CDF/PHYS/BOTTOM/CDFR/8396
Revision 4.0
May 10, 2007

Search for New Heavy Baryons Σ_b and Σ_b^* in the Two Track Trigger Data Using 1 fb^{-1}

I. Gorelov, D. Litvintsev, P. Maksimovic, M. Martin,
J. Pursley, M. Schmidt, S. Seidel, R. J. Tesarek, E. Vataga

Fermi National Accelerator Laboratory
Johns Hopkins University
University of New Mexico

May 10, 2007

Abstract

This note describes an unbiased search for a new bottom baryon Σ_b in the fully reconstructed decay mode $\Sigma_b^{(*)\pm} \rightarrow \Lambda_b^0\pi^\pm$. We report the sample selection, analysis strategy and cut optimization. We observe the four $\Lambda_b^0\pi$ resonant states with the following masses:

$$\mathrm{m}(\Sigma_b^-) = 5815.2 \pm 1.0 \text{ (stat) } \pm 1.7 \text{ (syst) MeV}/c^2$$
$$\mathrm{m}(\Sigma_b^+) = 5807.8^{+2.0}_{-2.2} \text{ (stat) } \pm 1.7 \text{ (syst) MeV}/c^2$$
$$\mathrm{m}(\Sigma_b^{*-}) = 5836.4 \pm 2.0 \text{ (stat) } ^{+1.8}_{-1.7} \text{ (syst) MeV}/c^2$$
$$\mathrm{m}(\Sigma_b^{*+}) = 5829.0^{+1.6}_{-1.8} \text{ (stat) } ^{+1.7}_{-1.8} \text{ (syst) MeV}/c^2$$

The hypothesis of these four states as a background fluctuation can be excluded at greater than the 5 σ level.

Contents

1 Motivation

Hadron colliders and e^+e^- machines provide a wealth of experimental data on bottom mesons. Yet at this time only one bottom baryon, the Λ_b^0, has been directly observed in the channels $\Lambda_b^0 \rightarrow \Lambda J/\psi$ and $\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$. The inclusive semileptonic decay of Λ_b^0 has also been observed in the form of of $\Lambda_c^+\ell^-$ correlations at LEP and the Tevatron [1]. Currently the CDF collaboration posseses the world's largest data sample of bottom baryons, due to a combination of two factors – the CDF displaced track trigger, and the $\sim 1\ \mathrm{fb}^{-1}$ of integrated luminosity delivered by the Tevatron. Using a sample of fully reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$ collected on the displaced track trigger, we search for the decay $\Sigma_b^{(*)\pm} \rightarrow \Lambda_b^0\pi^\pm$.

None of the $\Sigma_b^{(*)\pm}$ states have been established. In 1995 the DELPHI Collaboration claimed evidence for the two states Σ_b and Σ_b^*, reconstructed as broad peaks in the inclusive mass spectrum M("b-baryon jet",π). The "b-baryon jet" was identified as a jet associated with a displaced vertex and enriched with baryon content (protons or Λ's) [2]. The work remained unpublished as the result turned out to be non-reproducible [3].

Given the size of the fully reconstructed $\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$ sample in the CDF data collected through February 2006, and drawing from a positive experience with a similar study of B^{**} mesons performed on a sample of about 4,000 fully reconstructed B decays [4], we are optimistic about our ability to see Σ_b . We hope to establish the Σ_b and Σ_b^* states unambiguously, and to measure their masses and the isospin splitting within the charged members of the triplet.

2 Baryon Classification in HQET

The QCD treatment of quark-quark interactions significantly simplifies if one of the participating quarks is much heavier than the QCD confinement scale $\Lambda_{\mathrm{QCD}} \approx 400$ MeV/c^2. Therefore, it is convenient to calculate properties of heavy hadrons in the limit of an infinitely heavy quark mass and then consider corrections to this limit in terms of systematic $\Lambda_{\mathrm{QCD}}/\mathrm{m_Q}$ expansion, where $\mathrm{m_Q}$ is the mass of the heavy quark in this approximation. This approach is known as Heavy Quark Effective Theory and has been successfully applied to describe available experimental data on $Q\bar{q}$ mesons for the ground states and the lowest P-wave excitations in both the charm and b sectors.

In the limit of $\mathrm{m_Q} \rightarrow \infty$, the angular momentum and flavor of the light quark become good quantum numbers. Therefore, the HQET approach views a baryon made of one heavy quark and two light quarks as consisting of a heavy static color field surrounded by a cloud corresponding to the light diquark system. In SU(3) the two quarks are in diquark form $\bar{3}$ and 6 mutliplets according to the decomposition $3 \otimes 3 = \bar{3} \oplus 6$, leading to a generic scheme of baryon classification. Diquark states containing quarks in an antisymmetric flavor configuration, $[q_1, q_2]$, are called Λ-type whereas the states with diquarks containing quarks in flavor symmetric state, $\{q_1, q_2\}$, are called Σ-type.

Thus in the ground Λ-type state the light diquark has isospin $I = 0$ and $J_l^P = 0^+$, which when coupled with a heavy quark leads to a state with total $J^P = \frac{1}{2}$. Ground state Λ-type baryons decay weakly. The lowest lying P-wave excitations of baryon states have an added unit of angular momentum, $L_l = 1$. Then the excited diquark has $J_l^P = 1^-$, and the heavy baryon states have total spin-parity $J^P = \frac{1}{2}^-, \frac{3}{2}^-$, which are called $\Lambda_b^*(\frac{1}{2})$ and $\Lambda_b^*(\frac{3}{2})$ respectively. They decay strongly, phase space allowing, to $\Lambda_b^0 2\pi$ or $\Sigma_b^{(*)}\pi$.

In the ground Σ-type state the light diquark has isospin $I = 1$ and $J_l^P = 1^+$. Together with the heavy quark this leads to a doublet of baryons with $J^P = \frac{1}{2}^+$ (Σ_b) and $J^P = \frac{3}{2}^+$ (Σ_b^*). The ground state Σ-type baryons decay strongly to Λ-type baryons by emitting pions. In the limit $\mathrm{m_Q} \rightarrow \infty$, both spin doublets $\{\Sigma_b, \Sigma_b^*\}$ and $\{\Lambda_b^*(\frac{1}{2}), \Lambda_b^*(\frac{3}{2})\}$ would be exactly degenerate, since an infinitely heavy quark does not have a spin interaction with a light diquark system. As the heavy quark is not infinitely massive, there will be a hyperfine mass splitting between the doublet states. We define the hyperfine mass differences as $m(\Sigma_b^{*+}) - m(\Sigma_b^+) \equiv \Delta_{*+}$ and $m(\Sigma_b^{*-}) - m(\Sigma_b^-) \equiv \Delta_{*-}$.

There is also an isospin splitting between the $\Sigma_b^{(*)-}$ and $\Sigma_b^{(*)+}$ states. Due to this isopsin splitting, the hyperfine splitting is not expected to be the same for the Σ_b^+ and Σ_b^- states, *i.e.* $\Delta_{*+} \neq \Delta_{*-}$. Using the PDG mass values for the Σ_s system, there is a difference of 3.68 ± 0.64 MeV/c^2 between the mass differences m(Σ^+) - m($\Sigma(1385)^+$) and m(Σ^-) - m($\Sigma(1385)^-$). The difference is expected to be smaller in the Σ_b system due to the much heavier b quark mass. Scaling the mass difference in the Σ_s system by the ratio of the s quark mass (538 MeV/c^2) to the b quark mass (4.9 GeV/c^2), we expect $\Delta_{*+} - \Delta_{*-} = 0.40 \pm 0.07$ MeV/c^2 [5, 6].

There are a number of predictions for the masses and isospin splittings of these states made using HQET, non-relativistic and relativistic potential models, $1/\mathrm{N}_c$ expansion, sum rules and lattice QCD. Table 2 summarizes existing theoretical predictions. The general consensus is that $\Delta M(\Sigma_b) = M(\Sigma_b) - M(\Lambda_b^0)$ are typically in the range of 180 to 210 MeV/c^2 with Σ_b^* 10-40 MeV/c^2 above Σ_b.

reference	$M(\Lambda_b^0)$[MeV/c^2]	$M(\Sigma_b)$[MeV/c^2]	$M(\Sigma_b^*)$[MeV/c^2]
[7, 8]	5596	5859	5877
[9]	5640	5780	5820
[10]	$\leq$5630$\pm$30	$\geq M(\Lambda_b^0)$+168	
[11]	$\leq$5629	5670-5826	$\geq$ 5710
[12]	5580	5800	5841
[13]	5547	5714	5766
[14]	5620	5800	5820
[16]		$5844.0\pm8.9(\overline{M}(\Sigma_b))$	$23.8\pm1.6(\Sigma_b^*-\Sigma_b)$
[17]	$5623\pm5\pm4$ (exp. inp.)	5824.2 ± 9.0	5840.0 ± 8.8
[15]	$5640\pm50^{+30}_{-20}$	$5770\pm60\pm40$	$5780\pm60^{+40}_{-30}$
[19]	$5664\pm98^{+33}_{-46}$	$141\pm24^{+30}_{-22}(\Sigma_b-\Lambda_b^0)$	$22\pm10^{+7}_{-6}(\Sigma_b^*-\Sigma_b)$
[20]	5629-5663	5844-5871	5874-5900
[21]	5622	5805	5834

Table 1: Mass predictions for Λ_b^0 and Σ_b baryons

There are also three independent constraints relating the charm and bottom systems [22]:

$$M(\Lambda_b^0)-M(\Lambda_c)\approx\overline{M}(B)-\overline{M}(D)=3340\ \mathrm{MeV/c^2},$$

$$\overline{M}(\Sigma_b)-M(\Lambda_b)\approx\overline{M}(\Sigma_c)-M(\Lambda_c)=167\ \mathrm{MeV/c^2},$$

$$\frac{M(\Sigma_b^*)-M(\Sigma_b)}{M(\Sigma_c^*)-M(\Sigma_c)}\approx\frac{M(B^*)-M(B)}{M(D^*)-M(D)}=0.33,$$

where $\overline{M}$ denotes spin averaged mass, which cancels hyperfine interactions between the heavy quark and light degrees of freedom. $\overline{M}$ takes the form $(M(D)+3M(D^*))/4$ for ground state mesons and $(M(\Sigma)+2M(\Sigma^*))/3$ for spin $\frac{1}{2}$, $\frac{3}{2}$ baryons. Corrections to the first two equations are on the order of $\Lambda_{QCD}^2(\frac{1}{2\mathrm{m_c}}-\frac{1}{2\mathrm{m_b}})\sim 50\ \mathrm{MeV/c^2}$, while the corrections to the last equation are at the level of 25% [22].

We expect the natural width of $\Sigma_{c,b}$ baryons to be dominated by single pion transitions. Photon transitions of the type $\Sigma_{c,b}\rightarrow\Lambda_{c,b}\gamma$ are expected to have significantly smaller ($\sim$ 100 keV) partial widths than the single pion transition, and are thus negligible. The partial width of the P-wave one-pion transition depends on the available phase space. For charmed baryons, this partial width is given by the following equation [23]:

$$\Gamma_{\Sigma_c\rightarrow\Lambda_c^+\pi}=\frac{1}{6\pi}\frac{M_2}{M_1}\left|f_p\right|^2|\vec{p}|^3 \tag{1}$$

where $f_p=g_A/f_\pi$; $g_A=0.75$ is the constituent pion-quark coupling, and $f_\pi=92$ MeV is the pion decay constant. For the charmed baryons, $M_2=M_{\Lambda_c^+}$ and $M_1=M_{\Sigma_c}$. The momentum of the pion in the Σ_c center of mass (CM) frame is $\vec{p}$. This formula predicts widths for $\Sigma_c^{(*)}$ baryons which

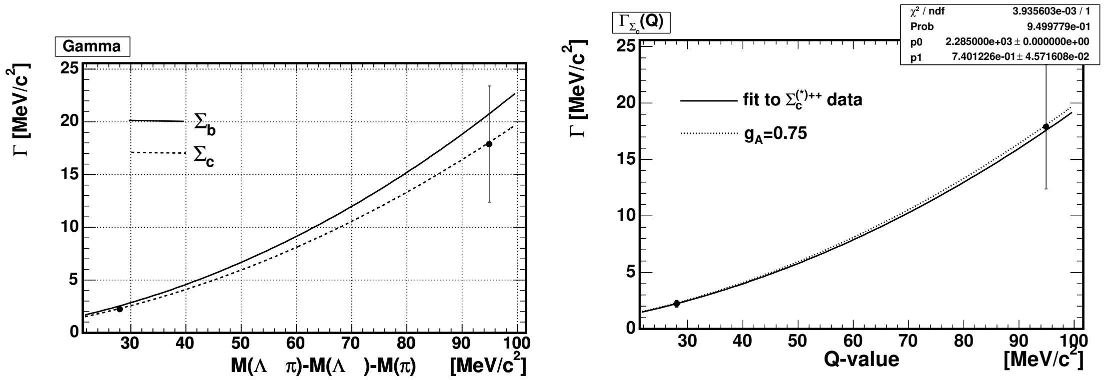


Figure 1: The plot on the left shows the natural width Γ of Σ_c baryons (dashed line) and Σ_b baryons (solid line) according to Eq. (1) as a function of the decay Q-value (defined as $Q = M_{\Sigma_{c,b}} - M_{\Lambda_{c,b}} - m_\pi$). Points with error bars show the PDG values for Σ_c and Σ_c^* [1]. The plot on the right shows the fit to PDG data on Σ_c with the parameter g_A floating.

are in excellent agreement with the PDG data [1], as can be seen in Fig. 2. Fitting the PDG data with the parameter g_A left free gives:

$$g_A = 0.75 \pm 0.05.$$

Eq. (1) is also valid for Σ_b baryons, by replacing M_2 with the Λ_b^0 mass and M_1 with the Σ_b mass. The momentum of the Σ_b pion in the CM frame ($\vec{p}$) is also precisely determined by the masses of the particles (Σ_b , Λ_b^0 , and π). For this analysis the mean value $g_A = 0.75$ is used to predict the Σ_b widths. The uncertainty obtained from the fit to Σ_c data is used to estimate systematic errors from fixing the Σ_b width to this equation (as described in Sec. 10.2.4). For the predicted Σ_b masses, we expect relatively narrow states – approximately 8 MeV/c^2 for Σ_b and 15 MeV/c^2 for Σ_b^* states. These narrow widths should make both states readily observable in the strong decays to $\Lambda_b^0\pi$ given sufficient resolution on the mass difference. The isospin splitting between the negative and positive partners of the isospin triplet, predicted to be large in some models (see Tab. 2), may also be possible to measure.

Baryons	[24], MeV/c^2	[25], MeV/c^2
$\Sigma_c^0 - \Sigma_c^{++}$	-3.0	-1.4
$\Sigma_b^- - \Sigma_b^+$	+7.1	+5.6
$\Sigma_c^{*0} - \Sigma_c^{*++}$	-2.7	+0.1
$\Sigma_b^{*-} - \Sigma_b^{*+}$	+6.5	+5.4

Table 2: Theoretical predictions of isospin mass splittings.

3 Analysis Strategy

Fig. 2 illustrates the expected multihadron topology of one Σ_b decay mode. As shown in the figure, the pion from Σ_b decay originates from the interaction region, also known as the primary vertex (P.V.).

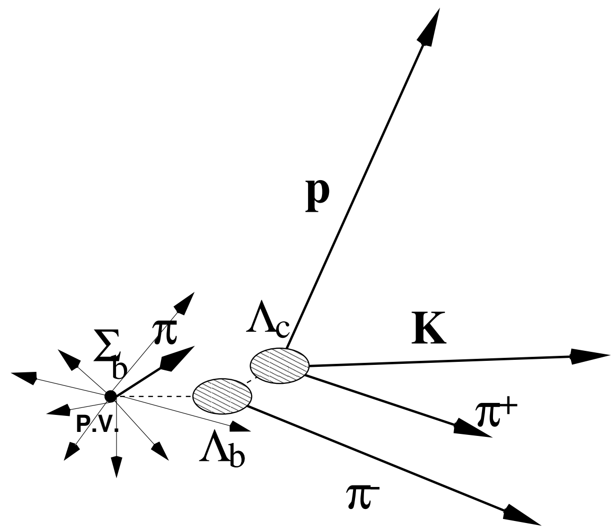


Figure 2: Sketch of the event topology of a Σ_b produced in the CDF detector.

Given the expected properties of Σ_b outlined in Section 2, we search for narrow resonances in the mass difference distribution of $Q = m(\Lambda_b^0\pi) - m(\Lambda_b^0) - m(\pi)$. Unless explicitly stated, Σ_b refers to both the $J = \frac{1}{2}$ ($\Sigma_b^{\pm}$) and the $J = \frac{3}{2}$ ($\Sigma_b^{*\pm}$) states.

The Λ_b^0 reconstruction is documented in another CDF note where the same sample is used to measure the Λ_b^0 mass [26]. The analysis strategy is as follows:

1. Strip a data sample of two track trigger events enriched in $\Lambda_b^0 \to \Lambda_c^+\pi^-$; $\Lambda_c^+ \to$ $pK^-\pi^+$ decays. [1]

2. Reconstruct $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays with maximum significance for the Λ_b^0 signal.

3. Add a fifth track and form the Q distribution, omitting the region where the Σ_b signals are expected. The blinded region is $0.03 < Q < 0.1$ GeV/c^2, which corresponds to a Σ_b mass region of $5.787 <$ m(Σ_b) < 5.857 GeV/c^2.

4. Estimate and understand backgrounds to any potential signal.

5. Estimate Q resolutions for the signal.

6. Once confidence in the estimated backgrounds is established, add the omitted region of the Q distribution and evaluate the significance of any possible signal.

[1] Unless otherwise stated all references to the specific charge combination imply the charge conjugate combination as well.

7. If a signal is observed, fit the Q distribution to estimate Σ_b yields and masses.

In analogy with B-meson hadronization chain we also define separate "$\Lambda_b^0\pi^-$" or SC and "$\Lambda_b^0\pi^+$" or OC combinations. As Λ_b^0 is neutral, the charge of the soft pion track determines the charge of the Σ_b , and there will be Σ_b signal for both positive and negative pions. $\Lambda_b^0\pi^-$ is defined as events where the Σ_b pion has the same charge as the pion from Λ_b^0 , and $\Lambda_b^0\pi^+$ events as events where the Σ_b pion charge is the opposite of the Λ_b^0 pion's charge. With these definitions, the SC contains all $\Sigma_b^{(*)-}$ and $\overline{\Sigma}_b^{(*)-}$ while OC contains $\Sigma_b^{(*)+}$ and $\overline{\Sigma}_b^{(*)+}$. Thus we expect to see two Σ_b peaks in SC and two in OC, and the masses are predicted to be slightly higher in SC than in OC (because Σ_b^- is expected to be heavier than Σ_b^+, as shown in the predictions in Table 2).

4 Data and Monte Carlo Samples

4.1 Data Sample

The present analysis is based on events collected by the CDF detector from February 2002 through February 2006, with an integrated luminosity of $\mathcal{L} = 1070 \pm 60$ pb^{-1} before good run criteria are applied. We apply the same good run criteria as the CDF B_s mixing analysis [27]; this criteria reduces the number of reconstructed Λ_b^0 by about 8%. The stripped sample described in detail in Ref. [26] is used for this analysis. This sample is a combination of the compressed all-hadronic datasets `xbhd0d`, `xbhd0h` and `xbhd0i`. We use the `B_CHARM_SCENA` and `B_CHARM_HIGHPT` flavours of the two displaced track trigger for this search. There also exists a dataset with data which pass on the `B_CHARM_LOWPT` trigger configurations. Events in this dataset which pass the `B_CHARM_LOWPT` but not the `B_CHARM_SCENA` trigger configuration may be added to this analysis in the future to increase the Λ_b^0 statistics. We also discovered that there are some duplicated events present in this sample – events in the run ranges from 186040 to 211539 which were processed twice and thus contain exactly the same Λ_b^0 candidates. The second instance of these events was removed from the analysis.

4.2 Λ_b^0 Reconstruction

The stripped sample was produced with very loose cuts and reconstructed with the Universal Finder [28]. [2] We added in the Layer-00 (L00) silicon hits and also used Event by Event Primary Vertexing. The tcl inputs for the Universal Finder track reconstruction are shown in Appendix A. To reconstruct a Λ_b^0 candidate, the selection module loops over three tracks, assumed to be proton, kaon, and pion candidates. These tracks must each pass basic track quality requirements:

- At least 5 hits each in 2 axial layers of the Central Outer Tracker (COT)
- At least 5 hits each in 2 stereo layers of the Central Outer Tracker (COT)
- At least 3 $r\phi$ hits in the Silicon Vertex Tracker (SVXII)

Additionally, each track must have an impact parameter (d_0) from the primary vertex of less than 0.1 cm and a transverse momentum of more than 500 MeV/c. The proton candidate is additionally required to have a transverse momentum greater than the pion candidate, and greater than 2.0 GeV/c. Once these tracks are selected, a kinematic fit is performed using `CTVMFT` to constrain the tracks to a common vertex. If this fit is successful, the following cuts are applied to this Λ_c^+ candidate:

- $\chi^2_{xy} < 30$
- $p_{\mathrm{T}}(\Lambda_c^+) > 4.3$ GeV/c
- $2.269 < |\mathrm{M}(pK\pi)| < 2.301$ GeV/c^2.

[2] We reconstructed Λ_b^0 from a xbhd0h subset of the data using three different packages: the Universal Finder, the JpsiXMods [29], and the BottomMods package [30]. No significant discrepancy in the Λ_b^0 yields or signal between the three methods was found.

If the above criteria are satisified, the program enters a loop over the fourth track. This track, assumed to be a pion candidate, must again pass track quality cuts and have an impact parameter < 0.1 cm, with a transverse momentum (p_{T}) of > 2.0 GeV/c. Another kinematic fit constrains this track along with the previous $pK\pi$ candidate vertex to form the Λ_b^0 candidate vertex. For this fit, the mass of the $pK\pi$ candidate is constrained to the Λ_c^+ mass from the PDG [1]. The requirements on this Λ_b^0 candidate are:

- $\chi^2_{xy} < 30$
- $4.8 < |\mathrm{M}(pK\pi\pi)| < 7.0$ GeV/c^2
- $p_{\mathrm{T}}(pK\pi\pi) > 6.0$ GeV/c
- $-0.007 < ct(\Lambda_c^+) < 0.028$ cm (w.r.t to Λ_b^0 vertex)
- $ct(\Lambda_b^0) > 0.025$ cm.

These are the basic requirements to reconstruct a Λ_b^0 candidate from the stripped sample, not the final analysis cuts. Cut optimization of the Λ_b^0 sample is described in detail in Ref. [26]. The final cuts determined by this optimization and applied in this analysis are listed in Table 3. Using these cuts on the stripped data sample gives a $\Lambda_b^0 \to \Lambda_c^+\pi^-$ yield of 2927 ± 58 candidates in the signal region m(Λ_b^0) $\in$ [5.565, 5.670] GeV/c^2, with the Λ_b^0 mass plot shown in Figure 3.

Variable	Cut value
	B_CHARM_SCENA
$p_{\mathrm{T}}(\pi_b^-)$	> 2 GeV/c
$p_{\mathrm{T}}(p)$	> 2 GeV/c
$p_{\mathrm{T}}(p)$	$>$ $p_{\mathrm{T}}(\pi^+)$
$p_{\mathrm{T}}(K^-)$	> 0.5 GeV/c
$p_{\mathrm{T}}(\pi^+)$	> 0.5 GeV/c
$ct(\Lambda_b^0)$	$> 250\,\mu$m
$ct(\Lambda_b^0)/\sigma_{ct}$	> 10
$\lvert d_0(\Lambda_b^0)\rvert$	$< 80\,\mu$m
$ct(\Lambda_c^+ \leftarrow \Lambda_b^0)$	$> -70\,\mu$m
$ct(\Lambda_c^+ \leftarrow \Lambda_b^0)$	$< 200\,\mu$m
$\lvert m(pK^-\pi^+) - m(\Lambda_c^+)_{PDG}\rvert$	< 16 MeV/c^2
$p_{\mathrm{T}}(\Lambda_b^0)$	> 6.0 GeV/c
$p_{\mathrm{T}}(\Lambda_c^+)$	> 4.5 GeV/c
Prob(χ^2_{3D}) of Λ_b^0 vertex fit	$> 0.1\%$

Table 3: Analysis cuts determined for Λ_b reconstruction.

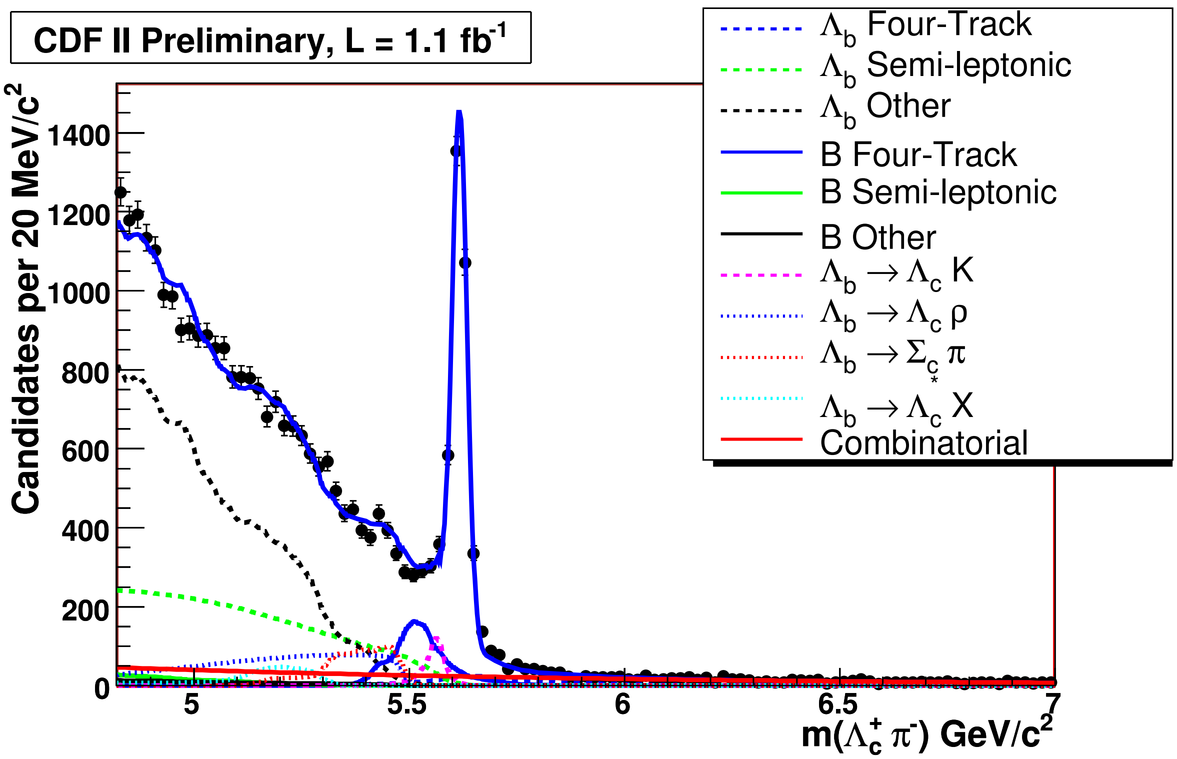


Figure 3: Λ_b^0 mass fit from Ref. [26]. The solid blue line is the total fit. The primary background components are listed in the legend and explained in detail in [26].

4.3 Monte Carlo Samples

Monte Carlo samples serve two purposes in this analysis: to measure the detector resolution for the Σ_b signal and to create templates for Σ_b background contributions. The different samples and their uses are listed briefly below:

- Various `BGenerator` B samples: These samples contain many different B decays and are reconstructed as Λ_b^0 to search for additional backgrounds in the Σ_b mass difference distribution. These are described in more detail in Sec. 6.1.

- $\Lambda_b^0 \to \Lambda_c^+\pi^-$ sample: This sample was generated with PYTHIA forcing the decay $\Lambda_b^0 \to \Lambda_c^+\pi^-$, with some of the Λ_b^0 coming from Σ_b decay. This is an msel = 5 sample, which only produces $b\bar{b}$ pairs through flavor creation. The PYTHIA default $p_{\rm T}$ spectrum of b baryons was used for generation, so the sample must be reweighted for Λ_b^0 $p_{\rm T}$. This sample is used to determine the Λ_b^0 hadronization (Ha) and the underlying event (UE) component of the Σ_b background, as described in Sec. 6.4.

- $\Sigma_b \to \Lambda_b^0\pi$ signal sample: This sample was generated with PYTHIA forcing the decay $\Sigma_b \to \Lambda_b^0\pi$ with zero intrinsic width for the four Σ_b states, and then forcing the Λ_b^0 decay $\Lambda_b^0 \to \Lambda_c^+\pi^-$. This is also an msel = 5 sample. The p_T spectrum of b baryons in the generation was corrected in accordance with Ref. [31]. This sample is used to measure detector resolution as described in Sec. 7.1.

Λ_b^0 reconstruction of these Monte Carlo samples is also performed with the Universal Finder, and follows the same procedure as described for data in Sec. 4.2.

4.4 Data-MC Comparisons

To estimate the Σ_b backgrounds coming from Λ_b^0 hadronization and the underlying event (described in detail in Sec. 6.4), we use a generic $\Lambda_b^0 \to \Lambda_c^+\pi^-$ PYTHIA sample. The PYTHIA default fragmentation and Λ_b^0 p_{T} spectrum were used in the generation. The default Λ_b^0 p_{T} spectrum in MC is harder than data, so we reweight the Monte Carlo for the Λ_b^0 p_{T}. To do this, we first normalize the MC to the same number of Λ_b^0 as in data. Then we plot the ratio of the data histogram to the MC histogram in bins of p_{T} and fit it to a linear function. The MC is then reweighted according to the following procedure:

1. For each event, the value of the reweighting variable (in this case the Λ_b^0 p_{T}) is determined.
2. The weight for this event is determined from the reweighting function. This is the value of the linear function for this value of the reweighting variable.
3. When filling histograms, the event is then weighted by the number calculated above.

We do not throw events away, but reweight all distributions with event-by-event weights. The Λ_b^0 p_{T} spectrum before and after reweighting are shown in Fig. 4. The fit parameters are given in Table 4.

The MC does not reproduce the data well for the soft tracks around the Λ_b^0 candidate, as shown in Fig. 5. For the same number of Λ_b^0 , the MC has fewer tracks, and the discrepancy is largest for low momentum tracks. To double check this, we also make data-MC comparison plots with $p_{\mathrm{T}} > 1$ GeV/c tracks. These comparison plots, available in Appendix B, do show better agreement between data and MC, although it is not good enough for our purposes.

Since we need the MC to match the data with no p_{T} cut on the tracks, we reweight the MC for the track p_{T} spectrum as well as the Λ_b^0 p_{T} spectrum. Track reweighting is done in the same way as for the Λ_b^0 p_{T}, by plotting the ratio of data to MC in bins of track p_{T} . The fit parameters are given in Table 4. After applying this weight to the remaining track histograms, the agreement between track quantities in data and MC becomes quite good. Figs. 5 through 11 show data to MC comparisons for track quantities both before and after reweighting for the track p_{T}.

Λ_b^0 p_T	Parameter Value	Error
p_0	1.30	0.06
p_1	-0.025	0.004
Fit Prob.	38%	–
Track p_T	Parameter Value	Error
p_0	1.73	0.06
p_1	-0.353	0.040
Fit Prob.	92%	–

Table 4: Parameter values for the linear functions used to reweight the Monte Carlo in Λ_b^0 p_T and track p_T .

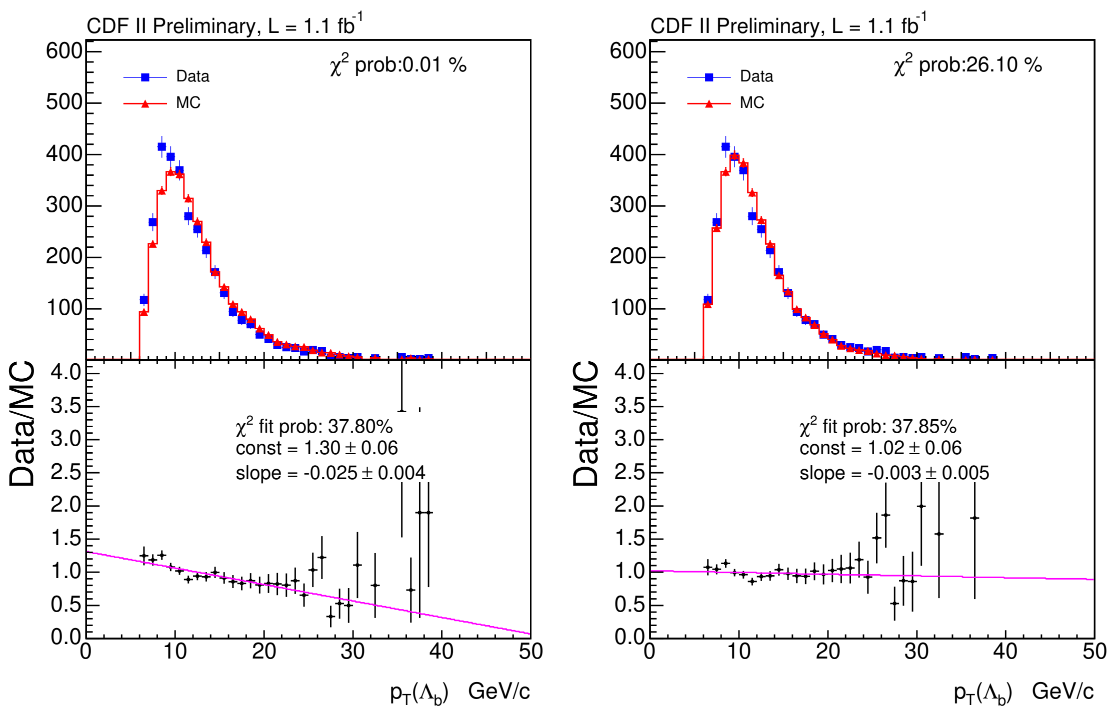


Figure 4: Data-MC comparison of Λ_b^0 p_T before (left) and after (right) reweighting for Λ_b^0 p_T . The linear fit to the left plot is used as the function to reweight the Monte Carlo. The right plot shows agreement with a straight line at +1.

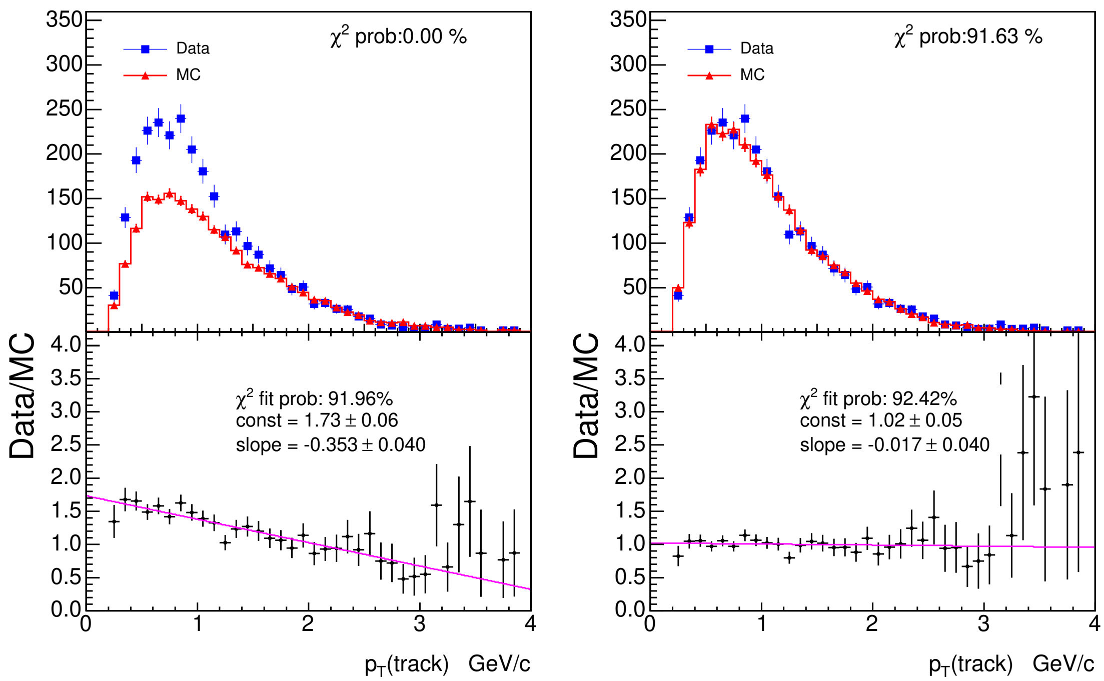


Figure 5: Data-MC comparison of track p_T before (left) and after (right) reweighting for track p_T . The linear fit to the left plot is used as the function to reweight the Monte Carlo. The right plot shows agreement with a straight line at +1.

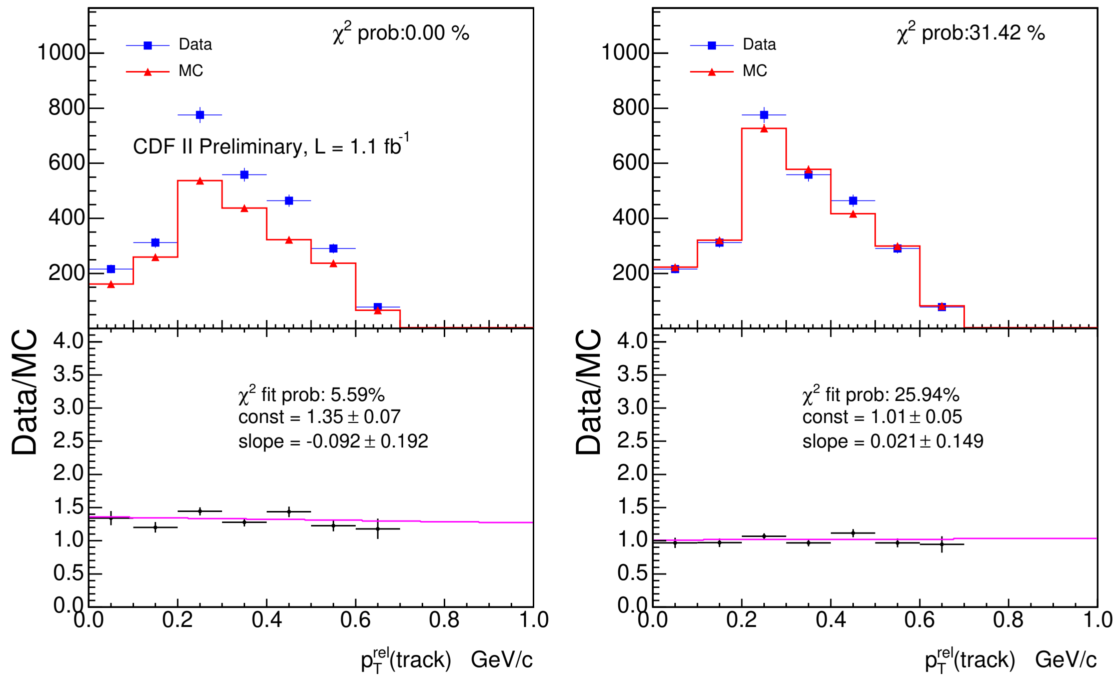


Figure 6: Data-MC comparison of track p_{T}^{rel} before (left) and after (right) reweighting for track p_{T} .

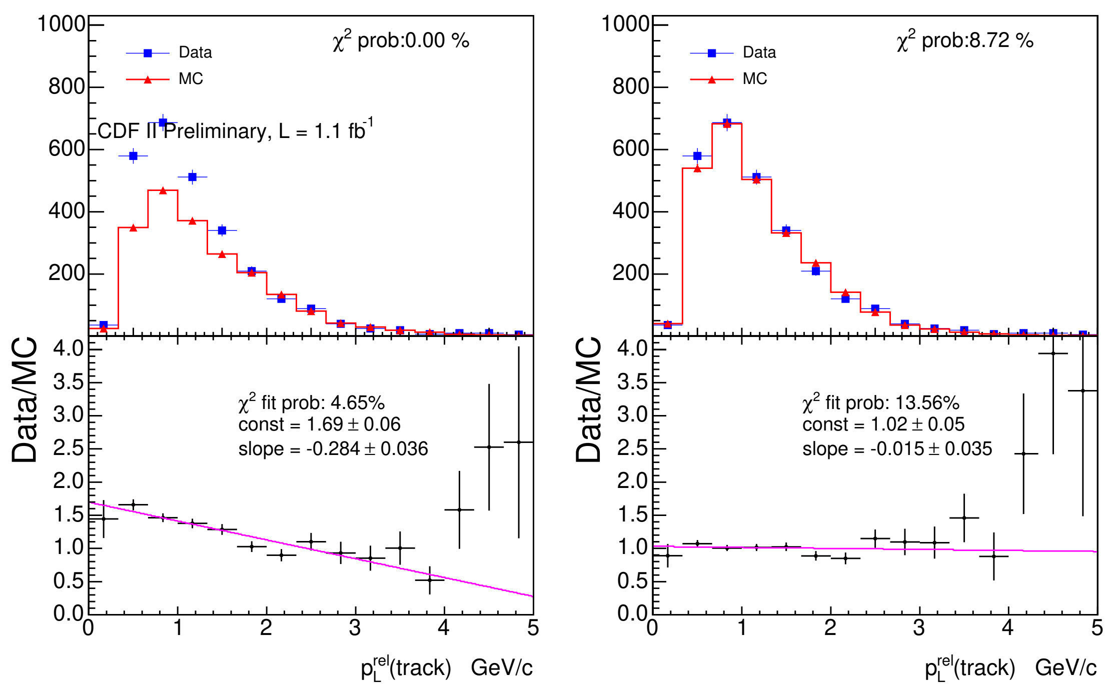


Figure 7: Data-MC comparison of track p_{L}^{rel} before (left) and after (right) reweighting for track p_{T} .

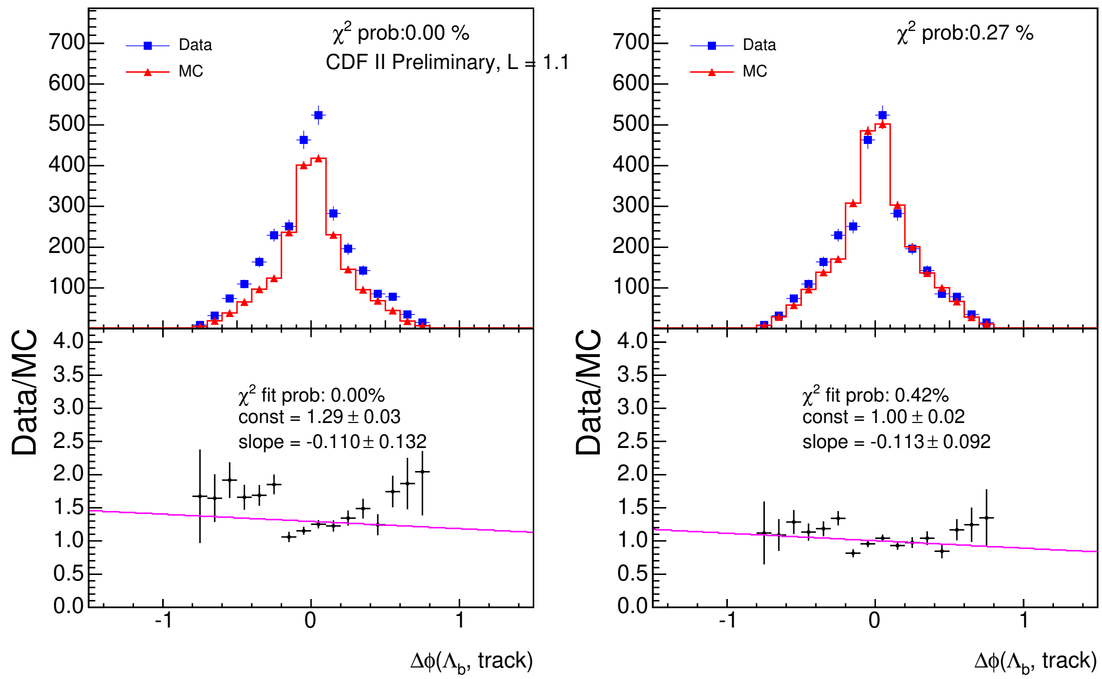


Figure 8: Data-MC comparison of $\Delta\phi$ between the track and the Λ_b^0 candidate before (left) and after (right) reweighting for track p_T .

