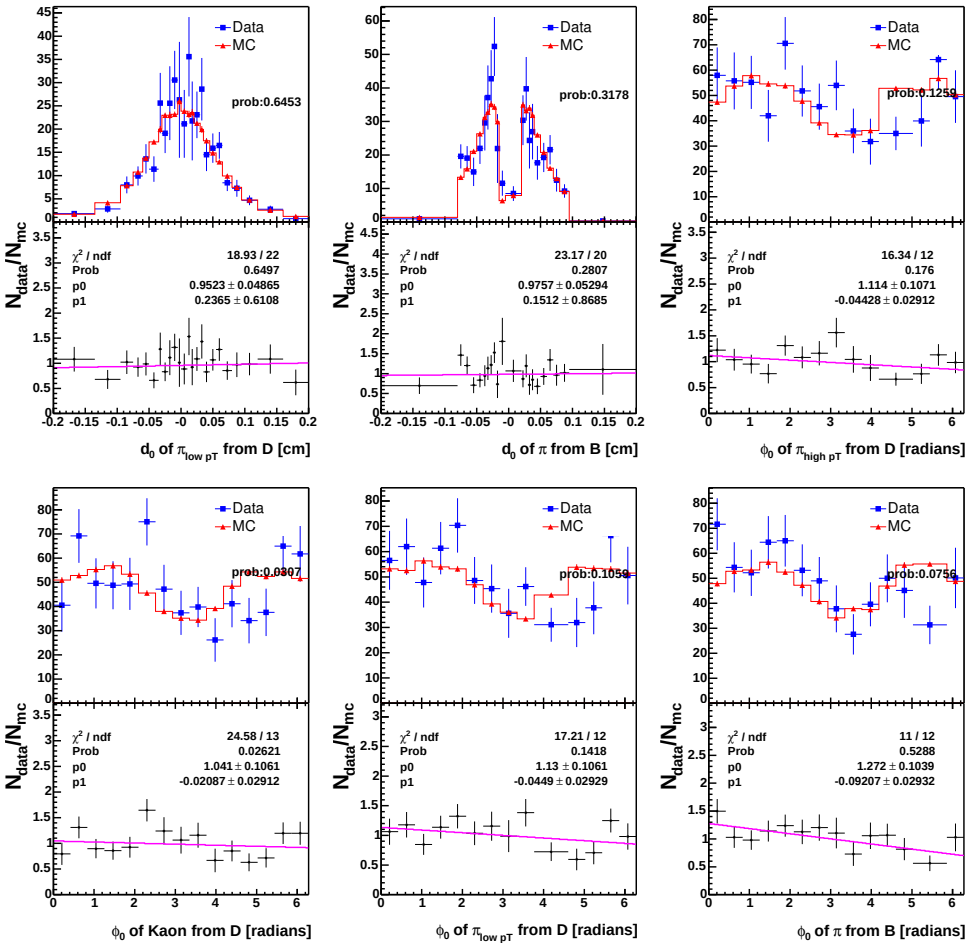


Figure 84: $\overline{B}^0 \to D^+\pi^-$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi_D^{low})$, $d_0(\pi_B)$, $\phi_0(\pi_D^{high})$, $\phi_0(K_D)$, $\phi_0(\pi_D^{low})$, $\phi_0(\pi_B)$