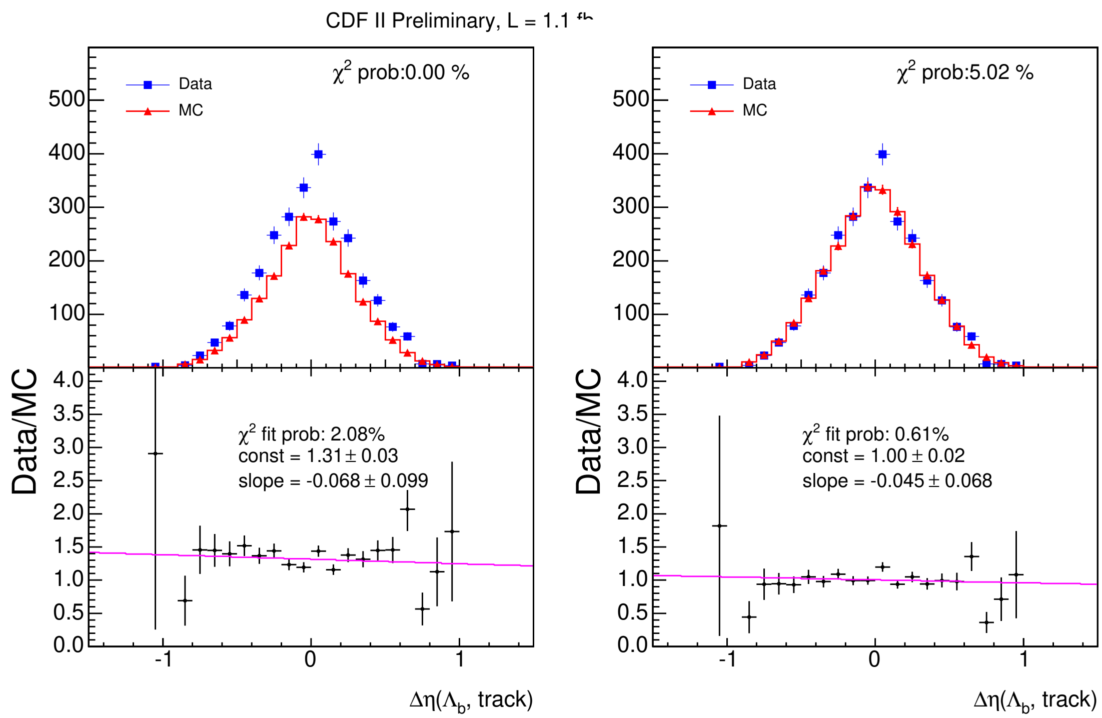


Figure 9: Data-MC comparison of $\Delta\eta$ between the track and the Λ_b^0 candidate before (left) and after (right) reweighting for track p_T .

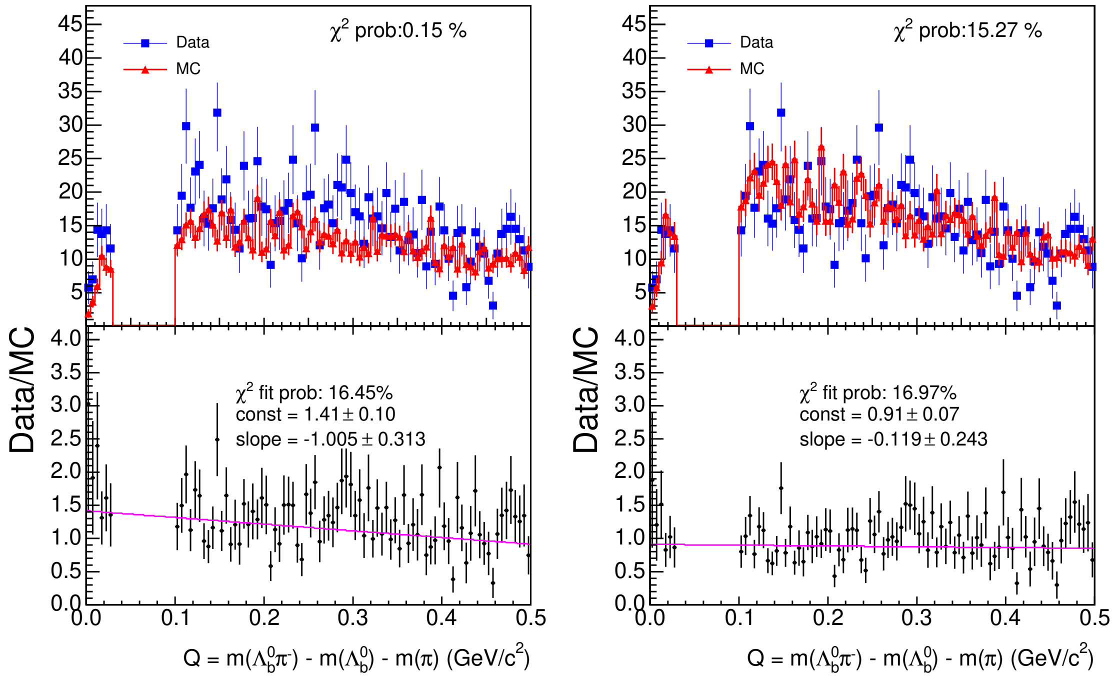


Figure 10: Data-MC comparison of the $\Lambda_b^0\pi^-$ Q distribution before (left) and after (right) reweighting for track p_{T} .

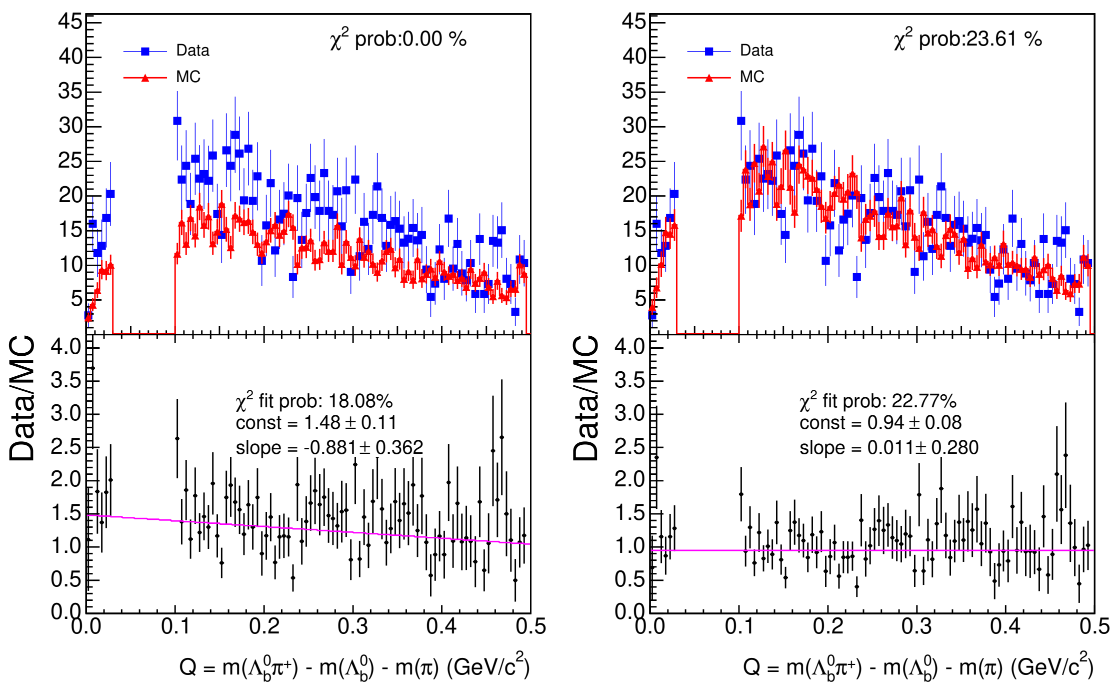


Figure 11: Data-MC comparison of the $\Lambda_b^0\pi^+ Q$ distribution before (left) and after (right) reweighting for track p_{T} .

5 Σ_b Reconstruction

After reconstructing the $\Lambda_c^+\pi^-$ mass, we select Λ_b^0 events in the mass window [5.565, 5.670] GeV/c^2. We then find a track originating from the region near the primary vertex, as depicted in Fig. 2, and perform another kinematic fit to constrain this track to a common vertex with the Λ_b^0 candidate. For this fit, the mass of the $pK\pi$ is again constrained to the Λ_c^+ mass although the combined $(pK\pi)\pi$ mass is not constrained to the Λ_b^0 mass. We also require that the Prob(χ^2_{3D}) of Σ_b vertex fit be $> 0.1\%$. We then form the $Q = m(\Lambda_b^0\pi) - m(\Lambda_b^0) - m(\pi)$ distribution. Initially, the only requirements on the Σ_b pion, denoted by π_{Σ_b} , are the default track quality cuts contained in defTracks [32]. This also includes silicon standalone tracks which do not pass the COT requirements.

Fig. 12 shows the resulting Q distributions for $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ with the Σ_b search region, $0.03 < Q < 0.1$ GeV/c^2, removed. To reduce backgrounds in the Q distributions, we search for a set of additional cuts optimized for π_{Σ_b} .

5.1 Optimization of the Σ_b selection criteria

The sources of background in the Σ_b reconstruction are described in detail in Sec. 6; the main components are:

- Λ_b^0 hadronization and underlying event ("Ha+UE")
- hadronization and underlying event of B meson reflections
- combinatorial background underneath the Λ_b^0 peak

Other backgrounds (*e.g.* from 5-track decays where one track is taken as the π_{Σ_b} candidate) are negligible, as shown in Section 6.1. Out of the three primary background sources, the first one ("Λ_b^0 Ha+UE") dominates by far.

We use the "Σ_b sidebands" outside of the blinded Σ_b signal region to represent the Σ_b background. These sidebands are defined as:

- **Lower Σ_b sideband:** $0 < Q < 30$ MeV/c^2
- **Upper Σ_b sideband:** $100 < Q < 500$ MeV/c^2

The signal for the optimization is taken from the Σ_b PYTHIA sample described in Sec. 4.3. To enhance the Σ_b signal with respect to the sources of background, we vary the minimum $p_T(\Sigma_b)$, as well as maximum $|d_0/\sigma_{d_0}|$ and minimum $\cos\theta^*$ of π_{Σ_b} candidate tracks. $\cos\theta^*$ is defined as the angle between π_{Σ_b} in the Σ_b rest frame and the direction of Σ_b in the lab frame.

We do not impose any p_T cut on the π_{Σ_b} candidate tracks, since, prior to optimization, about half of Σ_b events from our Monte Carlo sample lie below a 400 MeV/c^2 cut (Fig. 13). In addition, we want to keep the Q distribution in the signal region as level as possible, to reduce the systematic error on the masses of Σ_b states due to remaining uncertainties in the background shapes. A cut on $p_T(\pi_{\Sigma_b})$ would remove most of the candidates from the left corner of the Q distribution (the "left sideband"), and thus is a very poor choice of a selection variable. Instead we use $\cos\theta^*$, which is orthogonal to Q by definition but partially correlated with $p_T(\pi_{\Sigma_b})$. Increasing the minimal value of $\cos\theta^*$ of π_{Σ_b} candidates leaves the left corner of the Q distribution largely intact but significantly reduces the background in the signal region. As shown in Fig. 2, a true π_{Σ_b} originates from the primary vertex. By requiring the π_{Σ_b} candidate to be prompt, we reject poorly measured tracks and b-hadron daughters from the other b quark in the event, as well as tracks from spallation and other detector-related processes. Placing an upper limit on $|d_0/\sigma_{d_0}|$ of π_{Σ_b} suppresses these non-prompt components.

For a successful optimization, the sample of Σ_b signal events must be as close to real Σ_b events as possible. For this reason we apply two corrections to the PYTHIA Monte Carlo Σ_b signal sample: one for the $|d_0/\sigma_{d_0}|$ of the π_{Σ_b} candidate, and one for the Σ_b p_{T} distribution. We first compare the distributions of $|d_0/\sigma_{d_0}|$ for the upper Σ_b sideband region ($0.1 < Q < 0.5$ GeV/c^2) in data and Monte Carlo (Fig. 14, bottom and top, respectively) to see if the MC is correctly modeling the $|d_0/\sigma_{d_0}|$ distribution of candidate tracks. For each distribution, we fit the core Gaussian and obtain 1.12 ± 0.03 and 1.18 ± 0.03 in data and MC, respectively. While the two numbers are statistically compatible, we scale the MC down by 5% to account for a possible true discrepancy in the optimization. As a second correction to the PYTHIA Monte Carlo sample, we also reweight the $p_{\mathrm{T}}(\Sigma_b)$ spectrum from PYTHIA using the same functional form needed to make the $p_{\mathrm{T}}(\Lambda_b^0)$ spectrum from PYTHIA agree with data, as described in Sec. 4.4; the form is shown in Fig. 4. The $p_{\mathrm{T}}(\Sigma_b)$ spectrum is not known, so this reweighting is only an assumption made before Σ_b states are observed and their spectra measured explicitly. These two corrections to the PYTHIA MC, the $|d_0/\sigma_{d_0}|$ and $p_{\mathrm{T}}(\Sigma_b)$ reweighting, are both applied before performing the Σ_b optimization.

In the optimization, we use $\varepsilon(\Sigma_b)/\sqrt{B}$ as the score function, where $\varepsilon(\Sigma_b)$ is the efficiency to reconstruct candidates from the PYTHIA Σ_b Monte Carlo sample, and B is the total number of background events that pass the cuts. In the optimization we do not divide the events into "$\Lambda_b^0\pi^-$" and "$\Lambda_b^0\pi^+$" categories as we do later. We perform a simultaneous optimization of all three cuts ($p_{\mathrm{T}}(\Sigma_b)$, $|d_0/\sigma_{d_0}|$ and $\cos\theta^*$) using an iterative one-dimensional gradient algorithm with fixed step sizes. $p_{\mathrm{T}}(\Sigma_b)$ is optimized first while the other two variables are kept fixed; once the maximum of the score function is found, then $|d_0/\sigma_{d_0}|$ is optimized while the other two are kept fixed, and so on, until the score function is at the maximum with respect to all three cuts simultaneously. The algorithm uses discrete step sizes, so the optimal values of the cuts are rounded to the nearest step size. The step sizes are much smaller than the width of the maxima of the score function, and are listed in Tab. 5.

During this process, the ratio of the number of background events in the blinded search window given the number of observed events in the "Q sidebands" is kept fixed. This ratio is a function of the shape of the background: a blinded background distribution is fitted with a functional form from Eq. 2, which is then integrated inside and outside the search window. Since the cuts on both $p_{\mathrm{T}}(\Sigma_b)$ and $\cos\theta^*$ affect the shape of the Q distribution of the background events (although in different ways), assuming that this functional form is frozen is an approximation. However, the shape of the background may be fixed from a previous optimization; we thus perform several optimizations in a row until both the cuts and the background shape are stable. In the final optimization, the ratio of the expected number of background events in the search window and the number of events observed in the Q sidebands is ~ 0.21.

The result of the optimization is shown in Tab. 5, and on Fig. 15 and the $N-1$ scans are given in Figs. 16, 17, and 18. We find that $\cos\theta^*$ is the only variable which has non-negligible power to

Variable	Cut value	Step size
$p_{\mathrm{T}}(\Sigma_b)$	> 9.5 GeV/c	0.5 GeV/c
$\lvert d_0/\sigma_{d_0}\rvert$	< 3.0	0.25
$\cos\theta^*$	> -0.35	0.05

Table 5: The cuts determined for Σ_b reconstruction and the step sizes used by the optimization algorithm.

separate signal from the background. All maxima of the score function are broad, indicating that the optimization result is fairly stable, since slightly different choice of cuts would have yielded a very similar signal significance. After the optimization, we tighten the track quality cuts and apply the good run list. The Q distributions after all cuts are applied are shown in Fig. 19.

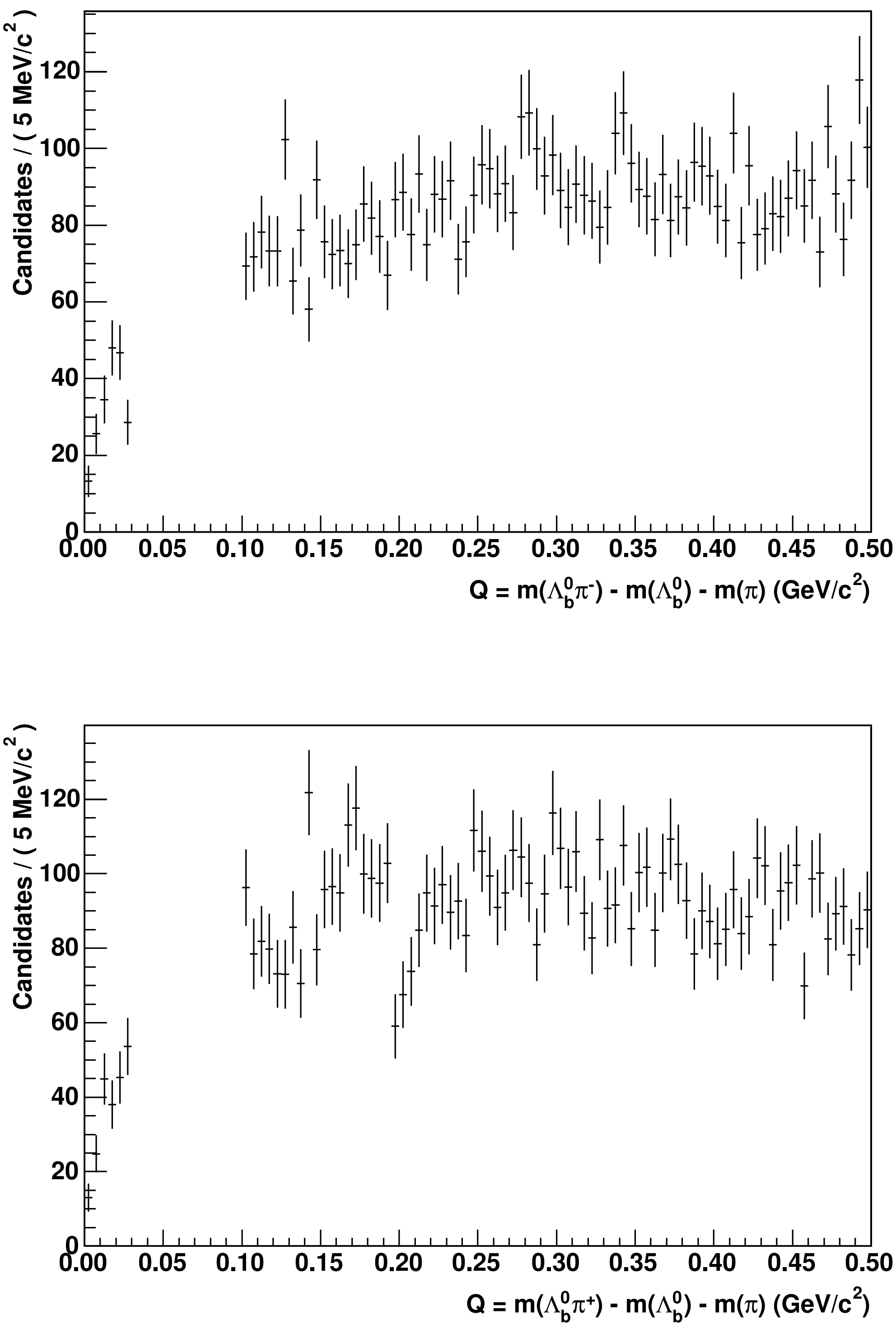


Figure 12: $\Lambda_b^0\pi^-$ (top) and $\Lambda_b^0\pi^+$ (bottom) Q distributions before applying optimized Σ_b cuts. For the distributions after optimized cuts are applied, see Fig. 19.

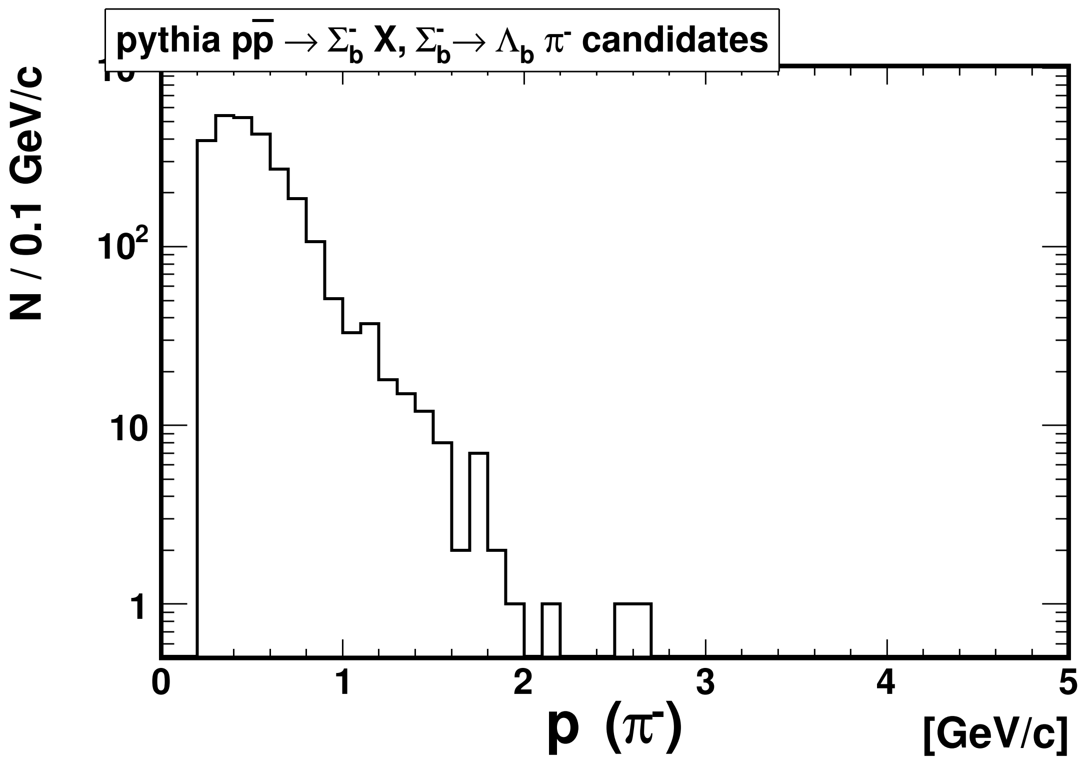


Figure 13: p_T distribution of the soft pion from Σ_b^- decay in PYTHIA Monte Carlo sample (after reweighting $p_T(\Lambda_b^0)$ distribution to match the one from data).

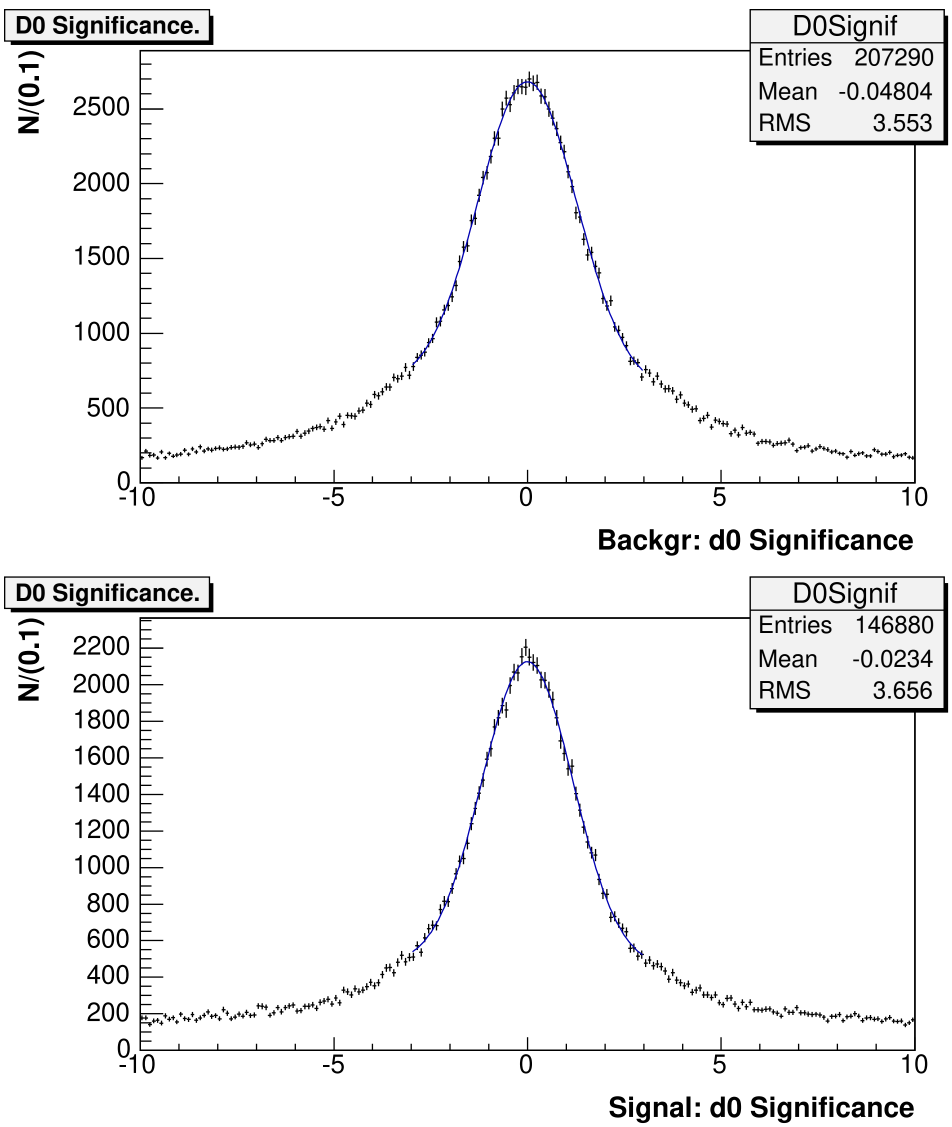


Figure 14: Distribution of d_0/σ_{d_0} of the π_{Σ_b} candidates from the upper Σ_b sideband ($0.1 < Q < 0.5$ GeV$/c^2$) in data (top) and Σ_b PYTHIA Monte Carlo (bottom). The sidebands in the Σ_b PYTHIA Monte Carlo are populated by Σ_b hadronization; these π_{Σ_b} candidates are prompt tracks, which is all that matters since we only compare the width of Gaussians centered at 0.

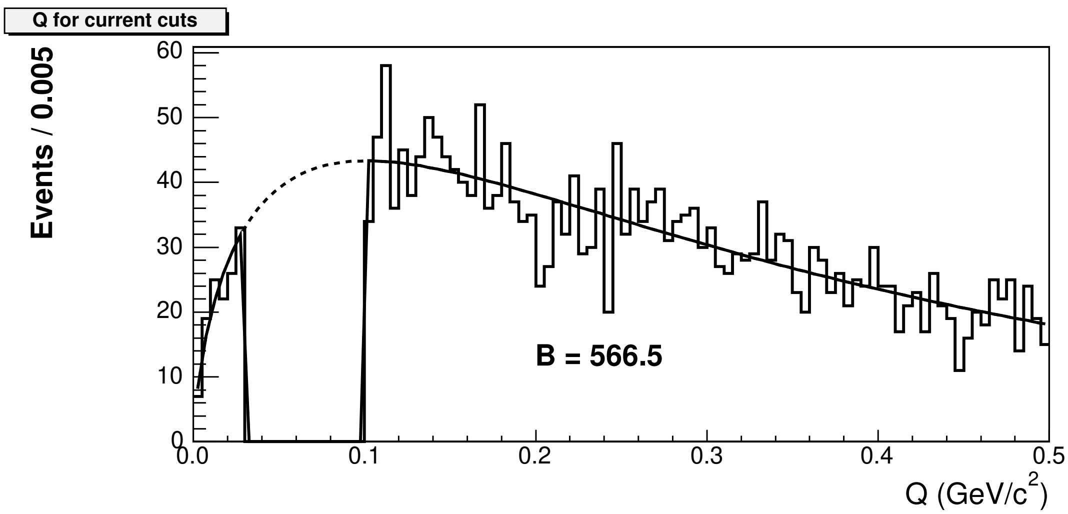


Figure 15: The result of the optimization procedure, with all background events combined. (The search window is blinded.) The total number of expected background events is 570; this number is different than the sum of Fig. 19 since the optimization was performed on ~ 922 pb^{-1} and in addition this distribution does not contain the good run list criterion as well as the final track quality cuts.

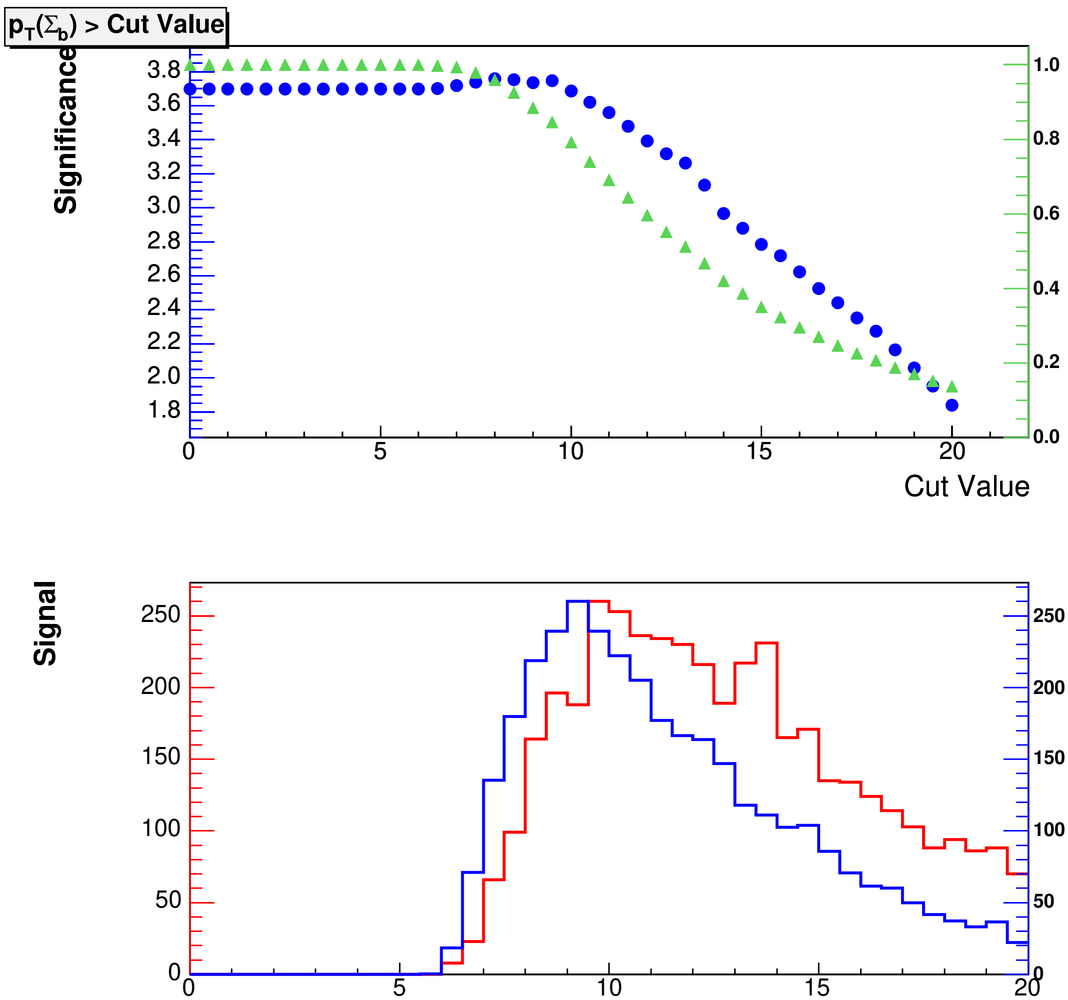
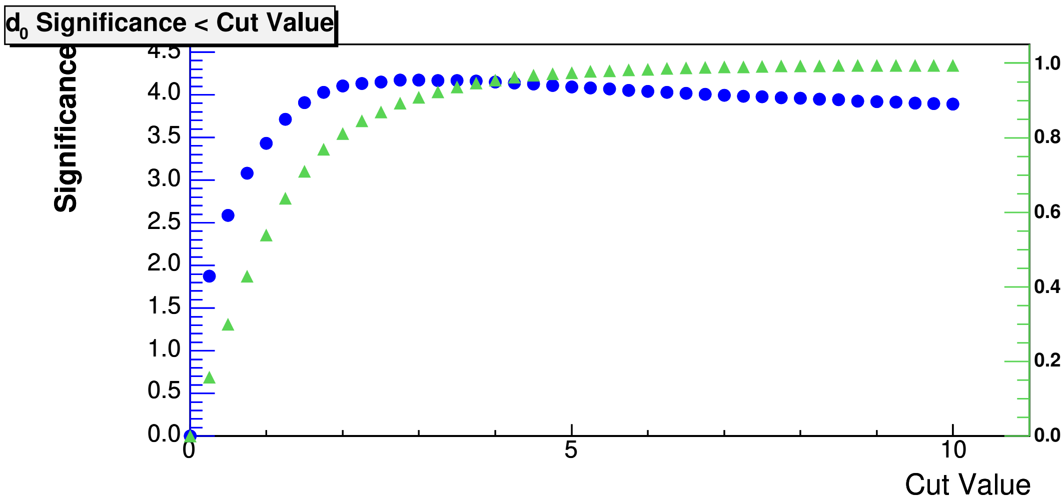


Figure 16: The "$N-1$" scan of the $p_{\mathrm{T}}(\Sigma_b)$ cut. Top: $\varepsilon(\Sigma_b)/\sqrt{B}$ as a function of $p_{\mathrm{T}}(\Sigma_b)$ (blue circles, left scale) and $\varepsilon(\Sigma_b)$ as a function of $p_{\mathrm{T}}(\Sigma_b)$ (green triangles, right scale). Bottom: distribution of $p_{\mathrm{T}}(\Sigma_b)$ for Σ_b signal (red histogram, left scale) and background in the Q sidebands (blue histogram, right scale). We cut at $p_{\mathrm{T}}(\Sigma_b) > 9.5$ GeV/c.



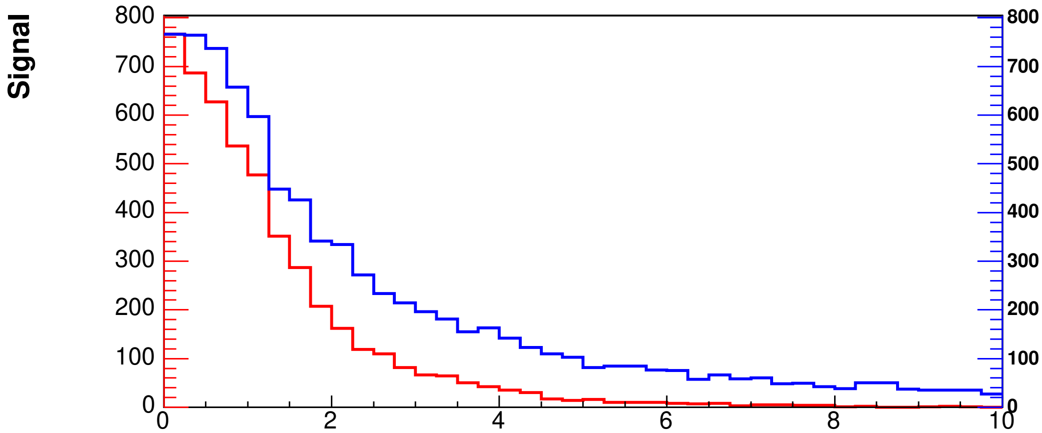


Figure 17: The "$N-1$" scan of the cut on $|d_0/\sigma_{d_0}|$ of the π_{Σ_b} candidates. Top: $\varepsilon(\Sigma_b)/\sqrt{B}$ as a function of $|d_0/\sigma_{d_0}|$ (blue circles, left scale) and $\varepsilon(\Sigma_b)$ as a function of $|d_0/\sigma_{d_0}|$ (green triangles, right scale). Bottom: distribution of $|d_0/\sigma_{d_0}|$ for Σ_b signal (red histogram, left scale) and background in the Q sidebands (blue histogram, right scale). We cut at $|d_0/\sigma_{d_0}| < 3.0$.

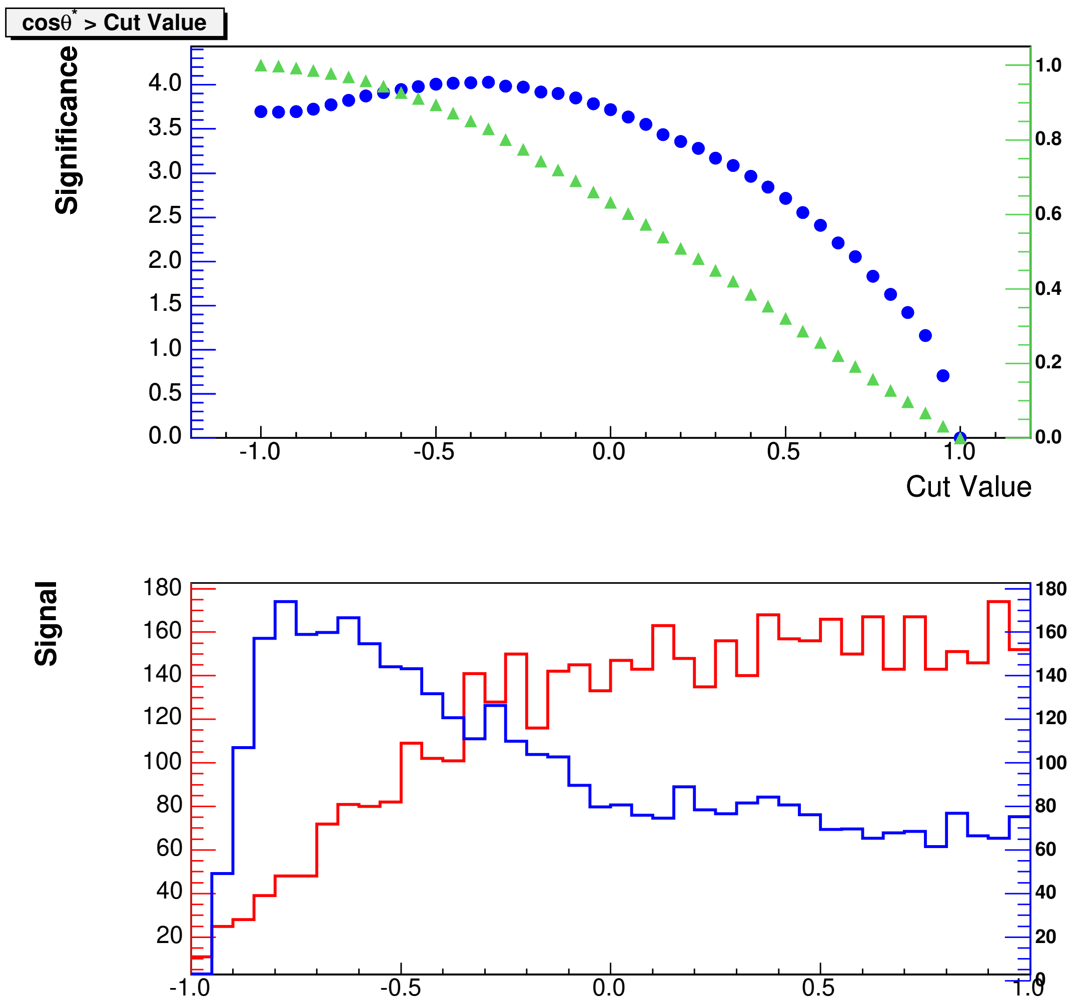
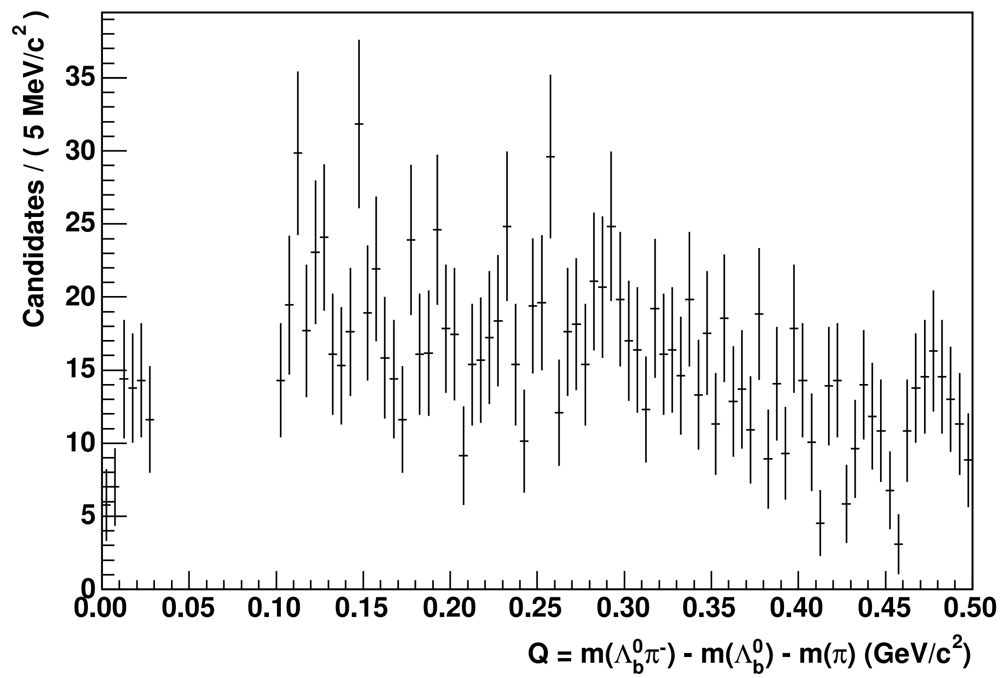


Figure 18: The "$N-1$" scan of the $\cos\theta^*$ cut. Top: $\varepsilon(\Sigma_b)/\sqrt{B}$ as a function of $\cos\theta^*$ (blue circles, left scale) and $\varepsilon(\Sigma_b)$ as a function of $\cos\theta^*$ (green triangles, right scale). Bottom: distribution of $\cos\theta^*$ for Σ_b signal (red histogram, left scale) and background in the Q sidebands (blue histogram, right scale). We cut at $\cos\theta^* > -0.35$.



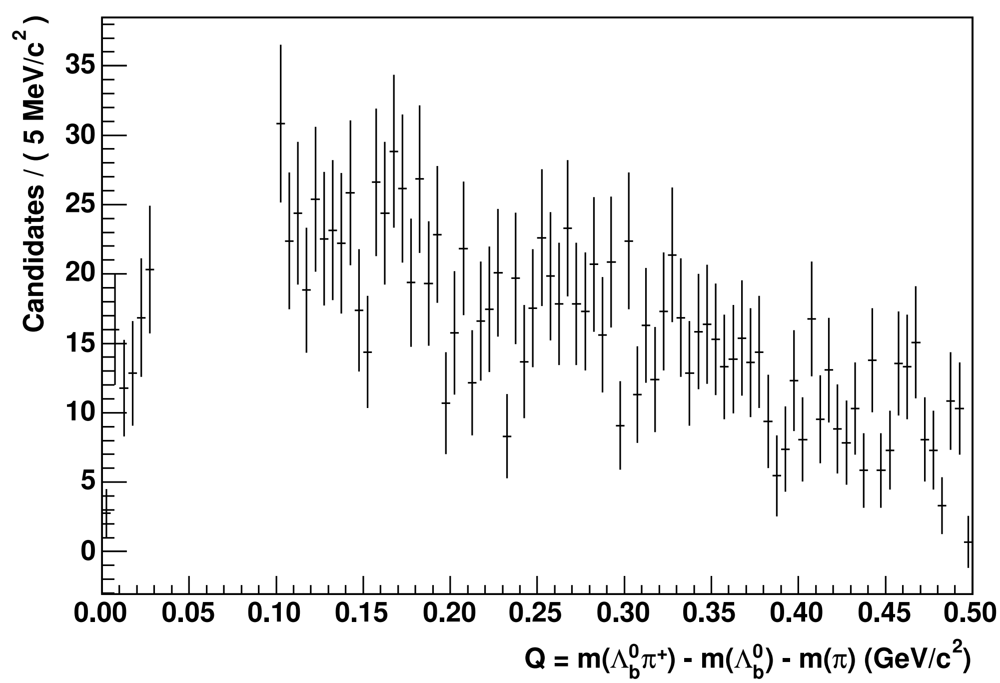


Figure 19: $\Lambda_b^0\pi^-$ (top) and $\Lambda_b^0\pi^+$ (bottom) Q distributions after the cuts listed in Table 5 are applied to the π_{Σ_b} candidate.

6 Fitting Backgrounds of the Q Distribution

The Σ_b fit is performed with the RooFit version 2.07 (included with ROOT 5.10/00 [33]). The various background components, which are described individually in the following sections, are fixed in the fit.

6.1 Other Backgrounds

The three main sources of Σ_b backgrounds are described in the following three sections. To check for other significant sources of backgrounds, such as from 5-track B decays where one track is taken as the π_{Σ_b} candidate, we looked at several large `BGenerator` B^+/B^0 Monte Carlo samples recently completed by an INFN/LBL/JHU joint effort, and reconstructed them as Σ_b candidates. The three available B decay modes are shown in Table 6, with the number of generated events and the number of Σ_b candidates passing the optimized analysis cuts in $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$. The number of candidates shown is for the entire range $Q \in [0.0, 0.5]$ GeV$/$c^2, and is counted after normalizing the samples to 1.1 fb^{-1} of data. Fig. 20 shows the Q distributions of the most significant background, which has only 16 events for SC and OC combined after normalization. These Q distributions show that these backgrounds are a negligible component of the total Σ_b background. Therefore, we do not consider these backgrounds in the fit.

Sample	Total events generated	SC Candidates	OC Candidates
$B^0 \to D^0\pi X$	1.5 billion	0	1
$B^0 \to D^-\pi X$	4.5 billion	4	12
$B^+ \to D^+\pi X$	4.5 billion	0	0

Table 6: Summary of the generic Monte Carlo samples considered in these background studies.