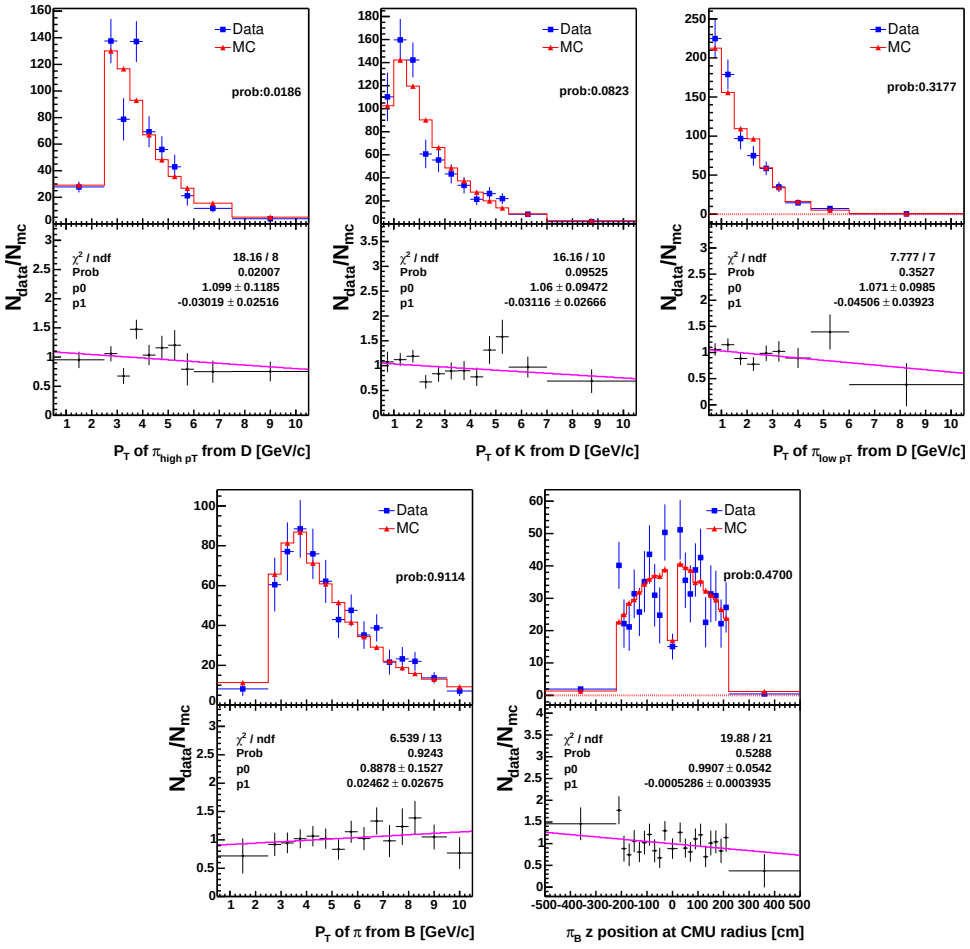


Figure 85: $\overline{B}^0 \rightarrow D^+\pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(\pi_D^{high})$, $P_T(K_D)$, $P_T(\pi_D^{low})$, $P_T(\pi_B)$ and extrapolated z position at CMU of π_B.

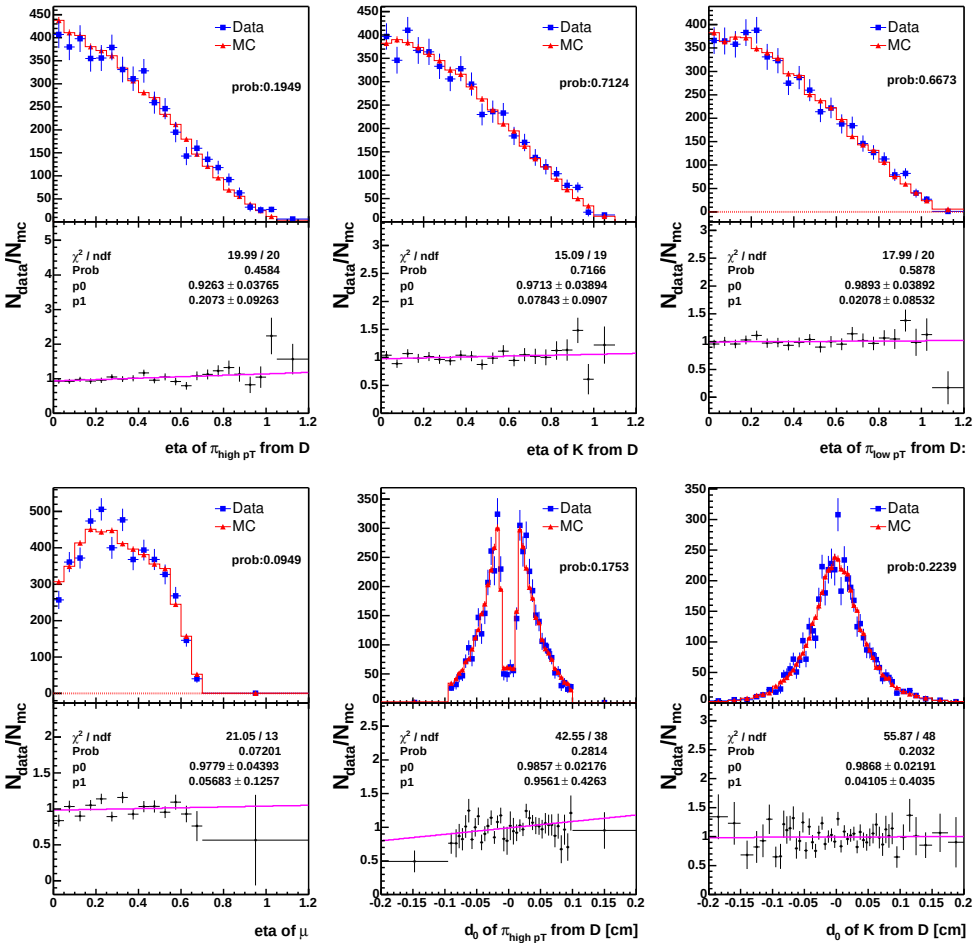


Figure 86: $\overline{B} \to D^+\mu^-X$ MC and data comparison: from the top left to the bottom right are: $\eta(\pi_D^{high})$, $\eta(K_D)$, $\eta(\pi_D^{low})$, $\eta(\mu)$, $d_0(\pi_D^{high})$, $d_0(K_D)$.