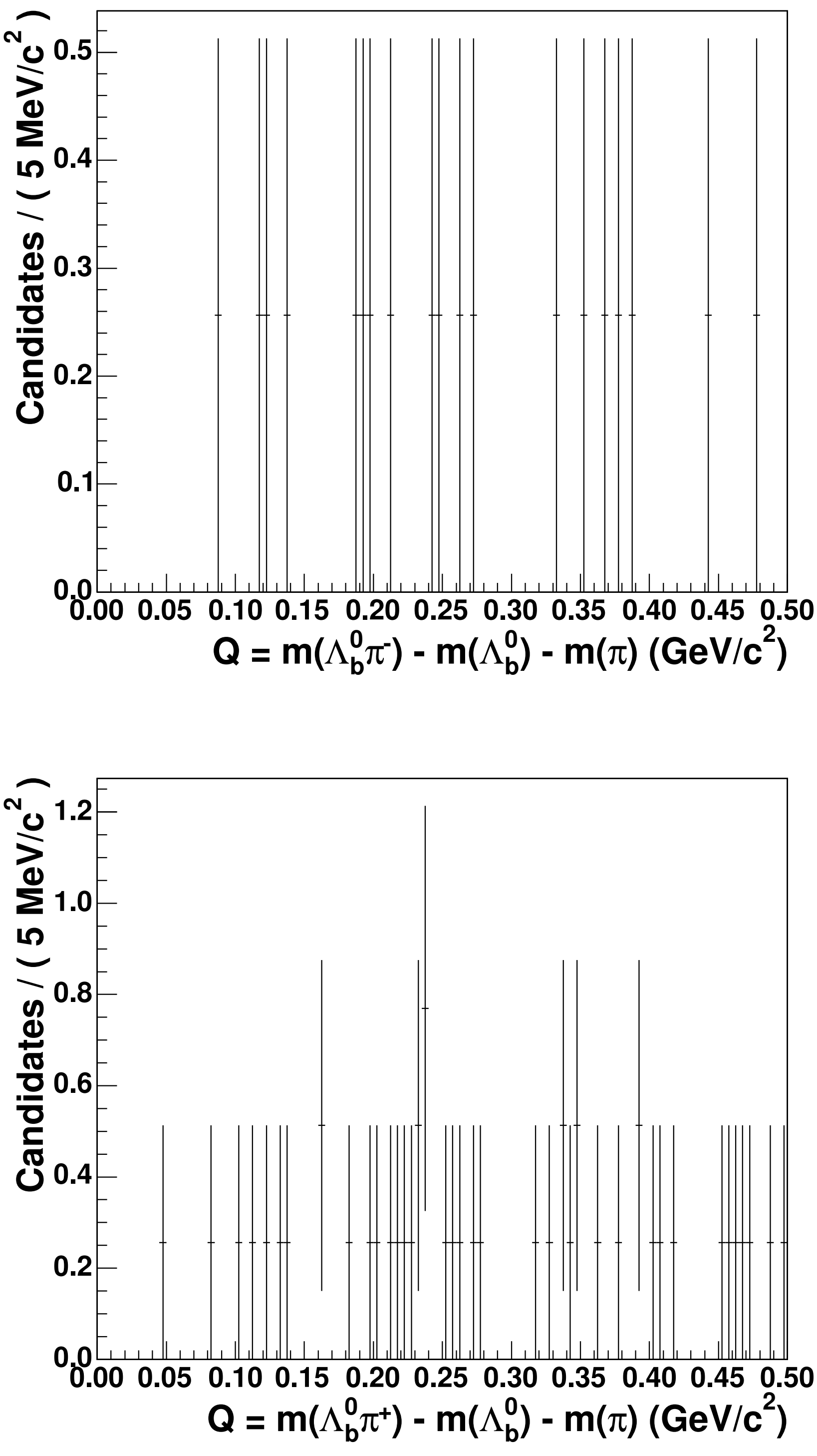


Figure 20: Normalized Q distributions for the $B^0 \to D^-\pi X$ sample after applying all Σ_b cuts. There is very little background in the Σ_b Q distribution due to these modes.

6.2 Combinatorial Background

The combinatorial background is taken from the high mass sideband, m(Λ_b^0) $\in$ [5.8, 7.0] GeV/c^2, of the Λ_b^0 mass fit, where it is fit with a decreasing exponential. The low mass sideband contains misreconstructed B decays, a background source which will be discussed in Sec. 6.3, so we use only the high mass sideband. The scale factor is the ratio of the area under the exponential for the signal region, m(Λ_b^0) $\in$ [5.565, 5.670] GeV/c^2, to the area under the sideband region. This gives a normalization factor value of 0.161 for the combinatorial background distributions.

Combinatorial background is fit with a function of the form

$$f(Q;\alpha,Q_{max},\gamma) = \left(\frac{Q}{Q_{max}}\right)^{\alpha} e^{-\frac{\alpha}{\gamma}\left(\left(\frac{Q}{Q_{max}}\right)^{\gamma}-1\right)} \tag{2}$$

Eq. (2) is a modified version of the function used to describe the backgrounds in Q spectra of B^{**} decays. It is know to fit well a variety of such shapes. In this parameterization, Q_{max} is the value of Q for which $f(Q;\alpha,Q_{max},\gamma)$ has a maximum for any α and γ; the variables α and γ regulate the shape of the function from 0 to Q_{max} and above Q_{max} respectively. Eq. (2) is not normalized to 1 – when building the PDF based on it, RooFit will do this automatically – however, it has the useful property that $f(Q_{max}) \equiv 1$ for any α and γ.

We perform a binnned likelihood fit of Eq. (2) to the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ combinatorial backgrounds is shown in Fig. 21. The values of the parameters are given in Table 7. Since the statistics are low, there are large fluctuations in the data and the fit parameters have very large errors. The number of events is then multiplied by the scale factor of 0.161 to give the correct normalization of the combinatorial background. Both the shape and the normalization of the combinatorial background are fixed in the fit on data.

	α	Q_{max}	γ	Events	Scaled Events
$\Lambda_b^0\pi^-$	0.5 ± 0.9	0.15 ± 0.10	1.3 ± 2.6	538 ± 33	87
$\Lambda_b^0\pi^+$	1.7 ± 2.9	0.12 ± 0.06	0.3 ± 1.3	528 ± 32	85

Table 7: Parameter values for the shape and normalization of the combinatorial background in both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$.

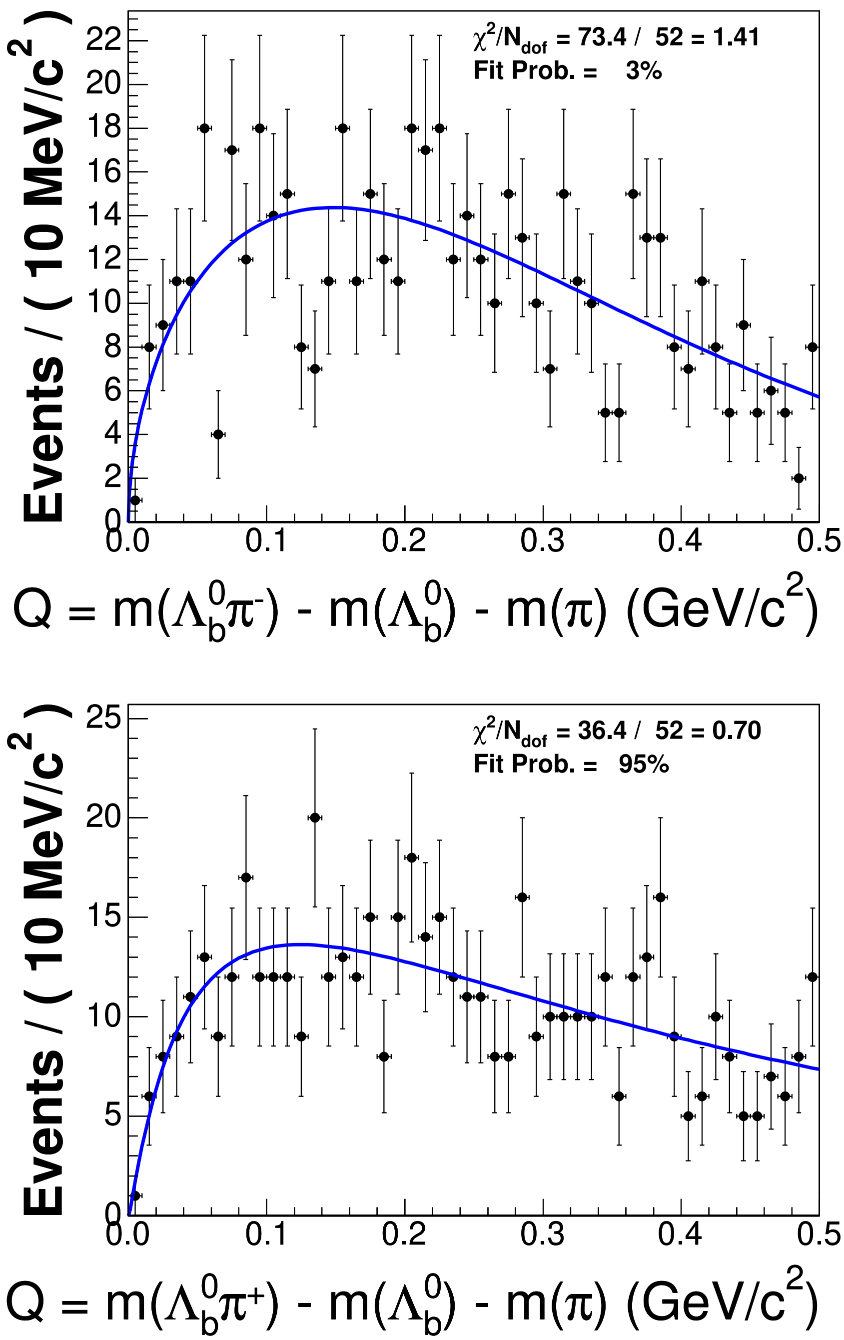


Figure 21: Fits to the combinatorial backgrounds. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$.

6.3 Physics Background

The physics background is primarily composed of real B^0 mesons misreconstructed as Λ_b^0 baryons. This includes B^0 mesons from $B^{**\pm}$ decay. Due to B^0 mixing, this background will be present in both the $\Lambda_b^0\pi^-$ and the $\Lambda_b^0\pi^+$ distributions.

To measure this background, we first reconstruct

$$B^0 \to D^-\pi^+; D^- \to \pi^- K^+\pi^-$$

By replacing one pion mass with the proton mass, we reconstruct the D^- as Λ_c and the B^0 as Λ_b^0 . All kinematic cuts from the Λ_b^0 analysis are then applied (*i.e.* p_{T} cuts, but not mass cuts) and we use those events to make a Σ_b candidate.

In a large B^0 PYTHIA MC sample with a B^{**} yield of 20%, $B^{**\pm}$ appear in the Σ_b mass range as two small peaks at $Q \sim 0.29$ and 0.31 GeV/c^2, far outside of the Σ_b blinded region. With the available 1 fb^{-1} of data, the number of true B^{**} passing the Λ_b^0 kinematic cuts is quite small, and thus the B^{**} peaks in the background are not visible. Consequently, we treat the physics backgrounds, including the B^{**} peaks, as smooth. At a later point, the B^{**} peaks may also be incorporated into the physics background by describing the full shape with a Keys PDF [34].

The normalization factor for these backgrounds is taken as the ratio of the number of B^0 events in the Λ_b^0 mass signal region to the number of B^0 events in the reconstructed B^0 sample. The number of B^0 in the Λ_b^0 signal region is taken from the total Λ_b^0 mass fit described in [26] and is about 260 ± 20 events. We do not yet have the full 1 fb^{-1} of B^0 data due to failed sections which are still being recovered. In the ~ 700 pb^{-1} of analyzed B^0 data we found 4570 ± 80 B^0 events, giving a normalization factor value of 0.056. The backgrounds are fit with the form from Eq. (2) for both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$, using a binned likelihood fit. The results are shown in Fig. 22, with the parameter values given in Table 8. Both the shape and normalization of the physics background is fixed in the fit to data.

	α	Q_{max}	γ	Events	Scaled Events
$\Lambda_b^0\pi^-$	0.35 ± 0.17	0.15 ± 0.03	1.8 ± 0.9	1936 ± 62	109
$\Lambda_b^0\pi^+$	0.8 ± 0.8	0.10 ± 0.02	0.7 ± 0.7	1871 ± 61	105

Table 8: Parameter values for the shape and normalization of the physics background in both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$.

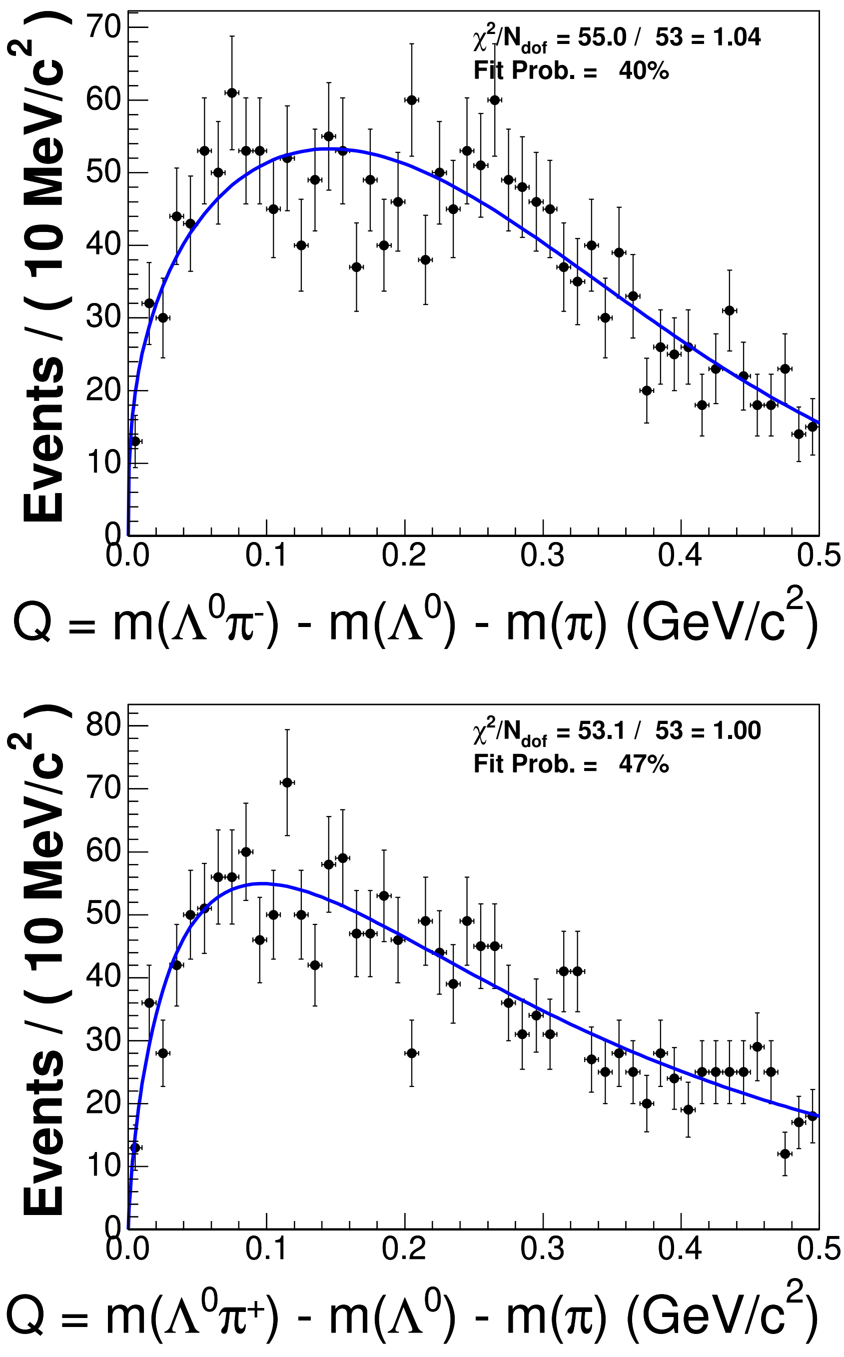


Figure 22: Fits to the physics background from B^0 data. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$.

6.4 Hadronization and Underlying Event Backgrounds

The majority of the Σ_b background is due to tracks from the Λ_b^0 fragmentation (*e.g.* hadronization of the b quark). There are also tracks from the rest of the underlying event. These two backgrounds are studied using the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ PYTHIA Monte Carlo sample described in Sec. 4.3. This sample does not contain physics or combinatorial backgrounds, and we remove all true Σ_b events, leaving behind only the two backgrounds of interest. In order to use this background as a fixed shape, the MC must model data well. The background shape is determined after the reweighting described in Sec. 4.4.

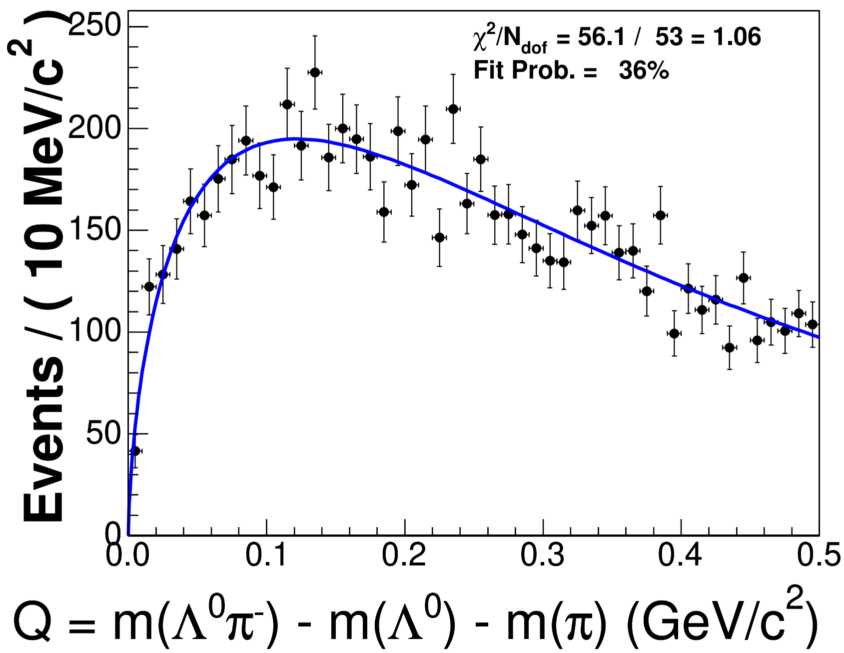
For this background, the normalization factor is the ratio of Λ_b^0 in the data to the number of Λ_b^0 in the MC sample. From the Λ_b^0 mass fit, the Λ_b^0 yield from $\Lambda_c^+\pi^-$ decays is 2927 ± 58 in the Λ_b^0 signal region. The number of Λ_b^0 in the MC sample is 14,064, giving a normalization factor value of 0.208. A binned likelihood fit of Eq. (2) is performed, and the results are shown in Fig. 23 and Table 9. Both the shape and normalization of the Ha + UE background is fixed in the fit to data.

	α	Q_{max}	γ	Events	Scaled Events
$\Lambda_b^0\pi^-$	0.66 ± 0.06	0.122 ± 0.005	0.73 ± 0.01	7560 ± 123	1572
$\Lambda_b^0\pi^+$	0.67 ± 0.25	0.11 ± 0.01	0.86 ± 0.31	7410 ± 122	1541

Table 9: Parameter values for the shape and normalization of the Λ_b^0 hadronization and underlying event background in both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$.

Fig. 24 shows the blinded Q distribution for our Λ_b^0 data, with the three backgrounds fitted for in these sections plotted on top. Also shown is the sum of the three backgrounds, which we expect to follow the Q sidebands. The sum of the three backgrounds agrees well with the shape of the data.

However, the reweighting of the track p_{T} directly affects the shape of the Q background which will be fixed in the fit. This introduces a systematic error, where we are dependent on the track p_{T} reweighting function for the background shape. The means of evaluating this as a systematic are described in the next section.



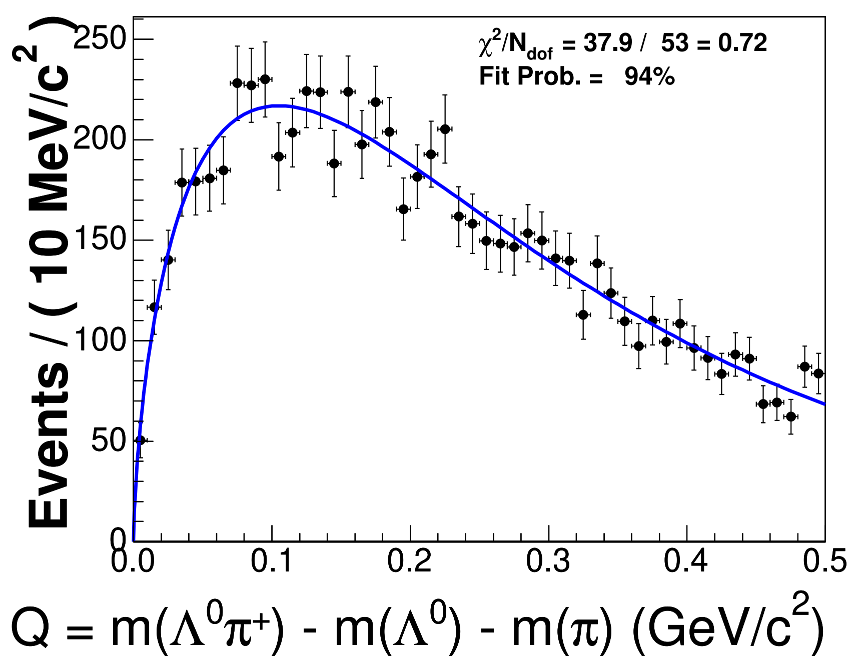


Figure 23: Fits to the Λ_b^0 hadronization and underlying event backgrounds, from PYTHIA. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$.

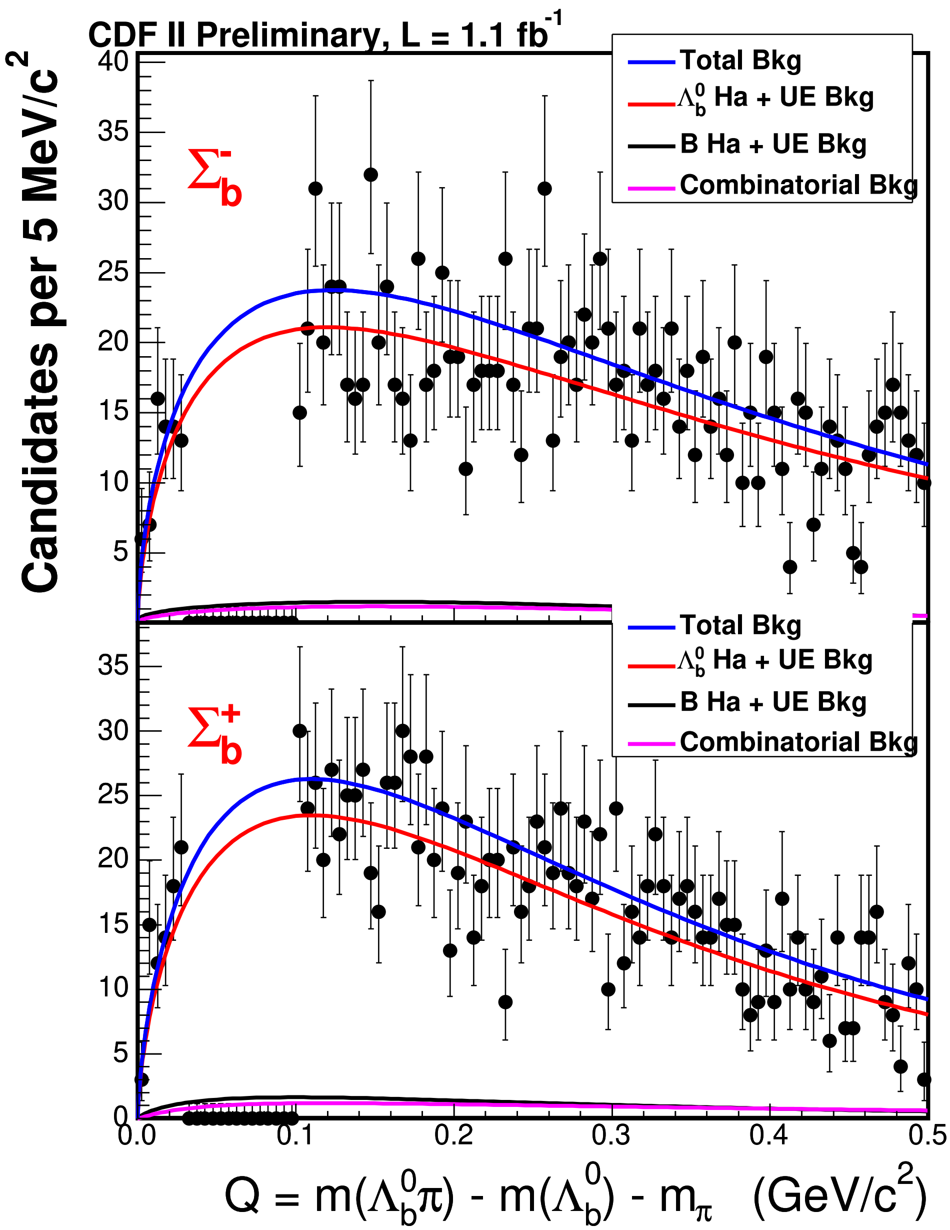


Figure 24: The three different background components described in Sec. 6 and their sum are all shown superimposed on the blinded Q distributions, with $\Lambda_b^0\pi^-$ on top and $\Lambda_b^0\pi^+$ on the bottom.

6.5 Systematic Evaluation of the PYTHIA Λ_b^0 Track Reweighting

As explained in Sec. 4.4, there were too few soft tracks around the Λ_b^0 for the PYTHIA Monte Carlo sample to match the Q distribution from data. Consequently, we took the ratio of the track p_T in data and in MC, fit that with a linear function, and reweighted the MC accordingly. However, there are errors on the track p_T distributions, and the ratio distribution may be biased in some systematic way.

To account for this, we reweighted the track p_T spectrum from data using the following procedure:

1. For each bin of the histogram, get the number of entries and the error.

2. Create a linear function $f(p_T) = 0.5p_T - 1$, which is equal to -1 at p_T = 0 and +1 at p_T = 4 GeV/c.

3. Fill a new histogram with the same entries as the original histogram, plus the value of the bin error multiplied by $f(p_T)$ evaluated at the central value of this p_T bin.

4. Use this as the data track p_T spectrum and perform the usual ratio to get a reweighting function.

This systematically moves the values within their error bars. The function in (2) above has the effect of reducing the number of soft tracks and increasing the number of hard (high p_T) tracks. We refer to systematics using this function as the "Reweighted Down" systematics. We also using a function $f(p_T) = -0.5p_T + 1$, which is equal to +1 at p_T = 0 and -1 at p_T = 4 GeV/c. This has the reverse effect of increasing the number of soft tracks, so we refer to this systematic as "Reweighted Up". The resulting linear fits to the data/MC ratio for these reweighted spectra are shown in Fig. 25 with fit parameters given in Tab. 10.

Parameter	Down 1 σ	Down 2 σ	Up 1 σ	Up 2 σ
p_0	1.59 ± 0.06	1.44 ± 0.06	1.89 ± 0.06	2.02 ± 0.07
p_1	-0.274 ± 0.041	-0.191 ± 0.041	-0.450 ± 0.039	-0.526 ± 0.043
Fit Prob.	69%	42%	92%	90%

Table 10: Parameter values for the linear functions used to reweight the Monte Carlo, after reweighting the track p_T spectrum in data either up or down by 1 and 2 σ.

Using the new linear fits to reweight the PYTHIA MC, we derive new background shapes for the Λ_b^0 Ha + UE background. These shapes are shown in Fig. 27 with fit parameters given in Tab. 11. The shape parameters change slightly from the default fit, although the only systematically shifted background parameter is the Q_{max} value. These alternate background shapes will be used on the fit to data (Sec. 9.3.4), and to evaluate systematic errors due to the PYTHIA Λ_b^0 track p_T reweighting (Sec. 10.2.2).

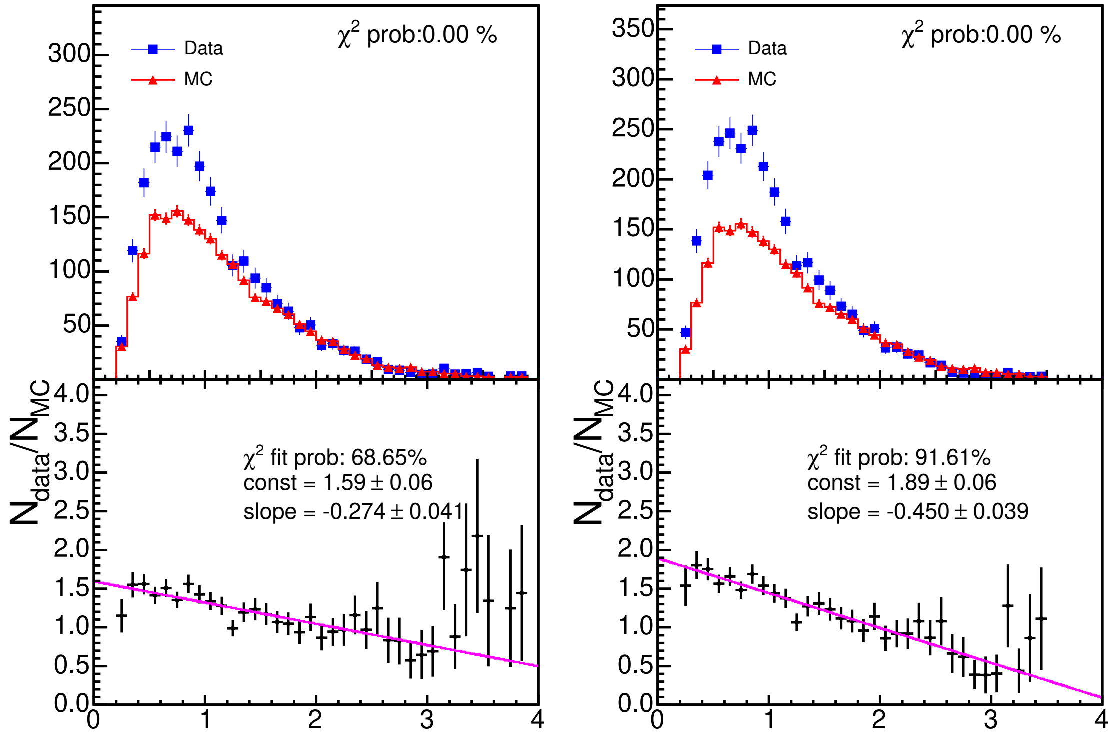


Figure 25: Functional form used to reweight the MC by track p_{T} . The histogram is the ratio of track p_{T} in data to MC and is fit to a linear function. The left histogram shows the data p_{T} spectrum reweighted down by a function $f(p_{\mathrm{T}}) = 0.5p_{\mathrm{T}} - 1$, while the right histogram has been reweighted up by a function $f(p_{\mathrm{T}}) = -0.5p_{\mathrm{T}} + 1$. This is referred to as reweighting up/down by $1\ \sigma$.

$\Lambda_b^0\pi^-$	Down $1\ \sigma$	Down $2\ \sigma$	Up $1\ \sigma$	Up $2\ \sigma$
α	0.66 ± 0.01	0.65 ± 0.005	0.67 ± 0.38	0.67 ± 0.003
Q_{max}	0.126 ± 0.007	0.133 ± 0.005	0.116 ± 0.018	0.112 ± 0.004
γ	0.72 ± 0.07	0.71 ± 0.005	0.73 ± 0.45	0.74 ± 0.003
Number of events	7270 ± 121	6990 ± 118	7840 ± 125	8120 ± 127
$\Lambda_b^0\pi^+$	Down $1\ \sigma$	Down $2\ \sigma$	Up $1\ \sigma$	Up $2\ \sigma$
α	0.68 ± 0.27	0.65 ± 0.24	0.67 ± 0.24	0.66 ± 0.22
Q_{max}	0.107 ± 0.011	0.111 ± 0.011	0.102 ± 0.010	0.100 ± 0.009
γ	0.85 ± 0.32	0.86 ± 0.31	0.87 ± 0.29	0.88 ± 0.27
Number of events	7100 ± 119	6780 ± 116	7720 ± 124	8030 ± 127

Table 11: Parameter values for the shape and normalization of the Λ_b^0 hadronization and underlying event background in both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$, after performing the reweighting described in the text (Sec. 6.5).

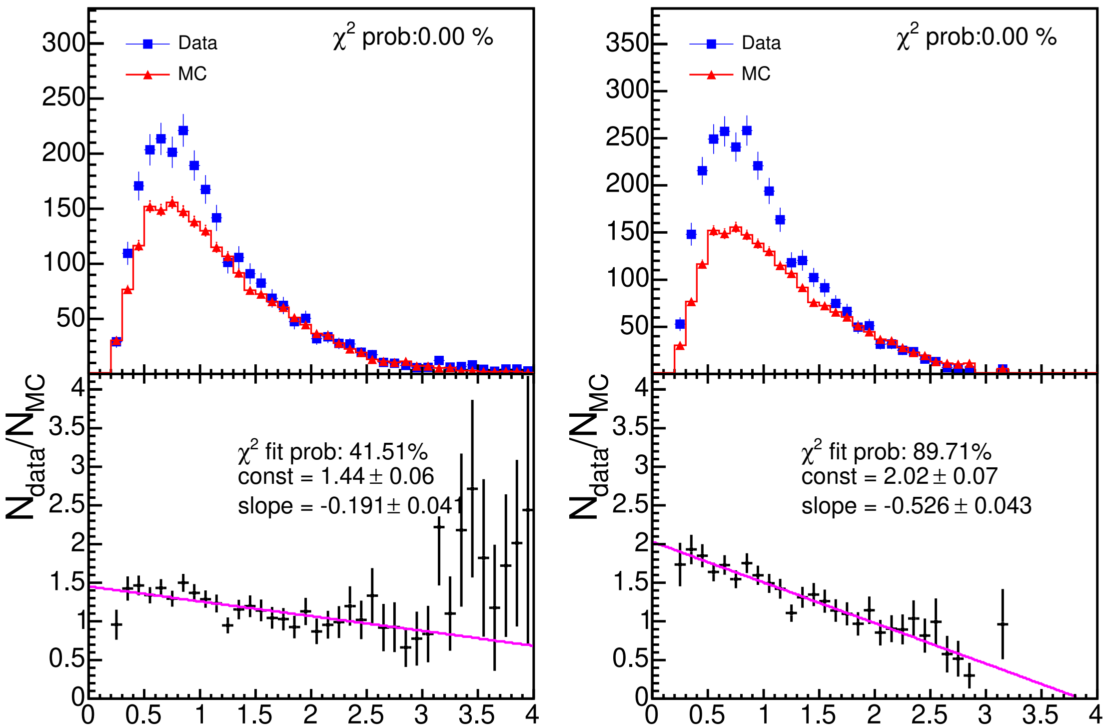


Figure 26: Functional form used to reweight the MC by track p_T . The histogram is the ratio of track p_T in data to MC and is fit to a linear function. The left histogram shows the data p_T spectrum reweighted down by a function $f(p_T) = 0.5p_T - 2$, while the right histogram has been reweighted up by a function $f(p_T) = -0.5p_T + 2$. This is referred to as reweighting up/down by 2 σ.

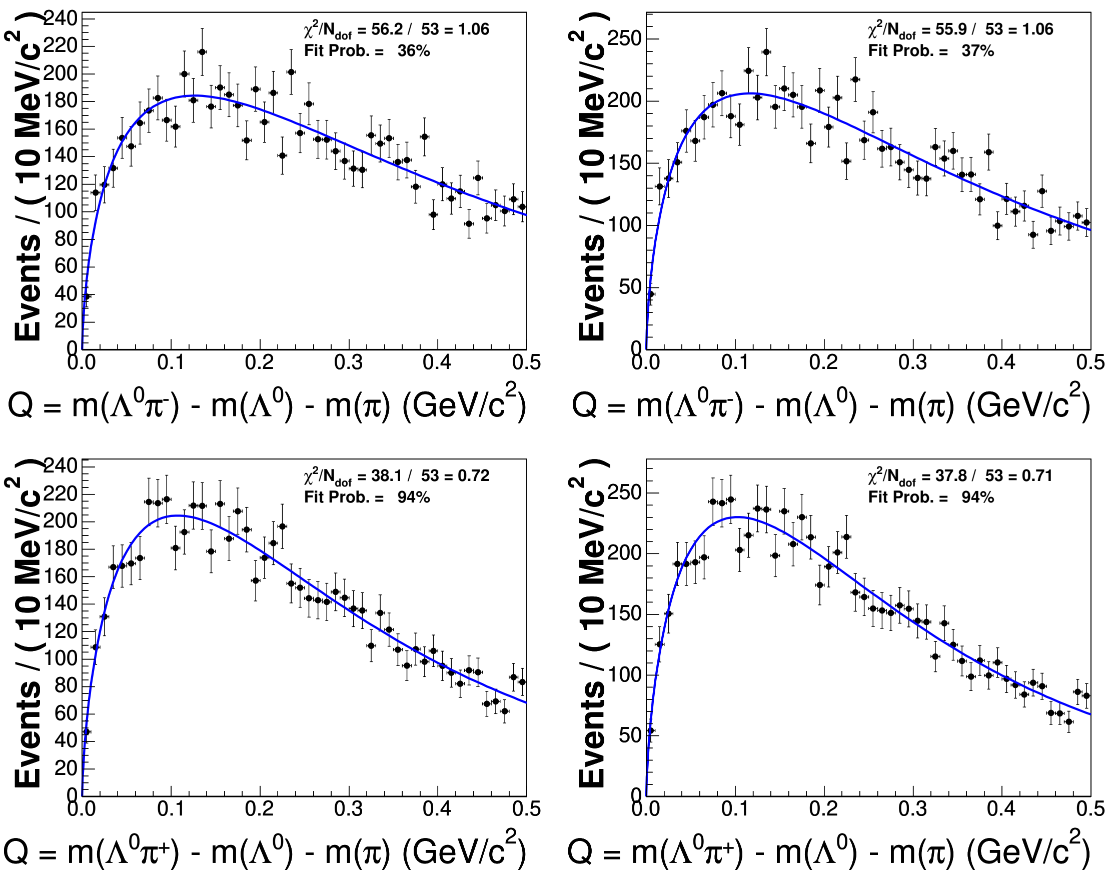


Figure 27: Fits to the Λ_b^0 hadronization and underlying event backgrounds, from PYTHIA. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$. The figures on the left are after performing the reweighting down by 1 σ, and on the right after reweighting up by 1 σ (as described in Sec. 6.5).

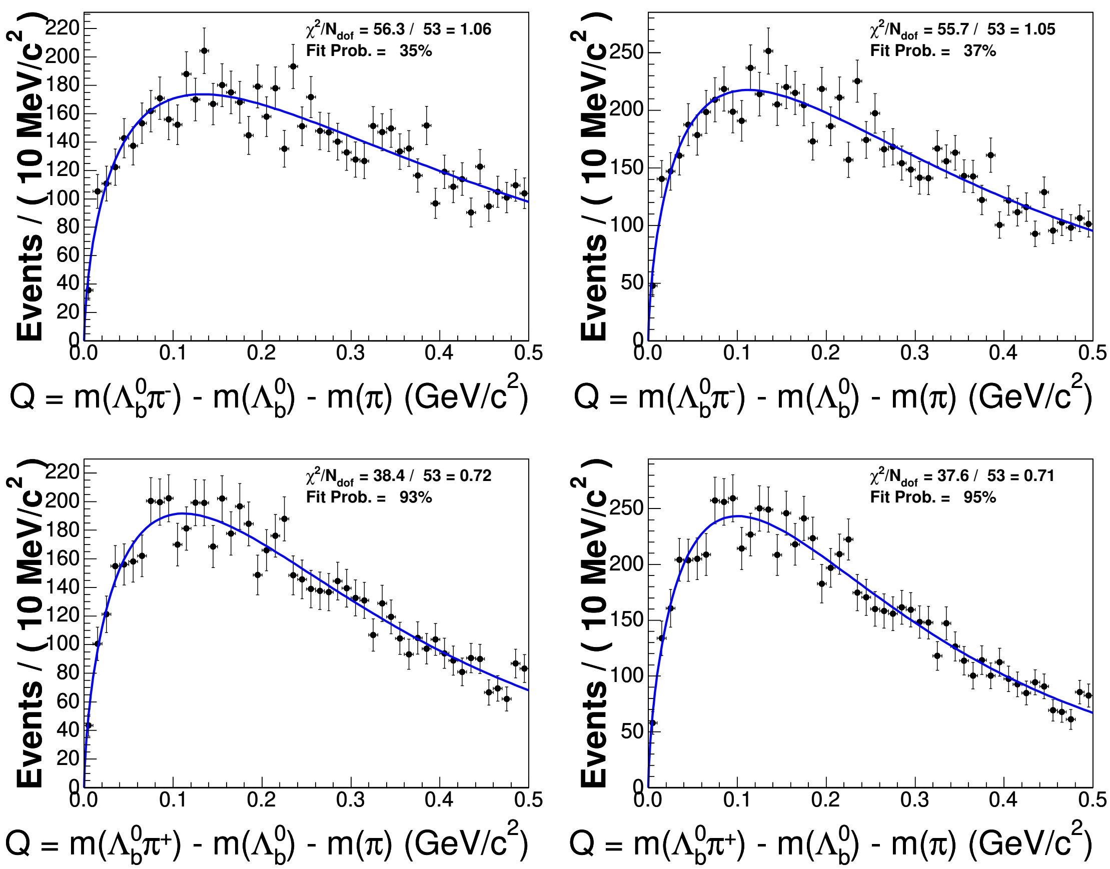


Figure 28: Fits to the Λ_b^0 hadronization and underlying event backgrounds, from PYTHIA. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$. The figures on the left are after performing the reweighting down by 2 σ, and on the right after reweighting up by 2 σ (as described in Sec. 6.5).

7 Description of the Σ_b Fit

As described in Sec. 6, the various Σ_b background components are fixed in the Σ_b Q fit. Only the Σ_b signal shapes are floating in the fit. As described in the following sections, we use a simultaneous unbinned negative log likelihood fit for the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$.

7.1 Q Detector Resolution

Each Σ_b state is described by a non-relativistic Breit-Wigner convoluted with a double Gaussian detector resolution. The Gaussian detector resolution is taken from the PYTHIA $\Sigma_b \rightarrow \Lambda_b^0\pi$ signal sample described in Sec. 4.3. Since the sample is generated with an input width of zero for all Σ_b states, the width of the reconstructed Q peaks is due to detector resolution. This resolution, with double Gaussian fit superposed, is shown for Σ_b^- and Σ_b^{*-} in Fig. 29 with the resolution values given in Tab. 12. The fit probability is low ($\sim 0.04\%$) due to very poorly measured tracks; these tracks form long non-Gaussian tails which cannot be described with a double Gaussian model. For the detector resolution in the fit we use the average values of σ_{narrow} = 1.17 MeV/c^2 and σ_{wide} = 3.0 MeV/c^2.

State	σ_{narrow} (MeV/c^2)	σ_{wide} (MeV/c^2)	Weight of narrow
Σ_b^-	1.14 ± 0.04	3.0 ± 0.2	0.92
Σ_b^{-*}	1.19 ± 0.03	3.0 ± 0.2	0.88

Table 12: Detector resolution values in Q for Σ_b^- and Σ_b^{*-}, taken from a PYTHIA signal only Monte Carlo sample.

7.2 Σ_b Fit Description

The expected Σ_b signal structure for $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ is described in Sec. 2. The signal structure is simple: two peaks in SC for $\Sigma_b^{(*)-}$ and two in OC for $\Sigma_b^{(*)+}$. The masses of $\Sigma_b^{(*)-}$ and $\Sigma_b^{(*)+}$ will be very similar, probably separated by only a few MeV as shown in Tab. 2. The Σ_b width will be fixed in the fit according to Eq. (1).

We expect the mass difference Σ_b^{*-} - Σ_b^- to be very similar to Σ_b^{*+} - Σ_b^+. Our original plan was to fit the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$events separately, each with an unbinned likelihood fit. However, due to the low statistics of our sample we fix the Σ_b - Σ_b^* mass difference to be the same for $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$. We do not expect these mass differences to be the same, as detailed in Sec. 2; the expected difference of 0.40 ± 0.07 MeV/c^2 will be used in Sec. 10.2.5 to quote a systematic error. So instead of fitting Σ_b^- and Σ_b^+ separately, we perform a simultaneous unbinned likelihood fit, fitting both at once. The seven floating variables in the simultaneous fit are:

- Σ_b^- Q value
- Σ_b^+ Q value
- Σ_b^* - Σ_b Q value

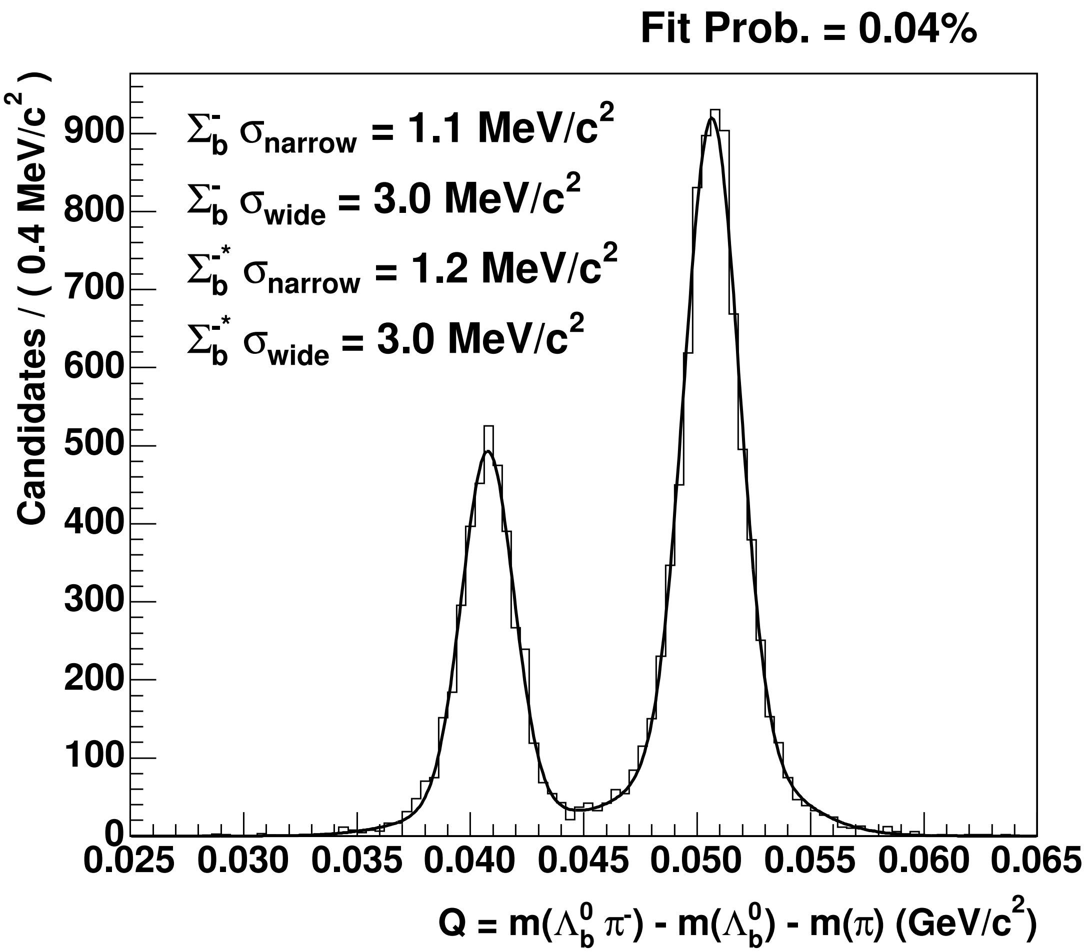


Figure 29: Detector resolutions in Q for Σ_b^- and Σ_b^{*-} in a Σ_b signal PYTHIA Monte Carlo. Each state is fit with two Gaussians, one narrow and one wide.