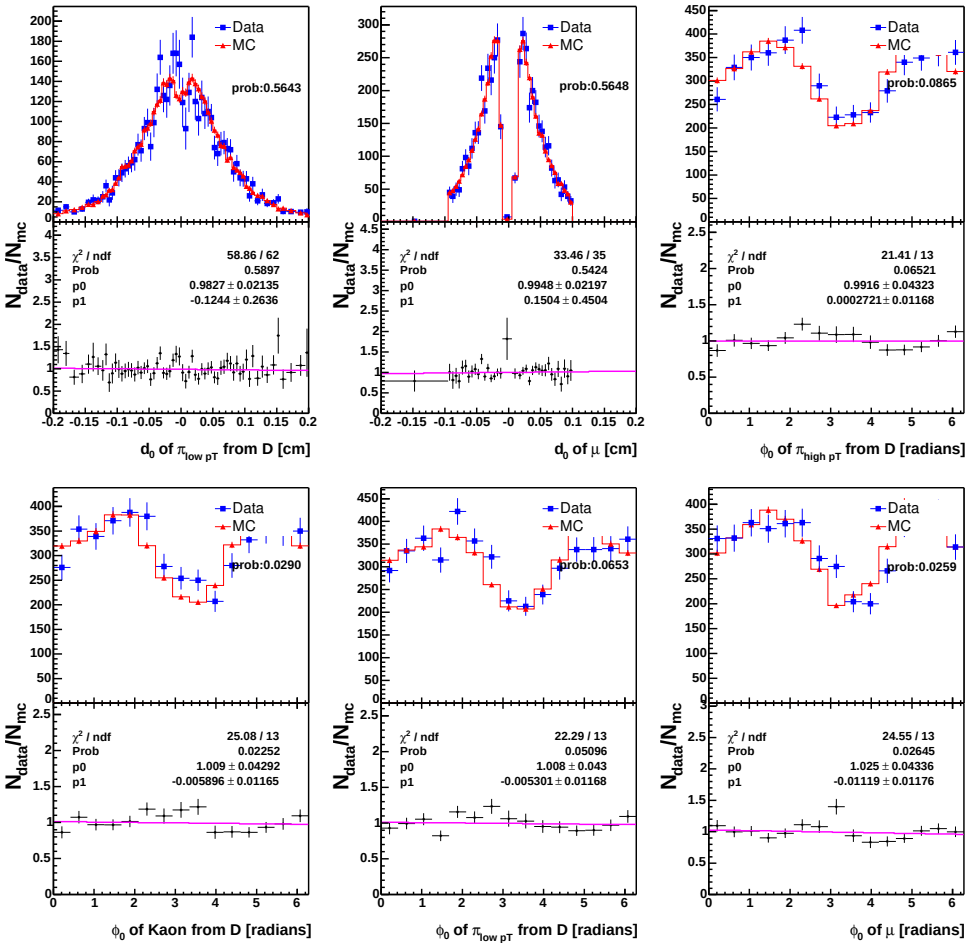


Figure 87: $\overline{B} \rightarrow D^{+}\mu^{-}X$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi_D^{low})$, $d_0(\mu)$, $\phi_0(\pi_D^{high})$, $\phi_0(K_D)$, $\phi_0(\pi_D^{low})$, $\phi_0(\mu)$

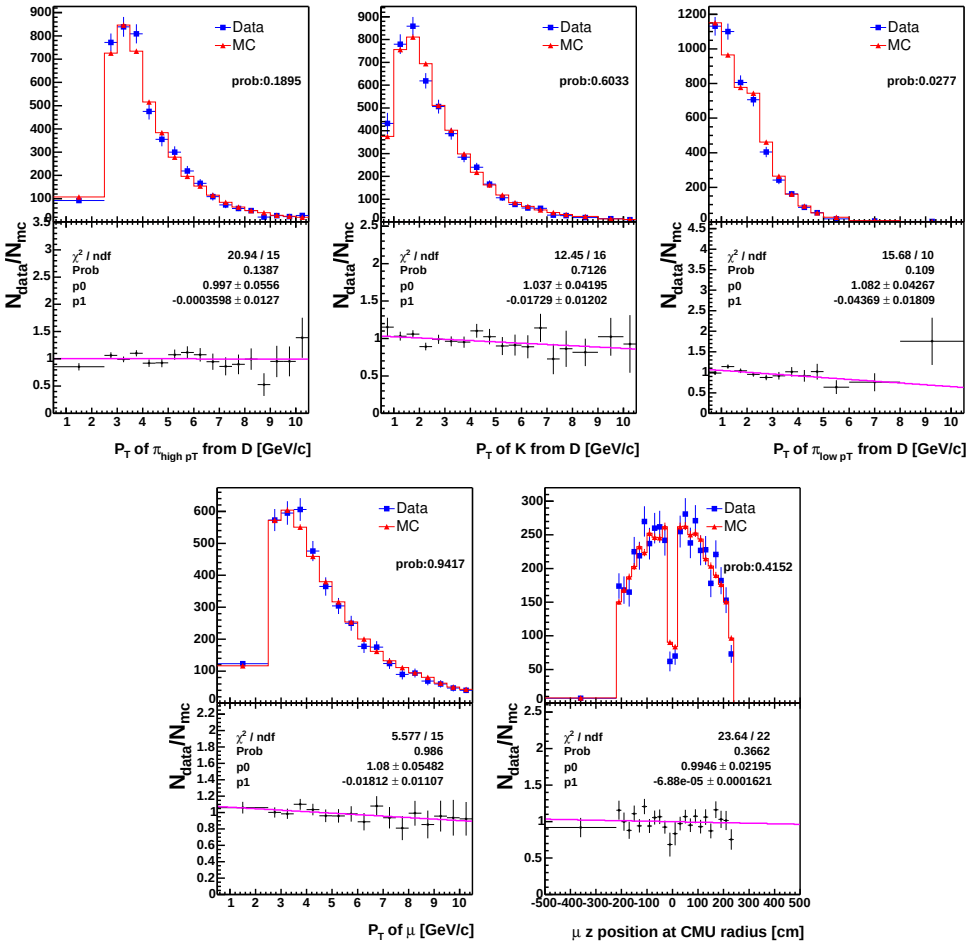


Figure 88: $\overline{B} \rightarrow D^{+}\mu^{-}X$ MC and data comparison: from the top left to the bottom right are: $P_T(\pi_D^{high})$, $P_T(K_D)$, $P_T(\pi_D^{low})$, $P_T(\mu)$, extrapolated z position at CMU of μ