- Number of events for each of the 4 Σ_b states

There are no constraints on the number of events in each state. The width for each state is constrained separately from Eq. (1).

7.3 Tests of the Σ_b Fit

Before unblinding the Σ_b signal region and attempting to fit for Σ_b in data, we tested the fitter on Toy Monte Carlo samples of a "fake" Σ_b signal. These tests, which are shown in Appendix C, are not applicable to the current fit for two reasons – they do not use the simultaneous fit, and the Σ_b width was erroneously taken to be on 0.8 MeV/c^2.

After unblinding and moving to a simultaneous fit, we performed additional tests of the fit. Using a preliminary fit to the Σ_b data, we chose the input parameters shown in Tab. 13. The Σ_b states for this set of input parameters are shown in Fig. 30. To measure the pulls on the seven

Signal parameters	Values
Σ_b^- Q (MeV/c^2)	56
Σ_b^- events	65
Σ_b^+ Q (MeV/c^2)	48
Σ_b^+ events	32
Σ_b^{*-} events	83
Σ_b^{*+} events	82
Σ_b^* - Σ_b Q (MeV/c^2)	22

Table 13: List of the input Σ_b signal parameters used to test the fitter.

floating parameters, we combine the measured signal parameters with the background shapes described in Sec. 6 and generate 2000 Toy MC experiments of $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$. We then fit each sample with the simultaneous fit. An example of one such Toy MC sample is shown in Fig. 31. The errors on the parameters are asymmetric, which can be seen in Fig. 32. When using the only the parabolic error, there are large tails on the pulls. The pulls in Fig. 32 are fit with a Gaussian only on the range [-2, 2], and the Gaussian parameters are shown in Tab. 14.

We then used the positive and negative errors from MINOS to calculate the pulls. For some of the fits, the errors on a parameter exceeded the limits of the parameter, and MINOS was unable to return an error. This typically happened when the number of events (usually the number of Σ_b^+ events, which is smaller than the other Σ_b peaks) has an error below zero. These fits were removed from the pull calculations. The resulting pulls are shown in Fig. 33 and the Gaussian parameters are given in Tab. 15. The pulls were fit both on the whole range [-5, 5] and on a limited range of [-2.5, 2.5]. The Gaussian parameters for both sets of fits are given in Tab. 15, but only the plots for the whole range are shown in Fig. 33.

There are two pulls in particular which deviate from the unit Gaussian: the Σ_b^+ Q and Σ_b^+ number of events. Qualitatively, we expect this because the Σ_b^+ is the smallest of the four Σ_b peaks and thus more sensitive to fluctuations.

We estimate the fitter bias on each parameter by fitting the raw difference between the input parameters and the final fit parameters for each Toy MC sample. These differences are shown in Fig. 34 and the fit results are given in Tab. 16. For all parameters, including Σ_b^+ Q and Σ_b^+ number of events, the fitter bias is much smaller than the expected statistical and systematic errors on that parameter. Thus we ignore the fitter bias, assuming it is adequately accounted for by systematic and statistical errors.

Signal parameters	Gaussian Mean	Gaussian σ
Σ_b^- Q	0.003 ± 0.04	1.04 ± 0.04
Σ_b^- events	-0.04 ± 0.04	1.06 ± 0.04
Σ_b^+ Q	0.04 ± 0.04	1.01 ± 0.04
Σ_b^+ events	0.05 ± 0.04	0.99 ± 0.04
Σ_b^{*-} events	-0.21 ± 0.04	1.06 ± 0.04
Σ_b^{*+} events	-0.09 ± 0.04	1.03 ± 0.04
Σ_b^* - Σ_b Q	-0.01 ± 0.05	1.13 ± 0.05

Table 14: Pulls on the floating parameters using only the parabolic errors from MIGRAD. The Gaussian is fit only on the range [-2, 2].

Signal parameters	Gaussian Mean	Gaussian σ
Fit on the whole range [-5, 5]		
Σ_b^- Q	0.07 ± 0.03	1.07 ± 0.02
Σ_b^- events	-0.02 ± 0.02	1.06 ± 0.02
Σ_b^+ Q	-0.02 ± 0.03	1.16 ± 0.02
Σ_b^+ events	-0.23 ± 0.02	0.94 ± 0.02
Σ_b^{*-} events	0.05 ± 0.03	1.08 ± 0.02
Σ_b^{*+} events	-0.03 ± 0.03	1.07 ± 0.03
Σ_b^* - Σ_b Q	-0.08 ± 0.02	1.06 ± 0.02
Fit on the limited range [-2.5, 2.5]		
Σ_b^- Q	0.05 ± 0.03	1.05 ± 0.02
Σ_b^- events	-0.03 ± 0.03	1.03 ± 0.02
Σ_b^+ Q	-0.003 ± 0.03	1.06 ± 0.02
Σ_b^+ events	-0.22 ± 0.02	0.95 ± 0.02
Σ_b^{*-} events	0.03 ± 0.03	1.09 ± 0.02
Σ_b^{*+} events	-0.04 ± 0.03	1.04 ± 0.02
Σ_b^* - Σ_b Q	-0.09 ± 0.03	1.03 ± 0.02

Table 15: Pulls on the floating parameters using the positive and negative errors from MINOS. The Gaussian is fit both on the whole range [-5, 5] and a limited range [-2.5, 2.5].

Signal parameters	Gaussian Mean	Gaussian σ
Σ_b^- Q (MeV/c^2)	0.05 ± 0.04	1.5 ± 0.03
Σ_b^- events	-0.6 ± 0.4	14.9 ± 0.3
Σ_b^+ Q (MeV/c^2)	-0.09 ± 0.05	2.0 ± 0.03
Σ_b^+ events	-3.0 ± 0.3	11.6 ± 0.2
Σ_b^{*-} events	0.4 ± 0.5	19.5 ± 0.3
Σ_b^{*+} events	-0.6 ± 0.4	18.0 ± 0.3
Σ_b^* - Σ_b Q (MeV/c^2)	-0.17 ± 0.05	1.9 ± 0.03

Table 16: Gaussian fit to the raw differences of the Toy MC samples. The mean value is an indication of the fitter bias on each parameter.

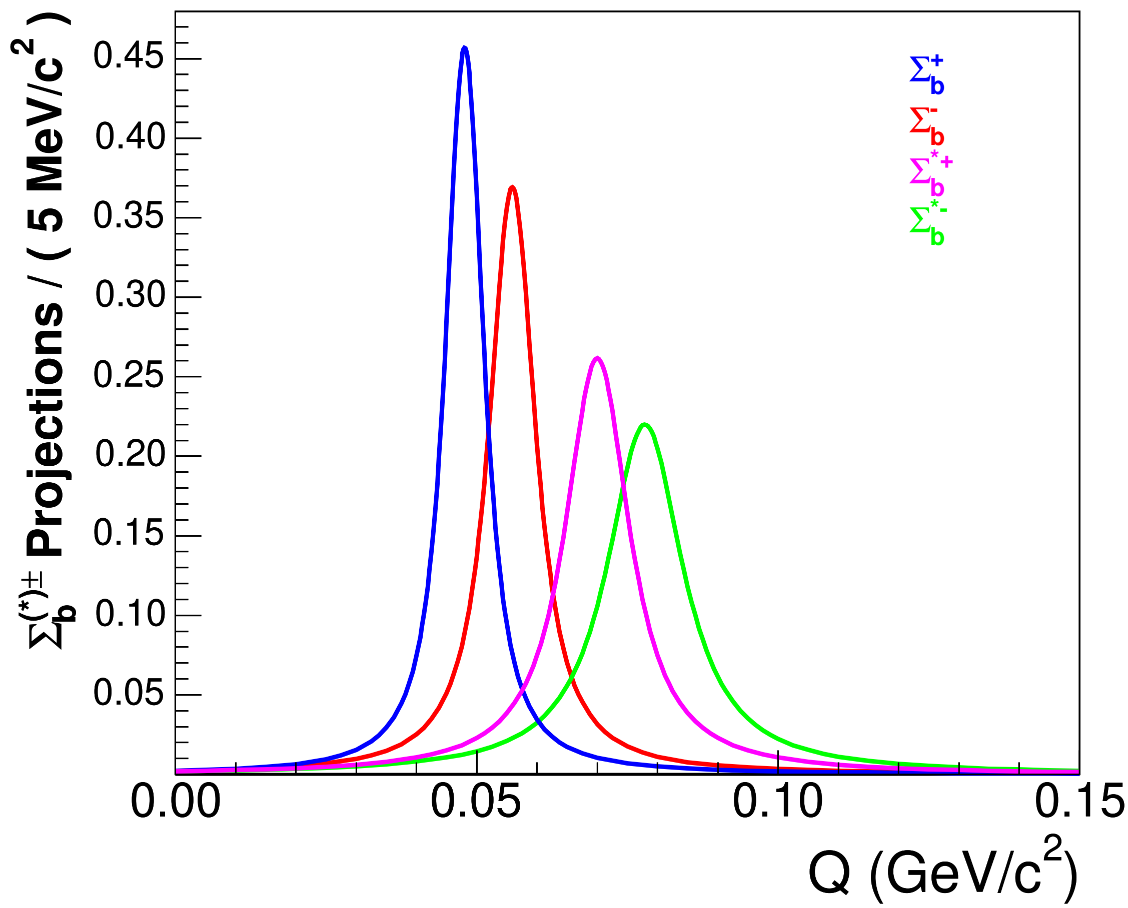
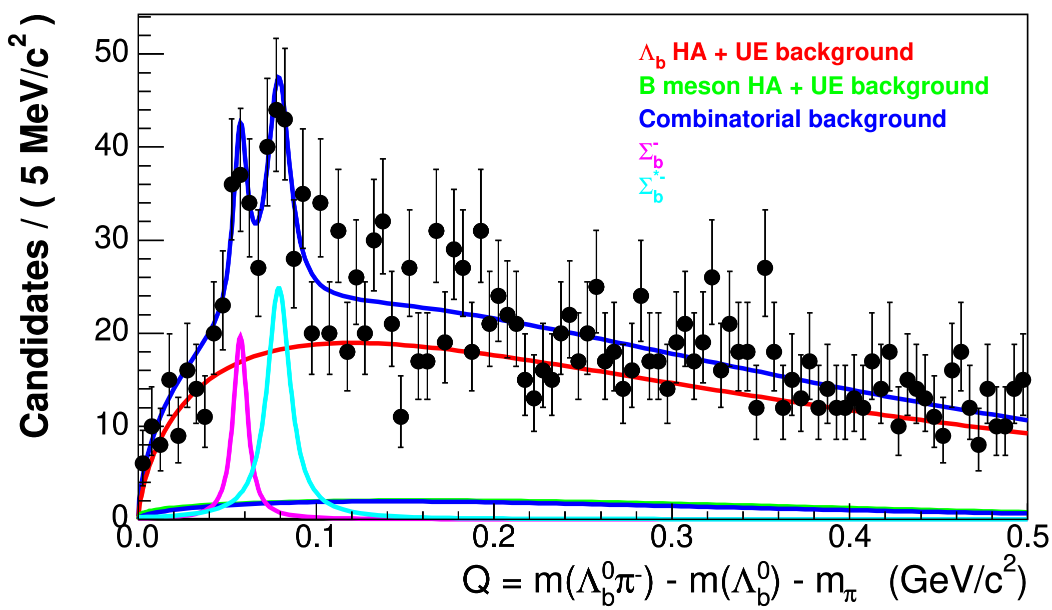


Figure 30: RooFit plot of the Σ_b signal structure for the input parameters given in Tab. 13. This plot shows $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ combined. Each peak is normalized to unity. The width of each peak is determined by its mean Q value, according to Eq (1).



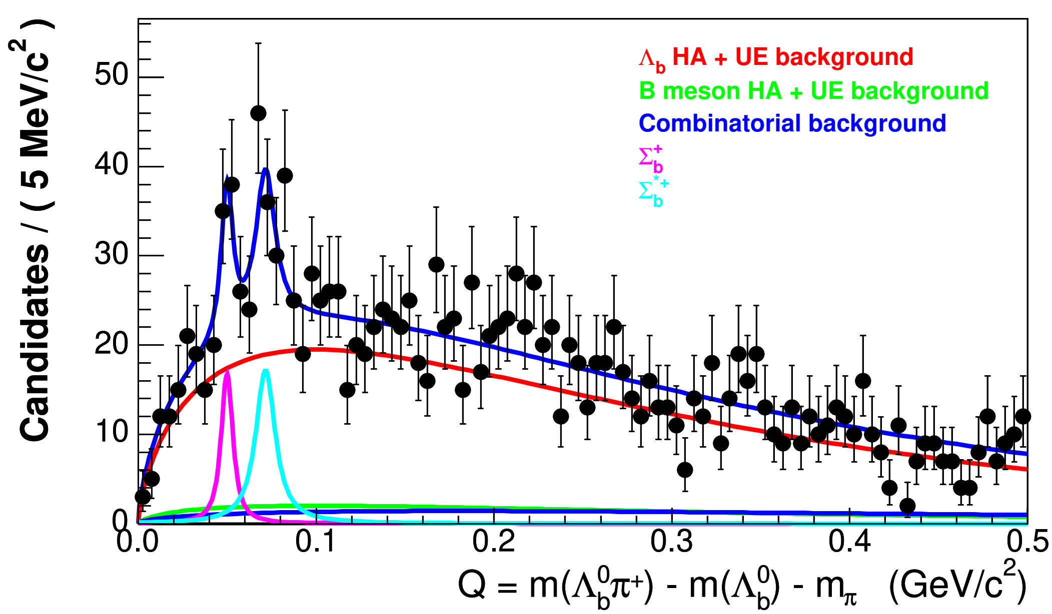


Figure 31: Plots of a Toy MC sample generated with the input signal parameters given in Tab. 13. The simultaneous unbinned fit to the Toy MC is also shown.

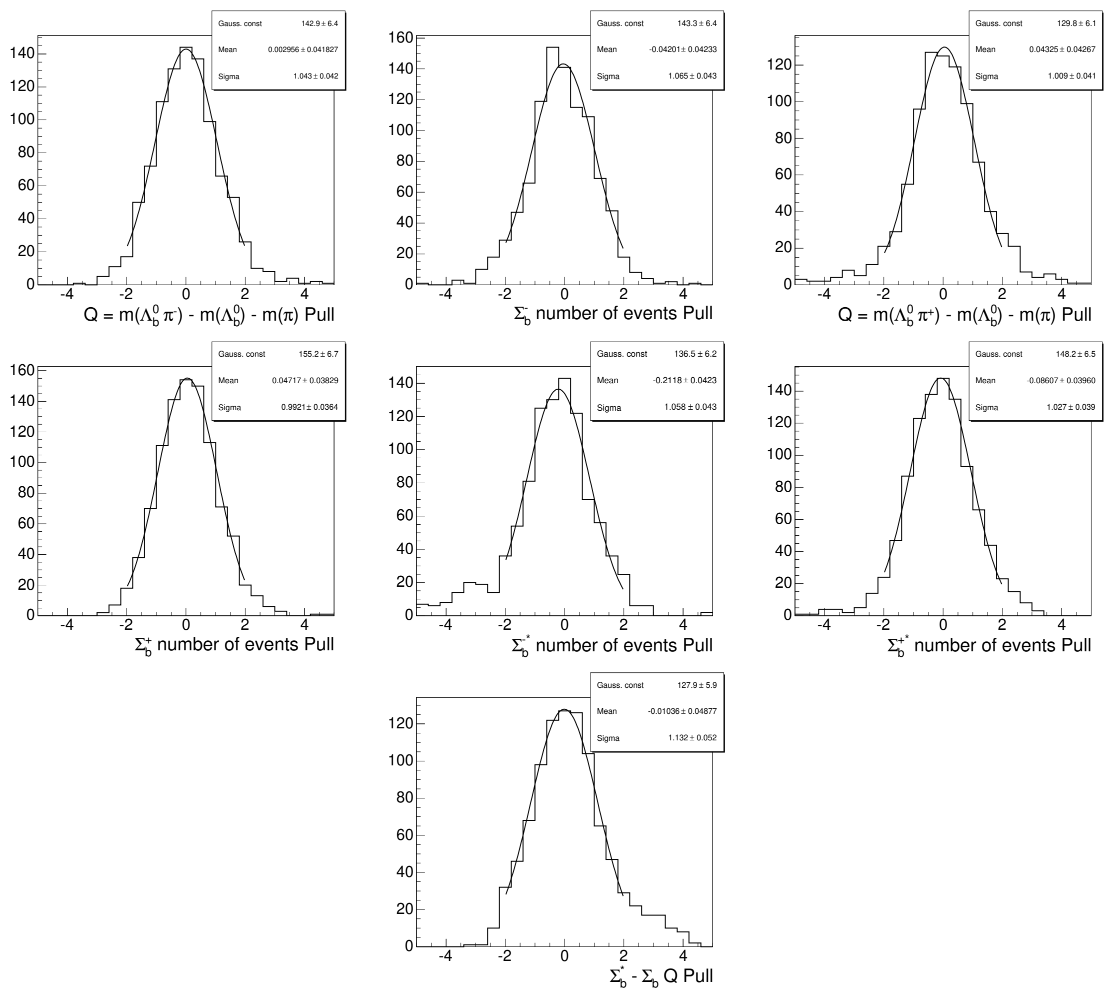


Figure 32: Pulls on all floating parameters in the fit, using only the parabolic errors returned from MIGRAD. The pull distributions are only fit on the range [-2, 2], and the results of the fits are listed in Tab. 14.

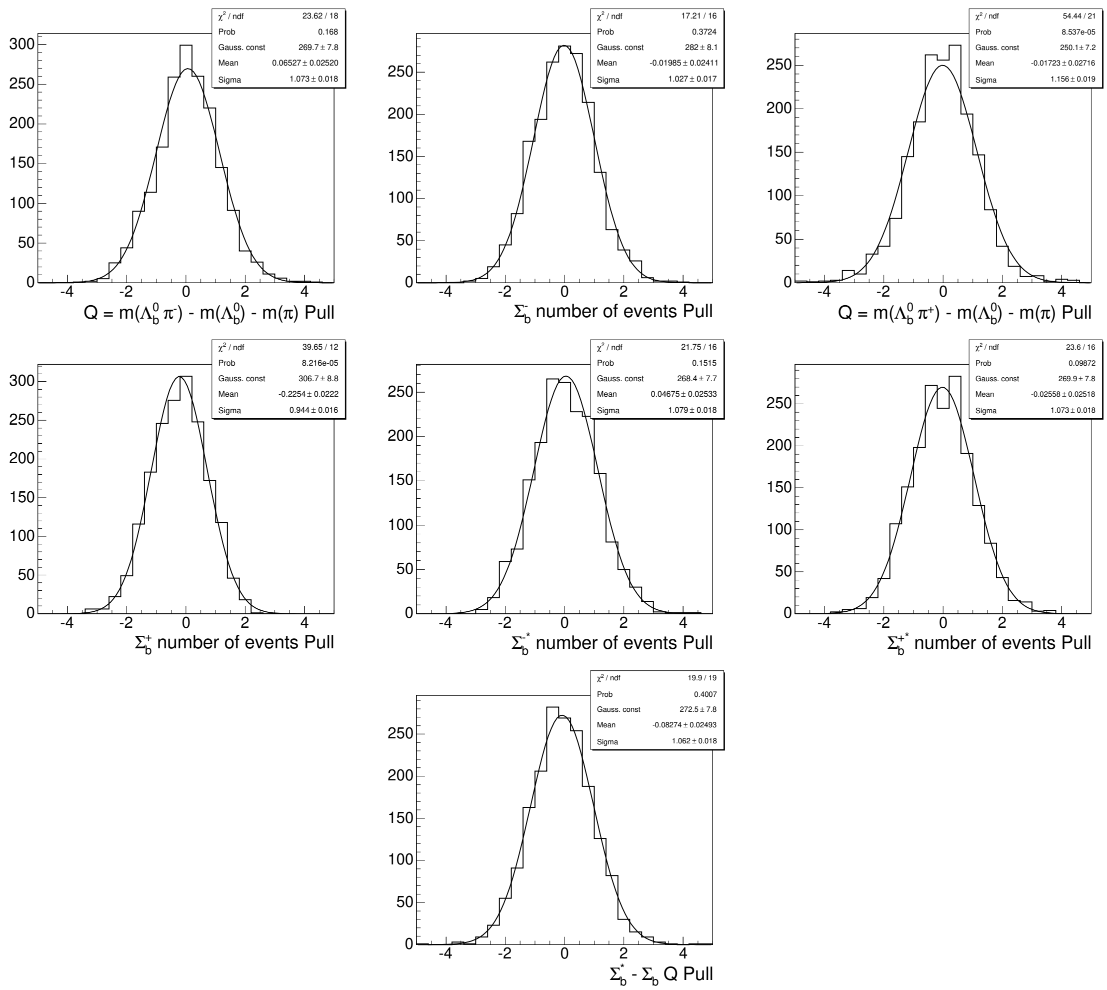


Figure 33: Pulls on all floating parameters in the fit using positive and negative errors returned from MINOS. The pull distributions are fit on the entire range [-5, 5], and the results of the fits are listed in Tab. 15.

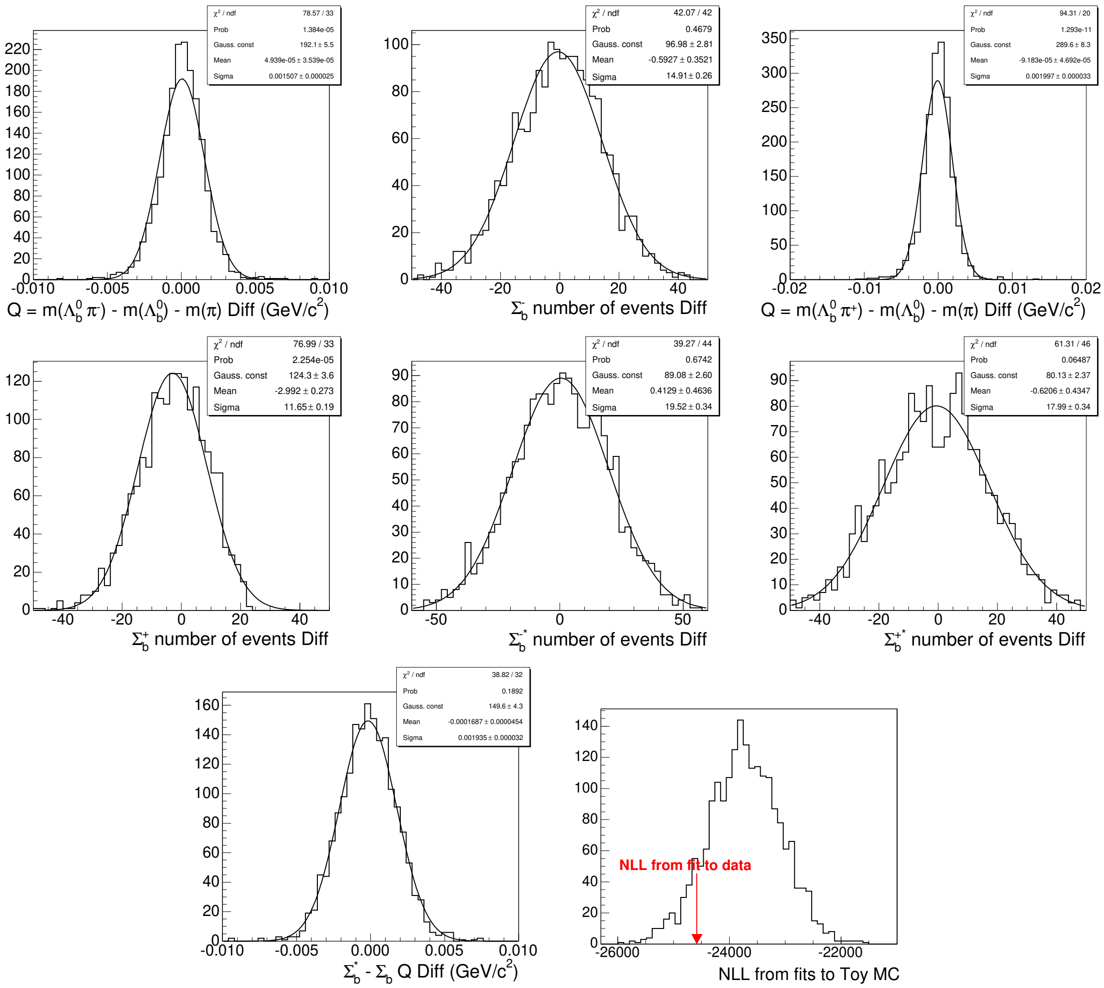


Figure 34: Raw differences for each floating parameter in the Toy MC samples. Each difference is fit with a Gaussian to measure the fitter bias on that parameter, and the results of the fits are listed in Tab. 16. Also shown is the distribution of NLLs from the ToyMC fits, with the NLL from the fit to data (as described in Sec. 9) superimposed.

8 Significance of the observed signal

There are several ways to evaluate the significance of the signal that might be observed.

8.1 Option 1: Bin Counting

We open the box, and define the "peaks". (We could even run the fit for a better estimation of the peak location.) Then we define 3σ around each peak – since the Σ_b signal peaks are fitted with the Voigtian PDF (*i.e.*, Breit-Wigner convoluted with a Gaussian resolution), we could use $\sqrt{\sigma_{\text{res}}^2 + 0.25 \cdot \Gamma^2}$ as the "σ". There are two Gaussians in the resolution function, and thus we use the weighted average of the σ_{res1} and σ_{res2} before computing the average "σ".

Then we collect all the bins within 3σ of each peak. And in this we treat SC and OC separately, since the peaks may be interleaved and thus we don't want to count the background of SC which is under a peak in SC, but not under a peak in OC. We integrate the background function in all the bins under 3σ of all peaks, and thus obtain "B".

For the signal ("S"), we look into the same set of bins, except that we subtract the integral of the background function in each bin from the bin content in that bin. The sum of all subtracted bin heights for all bins encompassed within 3σ of all peaks is S.

The significance of the observed "signal" is $S/\sqrt{S+B}$.

8.2 Option 2: p-value

We use the total number of fitted signal events (the sum of yields in all four peaks) as the figure of merit.

We run a series of Toy MC experiments based on the background components only, and feed it into the fitter which also includes the standard signal (namely all four Σ_b peaks). The fitter wakes up with the peaks at the locations averaged from the theoretical expectations (from Sec. 2), but is "greedy" in the sense that it opportunistically moves around until it finds a "bump". Thus the separate peaks can move around independently of each other. (This is not realistic for a proper Σ_b signal, but it is conservative since it overestimates the found signal.)

We repeat this many times on Toy MC data based on the background PDFs, and obtain a probability density $P_{\text{bkg}}(N_{\text{tot}})$ to observe at least N_{tot} events in all four peaks total.

We then fit the signal on data, obtain $N_{\text{tot}}^{\text{data}}$, and integrate $P_{\text{bkg}}(N_{\text{tot}})$ from $N_{\text{tot}}^{\text{data}}$ to infinity. That integral divided by the total number of Toy MC samples generated is the p-value.

8.3 Option 3: Likelihood ratio p-values

This option is less intuitive than bin-counting or using the sum of all signal events to define a p-value, but the likelihood contains all information about the fit. There is no need to define a signal region or incorporate the statistical error on the sum of all signal events.

The Likelihood Ratio (*LR*) is the ratio of the fit likelihood values for two different hypothesis. The value of the *LR* gives the likelihood of the first hypothesis over the second hypothesis. In the case of the null hypothesis (no signal, only background), the *LR* tells us how much better the signal hypothesis is at explaining the data than the null hypothesis.

The fit to data is done as a likelihood fit, with RooFit returning the negative log likelihood, -ln($\mathcal{L}$) or NLL, for each fit. Thus the Likelihood Ratio is defined as

$$LR = \frac{e^{-NLL_1}}{e^{-NLL_2}} = e^{NLL_2 - NLL_1} \tag{3}$$

where NLL_1 is the negative log likelihood returned by a fit of the alternate hypothesis (*eg* the null hypothesis) and NLL_2 is the negative log likelihood returned by the default Σ_b fit.

To compute the p-value based on the LR, we again generate a series of Toy MC experiments based on the background components only. The experiments are then fit both with the default fitter which includes the standard signal (namely all four Σ_b peaks), and the alternate background only fitter. We now obtain a probability density $P_{\mathrm{bkg}}(NLL_2 - NLL_1)$ to observe at least a likelihood difference of $NLL_2 - NLL_1$ for the four-peak fit on a background only sample.

We then use the measured negative log likelihood difference from the fit to data and integrate $P_{\mathrm{bkg}}(NLL_2 - NLL_1)$ from $(NLL_2 - NLL_1)^{\mathrm{data}}$ to infinity. That integral divided by the total number of Toy MC samples generated is the p-value. We can perform a similar procedure to find p-values for the two peak hypothesis (one Σ_b peak per charge combination) and for each individual Σ_b peak.

9 Unblinded Fit to Σ_b

After determining the methods to evaluate significance and fit the observed signal, we unblinded the Σ_b signal region.

9.1 Evaluating Signal Significance

We evaluate the signal significance via the methods outlined in Sec. 8.

9.1.1 Bin Counting

For the first option, bin counting, we subtract the background function from the signal histogram and count the number of events left in the signal region. Practically, this is done by filling a histogram for the background functions, where the value in each bin of the histogram is equal to the value of the function at the midpoint of the bin. It is this histogram which is then subtracted from the signal histogram. The excess of events in $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ are shown in Fig. 35. The number of events before and after subtraction in the signal region, $Q \in [0.03, 0.1]$ GeV/c^2, are shown in Table 17. This method yields a significance of 5.9σ for $\Lambda_b^0\pi^-$ alone and 4.5σ for $\Lambda_b^0\pi^+$ alone. The combined significance is 7.3σ.

Sample	Data events	Bkg events	Data excess over bkg
$\Lambda_b^0\pi^-$	406	288	118
$\Lambda_b^0\pi^+$	404	313	91

Table 17: Summary of the number of events in the Q signal region ($Q \in [0.03, 0.1]$ GeV/c^2) for the data and the predicted background.

9.1.2 *p*-value

For the p-value calculation using the total number of Σ_b signal events, we generated 335,688 Toy MC samples from only the background functions, with the number of Σ_b events set to zero, and then fit them with the default simultaneous fit. We then plotted the distribution of the total number of Σ_b events (the sum of events for all four states $\Sigma_b^{(*)\pm}$). This distribution is the left-hand plot in Fig. 36. The total number of Σ_b events in data is 262, and there were no Toy MC fits which found that number of events. To extract a p-value, we fit the tail of the distribution (from 120 events and above) with a decaying exponential of the form

$$p_0 \cdot e^{-(x-x_0)/p_1} \tag{4}$$

where $x_0 = 120$ is the starting point of the fit. This fit is shown on the right-hand plot of Fig. 36, with the parameters given in Tab. 18. To extract the p-value, we integrate the exponential function from 262 events up to infinity, and then divide this area by the total number of Toy MC samples. The integral is approximately 0.0284, which gives a p-value of $\sim 8.5 \times 10^{-8}$. This corresponds to a Gaussian significance of $\sim$5.24 σ.

In this case, the p-value is not a very good estimate of significance. It is impossible to generate enough Toy MC samples to reach the 9σ level which is indicated from the bin-counting measurement in the previous section. The large extrapolation of total events only confirms that the signal is of greater than 5σ significance. In this case the likelihood ratio is a more sensitive estimate of significance.

Parameter	Exponential fit
p_0	$(2.3 \pm 0.7) \times 10^{-6}$
p_1	-0.079 ± 0.002
Fit Prob	$\sim 41\%$
Integral above 262	0.0284
p-value	8.5×10^{-8}

Table 18: Fit parameters for a decaying exponential fit to the tail of the events distribution shown in Fig. 36.

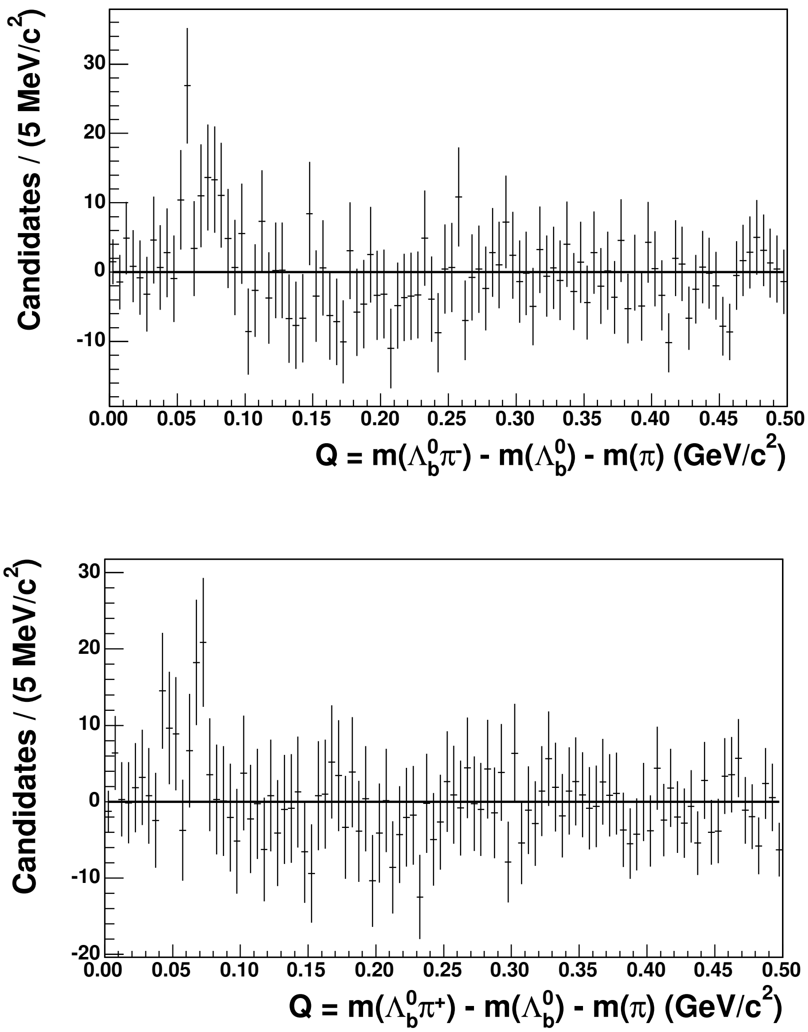


Figure 35: Distribution of the excess of data over the predicted backgrounds for $\Lambda_b^0\pi^-$ (top) and $\Lambda_b^0\pi^+$ (bottom). The backgrounds have been subtracted from the data.

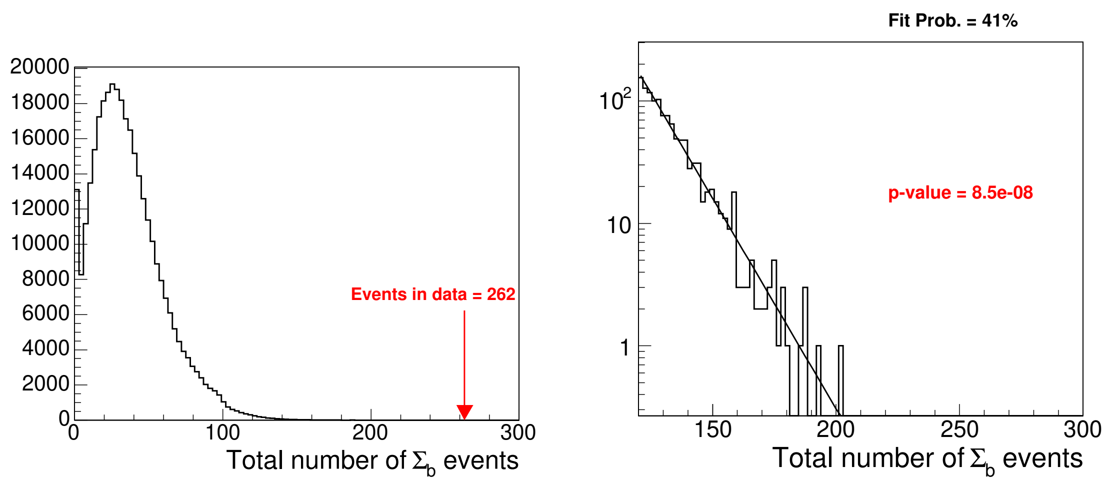


Figure 36: Total number of Σ_b events found from a fit to Toy MC samples generated with no Σ_b signal. The histogram on the left shows the distribution, while the histogram on the right shows only the tail of the distribution fit with a decaying exponential. This corresponds to a p-value of $\sim 8.5 \times 10^{-8}$.

9.1.3 Likelihood ratio

To estimate the strength of the 4 Σ_b peak hypothesis using the likelihood ratio, we study three alternate hypotheses:

(1) No signal (null hypothesis)

(2) One peak per $\Lambda_b^0\pi$ charge combination instead of two

(3) Three Σ_b peaks instead of four

The third hypothesis is used to test the strength of each Σ_b peak individually, apart from the other peaks.

One issue with this method is the systematic variations of the background and signal PDFs. These variations and the resulting uncertainty from each are described in Sec. 10, but how to account for these in each of our alternate fits? After discussion with the CDF statistics committee we have decided to profile the likelihood over the nuisance parameters, *i.e.* systematic variations. To do this, we had to parameterize each systematic variation as a constraint in the fit, which is added to the likelihood. For some of our systematic variations, this was straightforward. These variations and their term in the likelihood are:

- Σ_b intrinsic width, parameter $g_A = 0.75 \pm 0.05$. Likelihood term $\frac{1}{2}\left(\frac{g_A - 0.75}{0.05}\right)^2$
- $\Sigma_b^* - \Sigma_b$ isospin splitting, expected to be $0.40 \pm 0.07\ \mathrm{MeV/c^2}$. Likelihood term $\frac{1}{2}\left(\frac{\Delta_* - 0.0004}{0.00007}\right)^2$
- Λ_b^0 sample composition: propagate errors from the Λ_b^0 mass fit to determine errors on the normalization of each Σ_b background component. This results in 6 terms added to the likelihood, of the same Gaussian constraint form as the previous two.
  - Λ_b^0 hadronization: $\Lambda_b^0\pi^- = 1572 \pm 318$, $\Lambda_b^0\pi^+ = 1541 \pm 311$
  - B hadronization: $\Lambda_b^0\pi^- = 109 \pm 62$, $\Lambda_b^0\pi^+ = 105 \pm 60$
  - Combinatorial bkg: $\Lambda_b^0\pi^- = 87 \pm 45$, $\Lambda_b^0\pi^+ = 85 \pm 44$

The remaining nuisance parameters are not as straightforward. The detector resolution is modeled with a double Gaussian, and as a systematic we consider increasing the width of these Gaussians by 20%. To describe these nuisance parameters, we need an asymmetric function so that the parameters are not allowed to go below their default values. But we cannot use a discontinuous function in the likelihood. Instead we look for a function that increases rapidly below the default value, but increases like a Gaussian constraint above the default value. To do this, we fill a histogram with randomly generated events distributed according to the discontinuous function, with a large "wall" below the default value and a Gaussian-type distribution above the default value. We fit this histogram with several functions, and the one which worked best was of the form

$$\frac{1}{2}\left(\frac{x-\sigma}{a(1-b(x-\sigma))}\right)^2 \tag{5}$$

where σ is the default value of the Gaussian width and a and b are floating parameters we fit to the histogram. The fit of this function to the histograms for both values of σ is shown in Fig. 37, with parameters given in Tab. 19. The fits are not particularly good, but the functions show the properties we desire: rapid increase below σ and more gradual increase above σ. The two terms added to the likelihood for the detector resolution model are of the form of Eq. 5 with the fit values for a, b, and σ in place.

Parameter	Narrow Gaussian $\sigma = 1.2$ MeV/c^2	Wide Gaussian $\sigma = 3.0$ MeV/c^2
a	$(3.99 \pm 0.01) \times 10^{-6}$	$(2.59 \pm 0.01) \times 10^{-5}$
b	4214 ± 8	518 ± 2

Table 19: Parameter values for a fit of the form of Eq. (5) to the distributions for the detector resolution likelihood constraints shown in Fig. 37.

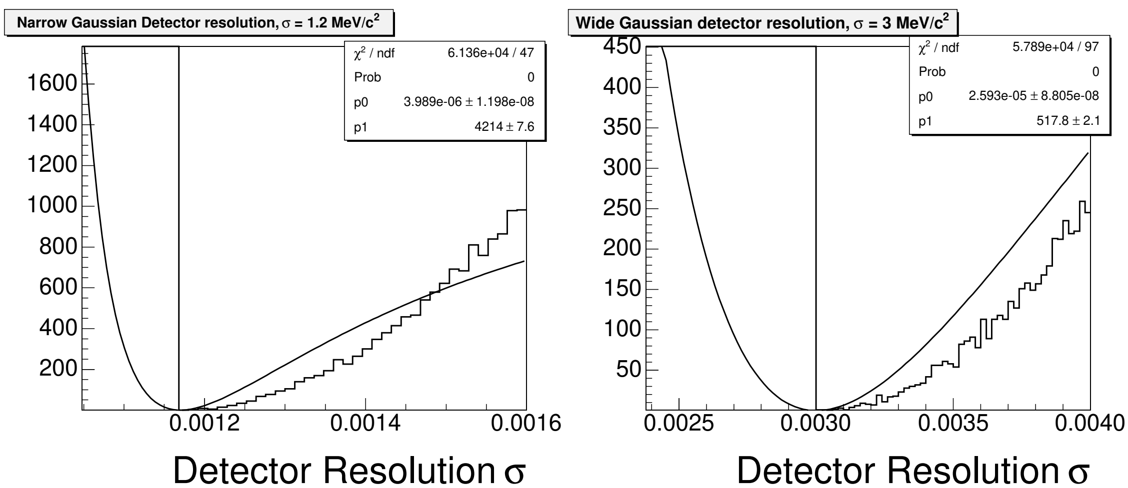


Figure 37: Plot of the likelihood constraints for the detector resolution parameter σ. On the left is the narrow Gaussian σ and on the right is the wide Gaussian σ. Both histograms are modeled by Eq. (5), with $p_0 = a$ and $p_1 = b$.

The final nuisance parameter to consider is the Λ_b^0 hadronization shape. The systematic variations we consider are described in Sec. 10.2.2. One variation involves shifting the number of events in the Λ_b^0 hadronization background; this nuisance parameter has already been taken care of by the Λ_b^0 sample composition. The effect on the Λ_b^0 hadronization shape from reweighting the Pythia can be described by the Q_{max} parameter, as shown in Sec. 6.5. Thus to describe the reweighting, we add two Gaussian constraints on the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ Q_{max} parameters: $\frac{1}{2}\left(\frac{Q_{max}-0.12}{0.01}\right)^2$ for $\Lambda_b^0\pi^-$, and $\frac{1}{2}\left(\frac{Q_{max}-0.11}{0.01}\right)^2$ for $\Lambda_b^0\pi^+$.

This gives a total of 12 constrained parameters added to the likelihood, bringing the number of floating parameters in the fit up to 19. As a check, we run this 19 parameter fit on the data and compare it to the default 7 parameter fit. The results are shown in Tab. 20, and there is almost no deviation in the final values for the 7 free floating parameters. However, this fit which took

Parameter	7 Param. Fit	19 Param. Fit	Nσ difference
$\Lambda_b^0\pi^-$ Comb. bkg norm	87 (fixed)	85 ± 45	–
$\Lambda_b^0\pi^+$ Comb. bkg norm	85 (fixed)	83 ± 44	–
$\Lambda_b^0\pi^-$ B Ha+UE norm	109 (fixed)	108 ± 61	–
$\Lambda_b^0\pi^+$ B Ha+UE norm	105 (fixed)	105 ± 59	–
$\Lambda_b^0\pi^-$ Λ_b^0 Ha+UE norm	1572 (fixed)	1497 ± 86	–
$\Lambda_b^0\pi^+$ Λ_b^0 Ha+UE norm	1541 (fixed)	1553 ± 84	–
$\Lambda_b^0\pi^-$ Λ_b^0 Ha+UE Q_{max}	0.12 (fixed)	0.123 ± 0.007	–
$\Lambda_b^0\pi^+$ Λ_b^0 Ha+UE Q_{max}	0.11 (fixed)	0.110 ± 0.005	–
Resolution (narrow Gauss)	0.00117 (fixed)	0.00117 ± 0.000006	–
Resolution (wide Gauss)	0.003 (fixed)	0.003 ± 0.00003	–
Isospin difference	0 (fixed), 0.0004 (expected)	0.00040 ± 0.00007	–
Width parameter	0.75 (fixed)	0.72 ± 0.05	–
Σ_b^- Q (MeV/c^2)	55.9 ± 1.0	56.0 ± 0.9	-0.1
Σ_b^- events	59 ± 14	63 ± 14	-0.3
Σ_b^+ Q (MeV/c^2)	48.5 ± 2.0	48.5 ± 1.9	0
Σ_b^+ events	32 ± 12	31 ± 12	0.08
Σ_b^{*-} events	69 ± 18	77 ± 19	-0.4
Σ_b^{*+} events	77 ± 17	74 ± 18	0.2
Σ_b^* - Σ_b Q (MeV/c^2)	21.2 ± 1.9	21.0 ± 1.8	0.1
NLL	-24160.4	-24162.5	–

Table 20: Result of the default Σ_b signal fit to data, with nuisance parameters included as constraints in the fit. The number of σ difference for the seven free floating parameters is shown in the last column.

30 seconds with 7 floating parameters now takes nearly 5 minutes to converge with 19 floating parameters.

The first step in evaluating the significance via likelihood ratio is to fit the data with the alternate hypotheses (null, two peak, and three peak). This is done with the appropriate constrained parameters added to the likelihood. For the null hypothesis, since there are no Σ_b signal peaks, the nuisance parameters related only to the Σ_b signal (Σ_b intrinsic width, isospin splitting, and detector resolution) are not added to the likelihood. The resulting likelihood ratios for each alternate hypothesis are given in Tab. 21. The one peak and zero peak fits are shown in Fig. 38.

To evaluate the p-value for each alternate hypothesis, we generate Toy MC samples from the alternate hypothesis. The nuisance parameter values are tossed before each generation. The Toy MC sample is then fit with both the alternate hypothesis and the default signal hypothesis, with the nuisance parameters included in both likelihoods. For the null hypothesis, which has an extremely large ΔNLL value, we expect that even in 10 million Toy MC samples we will never find one with as large a ΔNLL value as in data. In fact, most of these samples with have no sign of a Σ_b signal. Because the default fit with nuisance parameters takes nearly 5 minutes to converge, it would take a prohibitive amount of CPU to run as many Toy MC samples as we need if the default fit runs on each sample. After consultation with the CDF statistics committee, we

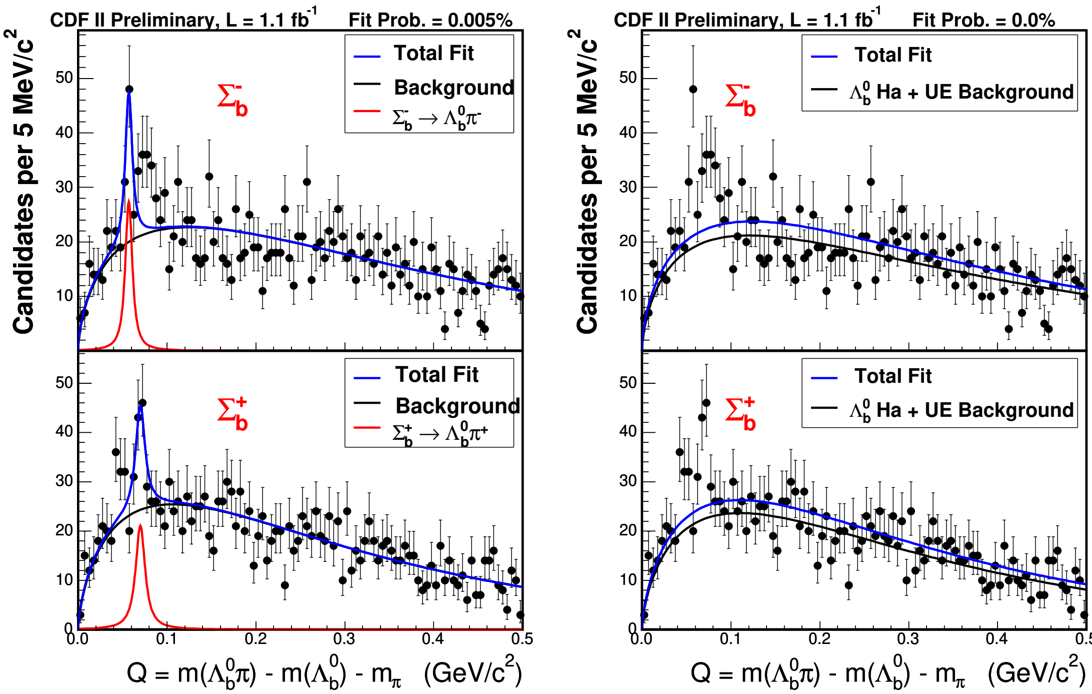


Figure 38: Fitting the signal with only one state per $\Lambda_b^0\pi$ charge combination (left), and with no signal at all (right).

decided to run the background only fit on each Toy MC sample generated with no signal. Then we calculate the χ^2 between the sample and the background only fit just in the Σ_b signal region of $Q \in [30,100]$ MeV/c^2. If the χ^2 is below a certain value, we can conclude that the background fit describes this sample well, and thus there is no point in running the default signal fit, because there is no sign of signal. This sample may then be added to the denominator of our p-value calculation.

To determine the appropriate χ^2 cut value, we generated 20,000 Toy MC samples from the default signal fit, and then fit them with the background only fit and calculated the χ^2. We also fit each sample with the default signal fit, so that we could calculate the ΔNLL value for each sample. We then plotted ΔNLL versus the background fit χ^2, as shown in Fig. 39. From this, we determine that all samples with a significant amount of signal (ΔNLL greater than or equal to the value found in data) should have a χ^2 between the sample and background only fit of greater than 45. There are only 11 fits with a $\chi^2 < 45$ and $\Delta NLL > 42$; running each of these 11 fits locally, we discovered that each was due to a failed fit, either the background fit or the default signal fit. Through this, we discovered that the error levels of the nuisance parameters were not being set properly, thus introducing a fit instability as the error level determines the step size during likelihood minimization. We were able to fix this by expliciting setting the error levels of each parameter. After rerunning the 11 fits with this fix, all converged properly and the ΔNLL of each fit was reduced to less than 40. Thus, we feel confident that using a χ^2 cut at 45 will not eliminate any Toy MC samples with potentially significant signal.

For the null signal hypothesis, we generated almost 12 million samples. Of those, $\sim 72,700$ failed the $\chi^2 < 45$ cut and ran the default signal fit. Even after fixing the error level of the con-

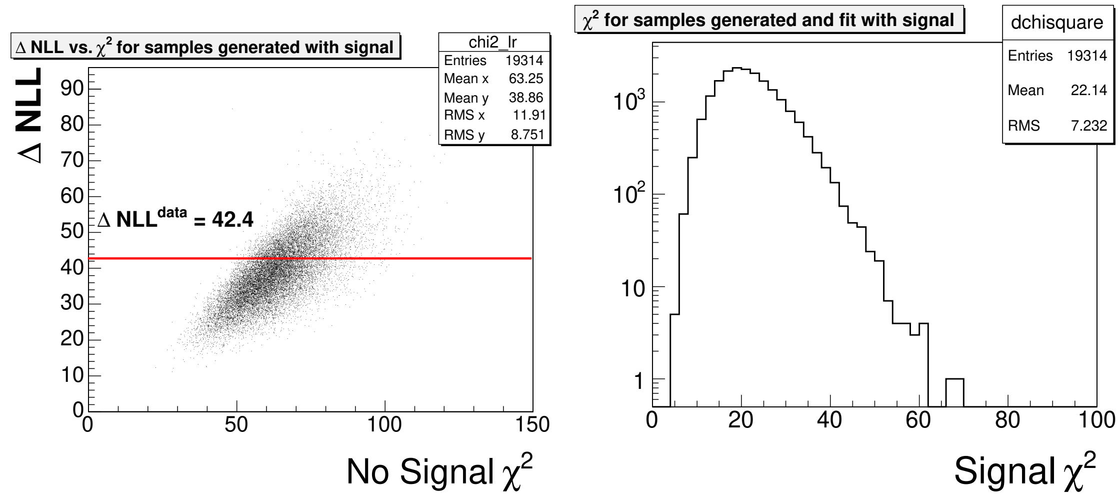


Figure 39: Left: Distribution of the ΔNLL versus the χ^2 of a background fit to Toy MC samples generated from the Σ_b signal distribution. The χ^2 is calculated only in the Σ_b signal region of $Q \in [30, 100]$ MeV/c^2. The ΔNLL is calculated between a background only fit and a default signal fit to the same sample. The ΔNLL value found in data is shown on the plot. The plot on the right shows the χ^2 distribution of the signal fit to the same Toy MC samples.

strained parameters, there were some failed signal fits to the Toy MC samples. When a fit fails, we would expect the fit status from Minuit to be returned as failed, with either an approximate error matrix or a non-positive definite error matrix. We removed all fits which did not end with a full, accurate error matrix. Some fits returned an accurate error matrix, but upon closer examination we found the expected distance to minimum (EDM) of the fit was very large. A properly minimized fit should have an EDM of order 10^{-5}. To remove these failed fits, we required an EDM of less than 0.01 on the default. However, even after these clearly failed fits were removed, some failed fits remained. In these fits, the error matrix was returned as full and accurate, the EDM was small, and yet the χ^2 between the signal fit and the Toy MC sample was large, on the order of several hundred in the signal region alone. Upon plotting some of these failed fits, we

Hypothesis	$(NLL_2 - NLL_1)^{\text{data}}$	p-value	Significance (σ)
Null	42.4	$< 8.3 \times 10^{-8}$ (7.0×10^{-14})	> 5.23 (7.40)
Two peaks	15.3	9.2×10^{-5}	3.74
No Σ_b^- Peak	11.7	3.2×10^{-4}	3.41
No Σ_b^+ Peak	3.9	9.0×10^{-3}	2.36
No Σ_b^{*-} Peak	10.8	6.4×10^{-4}	3.22
No Σ_b^{*+} Peak	11.3	6.0×10^{-4}	3.24

Table 21: Likelihood ratio p-values for the alternate signal hypotheses, where NLL_2 denotes the negative log likelihood of the alternate hypothesis and NLL_1 that of the four-peak fit. For the null hypothesis, no events were observed with the significance seen in data. The first null hypothesis p-value is a lower limit, while the p-value given in parenthesis is extrapolated from the ΔNLL distribution as shown in Fig. 40.

find that one of the four Σ_b peaks is a bogus spike with a hundred or more events where there is no peak in the sample.

To check that these fits are truly failed and not due to an instability in the fitter, we selected five seeds which resulted in a good signal fit and five seeds which resulted in a failed fit. We tried running these seeds in both RooFit v2.05 (the default for this analysis) and RooFit v2.07, on different machines. The fits, both good and failed, were the same in both versions on different machines. We ran the fits through valgrind and interactively through GDB, searching for memory leaks, overwrites, and debugging information. We found nothing, and both the good and failed fits finished with the same values in the debugging programs. For one good fit and two failed fits, we saved the generated Toy MC sample to a ROOT ntuple and tried running just the default fit over these samples. In all three cases, the standalone default fit finished with the same results as in the Toy MC framework. We ran with RooFit in verbose mode, printing out the step and FCN at every point in the minimization, and still found no errors. Finally, we converted our unbinned fits to binned fits, both binned maximum likelihood and binned χ^2, and ran over the three saved Toy MC samples. The good fit still converged, although the final values shifted slightly from the unbinned fit. The binned fits to the two failed fits either failed to converge, with a non-accurate error matrix and large EDM, or converged to about zero events in the Σ_b peaks.

From these exercises, we conclude the following:

- The "good" fits converged under every fit configuration, and are consist between binned and unbinned fits. Therefore, they truly are good fits.
- The failed fits which look like good fits (accurate error matrix, small EDM) are truly bad fits. A binned fit to the same sample will either fail or converge with no Σ_b peaks.
- The problem with the failed fits is not an instability of the Σ_b; if it was, the binned fits would fail in the same manner.
- The problem with the failed fits appears to be a bug in the computation of the unbinned NLL. When we do a scan of the likelihood as a function of the one "crazy" parameter (the number of events in the bogus spike), it is smooth with a clear minimum at the bogus value.

The authors of RooFit have been contacted to notify them of these problems. Although debugging work will continue, for this analysis it is enough to show that the fit itself is stable and to remove these failed fits.

To remove the last of the failed fits, we use a χ^2 cut on the default signal fit. From the χ^2 distribution of the signal fit on Toy MC samples generated with signal (Fig. 39, left) we choose a cut of $\chi^2 > 100$ to remove fits with a bogus peak. After applying these three cuts, we are left with ~ 17500 samples in the null hypothesis ΔNLL distribution. None of these Toy MC samples has close to the LR found in data; the largest value is $\Delta NLL \approx 24$. The p-value is calculated using the total number of generated samples, without removing the failed fits from the p-value denominator. These 12 million samples correspond to a p-value of $\sim 8.4 \times 10^{-8}$, or a significance of 5.23 σ, so we know the true p-value is less than this. We plot the distribution of $NLL_2 - NLL_1$ and extrapolate with a decaying exponential as in Sec. 9.1.2. The exponential function is fit only to the ΔNLL distribution above 12, as shown in Fig. 40. The resulting fit parameters and the integral above the ΔNLL found in data are given in Tab. 22.

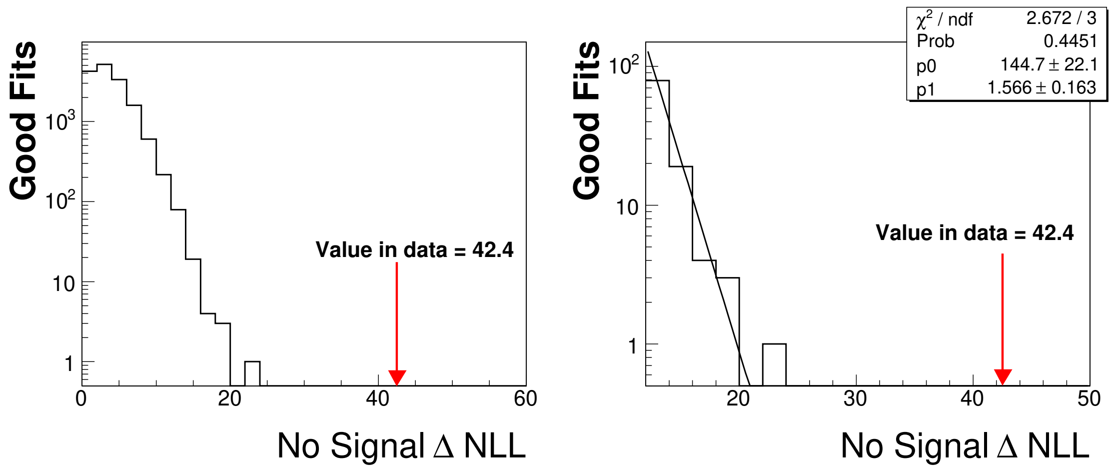


Figure 40: The ΔNLL distribution for the no signal Toy MC samples. While the ΔNLL value found in data is 42.4, the largest one found in the Toy MC samples was below 24, as shown on the left. We extrapolated the tail of the ΔNLL distribution out to the value found in data using a decaying exponential fit to the region $\Delta NLL > 12$, as shown on the right. The integral of this function above 42.4 is used to calculate the p-value.

For the two and three peak alternate hypotheses, we were able to generate enough Toy MC samples to find several with the same LR as data, so the p-values listed in Tab. 21 are easily calculated. However, we again had to remove failed fits using the same three cuts (accurate error matrix, EDM < 0.01, and default fit $\chi^2 < 100$) as for the null hypothesis. The ΔNLL distributions for the two and three peak hypotheses are shown in Figs. 41 and 42.

From this, we can conclude that the null hypothesis is excluded at the 5 σ level at least. Each of the four peaks is on the order of 3 σ significance on its own, except for the Σ_b^+ peak which is slightly weaker than the other three peaks.

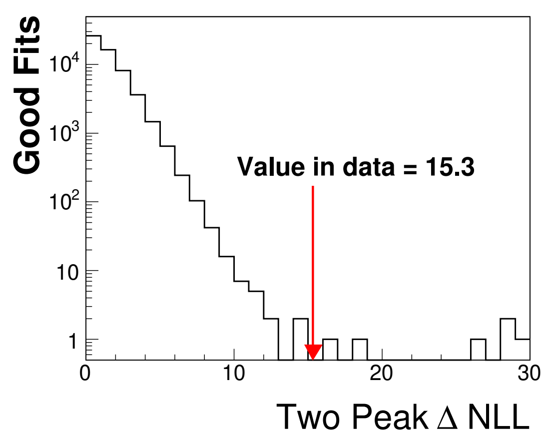


Figure 41: Distribution of the ΔNLL values for the fits using only one peak per $\Lambda_b^0\pi$ charge combination. The value of the ΔNLL in data is marked on the plot.

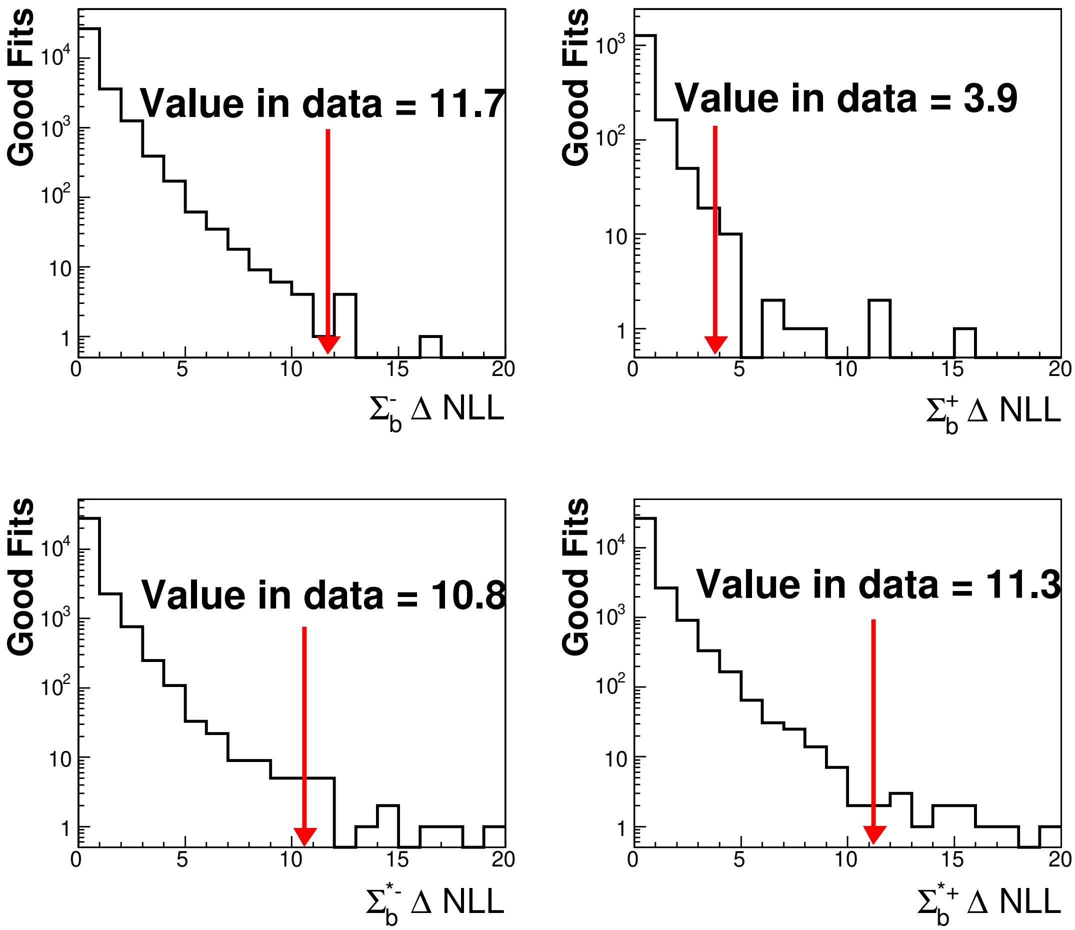


Figure 42: Distribution of the ΔNLL values for each of the fits using only three peaks. There are different numbers of Toy MC samples in each distribution. The value of the ΔNLL in data is marked on each plot.

Parameter	Exponential fit
p_0	145 ± 22
p_1	1.56 ± 0.16
Fit Prob	$\sim 45\%$
Integral above 42.4	8.4×10^{-7}

Table 22: Fit parameters for a decaying exponential fit to the ΔNLL distribution of background-only Toy MC samples shown in Fig. 40. The exponential is fit only to the tail of the distribution, $\Delta NLL > 12$.

Parameter	Value	Parabolic Error	MINOS Errors
Σ_b^- Q (MeV/c^2)	55.9	0.951	(+0.973, -0.950)
Σ_b^- events	59	14.2	(+14.6, -13.7)
Σ_b^+ Q (MeV/c^2)	48.5	1.97	(+1.98, -2.17)
Σ_b^+ events	32	12.1	(+12.5, -11.7)
Σ_b^{*-} events	69	17.6	(+18.0, -17.1)
Σ_b^{*+} events	77	16.8	(+17.3, -16.3)
Σ_b^* - Σ_b Q (MeV/c^2)	21.2	1.92	(+2.00, -1.94)
NLL	-24160.4	–	–

Table 23: Fit parameter and error values from the fit to data. We use positive and negative errors as the error range is asymmetric.

Parameter	Global	1	2	3	4	5	6	7
1 Σ_b^- Q	0.29310	1.000	0.162	0.151	-0.016	-0.122	-0.017	-0.212
2 Σ_b^- events	0.30083	0.162	1.000	-0.063	0.007	-0.246	0.007	0.088
3 Σ_b^+ Q	0.71368	0.151	-0.063	1.000	-0.052	0.029	-0.010	-0.712
4 Σ_b^+ events	0.18688	-0.016	0.007	-0.052	1.000	-0.003	-0.164	0.074
5 Σ_b^{*-} events	0.26287	-0.122	-0.246	0.029	-0.003	1.000	-0.003	-0.040
6 Σ_b^{*+} events	0.19997	-0.017	0.007	-0.010	-0.164	-0.003	1.000	0.080
7 Σ_b^* - Σ_b Q	0.72793	-0.212	0.088	-0.712	0.074	-0.040	0.080	1.000

Table 24: Correlation matrix for the fit to data. Only the Σ_b^+ Q value and the Σ_b^* - Σ_b Q value show high correlations.

9.2 Σ_b Fit Results

The results of the simultaneous $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ fit to Σ_b data is shown in Fig. 43, with the fit results in Tab. 23 and the correlation matrix given in Tab. 24. From the correlation matrix we see only two parameters which are highly correlated – the Σ_b^+ Q value and the Σ_b^* - Σ_b Q value. This is inescapable in the simultaneous fit; because the Σ_b^+ peak is much smaller than the other three Σ_b peaks, it relies on the Σ_b^* - Σ_b mass difference to fix its location. This correlation will be studied in more detail in Sec. 9.4.1.

The NLL given in Tab. 23 is also shown superimposed on the Toy MC NLL distribution in Fig. 34. As indicated by its position in this distribution, this is a good fit for this amount of data. This is also confirmed by the χ^2 goodness of fit test, which gives a fit probability of 79% for this fit to the region around the Σ_b signal, $Q \in [0, 200]$ MeV/c^2, which is shown in Fig. 44.

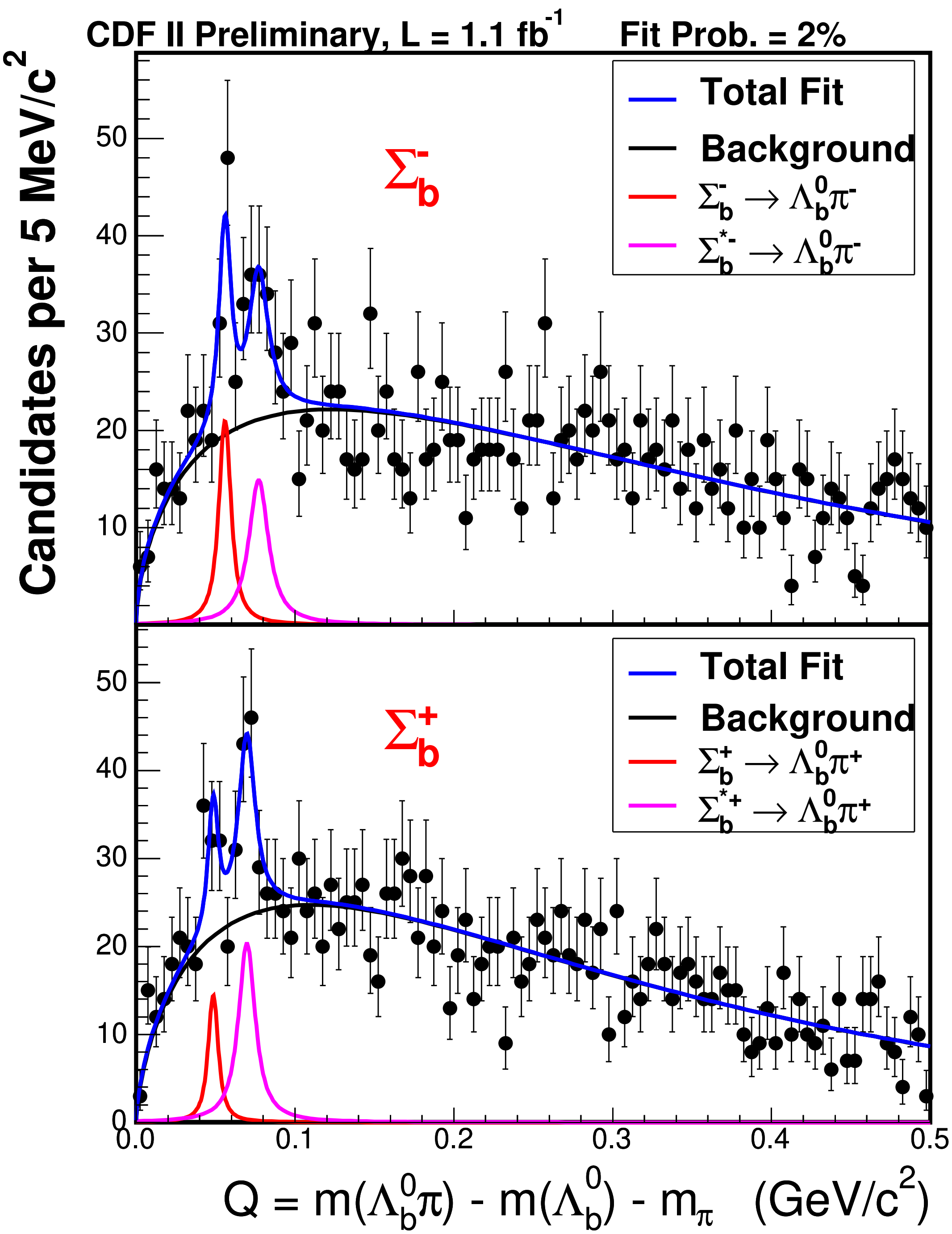


Figure 43: Simultaneous fit to the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ Σ_b signal.

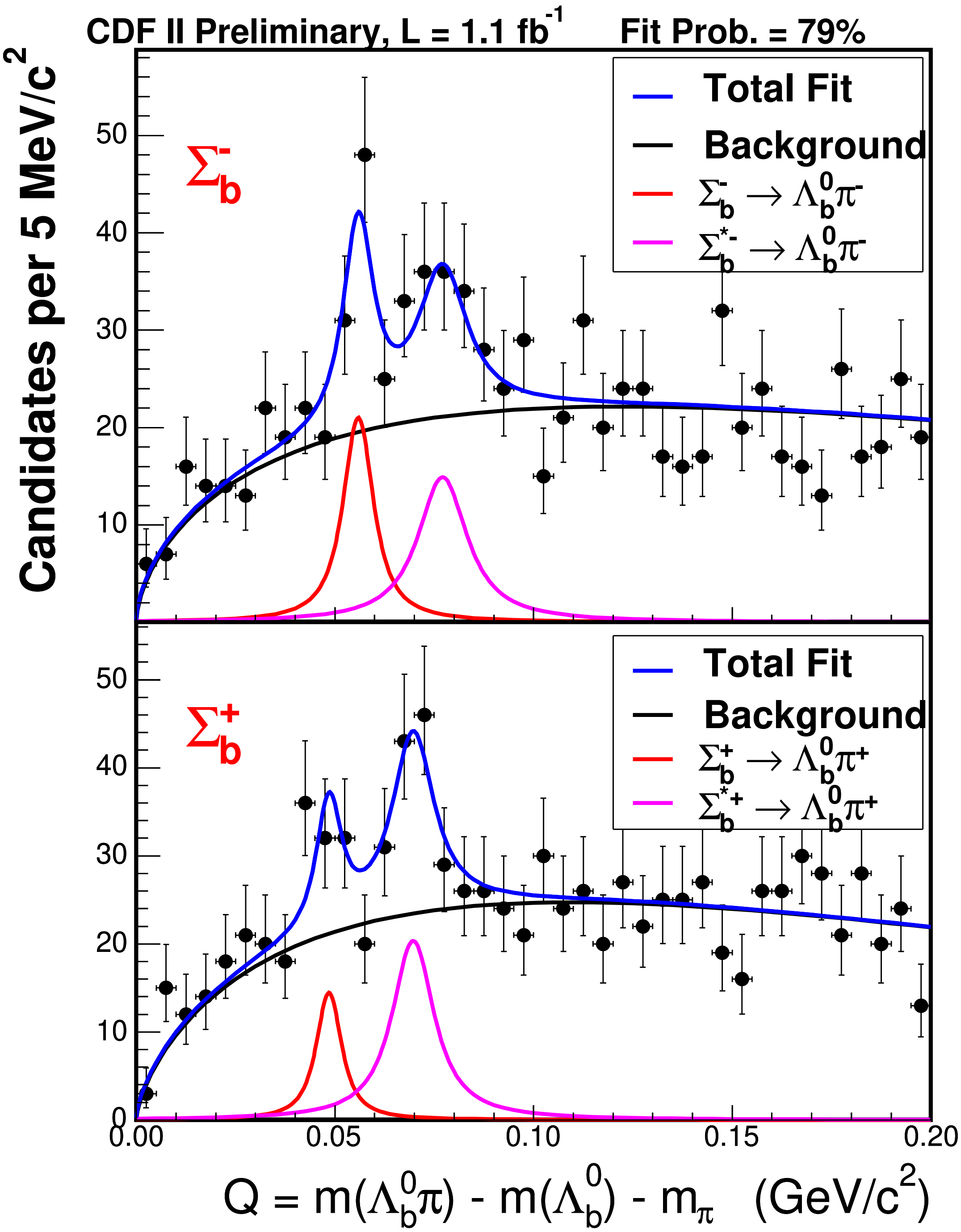


Figure 44: Simultaneous fit to the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ Σ_b signal, focusing on a smaller region around the signal peaks ($Q \in [0, 200]$ MeV$/c^2$).

Parameter	Value	Error	Diff. from default
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE Norm	1548	45	-24
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE Norm	1493	45	-48
Σ_b^- Q (MeV/c^2)	55.9	0.9	0
Σ_b^- events	63	14	4
Σ_b^+ Q (MeV/c^2)	48.5	2.0	0
Σ_b^+ events	32	12	0
Σ_b^{*-} events	77	18	8
Σ_b^{*+} events	76	17	-1
Σ_b^* - Σ_b Q (MeV/c^2)	21.2	1.9	0
NLL	-24162.0	–	-1.6

Table 25: Fit parameter and error values from the fit with the Λ_b^0 Ha + UE background normalization floating.

9.3 Alternate Σ_b Fits

The Λ_b^0 Ha + UE background was determined from a PYTHIA Monte Carlo sample which had to be reweighted twice – once for the Λ_b^0 p_{T} spectrum, and once for the p_{T} spectrum of tracks around the Λ_b^0 , as described in Sec. 6.4. Consequently, there are several systematics associated with this derived shape which we must check. These systematics are detailed in Sec. 10. In the following sections we also fit the data with the alternate Λ_b^0 Ha + UE descriptions, as a check to the value of the systematic assigned to each.

9.3.1 Floating Λ_b^0 Ha + UE Normalization

This normalization is fixed from the ratio of the number of Λ_b^0 in the MC sample to the number of Λ_b^0 in the data sample. As a check of this normalization, we fit the data with the Λ_b^0 Ha + UE normalization left floating in the fit, for both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$. The fit is shown in Fig. 45 and fit parameters are given in Tab. 25. The parameters, particularly the number of Ha + UE events, vary slightly from the default parameters.

9.3.2 Alternate Λ_b^0 Ha + UE Shape

We fit all the background shapes with Eq. (2). There are many other shapes we could have chosen, which may also cause a systematic effect. One of the other possible shapes which describes this background well is the RooFit $D^* - D^0$ background shape, which also drops steeply at a cut-off value and then levels out at high mass. This PDF consists of a cut-off parameter $dm0$ and three shape parameters. The fit of the $D^* - D^0$ shape to the Λ_b^0 Ha + UE backgrounds are shown in Fig. 46 with the fit parameters given in Tab. 26. The numbers of events are scaled 0.208 as explained in Sec. 6.4.

We then fit the data with the alternate $D^* - D^0$ background shape in place of the default shape. As shown in Fig. 47 and Tab: 27, there is virtually no shift in the fit parameters with the alternate shape, and the overall fit is slightly worse.

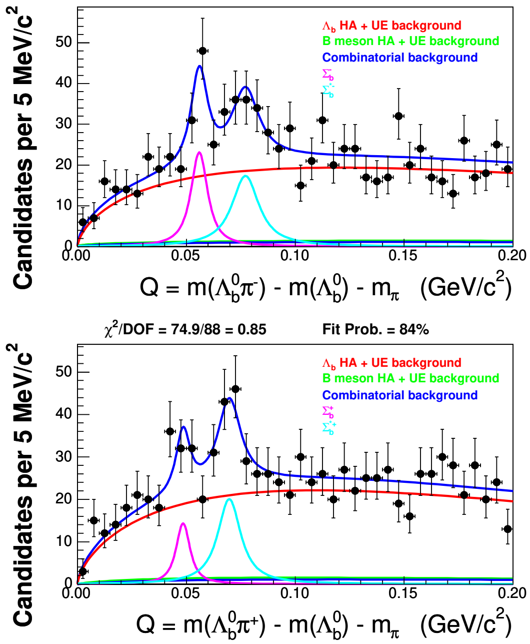


Figure 45: Fit to data with the Λ_b^0 Ha + UE background normalization floating.

9.3.3 Floating Λ_b^0 Ha + UE Shape

Reweighting the Λ_b^0 track $p_{\rm T}$ spectrum also changed the shape of the Λ_b^0 Ha + UE Q distribution. To consider the error on this derived shape, we allowed both the normalization and the fit parameters (α, γ, and Q_{max}) to float in the fit. The fit is shown in Fig. 47 and fit parameters are given in Tab. 28. Except for the number of Σ_b^{*-} events, the parameters are only very slightly different than the default fit to data.

$\Lambda_b^0\pi^-$	Parameter Value	Error
$dm0$	0.0028	0.0004
C	0.0855	0.0003
A	-0.467	0.001
B	-1.5×10^{-4}	4×10^{-5}
Number of events	7550	123
Scaled events	1570	–
$\Lambda_b^0\pi^+$	Parameter Value	Error
$dm0$	5.3×10^{-5}	7×10^{-6}
C	0.074	0.008
A	-0.37	0.02
B	-2.2×10^{-6}	5×10^{-7}
Number of events	7410	122
Scaled events	1541	–

Table 26: Parameter values for the shape and normalization of the Λ_b^0 Ha + UE background when fit with the RooFit $D^* - D^0$ background PDF.

Parameter	Value	Error	Diff. from default
Σ_b^- Q (MeV/c^2)	55.9	1.0	0
Σ_b^- events	58	14	-1
Σ_b^+ Q (MeV/c^2)	48.5	2.0	0
Σ_b^+ events	29	12	-3
Σ_b^{*-} events	68	18	-1
Σ_b^{*+} events	74	17	-3
Σ_b^* - Σ_b Q (MeV/c^2)	21.1	2.0	-0.1
NLL	-24157	–	3.4

Table 27: Fit parameter and error values from the fit with a RooFit $D^* - D^0$ function for the Λ_b^0 Ha + UE background shape.

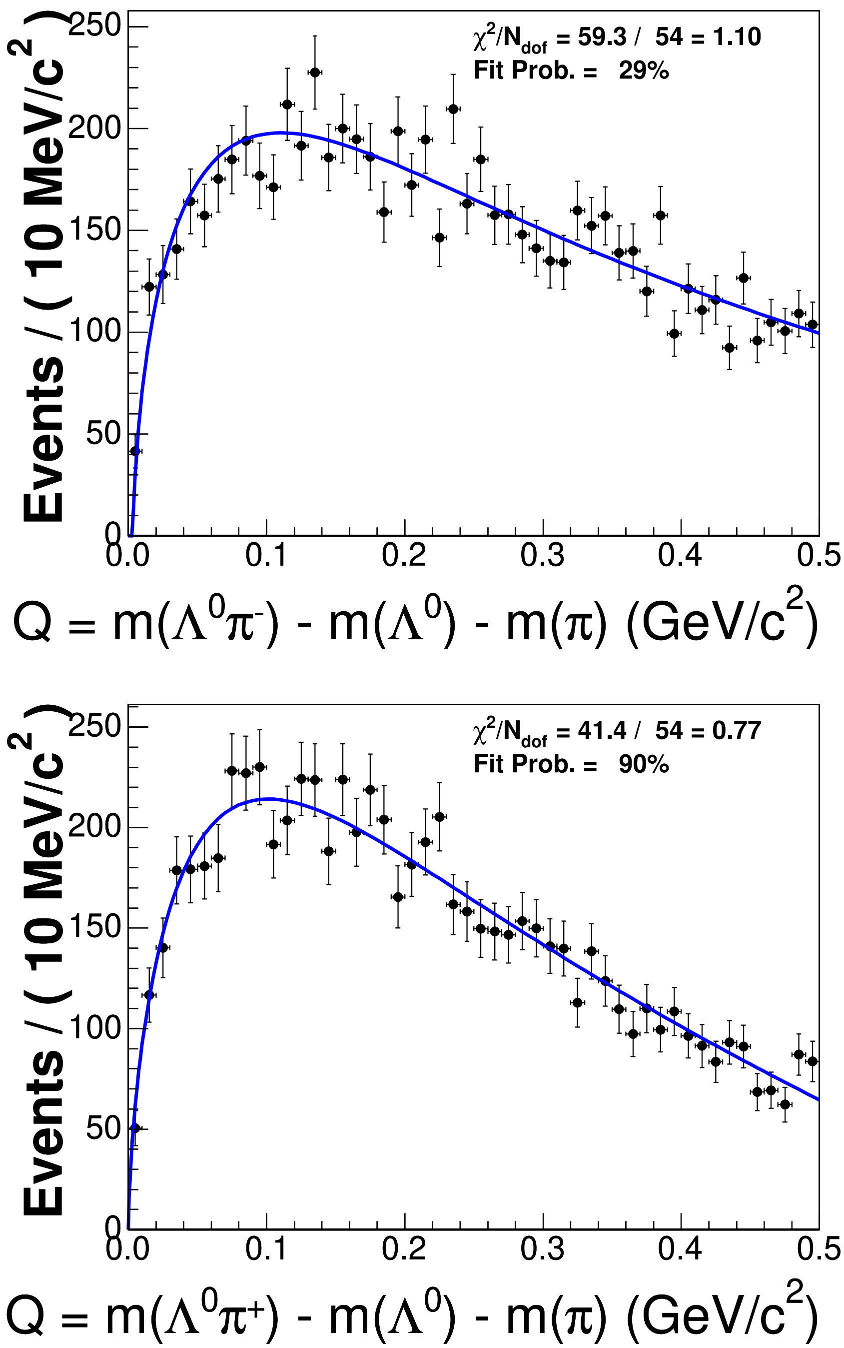


Figure 46: RooFit $D^* - D^0$ background PDF fits to the Λ_b^0 hadronization and underlying event backgrounds, from PYTHIA. Top: $\Lambda_b^0\pi^-$, bottom: $\Lambda_b^0\pi^+$.

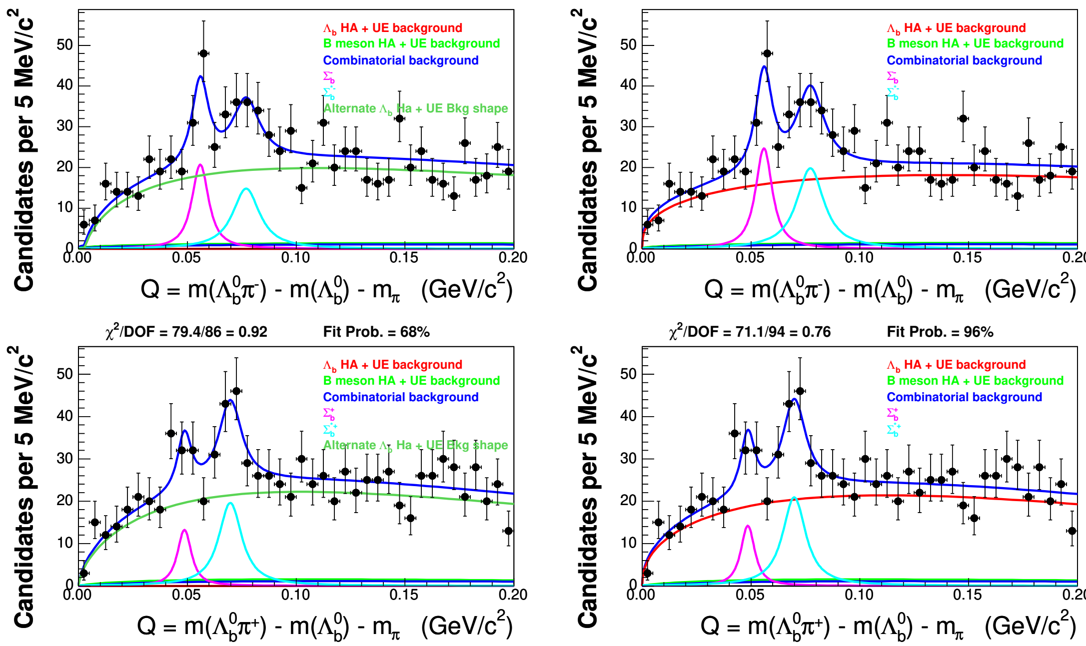


Figure 47: Fit to data with alternate Λ_b^0 Ha + UE background shapes. On the left, the default background shape is replaced with a RooFit $D^* - D^0$ function. On the right, the default shape is left floating in the fit.

Parameter	Value	Error	Diff. from default
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE Norm	1545	45	-27
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE α	0.45	0.01	-0.22
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE γ	1.1	0.1	0.24
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE Q_{max}	0.11	0.003	0
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE Norm	1477	54	-64
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE α	0.35	0.4	-0.31
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE γ	1.2	1.8	0.47
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE Q_{max}	0.14	0.07	0.02
Σ_b^- Q (MeV/c^2)	55.9	0.9	0
Σ_b^- events	67	18	8
Σ_b^+ Q (MeV/c^2)	48.5	1.9	0
Σ_b^+ events	32	12	0
Σ_b^{*-} events	89	28	20
Σ_b^{*+} events	80	18	3
Σ_b^* - Σ_b Q (MeV/c^2)	21.3	1.8	0.1
NLL	-24164.6	–	-4.2

Table 28: Fit parameter and error values from the fit with the Λ_b^0 Ha + UE background shape and normalization floating.

Reweighted down			
Parameter	Value	Error	Diff. from default
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE Norm	1485	44	-87
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE Norm	1553	45	12
Σ_b^- Q (MeV/c^2)	55.9	0.9	0
Σ_b^- events	66	15	7
Σ_b^+ Q (MeV/c^2)	48.5	2.0	0
Σ_b^+ events	30	12	-2
Σ_b^{*-} events	83	18	14
Σ_b^{*+} events	73	17	-4
Σ_b^* - Σ_b Q (MeV/c^2)	21.2	1.9	0
NLL	-24162.1	–	-1.7
Reweighted up			
Parameter	Value	Error	Diff. from default
$\Lambda_b^0\pi^-$ Λ_b^0 Ha + UE Norm	1499	45	-73
$\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE Norm	1564	46	23
Σ_b^- Q (MeV/c^2)	55.9	0.9	0
Σ_b^- events	60	14	1
Σ_b^+ Q (MeV/c^2)	48.7	2.1	0.2
Σ_b^+ events	26	12	-6
Σ_b^{*-} events	74	18	5
Σ_b^{*+} events	66	17	-11
Σ_b^* - Σ_b Q (MeV/c^2)	20.9	2.0	-0.3
NLL	-24159.7	–	0.7

Table 29: Fit parameter and error values from a different reweighting of the PYTHIA MC track p_T spectrum. The parameters for the “Reweight Down” shape are shown on top, while the parameters for the “Reweight Up” shape are on the bottom.

9.3.4 Reweighted Λ_b^0 Ha + UE Shape

Sec. 6.5 describes systematic variations to the function used to reweight the PYTHIA Λ_b^0 Monte Carlo sample for the track p_T , and the effect this has on the Λ_b^0 Ha + UE Q background shape. We use these alternate backgrounds (reweighted up and reweighted down) in a fit to data, to check the systematic effect these shapes have on the Σ_b measurement.

For these fits we also left the Λ_b^0 Ha + UE normalization floating. By construction, the normalization of the reweighted down shape was too low to match data, and the that of the reweighted up shape was too high, so we had to allow the normalization to float to match the data. The final fit parameters are given in Tab. 29 with the fits shown in Fig. 48.