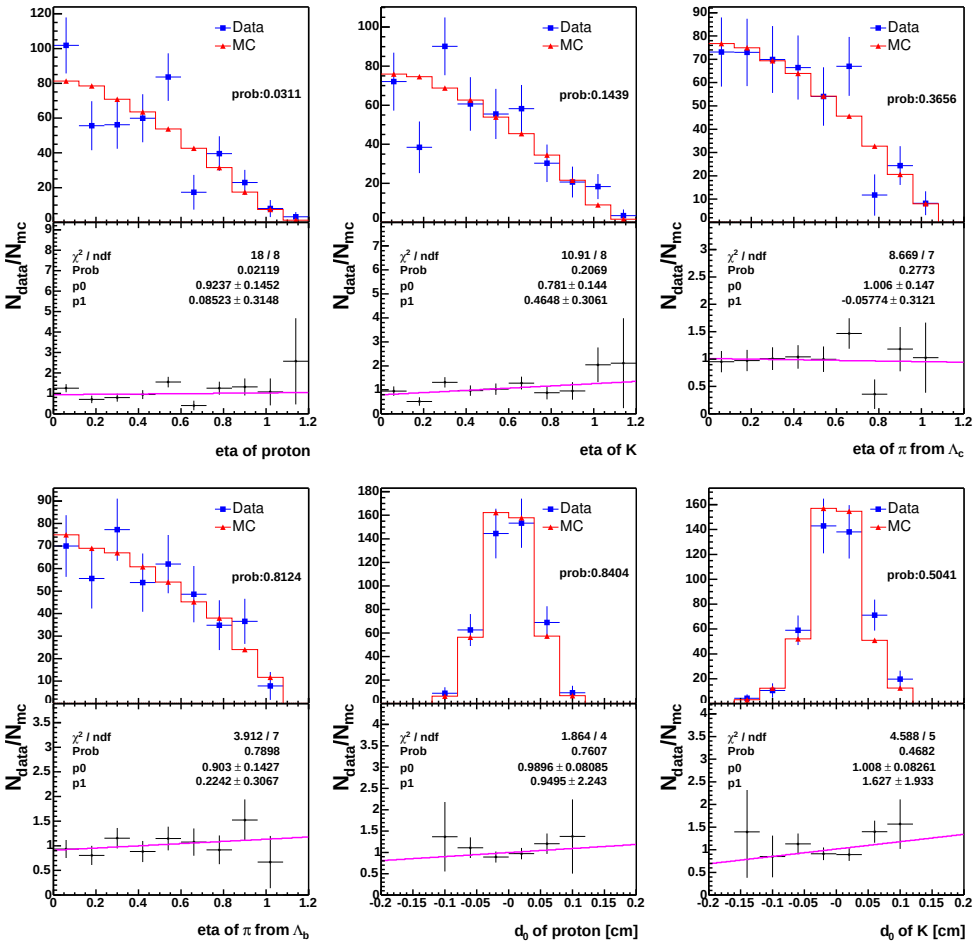


Figure 89: $\Lambda_b \to \Lambda_c^+ \pi^-$ MC and data comparison: from the top left to the bottom right are: $\eta(p)$, $\eta(K)$, $\eta(\pi_{\Lambda_c})$, $\eta(\pi_{\Lambda_b})$, $d_0(p)$, $d_0(K)$.