This shape reweighting turns out to be one of the largest systematic errors, as will be shown in Sec. 10.2.2. From the fits to data, we see large shifts in the numbers of Λ_b^0 Ha + UE and Σ_b^{*+} events.

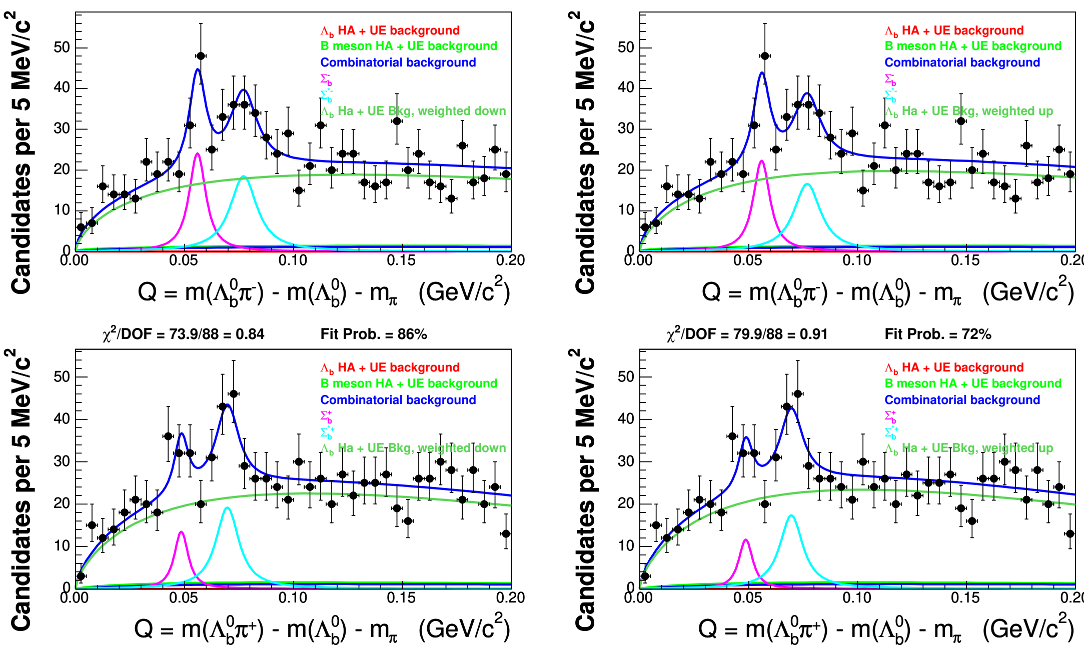


Figure 48: Fits to the data using a Λ_b^0 Ha + UE background shape derived from a different reweighting of the PYTHIA MC track p_{T} spectrum. On the left is the reweighted down shape; the reweighted up shape is shown on the right.

9.4 Likelihood Scans

To check for any other fit minima, we performed negative log likelihood (NLL) scans of each floating parameter. After the fit converges, we fix all parameters but one, and plot the value of the NLL as a function of that one floating parameter. The NLL scans over a 2.5σ range for all seven parameters are shown in Fig. 49, where the minimized NLL has been fixed to zero in all cases. Clearly there is only one minimum NLL for each parameter. However, as we already see from the comparison of parabolic errors to positive and negative errors, the NLL is asymmetric as a function of the parameters.

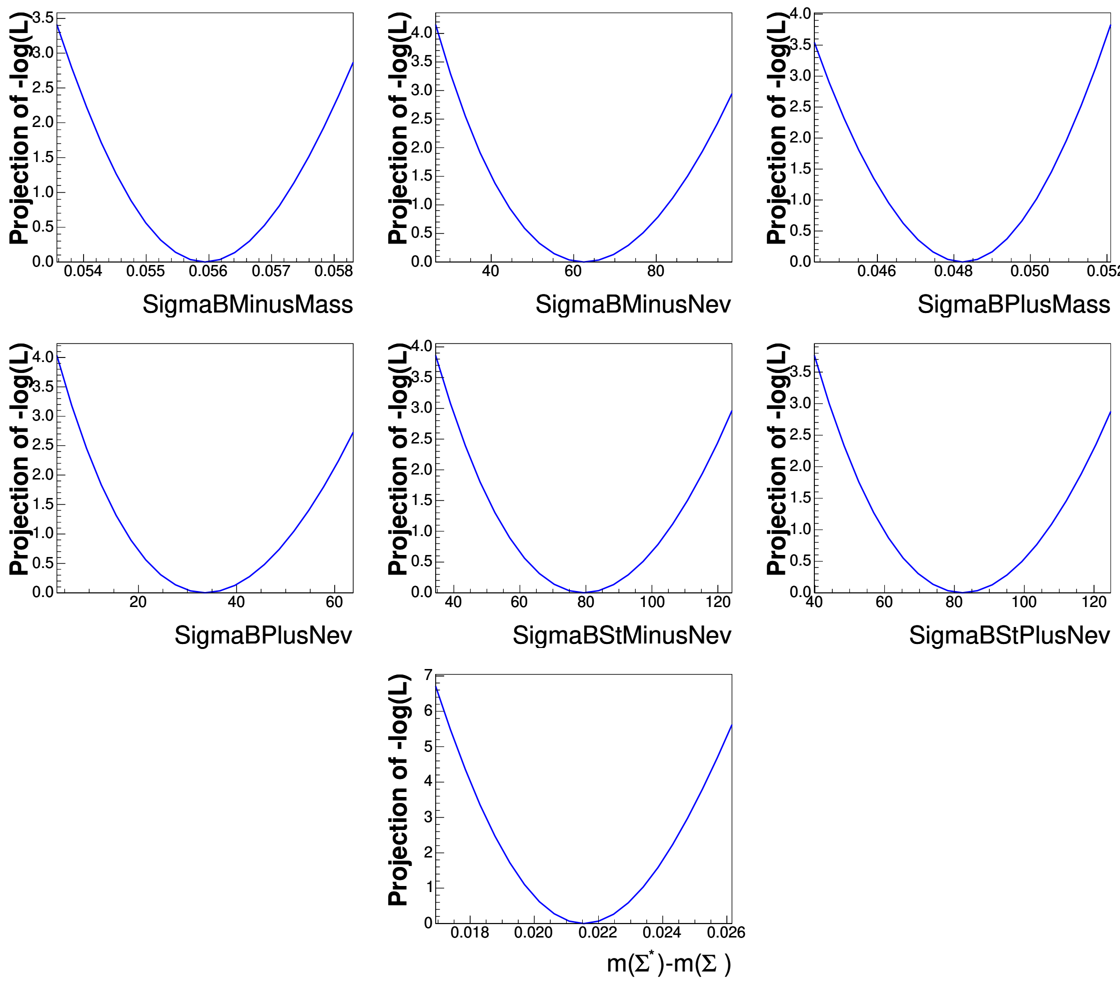


Figure 49: Negative log likelihood scans for the seven floating parameters in the fit to data.

9.4.1 Contour Plot

For the two highly correlated parameters, the Σ_b^+ Q value and the Σ_b^* - Σ_b Q value, we are interested in a two-dimensional contour plot. The 1, 2, and 3σ contours for these two variables are shown in Fig 50.

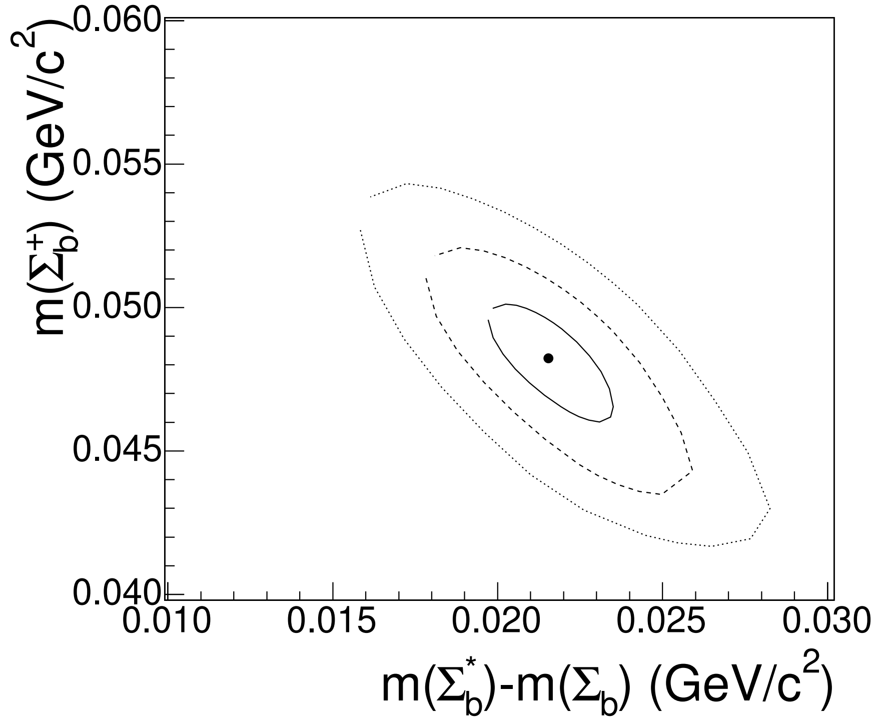


Figure 50: The 1, 2, and 3σ contours for the Σ_b^+ Q value and the Σ_b^* - Σ_b Q value parameters.

10 Systematic Error Estimates

There are two main components to the systematic errors on the mass measurement – mass scale systematics on the soft pion, and systematics on assumptions used in the fitter. Both sources of systematic errors are treated in detail below.

10.1 Mass Scale Systematics

Mass scale systematics are generally small, and due primarily to the fact that the calibration is only as precise as it has been tested out to. To determine mass scale errors, we reconstruct resonances of known negligible width such as the D^*. Using measurements of the D^*, Σ_c , and Λ_c^* resonances, we extrapolate the mass scale offset expected for the Σ_b Q values.

10.1.1 D^* Measurements

The $D^* - D^0$ mass difference is 145.421 ± 0.010 MeV/c^2 and the width of the D^* is 0.096 ± 0.022 MeV/c^2 [1], essentially zero. By fitting the $D^* - D^0$ distribution, we measure both the systematic errors on the mass measurement and the mass resolution – the difference in the measured $D^* - D^0$ mass gives the error on the mass measurement, while the width of the D^* resonance gives the mass resolution. For this study, we use the dataset `xbhd0d`, corresponding to 490 $\pm$ 30 pb^{-1} of data, reconstructed with the `BottomMods` package [30]. We also check the D^* resolution using a PYTHIA Monte Carlo sample of D^*. This sample was generated by the MIT group and uses default PYTHIA settings. Both data and MC are recosntructed for $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K^+\pi^-$.

Width systematics are additionally complicated by the use of soft pions, *e.g.* with p_T less than 400 MeV/c. If these tracks are not as well measured by the detector as tracks of higher momentum, they will have a worse detector resolution than higher momentum tracks. We divide the $D^* - D^0$ mass plot into bins of track p_T in order to compare the mass resolution of different momentum tracks. The bins used are:

- Track $p_T > 200$ MeV/c (all tracks)
- 200 MeV/c $<$ track $p_T <$ 400 MeV/c
- 400 MeV/c $<$ track $p_T <$ 600 MeV/c
- 600 MeV/c $<$ track $p_T <$ 800 MeV/c
- 800 MeV/c $<$ track $p_T <$ 1 GeV/c
- Track $p_T > 400$ MeV/c

The mass plots for all ranges in D^* data are shown in Figs. 51 - 56. Each mass plot is fit with a double Gaussian for the $D^* - D^0$ signal and the RooDstD0BG function from the RooFit library for the background. Mass and resolution parameters for all fits are listed in Tab. 30. As these values show, the resolution is the same within errors for tracks of all momentum, including those with $p_T < 400$ MeV/c. Thus the mass scale offset is taken as flat with respect to the p_T .

Track p_T range (MeV/c)	Mean	σ_{narrow}	σ_{wide}
Data			
Track $p_T > 200$	145.456	0.48 ± 0.01	1.16 ± 0.06
$200 <$ track $p_T < 400$	145.480	0.46 ± 0.03	1.24 ± 0.09
$400 <$ track $p_T < 600$	145.448	0.51 ± 0.02	1.31 ± 0.10
$600 <$ track $p_T < 800$	145.462	0.47 ± 0.02	1.13 ± 0.11
$800 <$ track $p_T < 1000$	145.460	0.46 ± 0.03	1.24 ± 0.13
Track $p_T > 400$	145.456	0.48 ± 0.01	1.16 ± 0.06
PYTHIA Monte Carlo			
Track $p_T > 200$	145.542	0.41 ± 0.01	1.07 ± 0.04
$200 <$ track $p_T < 400$	145.520	0.43 ± 0.02	1.58 ± 0.16
$400 <$ track $p_T < 600$	145.520	0.44 ± 0.01	1.29 ± 0.07
$600 <$ track $p_T < 800$	145.550	0.38 ± 0.01	1.04 ± 0.07
$800 <$ track $p_T < 1000$	145.550	0.38 ± 0.03	0.94 ± 0.09
Track $p_T > 400$	145.542	0.41 ± 0.01	1.06 ± 0.04

Table 30: Mass and resolution parameters for the $D^* - D^0$ fits in bins of track p_T , for both data and a PYTHIA Monte Carlo sample. Mean and σ are in units of MeV/c^2.

We then fit the D^* Monte Carlo sample with the same formula and in the same track p_T ranges as data. These fits are shown in Figs. 57 - 62, and the values of σ_{narrow} and σ_{wide} for each fit are given in the bottom of Tab. 30. In Tab. 31, we compute the percentage difference between data and Monte Carlo for each p_T range. These differences show that the MC typically underestimates the D^* width by about 15-20%. The discrepancy between data and MC is actually lowest in the low p_T ranges, indicating that MC models the resolution of the soft tracks as well as it models the resolution of the high p_T tracks.

The mass resolution from the Σ_b signal Monte Carlo sample is still used as the default. We increase the resolution by 20% to find a systematic error due to the MC underestimating the detector resolution; this systematic is described in Sec. 10.2.3.

Track p_T range (MeV/c)	σ_{narrow} Per. Diff.	σ_{wide} Per. Diff.
Track $p_T > 200$	-15%	-8%
$200 <$ track $p_T < 400$	-7%	27%
$400 <$ track $p_T < 600$	-14%	-2%
$600 <$ track $p_T < 800$	-19%	-8%
$800 <$ track $p_T < 1000$	-17%	-24%
Track $p_T > 400$	-15%	-9%

Table 31: Percentage difference between the D^* data and D^* PYTHIA Monte Carlo resolution values for the narrow and wide Gaussians. A negative sign indicates the MC has a smaller value for the detector resolution than data.

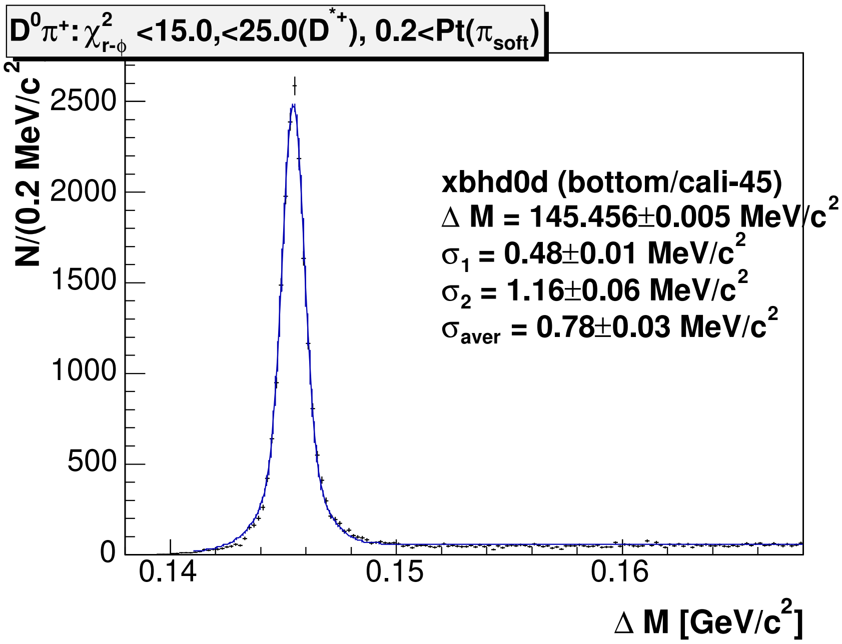


Figure 51: Fit to D^* peak using tracks with $p_\mathrm{T} > 200$ MeV/c (all tracks).

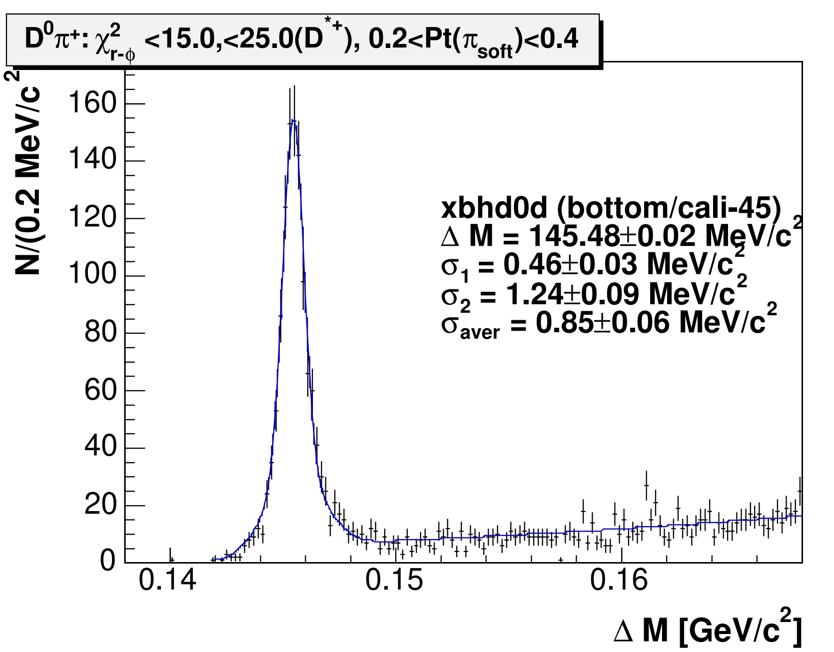


Figure 52: Fit to D^* peak using tracks with $200 < p_\mathrm{T} < 400$ MeV/c.

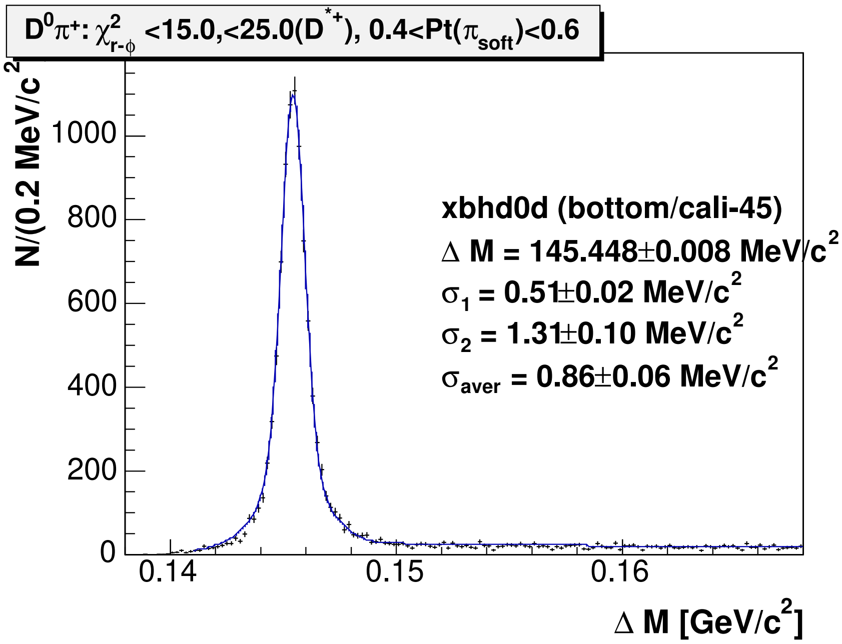


Figure 53: Fit to D^* peak using tracks with $400 < p_{\mathrm{T}} < 600$ MeV/c.

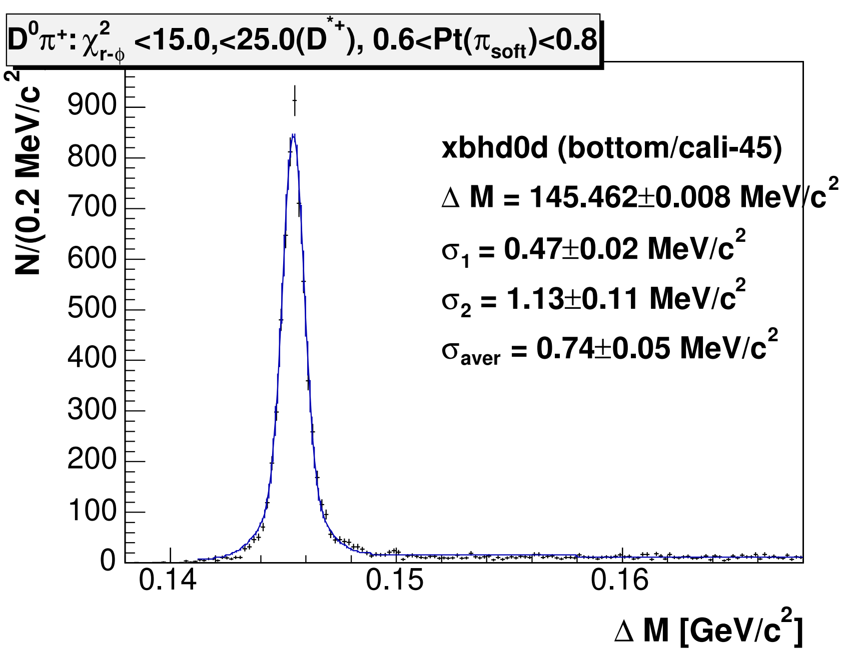


Figure 54: Fit to D^* peak using tracks with $600 < p_{\mathrm{T}} < 800$ MeV/c.

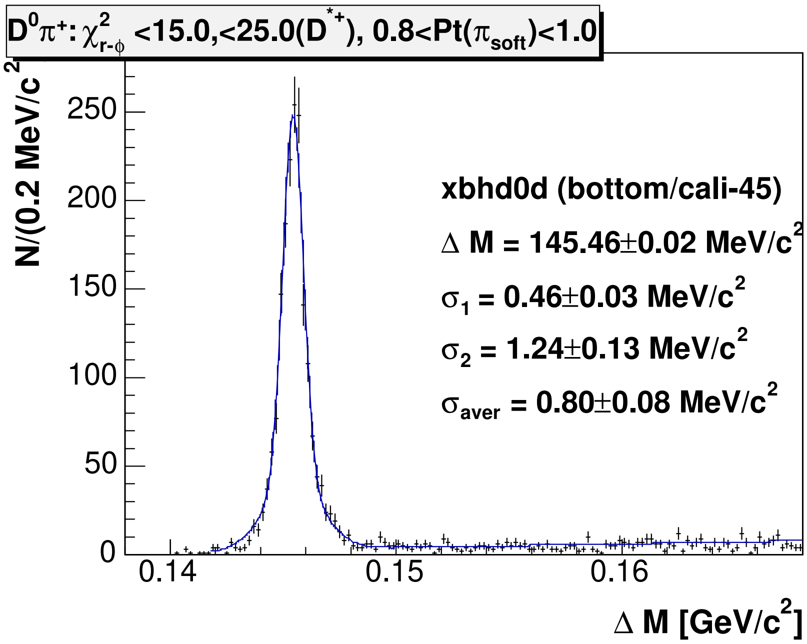


Figure 55: Fit to D^* peak using tracks with $800 < p_\mathrm{T} < 1000$ MeV/c.

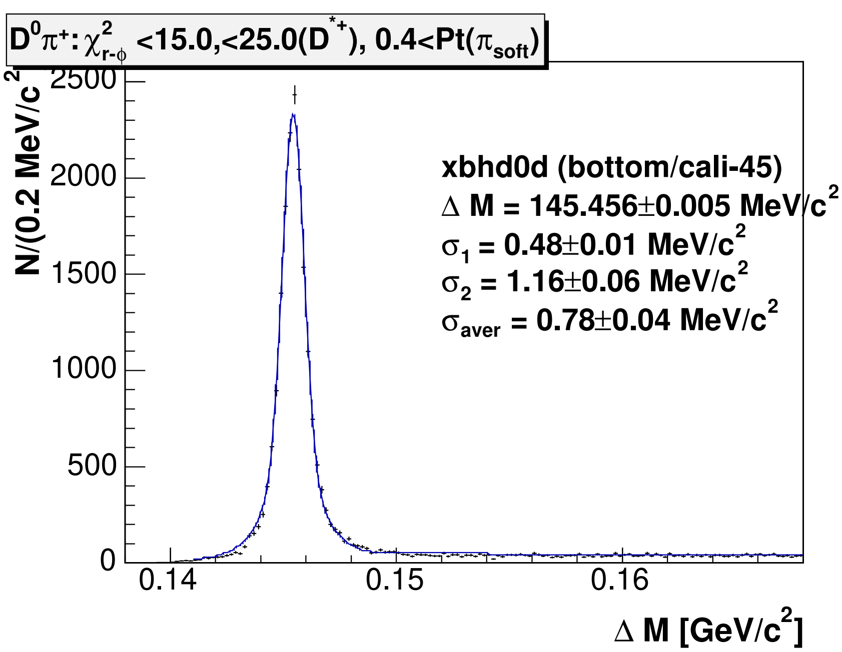


Figure 56: Fit to D^* peak using tracks with $p_\mathrm{T} > 400$ MeV/c.

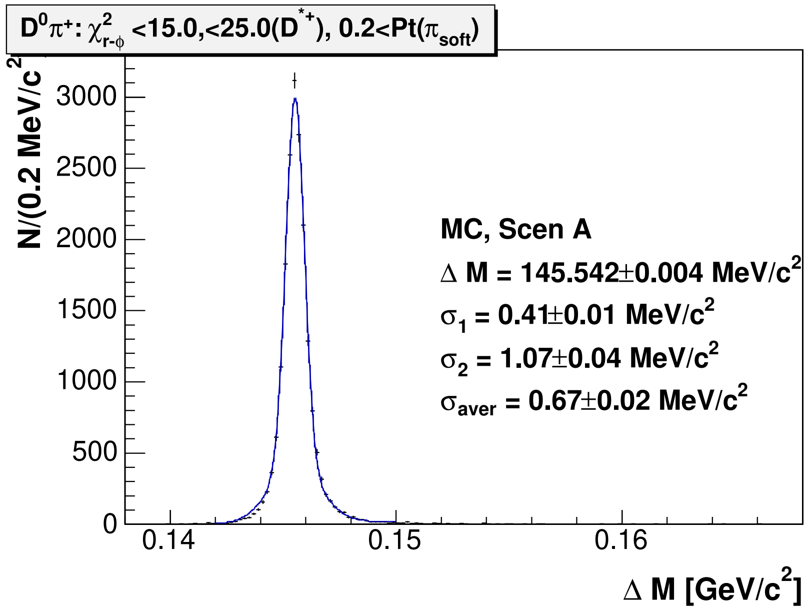


Figure 57: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $p_\mathrm{T} > 200\,\mathrm{MeV/c}$ (all tracks).

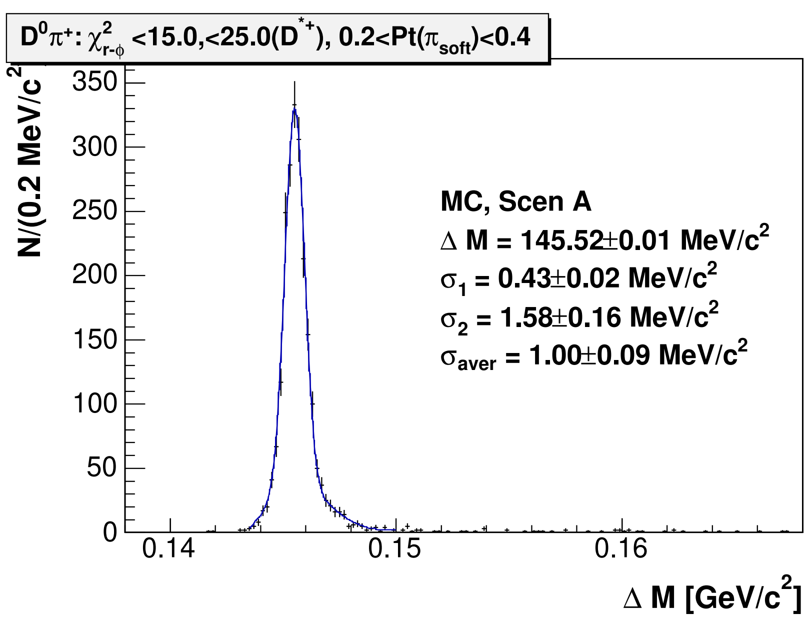


Figure 58: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $200 < p_\mathrm{T} < 400\,\mathrm{MeV/c}$.

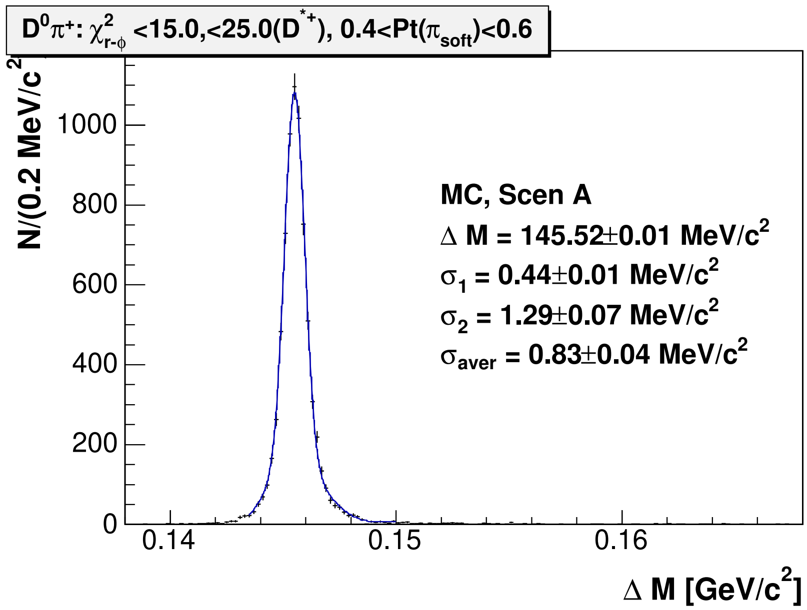


Figure 59: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $400 < p_{\mathrm{T}} < 600$ MeV/c.

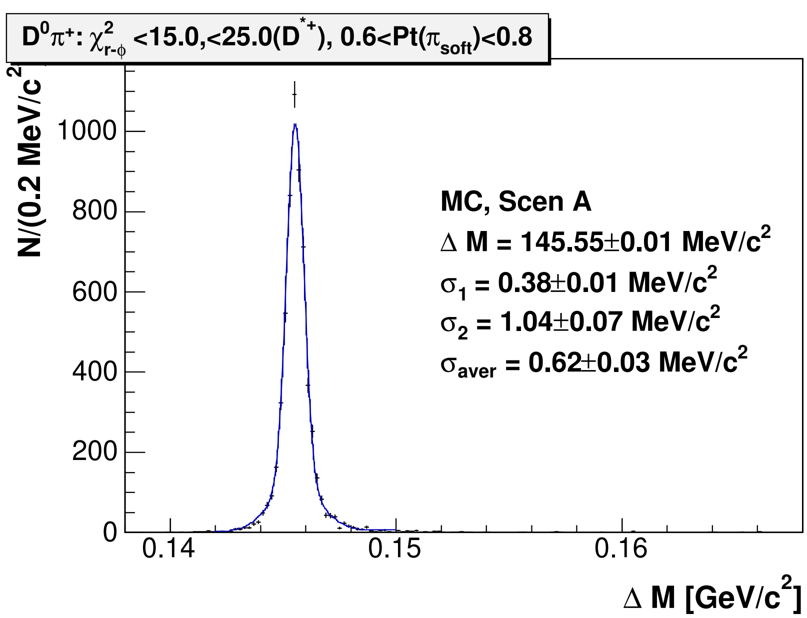


Figure 60: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $600 < p_{\mathrm{T}} < 800$ MeV/c.

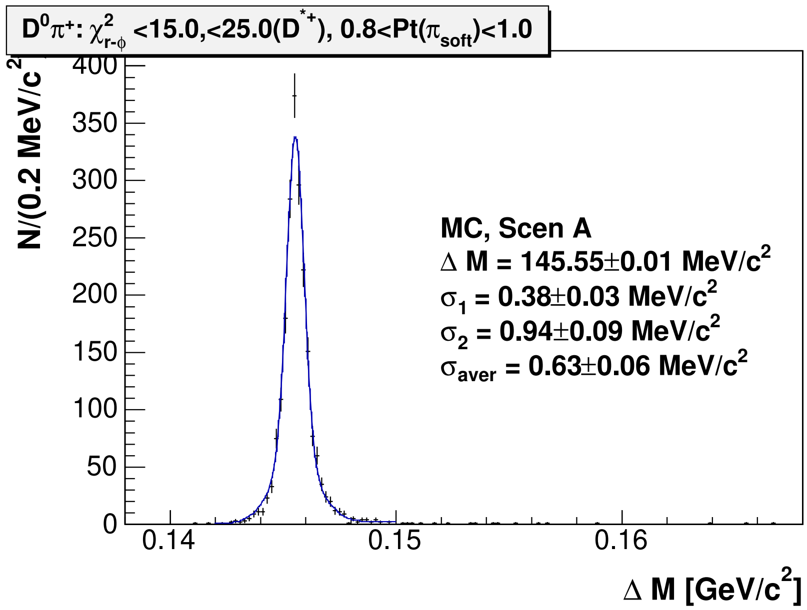


Figure 61: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $800 < p_\mathrm{T} < 1000$ MeV/c.

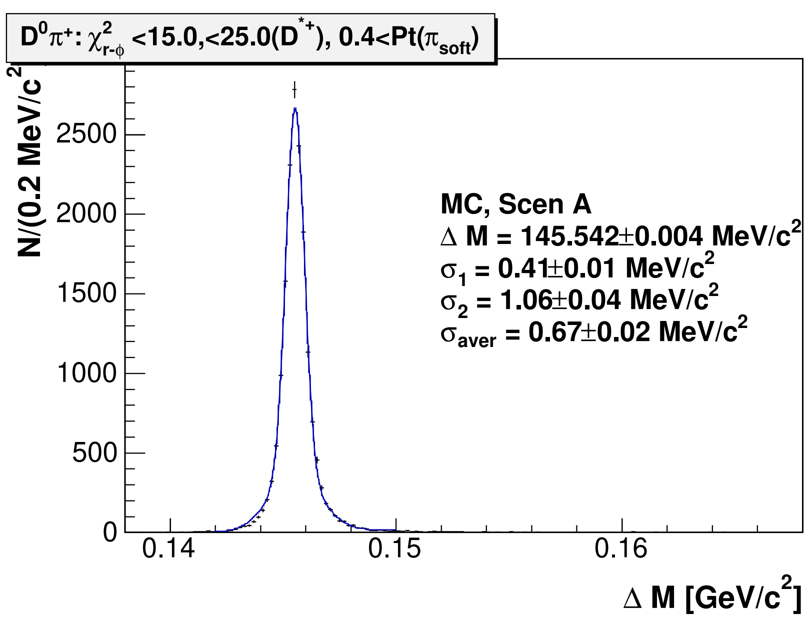


Figure 62: Fit to D^* peak in PYTHIA Monte Carlo using tracks with $p_\mathrm{T} > 400$ MeV/c.

10.1.2 Σ_c Measurements

The $\Sigma_c^{++} - \Lambda_c^+$ mass difference is 167.57 ± 0.13 MeV/c^2 while $\Sigma_c^0 - \Lambda_c^+$ mass difference is 167.29 $\pm$ 0.13 MeV/c^2 [1]. The width is much larger than for D^*; for Σ_c^{++} the width is 2.23 ± 0.30 MeV/c^2 and for Σ_c^0 the width is 2.2 ± 0.4 MeV/c^2. Thus we do not expect the width of the Σ_c to be a good measure of detector resolution. For this study, we use the datasets `xbhd0h` and `xbhd0i`, corresponding to 720 $\pm$ 40 pb^{-1} of data, also reconstructed with the `BottomMods` package [30].

Because the statistics on Σ_c are low compared to the D^* sample, we separate events into only three regions of track $p_{\rm T}$:

- Track $p_{\rm T} > 200$ MeV/c (all tracks)
- 200 MeV/c $<$ track $p_{\rm T} < 400$ MeV/c
- Track $p_{\rm T} > 400$ MeV/c

The mass difference peaks are then fit with a single Gaussian for the signal. In the case of Σ_c , tracks with $p_{\rm T}$ less than 400 MeV/c appear to have worse resolution, as evidenced in Figs. 63 - 65 and summarized in Tab. 32. However, there are very few events for tracks with $p_{\rm T} < 400$ MeV/c, so the error on the width is actually much larger as well. The width measurements between the $p_{\rm T}$ bins are still consistent within 2σ.

Track $p_{\rm T}$ range (MeV/c)	Σ_c^{++} Mean	Σ_c^{++} Width	Σ_c^0 Mean	Σ_c^0 Width
Track $p_{\rm T} > 200$	167.45	1.97 ± 0.12	167.38	1.78 ± 0.09
$200 <$ track $p_{\rm T} < 400$	167.51	2.95 ± 0.48	167.29	2.36 ± 0.37
Track $p_{\rm T} > 400$	167.48	1.78 ± 0.11	167.33	1.70 ± 0.09

Table 32: Mass and resolution parameters for the $\Sigma_c^{++} - \Lambda_c^+$ (left) and $\Sigma_c^0 - \Lambda_c^+$ (right) fits in bins of track $p_{\rm T}$. All are in units of MeV/c^2.

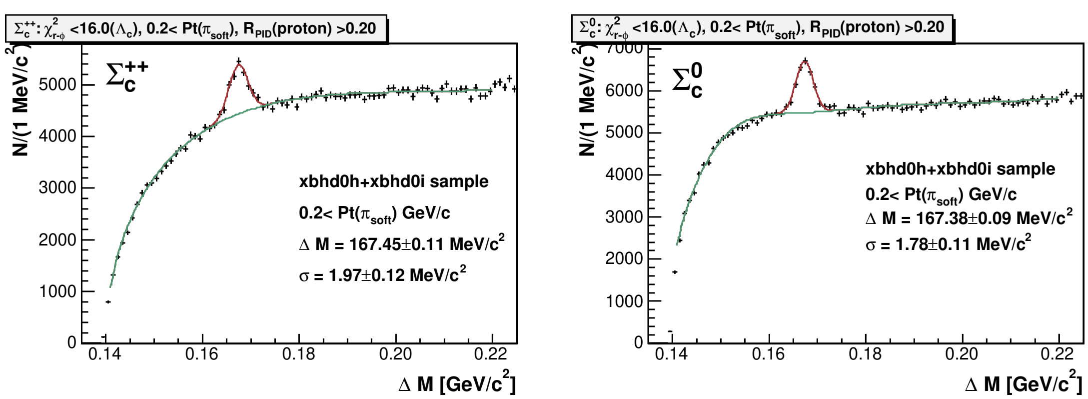


Figure 63: Fit to the $\Sigma_c^{++} - \Lambda_c^+$ (left) and $\Sigma_c^0 - \Lambda_c^+$ (right) peaks using all tracks with $p_{\rm T} > 200$ MeV/c.

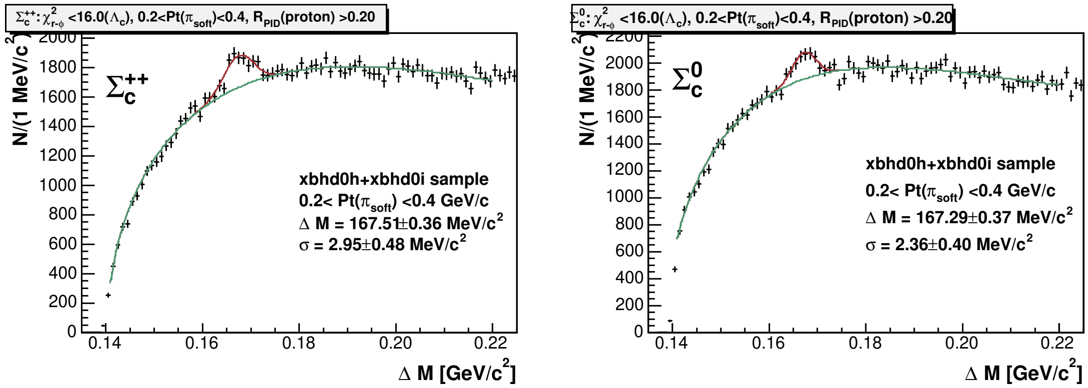


Figure 64: Fit to the $\Sigma_c^{++} - \Lambda_c^+$ (left) and $\Sigma_c^0 - \Lambda_c^+$ (right) peaks using only tracks with $p_T \in [200, 400]$ MeV/c.

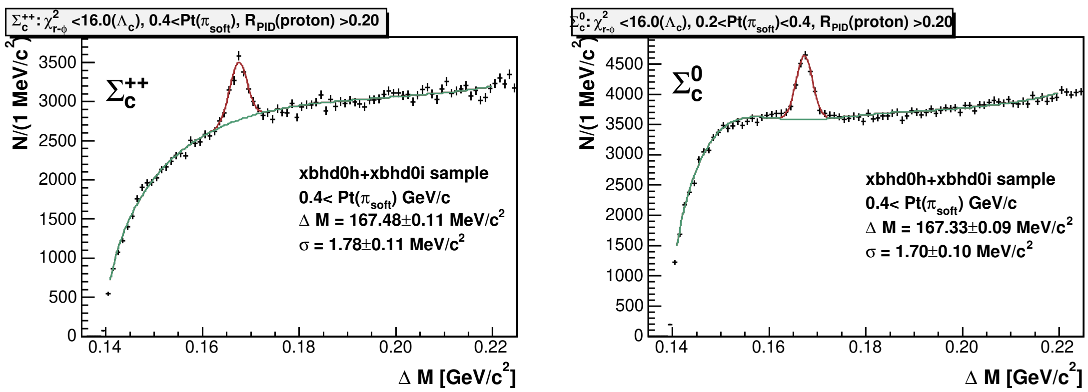


Figure 65: Fit to the $\Sigma_c^{++} - \Lambda_c^+$ (left) and $\Sigma_c^0 - \Lambda_c^+$ (right) peaks using only tracks with $p_T > 400$ MeV/c.

10.1.3 Λ_c^* Measurements

The Λ_c^{*+} states, $\Lambda_c^*(2593)$ and $\Lambda_c^*(2625)$, decay to $\Lambda_c^+\pi^+\pi^-$. The $\Lambda_c^*(2593) - \Lambda_c^+$ mass difference is 308.9 ± 0.6 MeV/c^2 while $\Lambda_c^*(2625) - \Lambda_c^+$ mass difference is 341.7 ± 0.6 MeV/c^2 [1]. The widths are again larger than for D^*; for $\Lambda_c^*(2593)$ the world average width is $3.6^{+2.0}_{-1.3}$ MeV/c^2 and for $\Lambda_c^*(2625)$ the width is < 1.9MeV/c^2. For this study, we use only the dataset `xbhd0h`, corresponding to 355 ± 40 pb^{-1} of data, also reconstructed with the `BottomMods` package [30]. Due to the low statistics, we do not separate events into different bins of track p_T .

The mass difference peaks are then fit with a single Gaussian for the signal, with the result shown in Fig. 66 and summarized in Tab. 33.

10.1.4 Determination of Mass Scale Systematic Errors

To determine the systematic error due to calibration of the energy scale, we compare the measured masses of the D^*, Σ_c , and Λ_c^* particles with the world average values quoted in the PDG. For these decays which release little kinetic energy, the figure of merit is the Q-value; this is defined

Λ_c^* State	Mean ΔM (MeV/c^2)	Width σ (MeV/c^2)	Yield	Fit Prob.
$\Lambda_c^*(2593)$	308.0 ± 0.9 (stat.)	2.6 ± 1.0 (stat.)	126 ± 42(stat.)	69%
$\Lambda_c^*(2625)$	342.0 ± 0.3 (stat.)	2.2 ± 0.3 (stat.)	420 ± 49(stat.)	69%

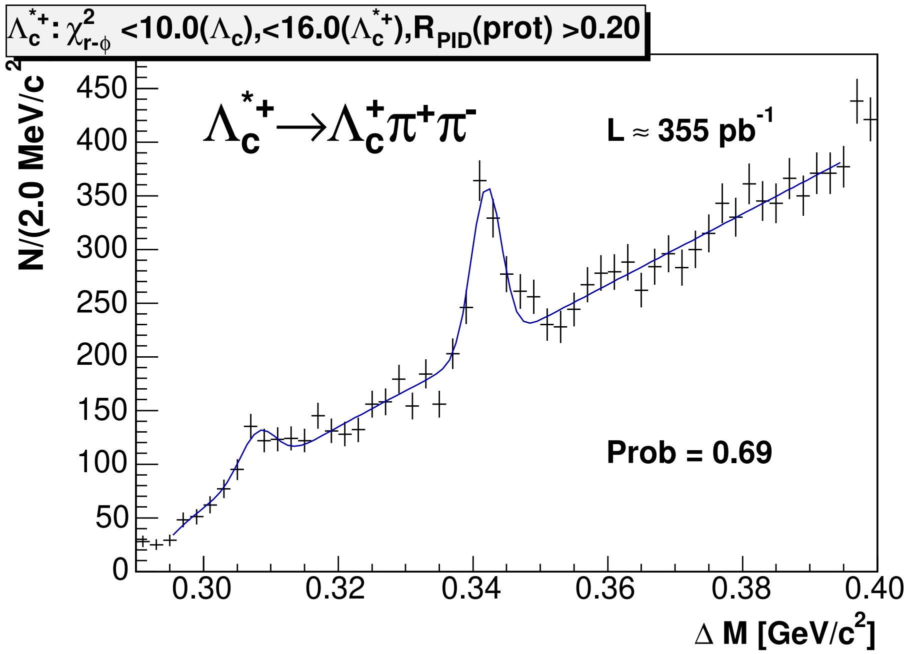
Table 33: Fit parameters for the Λ_c^* states shown in Fig. 66.

as the ΔM value less the pion mass (or two pion masses, in the case of the Λ_c^*). In a previous analysis, it has been shown that the systematic error on this Q value may be approximated as linear, $\delta Q = a \cdot Q + \delta m(Q = 0)$ [35]. We thus plot the mass differences between the CDF and PDG mass measurements as a function of the Q value of the decays, and fit the graph to a linear function. This function is then evaluated at the Σ_b Q value to give the systematic error.