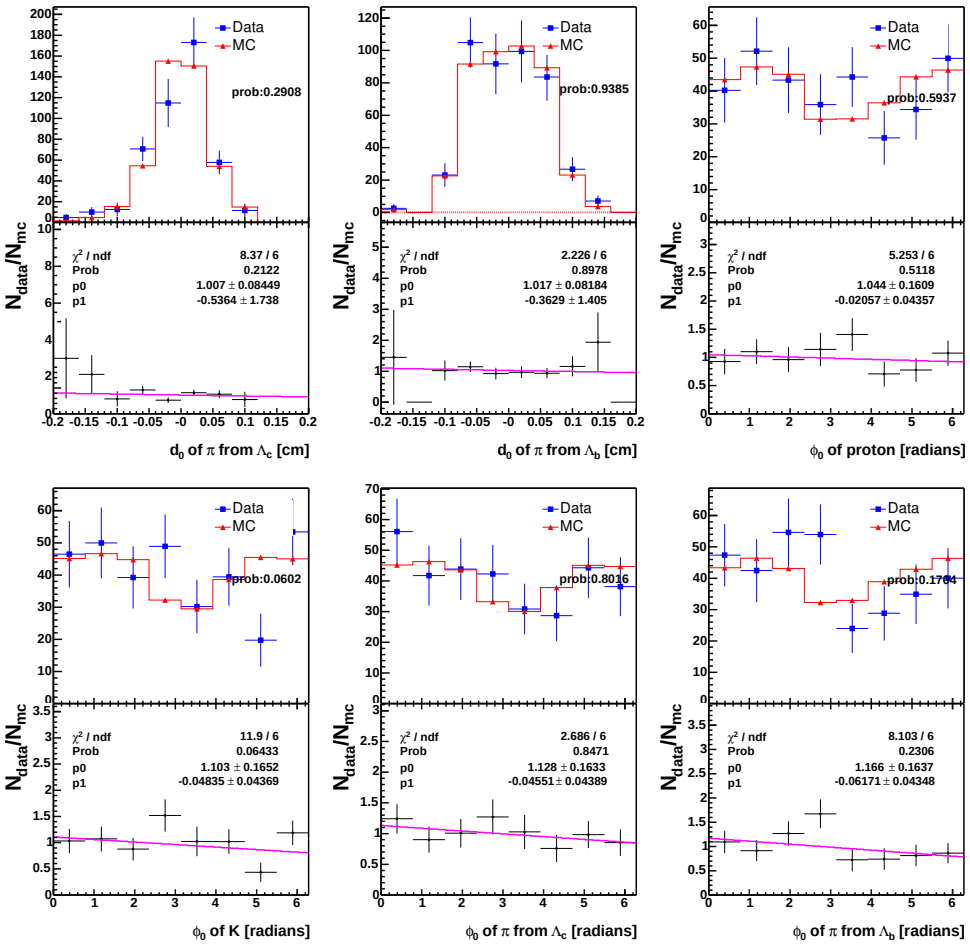


Figure 90: $\Lambda_b \to \Lambda_c^+ \pi^-$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi_{\Lambda_c})$, $d_0(\pi_B)$, $\phi_0(p)$, $\phi_0(K)$, $\phi_0(\pi_{\Lambda_c})$, $\phi_0(\pi_{\Lambda_b})$

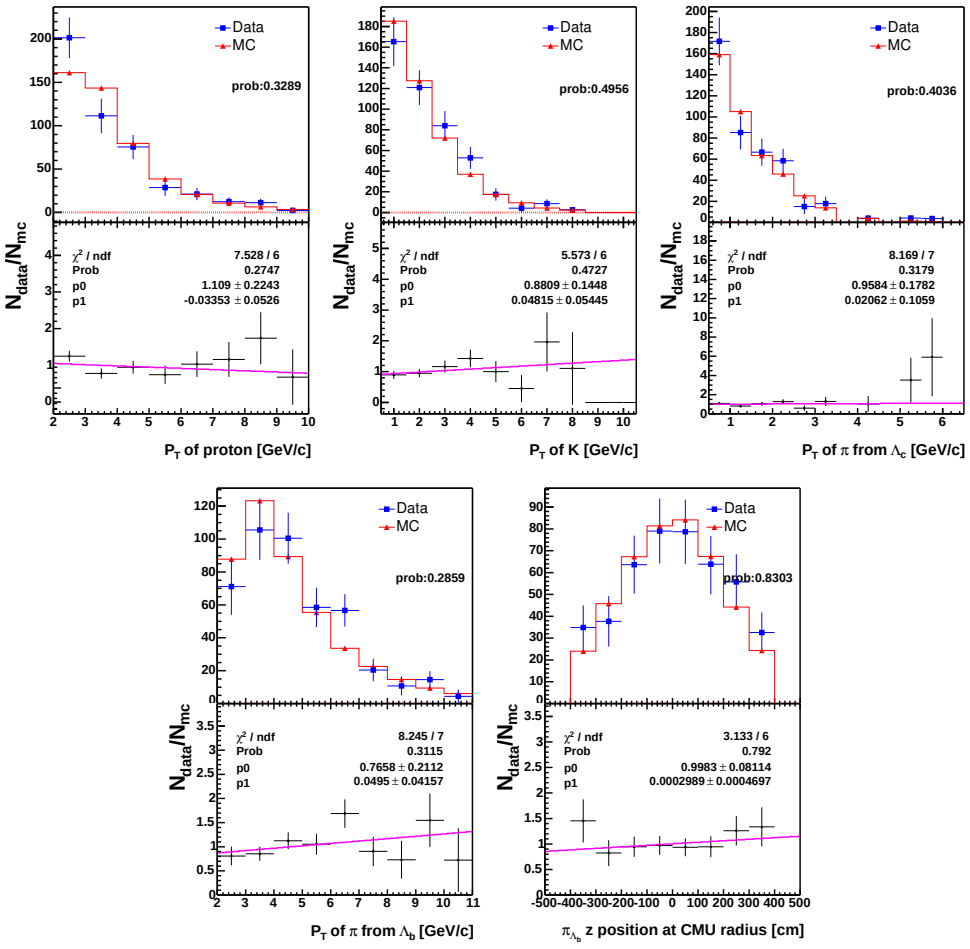


Figure 91: $\Lambda_b \to \Lambda_c^+ \pi^-$ MC and data comparison: from the top left to the bottom right are: $P_T(p)$, $P_T(K)$, $P_T(\pi_{\Lambda_c})$, $P_T(\pi_{\Lambda_b})$ and extrapolated z position at CMU of π_{Λ_b}.