In avoid correlations in the fit function, we introduce an offset of the fit variable equal to the Σ_b Q value. For example, the Σ_b^+ Q value is 48.2 MeV/c^2, so the fit takes the form $\delta Q = a \cdot (Q - 48.2) + \delta m(Q = 0)$. In this case, the intercept $\delta m(Q = 0)$ is the bias on the Q value. The mass difference graph is shown in Fig. 67 with four fits, one for each Σ_b state, and the fit parameters are given in Tab. 34. As the intercept is much smaller than its statistical error, we ignore this fit bias and take the error on the intercept as the mass scale systematic error. For the mass difference, Σ_b^* - Σ_b Q, we take the slope of the line and multiply it by the mass difference value. As the slope is smaller than its statistical error, we take the error of 0.004 MeV/c^2 and multiply it by the mass difference of 21.2 MeV/c^2. This results in a relative mass shift of ~ 0.09 MeV/c^2 which we round up to 0.1 MeV/c^2 for the mass scale systematic error on the Σ_b^* - Σ_b Q value.

Particle	Q	Slope	Intercept	Fit Prob.
Σ_b^+	48.2	-0.001 ± 0.004	-0.006 ± 0.19	58%
Σ_b^-	55.9	-0.001 ± 0.004	-0.01 ± 0.22	58%
Σ_b^{*+}	69.7	-0.001 ± 0.004	-0.03 ± 0.28	58%
Σ_b^{*-}	77.4	-0.001 ± 0.004	-0.03 ± 0.32	58%

Table 34: Fit parameters and mass scale systematic errors for each of the Σ_b states. All are in units of MeV/c^2.



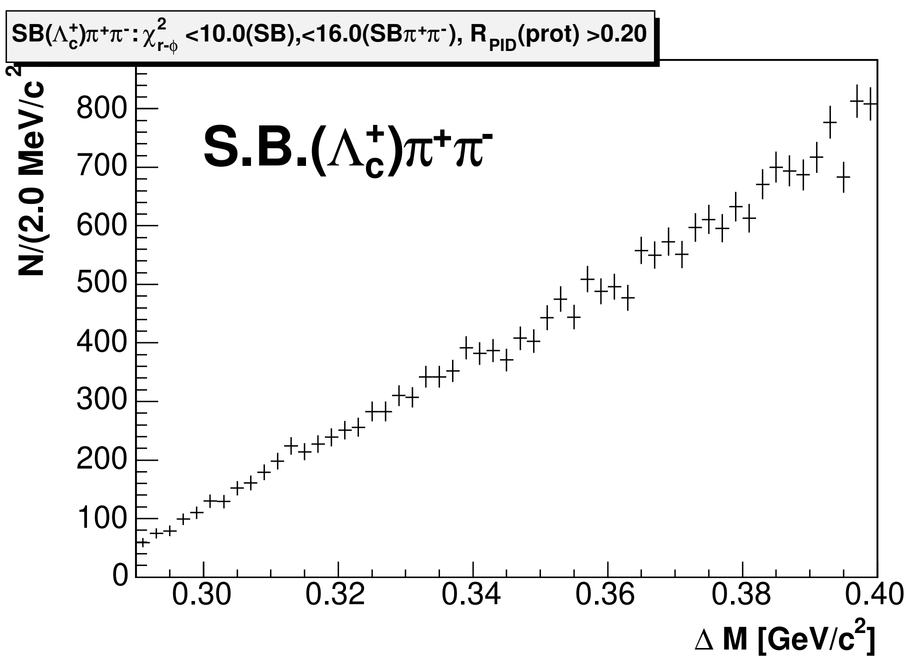


Figure 66: Top: fit to the $\Lambda_c^*(2593)$ and $\Lambda_c^*(2625)$ states. Bottom: invariant mass difference plot in the Λ_c^+ mass sideband region, where no Λ_c^* signal is seen.

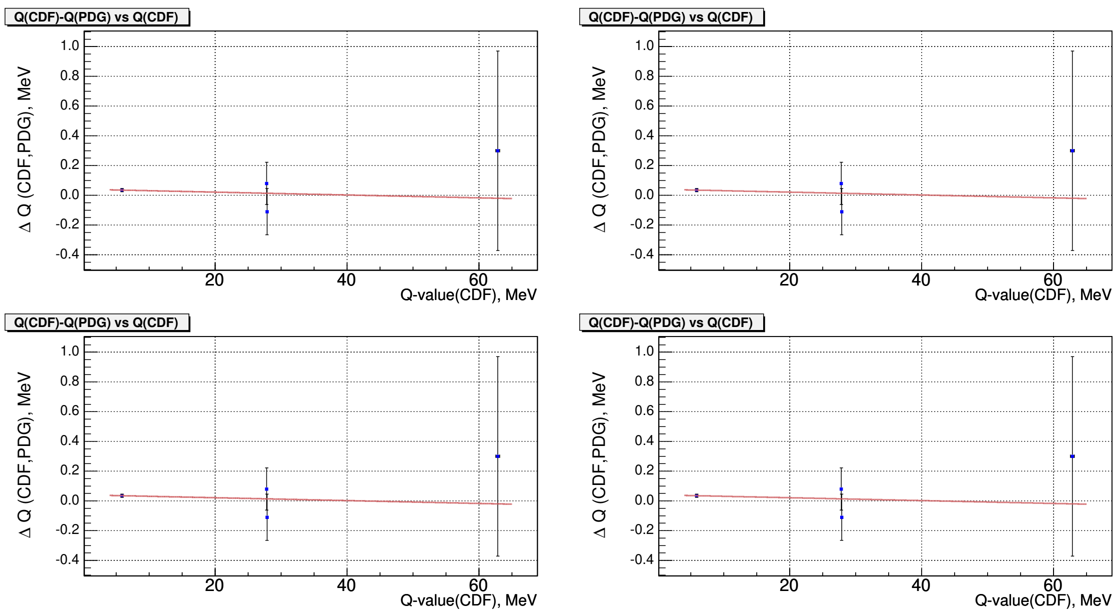


Figure 67: Graph of the D^*, Σ_c , and Λ_c^* mass differences between the CDF measurements and the PDG values, plotted versus the Q value of each decay. The graph is fitted with four linear functions, one for each Σ_b state, to determine the mass scale systematic error at each Σ_b Q value.

10.2 Fit Systematics

In addition to mass scale systematics, there are systematic errors which are built into the assumptions of the Σ_b fit. One example is the fixed shapes used for most of the Σ_b backgrounds. We evaluate these errors using the same method as was done for the B^{**} analysis. For the parameters associated with an individual systematic, we generate Toy MC samples where these parameters are varied. Then we fit the sample with both the default fit and the fit with varied parameters. For the Σ_b parameters of interest, we take the difference between their values in the varied fit and the default fit. This difference, caused by the systematic variation, constitutes the associated systematic error. After generating 500 of these Toy MC samples, we plot the difference for each sample where the fits both converged. That difference is fit with a Gaussian, and the mean of the Gaussian is taken as the systematic shift due to that particular systematic error.

For the number of $\Sigma_b^{(*)\pm}$ events, there are large systematic shifts due to the Λ_b^0 Ha + UE background shape, where the error changes predominantly in one direction. Thus we quote both positive and negative systematic errors on all quantities as well.

The following sections each describe a source of systematic error in the fit, the variations used to determine the error due to that systematic, and the error due to that systematic. We also evaluate the error on the $\Sigma_b^{*\pm}$ Q value, which is equivalent to the $\Sigma_b^{\pm} + (\Sigma_b^{*} - \Sigma_b)$ Q value, in order to quote an accurate systematic error on the absolute $\Sigma_b^{*\pm}$ masses.

10.2.1 Λ_b^0 Sample Composition

The normalizations of the three Σ_b backgrounds, described in Sec. 6, are all determined from the Λ_b^0 mass fit from Ref. [26]. There is some error associated with this mass fit, so the actual normalizations of the background components may be slightly different from what we used. To evaluate this as a systematic, we shifted the relative normalizations of the background components in the fit. As the Λ_b^0 Ha + UE background is by far the largest background, we took some number of events away from this background and added them in equal numbers to the two smaller backgrounds, the combinatorial and the B meson hadronization.

We first tried taking 100 events from Λ_b^0 Ha + UE background and adding 50 to each of the smaller backgrounds. This showed very little systematic effect, as seen in Fig. 68. We then subtracted 200 events from Λ_b^0 Ha + UE background and adding 100 to each of the smaller backgrounds. This also had a fairly small effect, as seen in Fig. 69. Finally, we took 400 events from Λ_b^0 Ha + UE background and adding 200 to each of the smaller backgrounds, more than doubling the normalizations of the two smaller backgrounds. The systematic plots for this are shown in Fig. 70. The systematic shifts for each of the seven floating parameters in each case are shown in Tab. 35. The errors for the third case, where 400 events are removed from the Λ_b^0 Ha + UE background, will be taken as the systematic error due to Λ_b^0 sample composition because they are the largest. This is a relatively small systematic error, indicating that our result is not sensitive to the Λ_b^0 sample composition and is thus primarily independent of the Λ_b^0 mass fit.

Parameter	100 events	200 events	400 events
Σ_b^- Q (MeV/c^2)	-0.009 ± 0.001	-0.014 ± 0.001	-0.029 ± 0.002
Σ_b^- events	0.143 ± 0.004	0.322 ± 0.004	0.68 ± 0.01
Σ_b^+ Q (MeV/c^2)	0.007 ± 0.001	0.014 ± 0.002	0.027 ± 0.003
Σ_b^+ events	0.84 ± 0.01	1.61 ± 0.02	3.30 ± 0.04
Σ_b^{*-} events	0.099 ± 0.004	0.21 ± 0.01	0.39 ± 0.01
Σ_b^{*+} events	1.83 ± 0.01	3.66 ± 0.02	7.28 ± 0.03
Σ_b^* - Σ_b Q (MeV/c^2)	0.015 ± 0.001	0.029 ± 0.002	0.052 ± 0.005
Σ_b^{*-} Q (MeV/c^2)	0.007 ± 0.001	0.014 ± 0.002	0.017 ± 0.003
Σ_b^{*+} Q (MeV/c^2)	0.023 ± 0.001	0.043 ± 0.002	0.089 ± 0.003

Table 35: Systematic errors on each parameter due to shifting a given number of events from the Λ_b^0 Ha + UE background to the two smaller backgrounds.

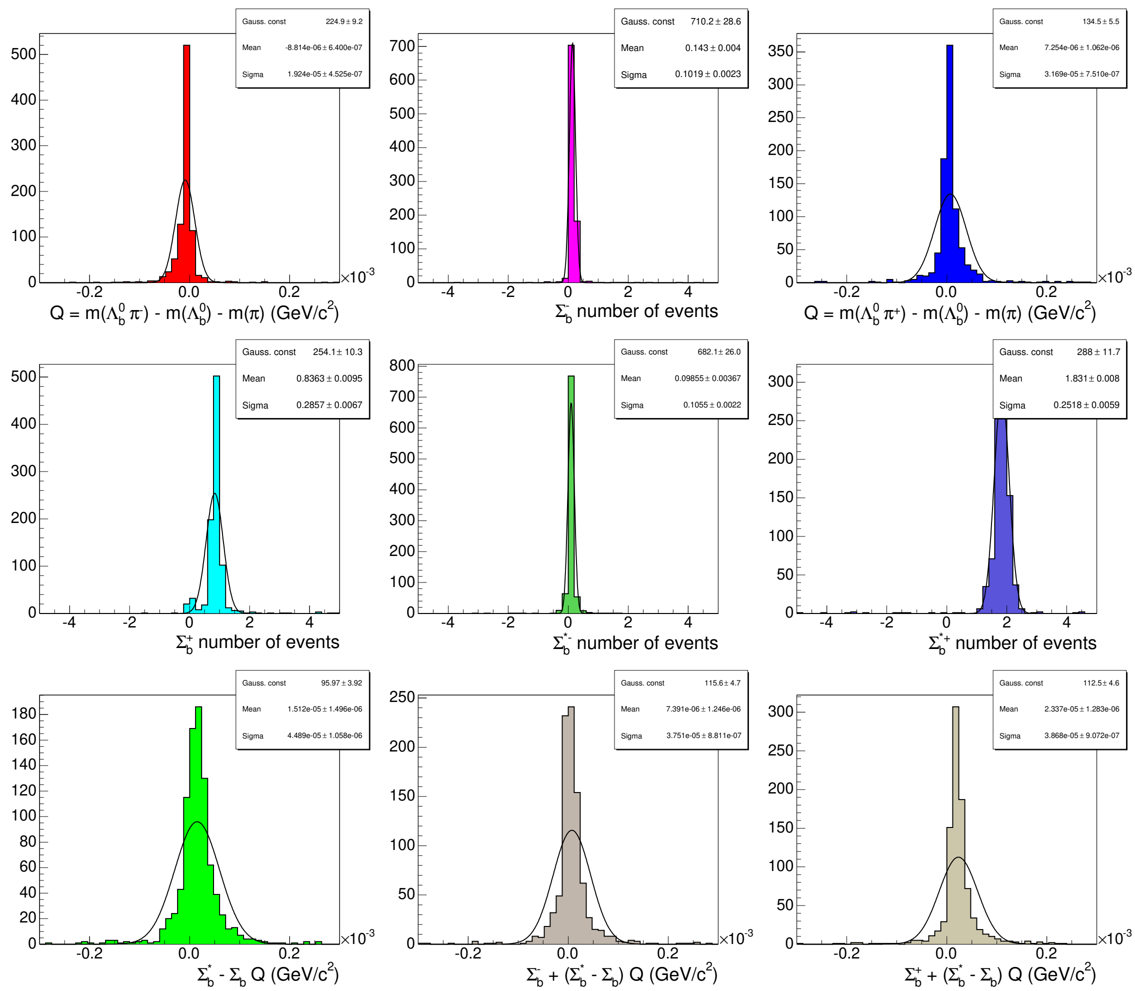


Figure 68: Plots of the systematic shift in each floating parameter for Toy MC samples where 100 events have been transferred from the Λ_b^0 Ha + UE background to the two smaller backgrounds.

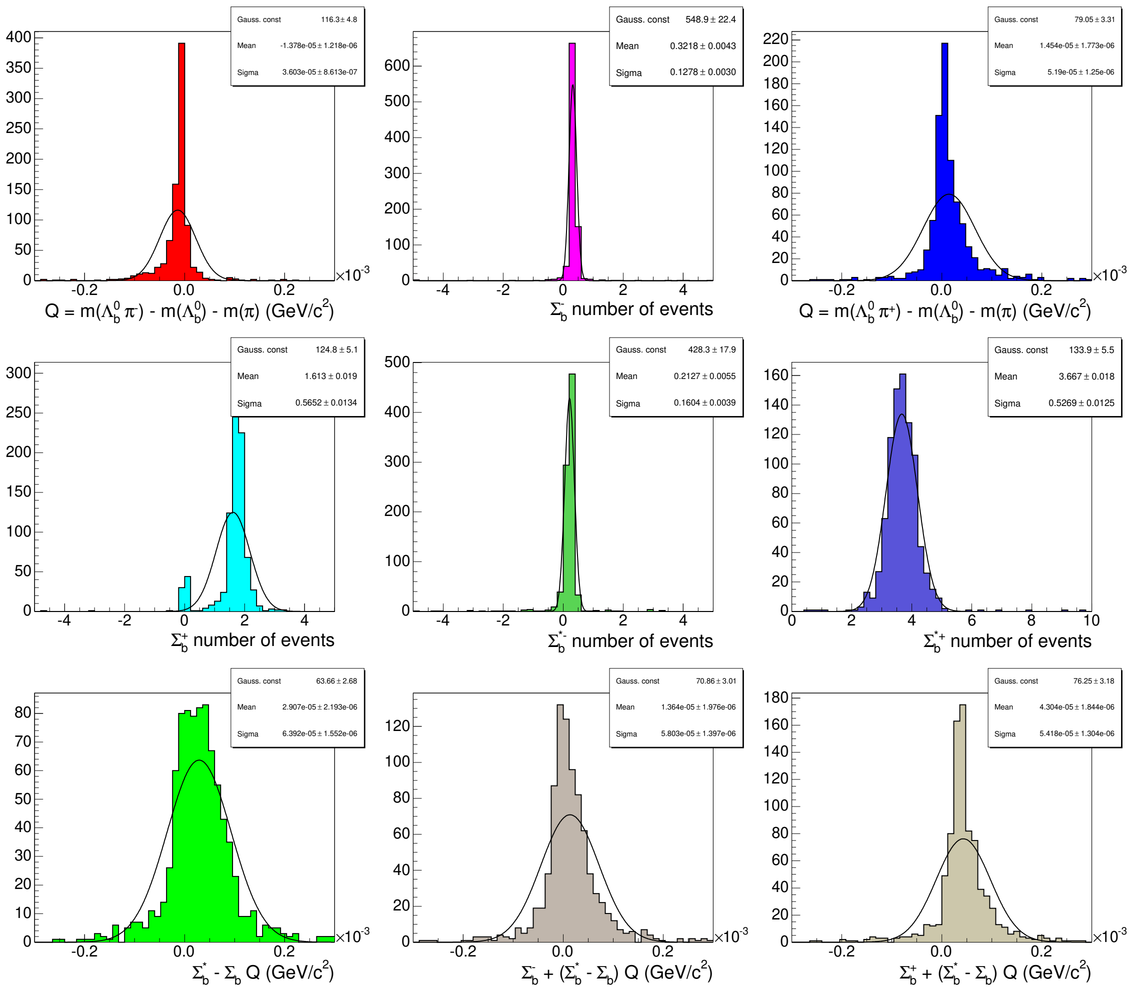


Figure 69: Plots of the systematic shift in each floating parameter for Toy MC samples where 200 events have been transferred from the Λ_b^0 Ha + UE background to the two smaller backgrounds.

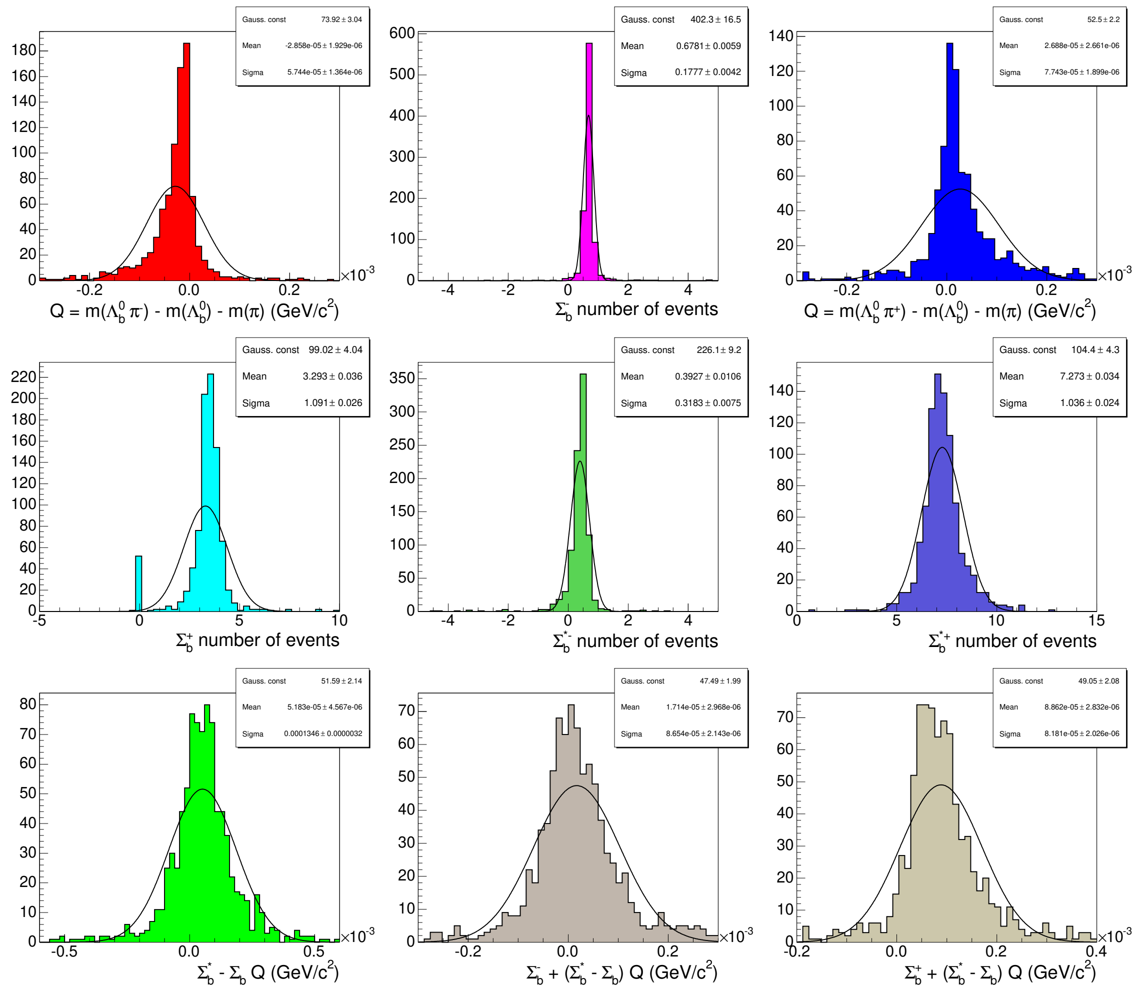


Figure 70: Plots of the systematic shift in each floating parameter for Toy MC samples where 400 events have been transferred from the Λ_b^0 Ha + UE background to the two smaller backgrounds.

10.2.2 Λ_b^0 Hadronization Background

As explained in Sec. 9.3, there are several systematics associated with the derivation of the Λ_b^0 Ha + UE background shape. The evaluation of these systematics is detailed below.

The first is the normalization of the derived shape. This normalization comes from the ratio of the number of Λ_b^0 in the MC sample to the number of Λ_b^0 in the data sample. The number of Λ_b^0 in data is given by the Λ_b^0 mass fit, which we already tested with the Λ_b^0 sample composition systematic. However, if the number of Λ_b^0 in MC is off even after reweighting the Λ_b^0 p_T spectrum, the normalization could also be incorrect. To test this, we kept the shape of the Λ_b^0 Ha + UE background fixed but shifted the number of events. From the fit to the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ backgrounds, we found the number of events to be 1409 ± 43 and 1379 ± 42 respectively. For the Toy MC samples, we generated with both $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ Λ_b^0 Ha + UE backgrounds either increased by 43 events or decreased by 43 events. The resulting systematic shifts are shown in Figs. 71 and 72, and given in Tab. 36. The effect on the mass measurements is minimal, but there is some systematic shift of the number of events. We take the largest value as the systematic on each parameter.

Parameter	Λ_b^0 Ha + UE Norm. - 43	Λ_b^0 Ha + UE Norm. + 43
Σ_b^- Q (MeV/c^2)	-0.002 ± 0.001	0.009 ± 0.002
Σ_b^- events	2.23 ± 0.01	-2.20 ± 0.01
Σ_b^+ Q (MeV/c^2)	-0.013 ± 0.002	0.013 ± 0.002
Σ_b^+ events	2.07 ± 0.01	-2.09 ± 0.01
Σ_b^{*-} events	4.77 ± 0.02	-4.72 ± 0.03
Σ_b^{*+} events	4.81 ± 0.02	-4.80 ± 0.03
Σ_b^* - Σ_b Q (MeV/c^2)	0.135 ± 0.003	-0.129 ± 0.002
Σ_b^{*-} Q (MeV/c^2)	0.133 ± 0.003	-0.133 ± 0.003
Σ_b^{*+} Q (MeV/c^2)	0.122 ± 0.003	-0.128 ± 0.003

Table 36: Systematic errors on each parameter due to changing the number of events in the Λ_b^0 Ha + UE background by the given amount.

A second systematic comes more directly from the shape used to fit the Λ_b^0 Ha + UE background; an alternate shape is described in Sec. 9.3.2. To calculate this shape systematic, we generate Toy MC samples with the $D^* - D^0$ background and then fit with the default background shape. The systematic shifts are shown in Fig. 73 and given in Tab. 37. As expected, there is very little shift in the Q measurements and a small effect on the number of events.

The third and final systematic relating to the Λ_b^0 Ha + UE background is the effect of reweighting the PYTHIA Monte Carlo track p_T spectrum to agree with data. The procedure for getting the "Reweighted Down" and "Reweighted Up" shapes has been described in Sec. 6.5, and the fit of these alternate shapes to data is shown in Sec. 9.3.4. To estimate the systematic shift associated with this reweighting, we took the background shape from Tab 11 and the normalization from the fit to data, and generated Toy MC using this as the background shape. The resulting systematic shifts are shown in Figs. 74 and 75, and summarized in Tab. 38. As expected, this is one of the dominant systematic errors.

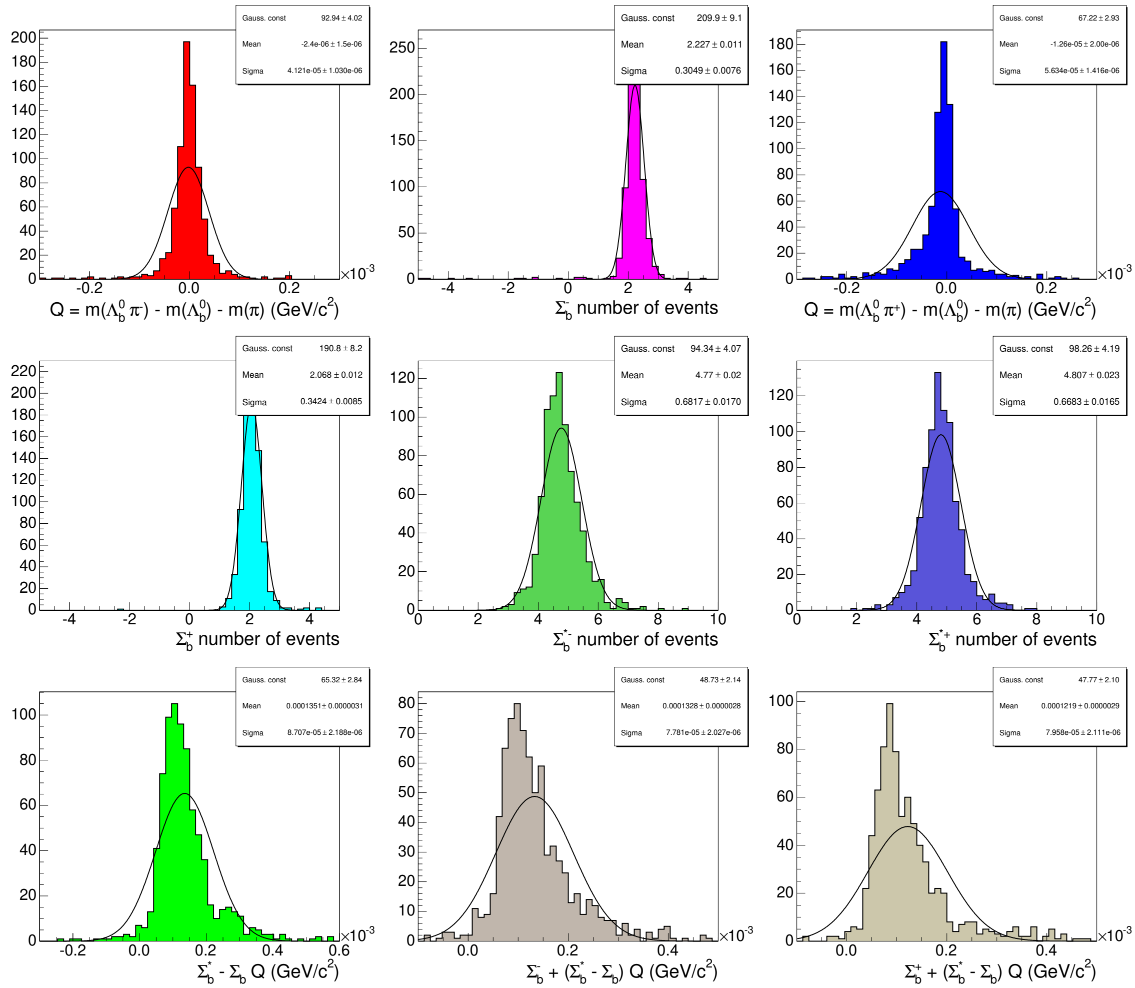


Figure 71: Plots of the systematic shift in each floating parameter for Toy MC samples where the Λ_b^0 Ha + UE background normalization has been decreased by 43 events.

Parameter	Systematic Shift
Σ_b^- Q (MeV/c^2)	-0.011 ± 0.001
Σ_b^- events	0.268 ± 0.004
Σ_b^+ Q (MeV/c^2)	0.013 ± 0.001
Σ_b^+ events	1.16 ± 0.01
Σ_b^{*-} events	0.326 ± 0.004
Σ_b^{*+} events	2.76 ± 0.02
Σ_b^* - Σ_b Q (MeV/c^2)	0.038 ± 0.002
Σ_b^{*-} Q (MeV/c^2)	0.029 ± 0.002
Σ_b^{*+} Q (MeV/c^2)	0.053 ± 0.002

Table 37: Systematic errors on each parameter due to using a different shape for the Λ_b^0 Ha + UE background.

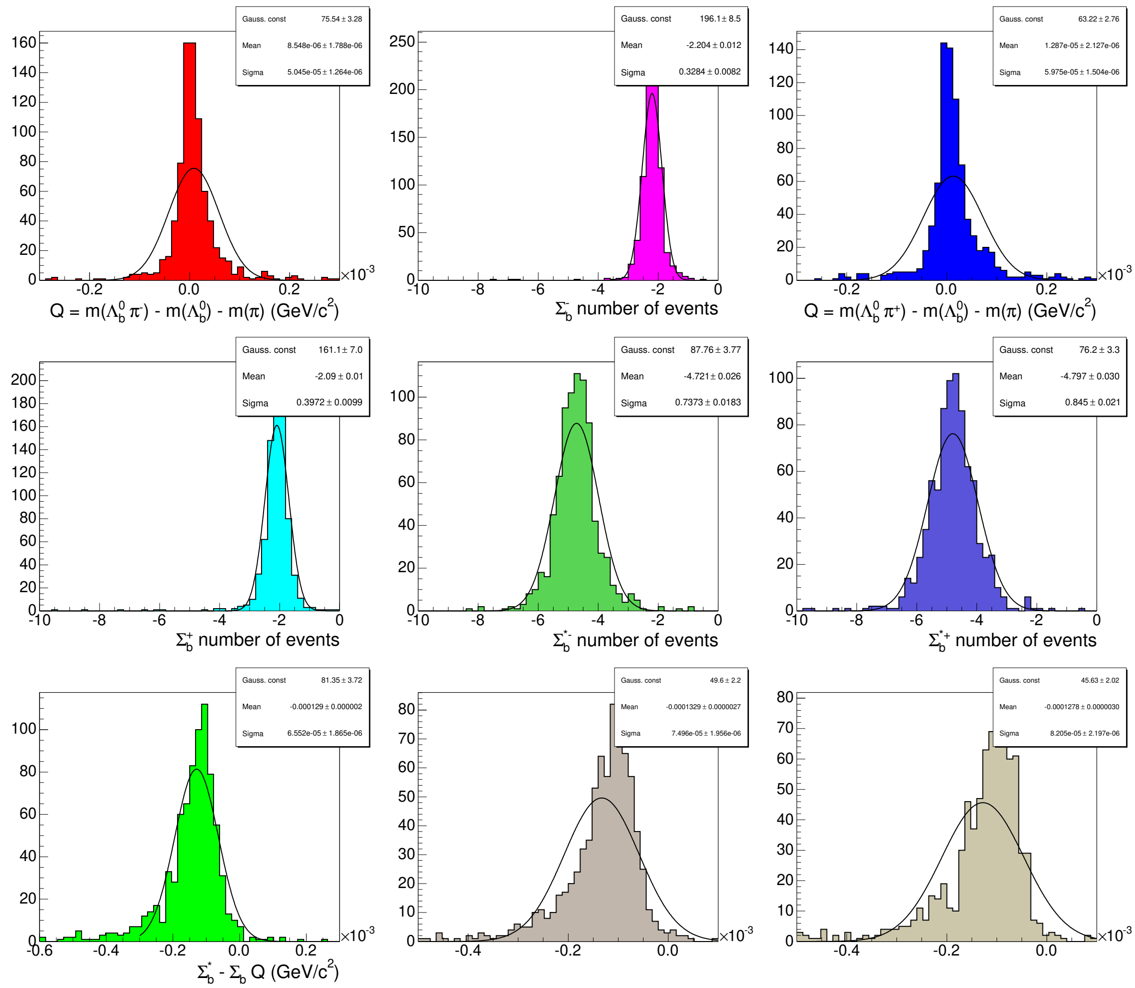


Figure 72: Plots of the systematic shift in each floating parameter for Toy MC samples where the Λ_b^0 Ha + UE background normalization has been increased by 43 events.

Parameter	Λ_b^0 Ha + UE Reweight. Down	Λ_b^0 Ha + UE Reweight. Up
Σ_b^- Q (MeV/c^2)	-0.0004 ± 0.002	0.038 ± 0.001
Σ_b^- events	7.38 ± 0.02	1.84 ± 0.01
Σ_b^+ Q (MeV/c^2)	-0.112 ± 0.004	-0.050 ± 0.003
Σ_b^+ events	2.32 ± 0.01	-1.78 ± 0.01
Σ_b^{*-} events	14.7 ± 0.04	5.23 ± 0.02
Σ_b^{*+} events	4.58 ± 0.01	-2.88 ± 0.01
Σ_b^* - Σ_b Q (MeV/c^2)	0.314 ± 0.006	0.078 ± 0.004
Σ_b^{*-} Q (MeV/c^2)	0.317 ± 0.005	0.117 ± 0.004
Σ_b^{*+} Q (MeV/c^2)	0.166 ± 0.004	0.022 ± 0.002

Table 38: Systematic errors on each parameter due to changes in the reweighting of the PYTHIA Monte Carlo track p_T spectrum.

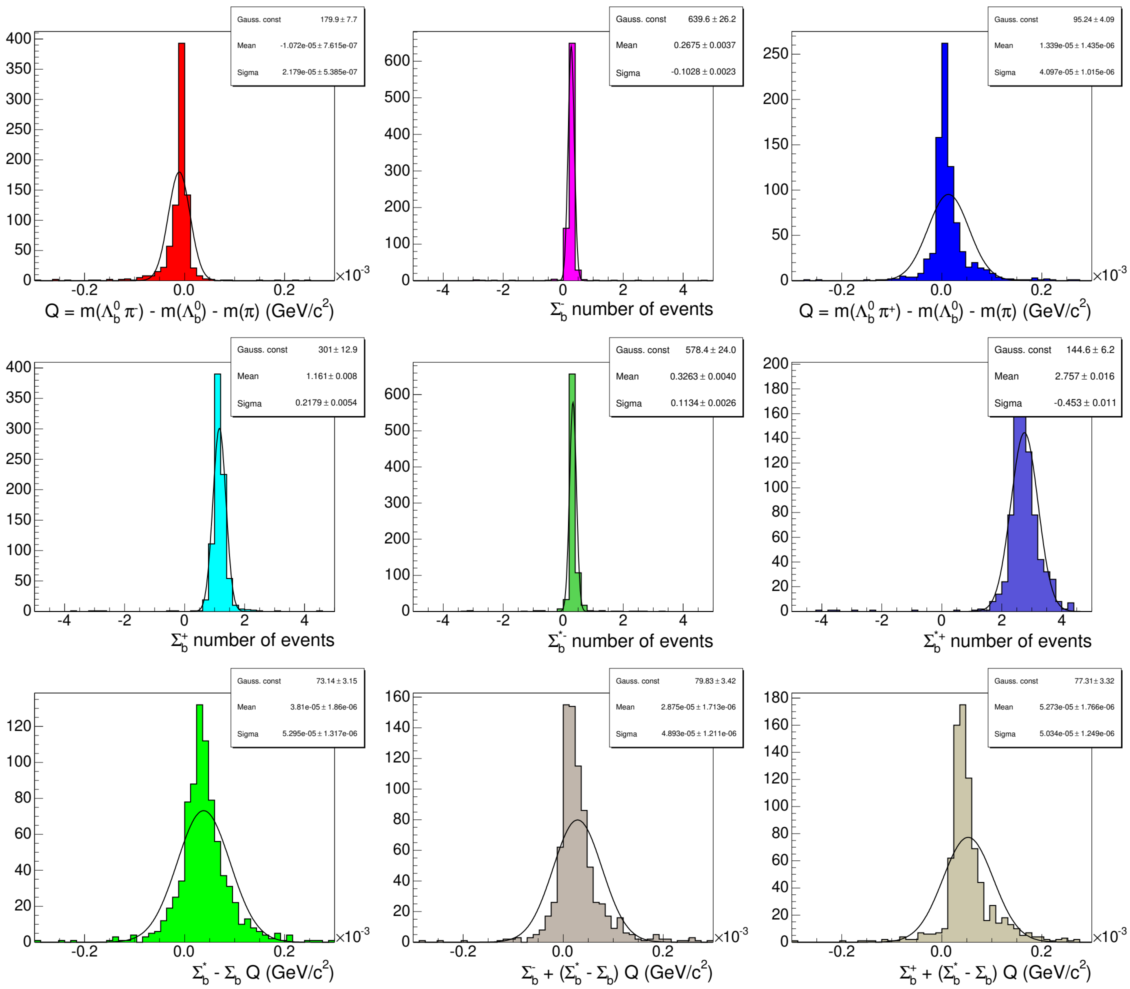


Figure 73: Plots of the systematic shift in each floating parameter for Toy MC samples where the Λ_b^0 Ha + UE background was generated with the RooFit $D^* - D^0$ background shape.

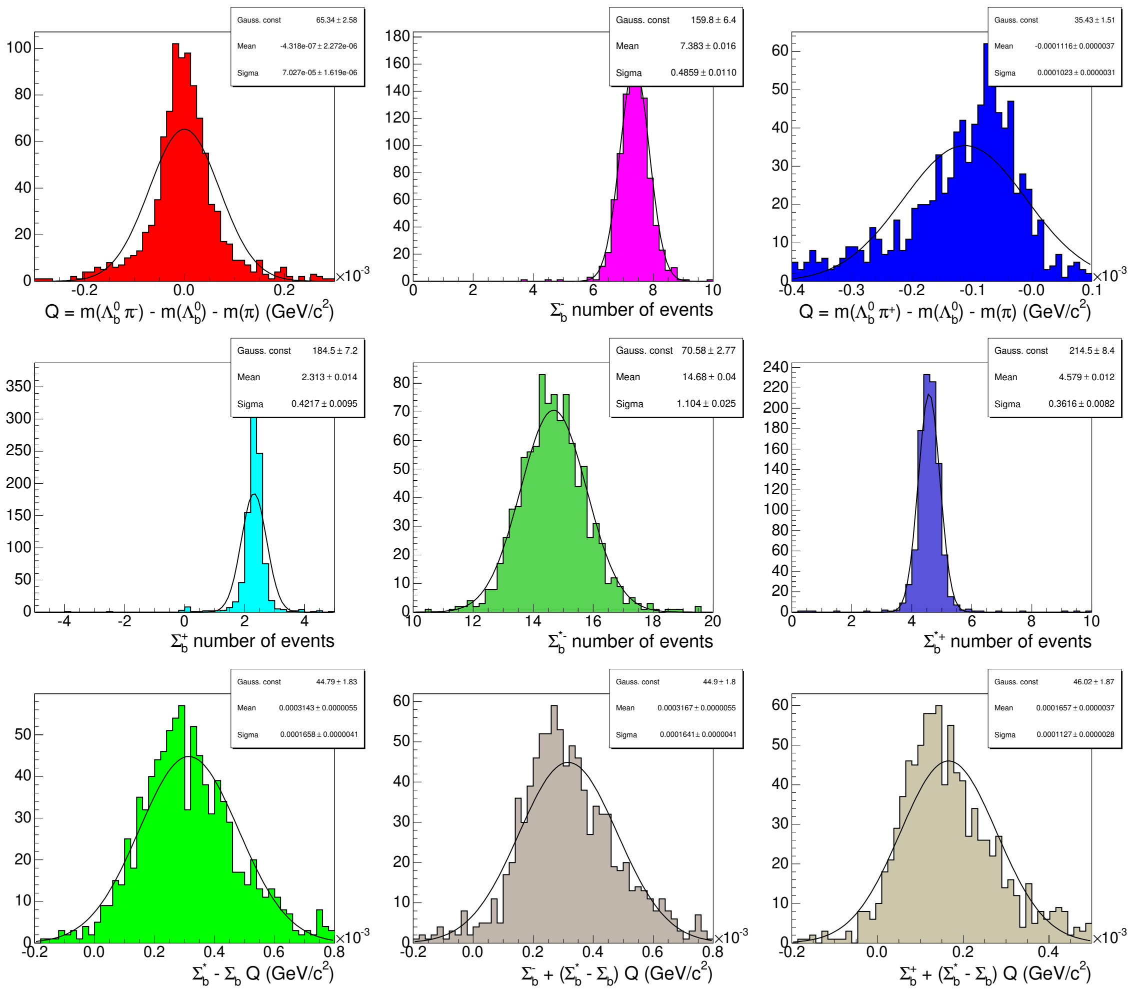


Figure 74: Plots of the systematic shift in each floating parameter for Toy MC samples where the soft tracks of the Λ_b^0 Ha + UE background have been weighted down.

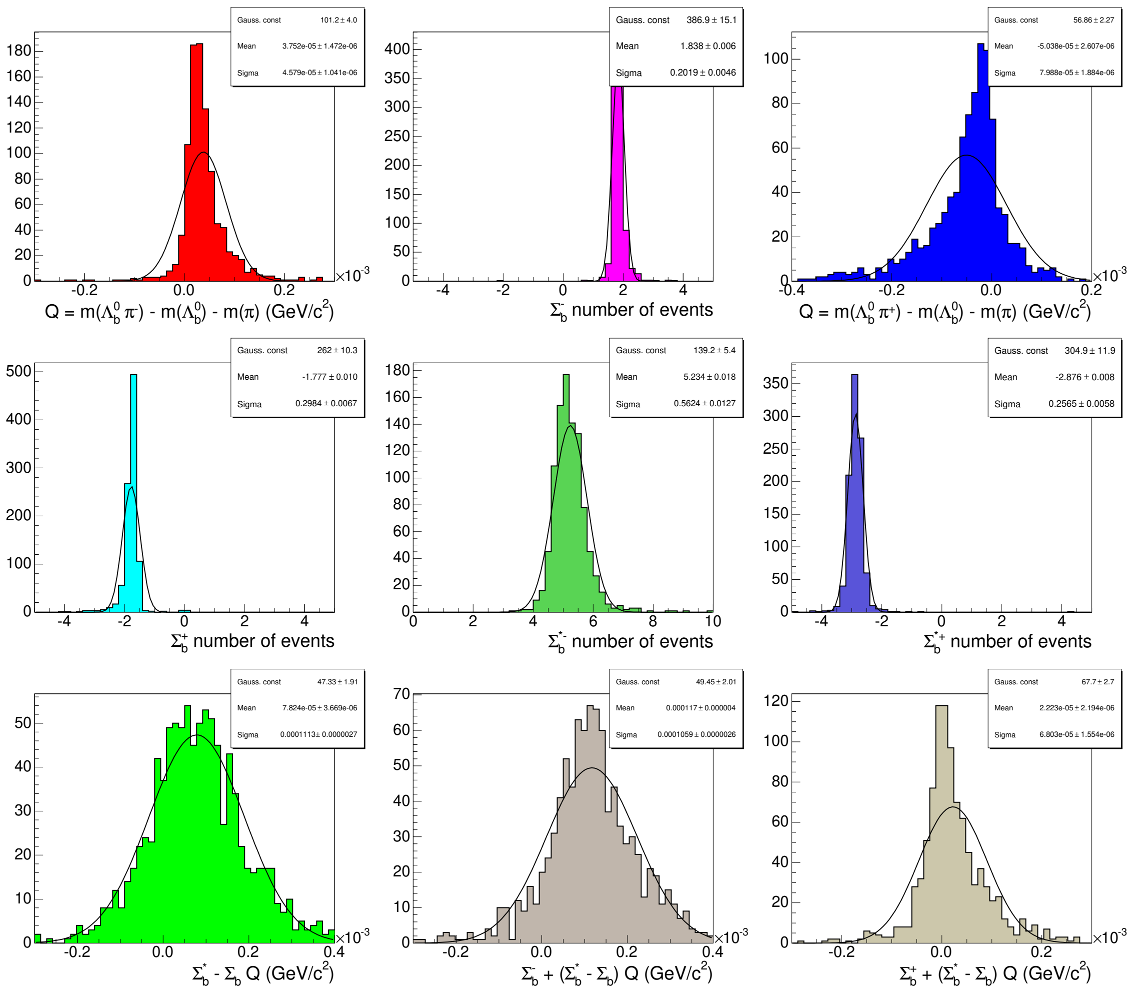


Figure 75: Plots of the systematic shift in each floating parameter for Toy MC samples where the soft tracks of the Λ_b^0 Ha + UE background have been weighted up.