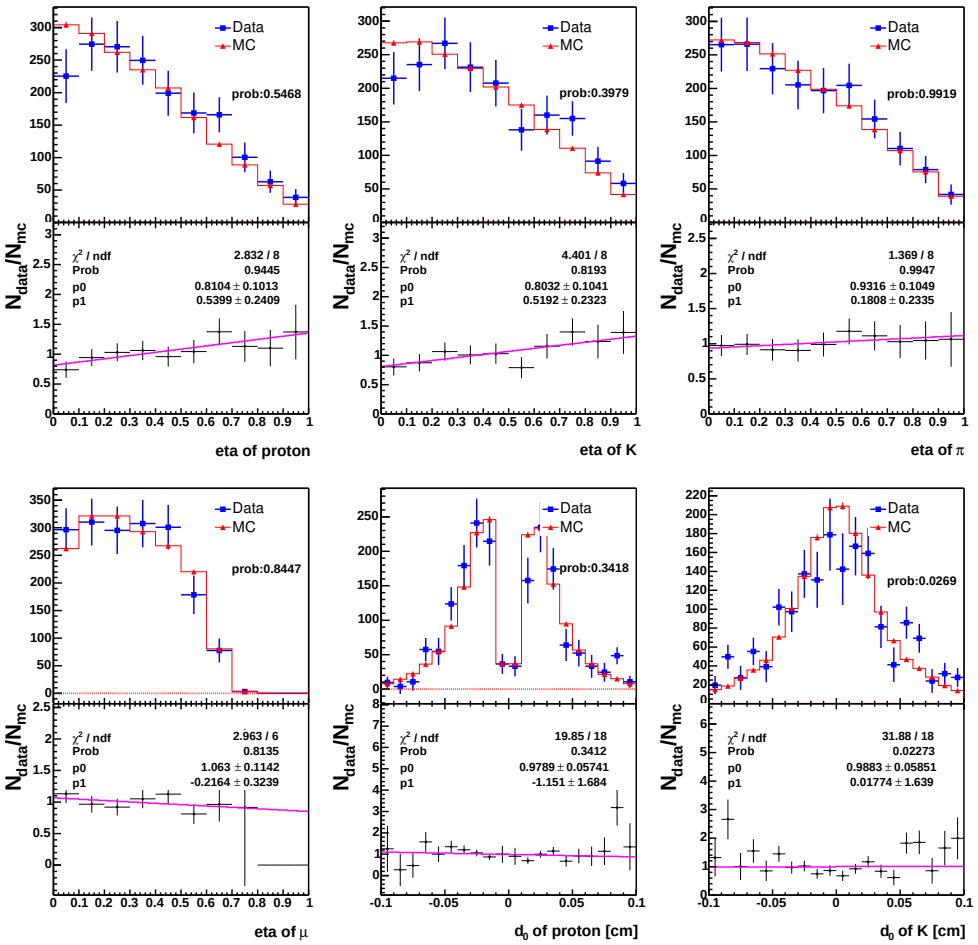


Figure 92: $\overline{B} \to \Lambda_c^+ \mu^- X$ MC and data comparison: from the top left to the bottom right are: $\eta(p)$, $\eta(K)$, $\eta(\pi)$, $\eta(\mu)$, $d_0(p)$, $d_0(K)$.

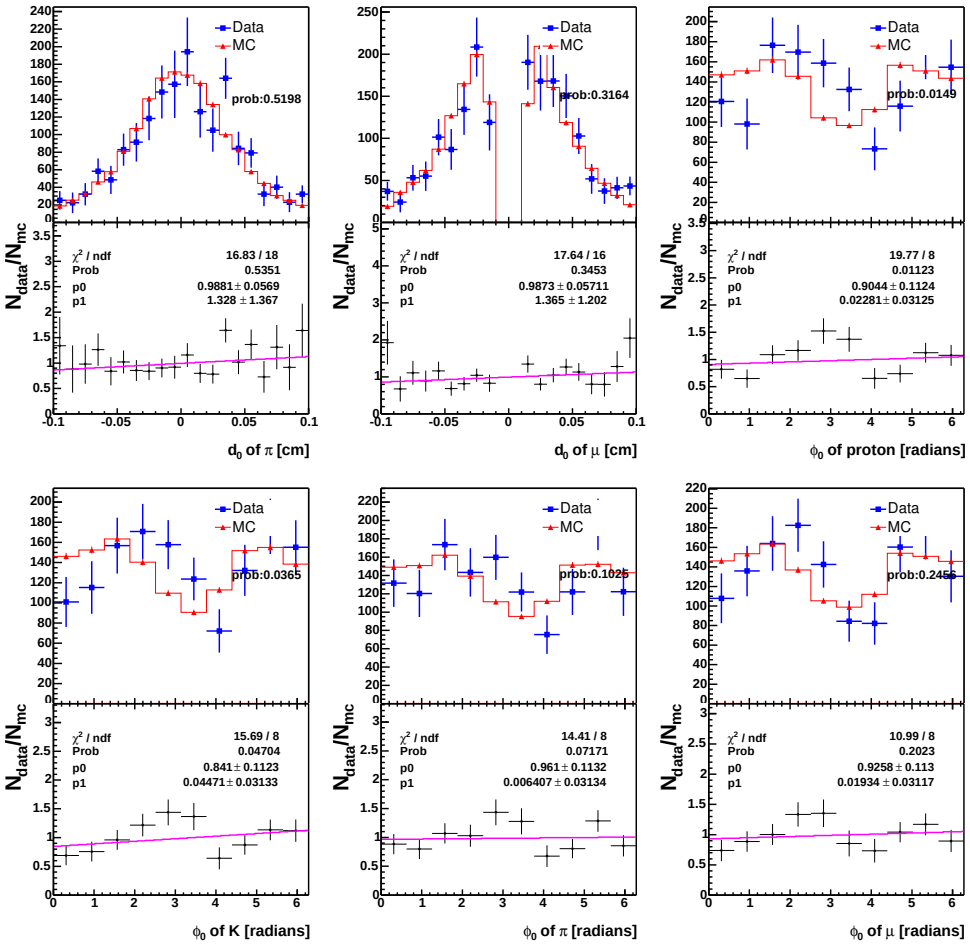


Figure 93: $\overline{B} \to \Lambda_c^+ \mu^- X$ MC and data comparison: from the top left to the bottom right are: $d_0(\pi)$, $d_0(\mu)$, $\phi_0(p)$, $\phi_0(K)$, $\phi_0(\pi)$, $\phi_0(\mu)$

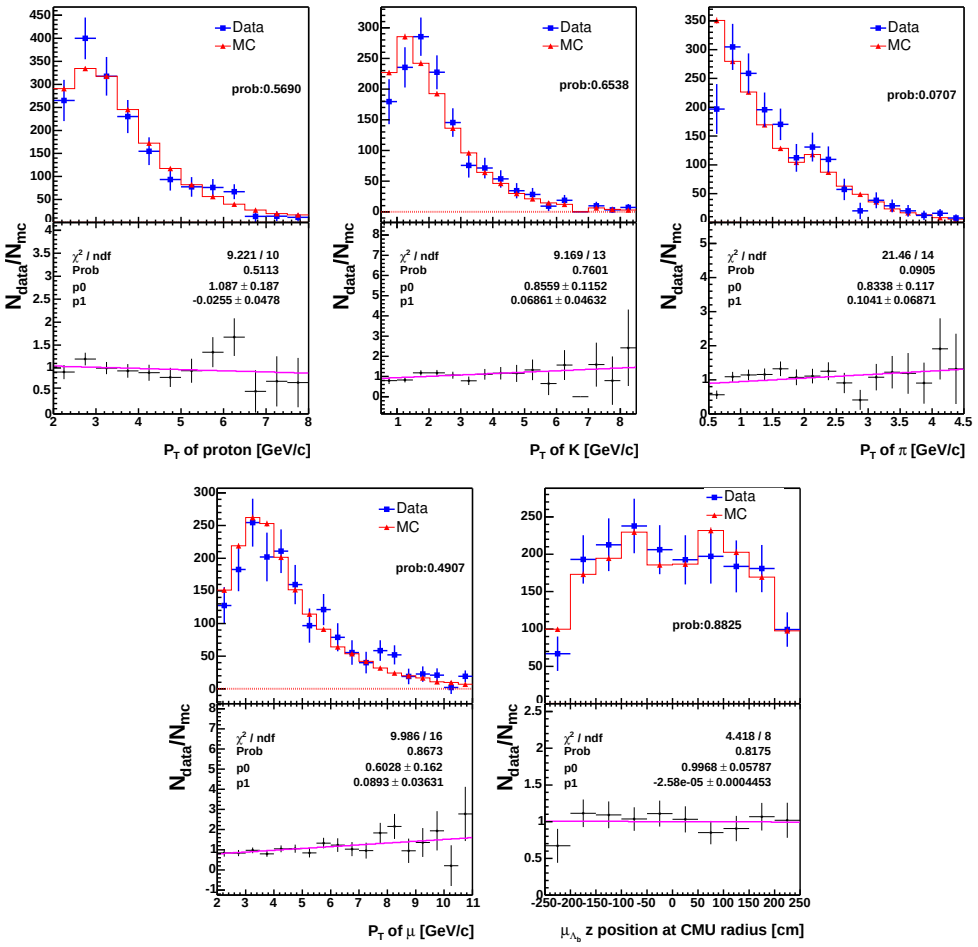


Figure 94: $\overline{B} \to \Lambda_c^+ \mu^- X$ MC and data comparison: from the top left to the bottom right are: $P_T(p)$, $P_T(K)$, $P_T(\pi)$, $P_T(\mu)$, extrapolated z position at CMU of μ

References

[1] S.-S. Yu, R. Tesarek, D. Litvintsev, J. Heinrich, and N. Lockyer, *A Correction for the Cross-Section Ratio* $\frac{\sigma_{\Lambda_b}}{\sigma_{B^0}}$, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/7558, University of Pennsylvania,Fermilab (2005).

[2] S.-S. Yu, R. Tesarek, N. Lockyer, and D. Litvintsev, *Description of the Datasets for the Ratio of Semileptonic to Hadronic Partial Decay Width*, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/6979, University of Pennsylvania,Fermilab (2004).

[3] I. Furic, *Measurement of the Ratio of Branching Fractions* $\mathcal{B}(B_s^0 \to D_s^- \pi^+)/\mathcal{B}(B^0 \to D^- \pi^+)$ *at CDF-II*, CDF Thesis CDF/THESIS/BOTTOM/PUBLIC/7352, Massachusetts Institute of Technology (2004).

[4] Y. Le, M. Martin, and P. Maksimović, *Observation of* $\Lambda_b \to \Lambda_c^+ \pi^-$ *and the Measurement of* $f_{\Lambda_b}\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-)/f_{B^0}\mathcal{B}(B^0 \to D^- \pi^+)$, CDF Internal Note CDF/ANAL/BOTTOM/CDFR/6396, Johns Hopkins University (2004).

[5] A. Belloni, J. Piedra *et al.*, *Unbinned Likelihood Fit for B0 Mixing in Fully Reconstructed Decays*, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/7388 (2004), see Chapter 11.

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

[7] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, Phys. Rev. **D27**, 105 (1983).

[8] C. Caso *et al.*, Phys. Let. B (2004), see Monte Carlo techniques, page 289.

[9] Ken Bloom and David Dagenhart, *Muon-Reconstruction Efficiency for Winter 2003 Conferences*, CDF Internal Note CDF/ANAL/MUON/CDFR/6347 (2003).

[10] S.Giagu, M.Rescigno, *et al..*, *BR ratios and direct CP violation in Cabibbo supressed decays of* D^0, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/6391 (2003).

[11] M. Herndon, *et al..*, *Proton XFT Efficiency Estimate for* $\Lambda_b \to \Lambda_c^+ \pi^-$ *Analysis*, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/7301 (2004).

[12] M.-Q. Huang, H.-Y. Jin, J. G. Korner, and C. Liu (2005), `hep-ph/0502004`.

[13] D. Litvintsev, S.-S. Yu, and R. Tesarek, *Observation of* Λ_c^* *and* Σ_c *in semileptonic* Λ_b *decays*, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/7546, Fermilab, University of Pennsylvania (2005).

[14] R. Tesarek and S.-S. Yu, *A Study of Backgrounds to the Decay* $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/6599, Fermilab, University of Pennsylvania (2003).

[15] R. Tesarek, S.-S. Yu, and D. Litvintsev, *A Study of Backgrounds to the decay* $\Lambda_b \to \Lambda_c^+ \mu^- \overline{\nu}_\mu$, CDF Internal Note CDF/ANAL/BOTTOM/CDFR/7545, Fermilab, University of Pennsylvania (2005).

[16] CDF Collaboration, Phys. Rev. D **68**, 091101 (2003).

[17] D. Litvintsev, *CDF note in preparation*, Tech. Rep.

[18] D. Acosta *et al.* (CDF II) (2004), `hep-ex/0406078`.

[19] K. Lannon and K. Pitts, *Bottom Quark Production Using PYTHIA and HERWIG*, CDF Internal Note CDF/PHYS/BOTTOM/CDFR/6253, University of Illinois at Urbana-Champaign (2003).

[20] G. Gomez-Ceballos *et al.*, *B Monte Carlo homepage*, Tech. Rep. (2004), see http://www-cdf.fnal.gov/internal/physics/bottom/b-montecarlo/db/g020.txt.

[21] CDF Collaboration, Phys. Rev. D **65**, 052006 (2002).

[22] D. Acosta *et al.* (CDF), Phys. Rev. Lett. **91**, 241804 (2003), hep-ex/0307080.

[23] D. Acosta *et al.* (CDF) (2004), hep-ex/0412071.

[24] D. Acosta *et al.* (CDF), *Measurements of bottom anti-bottom azimuthal production correlations in proton antiproton collisions at s**(1/2) = 1.8-TeV*, Tech. Rep. (2004), hep-ex/0412006.

[25] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. **B592**, 1 (2004).

[26] M. Martin *et al.*, *Evaluation of fit systematics due to background shapes in* $\Lambda_b \to \Lambda_c^+ \pi^-$, Tech. Rep. CDF/DOC/BOTTOM/PUBLIC/6953, Johns Hopkins University (2004).

[27] A. Korn, G. Bauer and C. Paus, *Update on Calibration of Energy Loss and Magnetic Field using* J/ψ *Events in Run II*, CDF Internal Note CDF/DOC/BOTTOM/CDFR/6355, Massachusetts Institute of Technology (2003).

[28] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini CERN-DD/EE/84-1.

[29] D. Michael (1994), nuMI-NOTE-BEAM-0019.

[30] R. H. Dalitz, Phil. Mag. **44**, 1068 (1953).

[31] E. M. Aitala *et al.* (E791), Phys. Lett. **B471**, 449 (2000), hep-ex/9912003.

[32] A. K. Leibovich and I. W. Stewart, Phys. Rev. **D57**, 5620 (1998), hep-ph/9711257.

[33] M.-Q. Huang, J.-P. Lee, C. Liu, and H. S. Song, Phys. Lett. **B502**, 133 (2001), hep-ph/0012114.

[34] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Lett. **B586**, 337 (2004), hep-ph/0312319.

[35] J. Abdallah *et al.* (DELPHI), Phys. Lett. **B585**, 63 (2004), hep-ex/0403040.

[36] I. Dunietz, Phys. Rev. **D58**, 094010 (1998), hep-ph/9805287.

[37] P. Migliozzi, G. D'Ambrosio, G. Miele, and P. Santorelli (CHORUS), Phys. Lett. **B462**, 217 (1999), hep-ph/9906219.

[38] G. Gómez-Ceballos, *Private communications.*