10.2.3 Detector Resolution

The detector resolution is described by two Gaussians of fixed width. This resolution was taking from a realistic MC sample, as described in Sec. 7.1. For the low momentum tracks, the MC may underestimate the detector resolution. We account for this systematic by generating Toy MC samples with the detector resolution widths increased by 20% (σ_{narrow} = 1.4 MeV/c^2 and σ_{wide} = 3.6 MeV/c^2). The effects on the floating parameters are shown in Fig. 76. The systematic shifts are summarized in Table 39. The effect of this systematic is very small.

Parameter	Systematic Shift
Σ_b^- Q (MeV/c^2)	-0.011 ± 0.002
Σ_b^- events	0.34 ± 0.01
Σ_b^+ Q (MeV/c^2)	-0.014 ± 0.003
Σ_b^+ events	0.25 ± 0.01
Σ_b^{*-} events	0.08 ± 0.01
Σ_b^{*+} events	0.21 ± 0.01
Σ_b^* - Σ_b Q (MeV/c^2)	0.016 ± 0.003
Σ_b^{*-} Q (MeV/c^2)	0.003 ± 0.002
Σ_b^{*+} Q (MeV/c^2)	0.001 ± 0.002

Table 39: Systematic errors on each parameter due to underestimating the detector resolution.

10.2.4 Σ_b Natural Width

The width of each Σ_b peak is calculated from its mean Q value using Eq (1). This equation depends on a parameter g_A; from the fit to the Σ_c width, this parameter is measured to be

$$g_A = 0.75 \pm 0.05$$

To incorporate the error on g_A, we substituted the upper and lower values (g_A = 0.70 and g_A = 0.80) into the natural width equation and ran Toy MC samples from that. The resulting systematic distributions are shown in Fig. 77 for g_A = 0.70 and Fig. 78 for g_A = 0.80. The mean values are given in Tab. 40 for both cases. For every parameter, we see that the systematic shift changes sign when the value of g_A changes from low to high, as expected. As the final value for this systematic, we take whichever shift is largest.

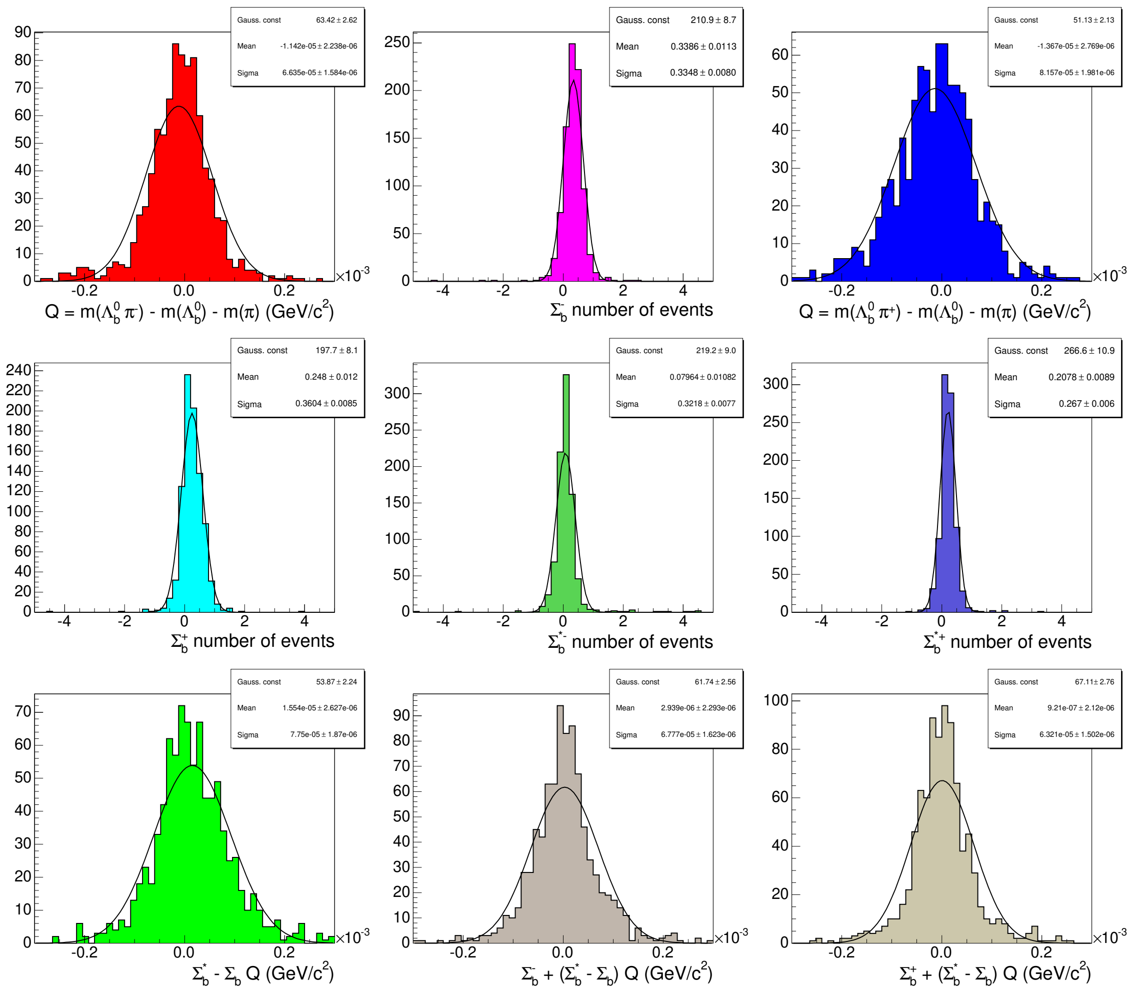


Figure 76: Plots of the systematic shift in each floating parameter for Toy MC samples where the detector resolution widths have been increased by 20%.

Parameter	$g_A = 0.70$	$g_A = 0.80$
Σ_b^- Q (MeV/c^2)	-0.005 ± 0.004	0.009 ± 0.005
Σ_b^- events	-3.44 ± 0.06	3.36 ± 0.07
Σ_b^+ Q (MeV/c^2)	0.012 ± 0.004	-0.021 ± 0.005
Σ_b^+ events	-1.97 ± 0.05	1.80 ± 0.05
Σ_b^{*-} events	1.65 ± 0.04	-1.67 ± 0.04
Σ_b^{*+} events	0.82 ± 0.03	-0.79 ± 0.03
Σ_b^* - Σ_b Q (MeV/c^2)	-0.074 ± 0.004	0.072 ± 0.004
Σ_b^{*-} Q (MeV/c^2)	-0.073 ± 0.004	0.082 ± 0.005
Σ_b^{*+} Q (MeV/c^2)	-0.064 ± 0.003	0.053 ± 0.003

Table 40: Systematic errors on each parameter due to modifying the natural width calculation.

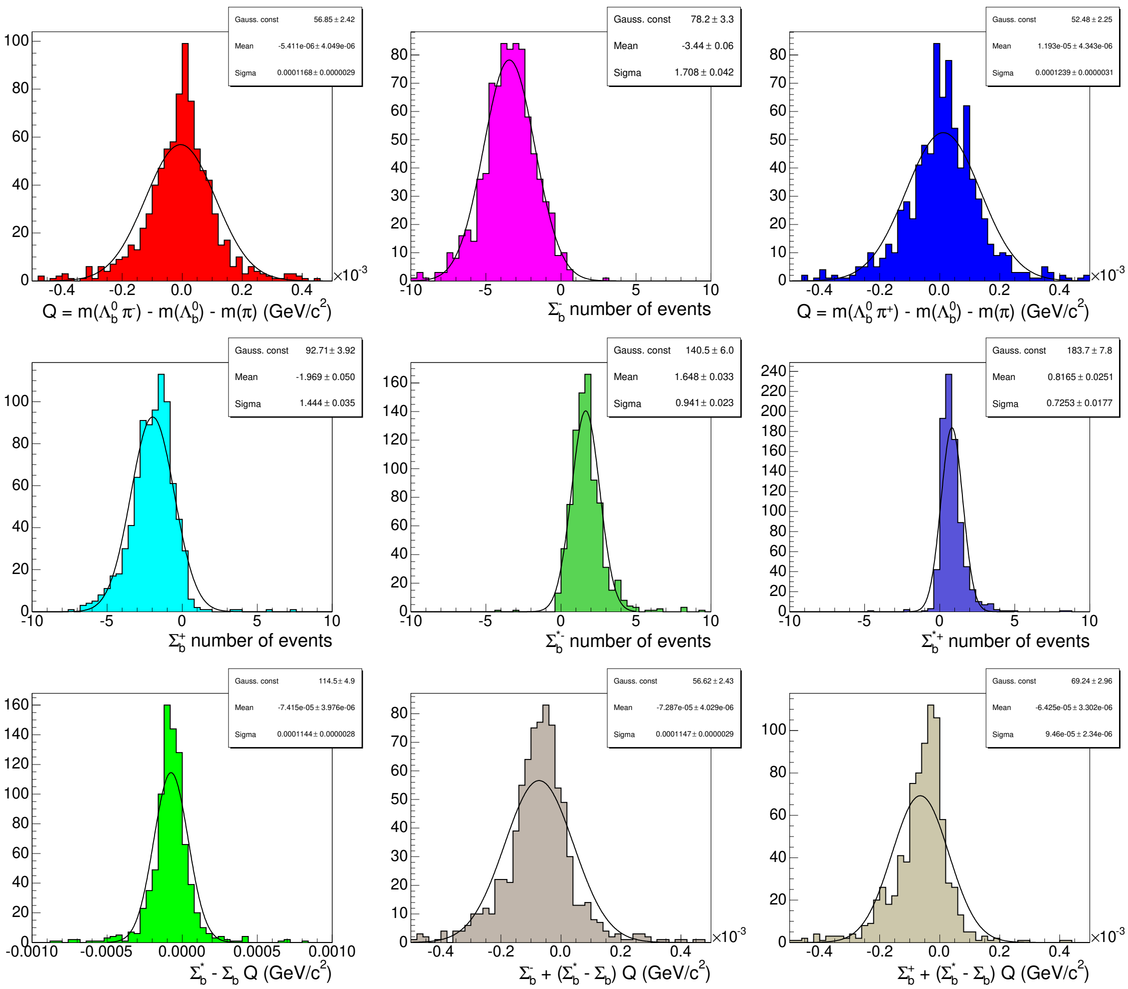


Figure 77: Plots of the systematic shift in each floating parameter for Toy MC samples where the width parameter g_A is set to 0.70.

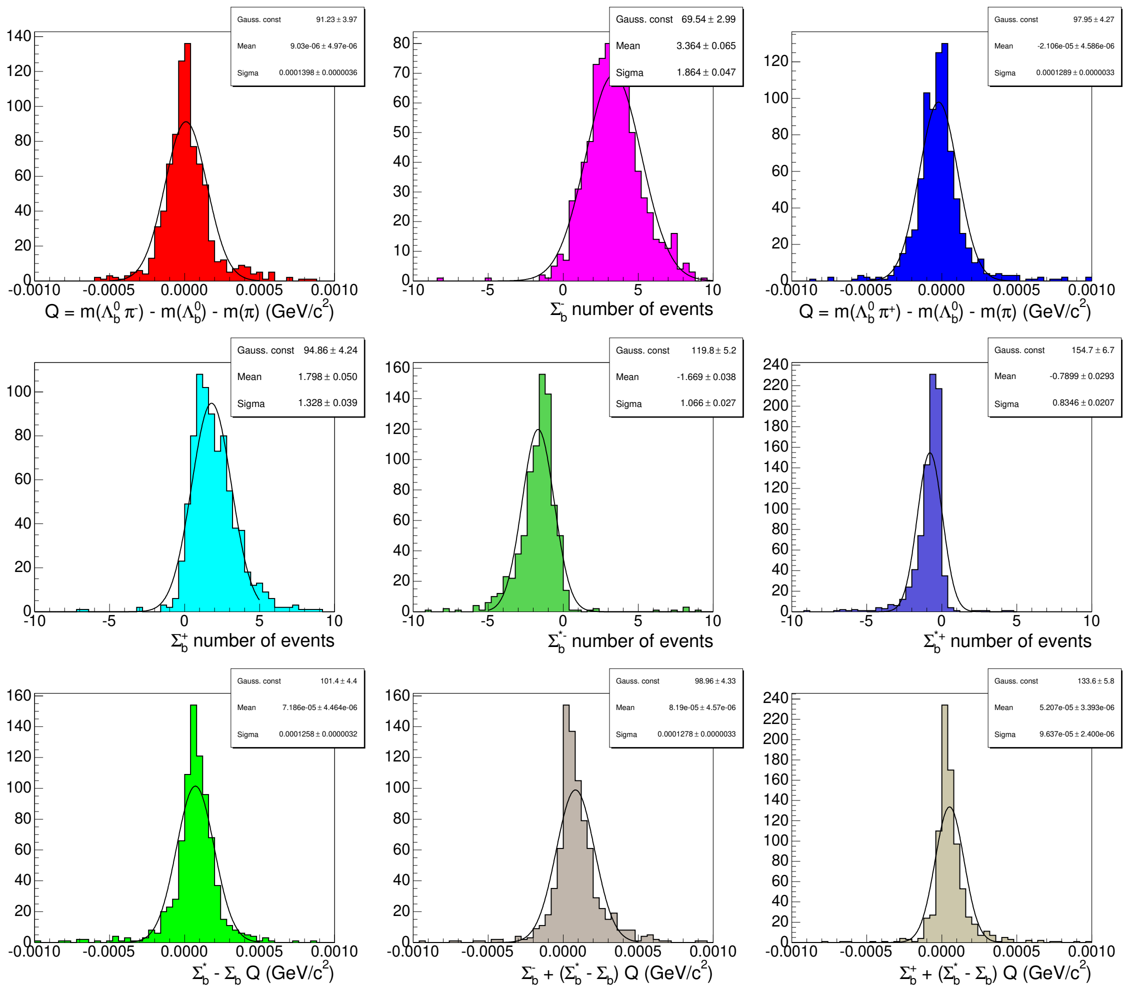


Figure 78: Plots of the systematic shift in each floating parameter for Toy MC samples where the width parameter g_A is set to 0.80.

10.2.5 Σ_b^* - Σ_b mass difference

With isospin splitting, the mass differences $m(\Sigma_b^{*+}) - m(\Sigma_b^+) \equiv \Delta_{*+}$ and $m(\Sigma_b^{*-}) - m(\Sigma_b^-) \equiv \Delta_{*-}$ are not expected to have the same value, as described in Sec. 2. However, due to the low statistics of our sample we constrained these mass differences to the same value in the fit, namely $m(\Sigma_b^{*+}) - m(\Sigma_b^+) = m(\Sigma_b^{*-}) - m(\Sigma_b^-) \equiv \Delta_*$.

To evaluate a systematic error on this assumption, we use the prediction $\Delta_{*+} - \Delta_{*-} = 0.40 \pm 0.07$ MeV/c^2 [5, 6]. Taking the worst case scenario, we used $\Delta_{*+} = \Delta_{*-} + 0.5$ MeV/c^2 and ran Toy MC samples for this configuration. The resulting systematic distributions are shown in Fig. 79 and the mean values are given in Tab. 41. As expected, the shift is negligible for the numbers of events, but does have an effect on the masses. This is the largest single error for the mass values, particularly the $\Sigma_b^{*\pm}$ Q values.

Parameter	$\Delta_{*+} = \Delta_{*-} + 0.5$ MeV/c^2
Σ_b^- Q (MeV/c^2)	0.060 ± 0.002
Σ_b^- events	-0.084 ± 0.006
Σ_b^+ Q (MeV/c^2)	-0.107 ± 0.003
Σ_b^+ events	-0.004 ± 0.008
Σ_b^{*-} events	-0.16 ± 0.01
Σ_b^{*+} events	0.16 ± 0.01
Σ_b^* - Σ_b Q (MeV/c^2)	-0.260 ± 0.004
Σ_b^{*-} Q (MeV/c^2)	-0.184 ± 0.003
Σ_b^{*+} Q (MeV/c^2)	-0.390 ± 0.003

Table 41: Systematic errors on each parameter due to using the same mass difference (Δ_*) for the positive (Δ_{*+}) and negative (Δ_{*-}) hyperfine mass splittings.

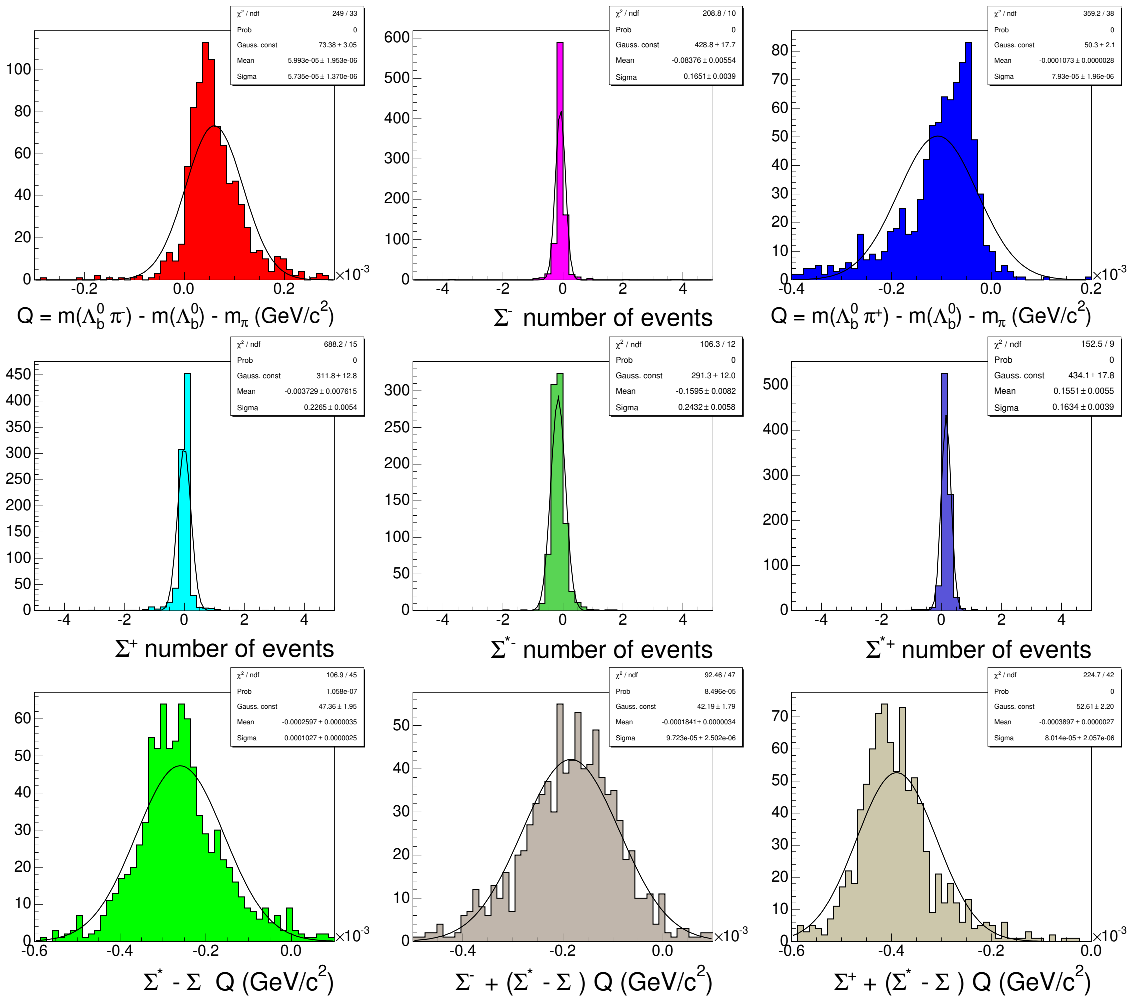


Figure 79: Plots of the systematic shift in each floating parameter for Toy MC samples where the positive mass difference, Δ_{*+}, is shifted up from the negative mass difference, Δ_{*-}, by 0.5 MeV/c^2.

10.3 Systematics: Summary

Table 42 lists all sources of systematic error on the four measured Σ_b quantities and the absolute value assigned to that systematic in the preceeding sections.

Parameter	Mass Scale	Λ_b^0 Comp.	Λ_b^0 Norm.	Λ_b^0 Shape	Reweight	Reso.	Σ_b Width	Δ_*	Total
Σ_b^- Q	0.22	0.0	0.009	0.0	0.04	0.0	0.009	0.06	0.23
	-0.22	-0.03	-0.002	-0.011	-0.0004	-0.011	-0.005	0.0	-0.22
Σ_b^- events	0.0	0.7	2.2	0.3	7.4	0.3	3.4	0.0	8.5
	0.0	0.0	-2.2	0.0	0.0	0.0	-3.4	-0.08	-4.1
Σ_b^+ Q	0.19	0.03	0.013	0.013	0.0	0.0	0.01	0.0	0.19
	-0.19	0.0	-0.013	0.0	-0.11	-0.014	-0.02	-0.11	-0.25
Σ_b^+ events	0.0	3.3	2.1	1.2	2.3	0.3	1.8	0.0	5.0
	0.0	0.0	-2.1	0.0	-1.8	0.0	-2.0	-0.004	-3.4
Σ_b^{*-} events	0.0	0.4	4.8	0.3	14.7	0.1	1.7	0.0	15.6
	0.0	0.0	-4.7	0.0	0.0	0.0	-1.7	-0.16	-5.0
Σ_b^{*+} events	0.0	7.3	4.8	2.8	4.6	0.2	0.8	0.16	10.3
	0.0	0.0	-4.8	0.0	-2.9	0.0	-0.8	0.0	-5.7
Σ_b^* - Σ_b Q	0.10	0.05	0.14	0.04	0.32	0.02	0.07	0.0	0.38
	-0.10	0.0	-0.13	0.0	0.0	0.0	-0.07	-0.26	-0.32
Σ_b^{*-} Q	0.28	0.02	0.13	0.03	0.32	0.003	0.08	0.0	0.45
	-0.28	0.0	-0.13	0.0	0.0	0.0	-0.07	-0.184	-0.37
Σ_b^{*+} Q	0.32	0.09	0.12	0.05	0.17	0.001	0.05	0.0	0.40
	-0.32	0.0	-0.13	0.0	0.0	0.0	-0.06	-0.39	-0.52

Table 42: Table of all systematic errors on the Σ_b measurement. All Q values are in units of MeV/c^2.

11 Conclusions

We have observed $\Sigma_b^{(*)\pm}$ in 1 fb^{-1} of data collected by the CDF detector. The widths predicted by Eq. (1) are in agreement with our data. We measure the Q values of Σ_b^- and Σ_b^+, and the Σ_b^*-Σ_b mass difference to be:

- m(Σ_b^-) - m(Λ_b^0) - m(π) = 55.9 ± 1.0 (stat) ± 0.2 (syst) MeV/c^2
- m(Σ_b^+) - m(Λ_b^0) - m(π) = $48.5^{+2.0}_{-2.2}$ (stat) $^{+0.2}_{-0.3}$ (syst) MeV/c^2
- m(Σ_b^{*-}) - m(Σ_b^-) = m(Σ_b^{*+}) - m(Σ_b^+) = $21.2^{+2.0}_{-1.9}$ (stat) $^{+0.4}_{-0.3}$ (syst) MeV/c^2

To go from Q values to absolute masses for $\Sigma_b^{\pm}$, we add the PDG value of the Λ_b^0 and pion masses. However, there is an error on the mass of the Λ_b^0 which must also be added to the systematic error on the measurement. Using the best CDF mass measurement for Λ_b^0 , which is m(Λ_b^0) = 5619.7 ± 1.2 (stat) ± 1.2 (syst) MeV/c^2 [35], the absolute mass values and errors are

- m(Σ_b^-) = 5815.2 ± 1.0 (stat) ± 1.7 (syst) MeV/c^2
- m(Σ_b^+) = $5807.8^{+2.0}_{-2.2}$ (stat) ± 1.7 (syst) MeV/c^2

To quote the absolute masses for $\Sigma_b^{*\pm}$, we have already calculated the systematic errors. The statistical errors must also be calculated, taking into account the correlations between the Σ_b and ($\Sigma_b^* - \Sigma_b$) parameters. From the error matrix output of the fit to data, the external error between Σ_b^- and ($\Sigma_b^* - \Sigma_b$) is -3.891×10^{-7} GeV/c^2, while the error between Σ_b^+ and ($\Sigma_b^* - \Sigma_b$) is -2.703×10^{-6} GeV/c^2. Using this along with the errors on the parameters yields absolute mass values with errors of

- m(Σ_b^{*-}) = 5836.4 ± 2.0 (stat) $^{+1.8}_{-1.7}$ (syst) MeV/c^2
- m(Σ_b^{*+}) = $5829.0^{+1.6}_{-1.8}$ (stat) $^{+1.7}_{-1.8}$ (syst) MeV/c^2

The number of events for each state are

- N(Σ_b^-) = 59^{+15}_{-14} (stat) $^{+9}_{-4}$ (syst)
- N(Σ_b^+) = 32^{+13}_{-12} (stat) $^{+5}_{-3}$ (syst)
- N(Σ_b^{*-}) = 69^{+18}_{-17} (stat) $^{+16}_{-5}$ (syst)
- N(Σ_b^{*+}) = 77^{+17}_{-16} (stat) $^{+10}_{-6}$ (syst)

While the measurement of the Σ_b absolute mass values are limited by statistical errors and the error on the Λ_b^0 mass equally, the measurement of the Σ_b Q values is clearly statistically limited and will benefit greatly from the addition of more data.

A Universal Finder Reconstruction

```
if { $UFIND_DO_PID == 1 } {
  mod enable  CT_TrackingModule
  talk CT_TrackingModule
       NoTracking set t
  exit

  module enable TofModule
  set UFIND_PID_MODULES "CT_TrackingModule TofModule"
} else {
  mod disable CT_TrackingModule
  mod disable TofModule
  set UFIND_PID_MODULES ""
}

talk CalibrationManager
  ProcessName set PROD_PHYSICS_CDF
      PassName set 16
exit

talk PuffModule
   puffOnly set StorableRun2SiStripSet SiClusterSet SiHitSet
exit

mod enable  Prereq
talk        Prereq
   if { $UFIND_IS_REALMC == 1 } {
      bankType set "Simulated Trigger Bank"
   }
   bankType list
   L1Accept       set true
   L2Accept       set true
   L3Accept       set false
   L3TriggerNames set B_CHARM B_CHARM_L1_DPS
   show
exit

talk GeometryManager
   #  TOF geometry model set to Survey
   TofGeometryMenu
     # Other options: Nominal, Naive
     GeometryModel set Survey
```

```
   exit

   #--- Misalignment in MC set in SiliconGeometryMenu
   SiliconGeometryMenu
     L00Alignment       set true
     AlignmentPrint     set 3
     AlignmentSource set default
     #--- Simulation of Si passive material
     BuildPassive set t
     CreatePhantomLayer set true
     PhantomLayerRmin set 14.8  14.8  14.8  14.8  14.8  20.5   20.5   20.8  20.5  20.5
     PhantomLayerZmin set -100  -45   -15   15.1  46    -100   -45   -15    15.1  46
     PhantomLayerZmax set -46   -15.1 15    45    100   -46    -15.1  15    45    100
     PhantomLayerThickness set 2.8 0.9 0.8 0.9 2.8 0.38 0.524  0.232  0.524  0.38
     PhantomLayerMaterial set BERYLLIUM BERYLLIUM BERYLLIUM BERYLLIUM BERYLLIUM
BERYLLIUM BERYLLIUM BERYLLIUM BERYLLIUM BERYLLIUM
     PhantomLayerContainer set SVCC SVCC SVCC SVCC SVCC ISLC ISLC ISLC ISLC ISLC
  exit
exit


#--- No use for TrackSelectorModule in QuickSim
talk TrackSelectorModule
   #--- Cuts for the Std track views
   trackCutManager
      cut errCut  = HasHelixFit && PhysicalError
      cut ptCutL  = Pt > 1.3
      cut myptCut = Pt > 0.5
      cut etaCut  = AbsEta < 2.0
      # 2 SL AX (5 hits), 2 SL ST (5 hits)
      cut cotCut  = HasCOTHits 0 0 2 2 5 5
      # 3 AX Si hits
      cut svxCut  = HasSVXIIHitLayers 3 0 0 0
      #
      cut K_std   = etaCut && svxCut && cotCut && myptCut && errCut
      cut pi_std  = etaCut && svxCut && cotCut && myptCut && errCut
      cut pi_soft = etaCut && errCut
      cut p_std   = errCut && svxCut && cotCut && Pt > 2.0 && AbsEta < 1.3
      cut mu_std  = etaCut && svxCut && cotCut && ptCutL  && errCut
   exit

   inputTracks set defTracks

   #--- Track Refitting
```

```
   trackRefit
      refitTracks set true
      if { $USEL00 == 1 } {
  dropL00     set false
  L00Refit    set true
      } else {
  dropL00     set true
  L00Refit    set false
      }
      dropISL              set false
      #--- Refit method (KAL vs G3X)
      refitMethod          set KAL
      #--- Rescale COT covariance matrix
      rescaleCOTCov        set true
      curvCOTFactor        set 21.72
      d0COTFactor          set 11.57
      phi0COTFactor        set 14.64
      lambdaCOTFactor      set 1.544
      z0COTFactor          set 1.71
      show
   exit
exit

if { $UFIND_DO_PID == 1 } {
    set doPid true
} else {
    set doPid false
}

talk TrackAssocModule
   #--- Pions
   pi_Std_Assoc
     enabled     set t
     doSvt       set t
     doDedx      set $doPid
     doTof       set $doPid
     correctDedx set $doPid
   exit

   #--- Soft pions
   pi_Soft_Assoc
     enabled     set t
     doSvt       set t
     doDedx      set $doPid
```

```
    doTof       set $doPid
    correctDedx set $doPid
  exit

  #--- Kaons
  K_Std_Assoc
    enabled     set t
    doSvt       set t
    doDedx      set $doPid
    doTof       set $doPid
    correctDedx set $doPid
  exit

  #--- Protons
  p_Std_Assoc
    enabled     set t
    doSvt       set t
    doDedx      set $doPid
    doTof       set $doPid
    correctDedx set $doPid
  exit

  #--- Muons
  mu_Std_Assoc
     doMuons set true
     doSvt   set true
     doDedx  set false
     doTof   set false
  exit

exit
```

B Additional Data-MC Comparison Plots

Continued from Section 4.4: Data-MC comparison plots when a $p_{\mathrm{T}} > 1$ GeV$/c$ cut is made on the Σ_b track. The MC has been reweighted for Λ_b^0 p_{T} , but not for the track p_{T} .

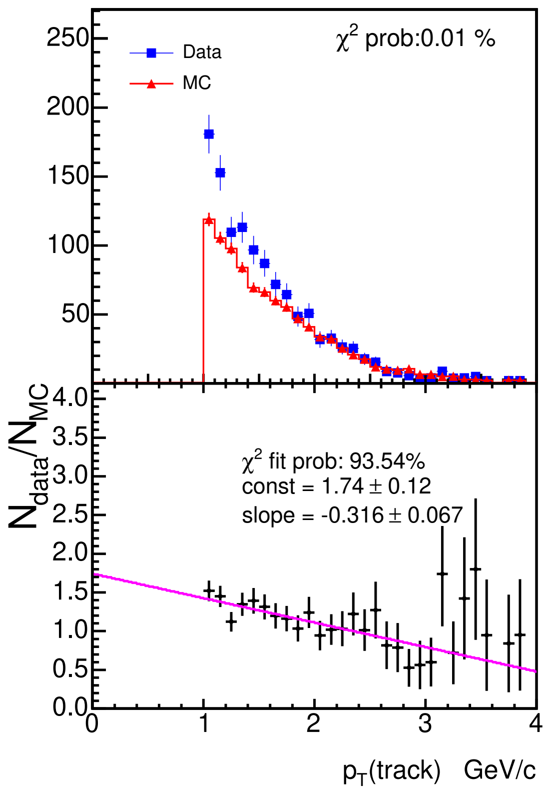


Figure 80: With a p_{T} (track) > 1 GeV$/c$ cut: Data-MC comparison of track p_{T} .

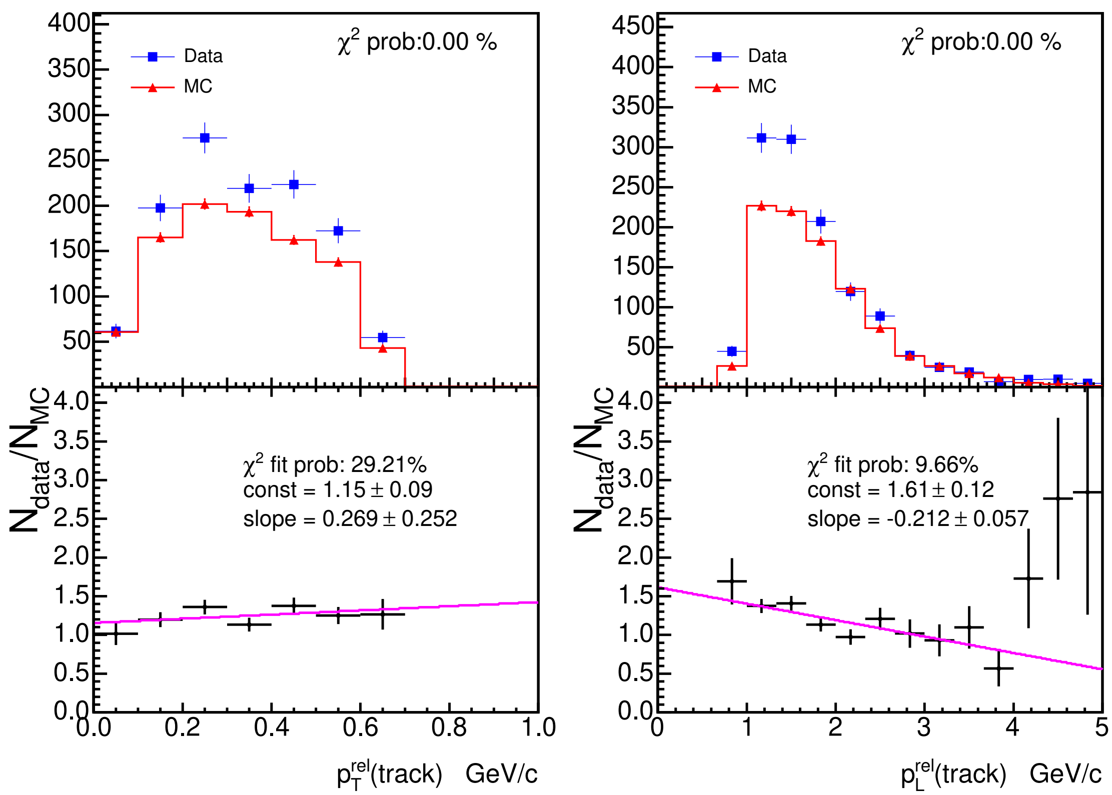


Figure 81: With a p_{T} (track) > 1 GeV$/c$ cut: Data-MC comparisons of track p_{T}^{rel} and p_{L}^{rel}.

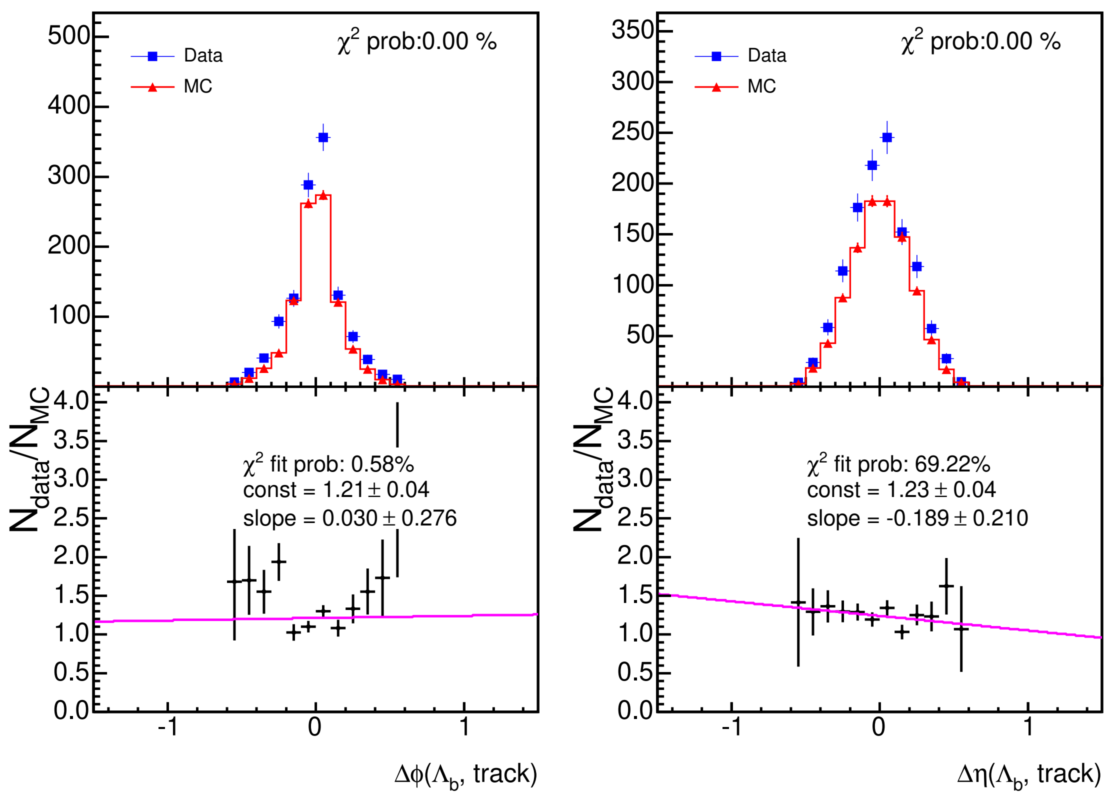


Figure 82: With a p_{T} (track) > 1 GeV$/c$ cut: Data-MC comparisons of $\Delta\phi$ and $\Delta\eta$ between the track and the Λ_b^0 candidate.

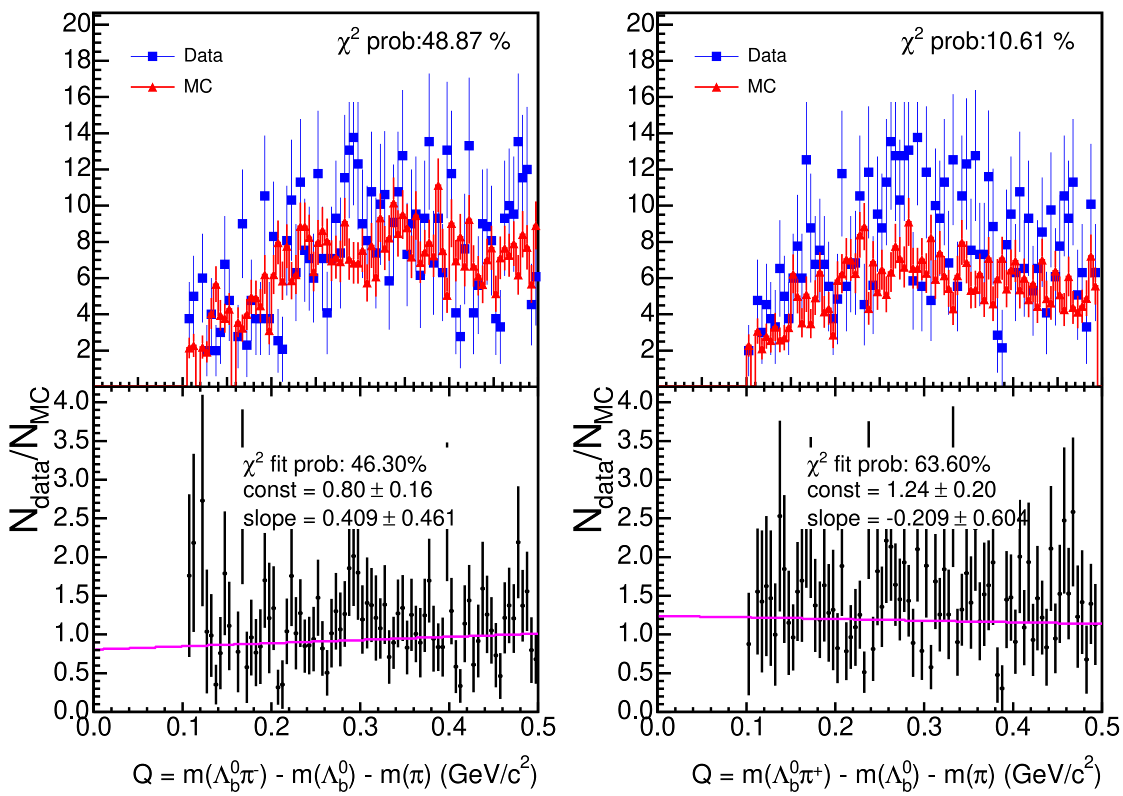


Figure 83: With a p_{T} (track) > 1 GeV$/c$ cut: Data-MC comparisons of the $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ Q distributions.

C Blind Tests of Σ_b Fit

Before unblinding the Σ_b signal region and attempting to fit for Σ_b in data, we test the fitter on Toy Monte Carlo samples of a "fake" Σ_b signal. Using the theoretical expectations outlined in Sec. 2, we chose the Σ_b input parameters shown in Table 43. Due to an error in our original calculation of the Σ_b width, we used an input Σ_b width of only 0.8 MeV/c^2. We also assumed the same arbitrarily chosen number of input events for Σ_b and Σ_b^*. The signal functions for this set of input parameters are shown in Fig. 84. Combining this signal shape with the background shapes described in Sec. 6, we generated 1000 Toy MC experiments for $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ separately and fit each one. An example of a Toy MC sample is shown in Fig. 85. We then calculate the pulls on the four floating variables in each fit:

- Σ_b mass
- Σ_b^* - Σ_b mass difference
- Number of Σ_b events
- Number of Σ_b^* events

For this fit, $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ are fit separately, so there is no constraint on the Σ_b^* - Σ_b mass difference for Σ_b^- and Σ_b^+. The pulls on the floating variables are shown in Figs. 86 through 89. Except for the number of Σ_b events in the $\Lambda_b^0\pi^-$, all pulls agree with the unit Gaussian. For this one exception, there are some events in the positive tail of the Gaussian which need to be looked into. It may be that these fits did not converge well for some reason.

Signal parameters	Values (MeV/c^2)
m(Σ_b^-) - m(Λ_b^0) - m(π)	45
$\Gamma(\Sigma_b)$	0.8
m(Σ_b^*) - m(Σ_b)	15
m($\Sigma_b^{(*)-}$) - m($\Sigma_b^{(*)+}$)	6
N(Σ_b) = N(Σ_b^*)	40

Table 43: List of the input Σ_b signal parameters used to test the fitter.

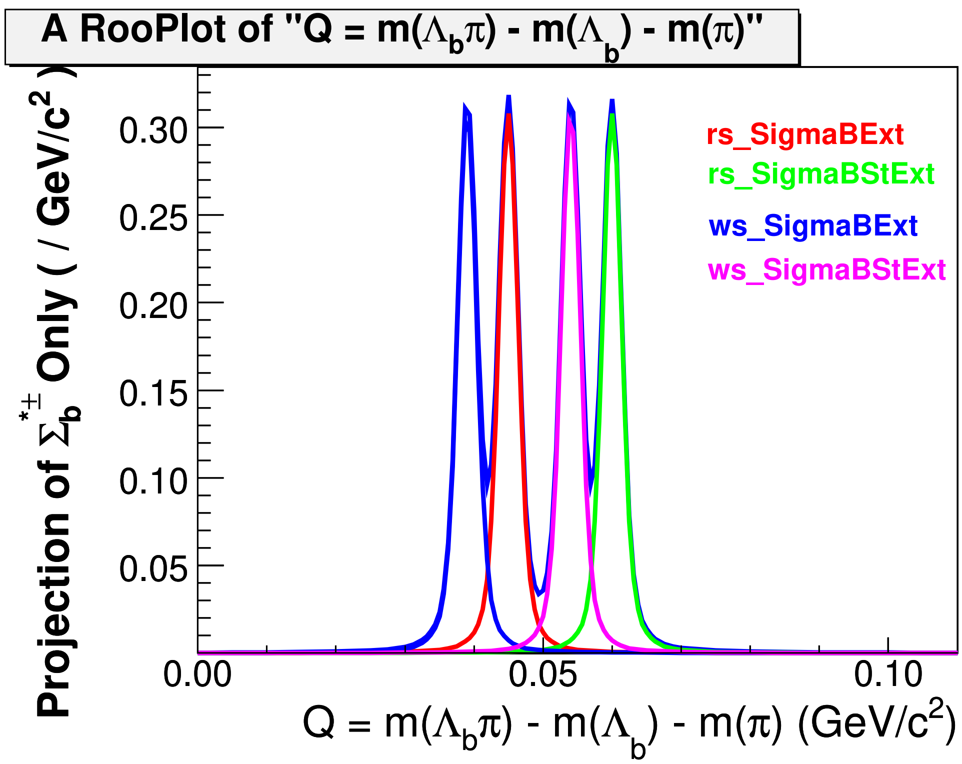


Figure 84: RooFit plot of the Σ_b signal structure for the input parameters given in Table 43. This plot shows $\Lambda_b^0\pi^-$ and $\Lambda_b^0\pi^+$ combined.

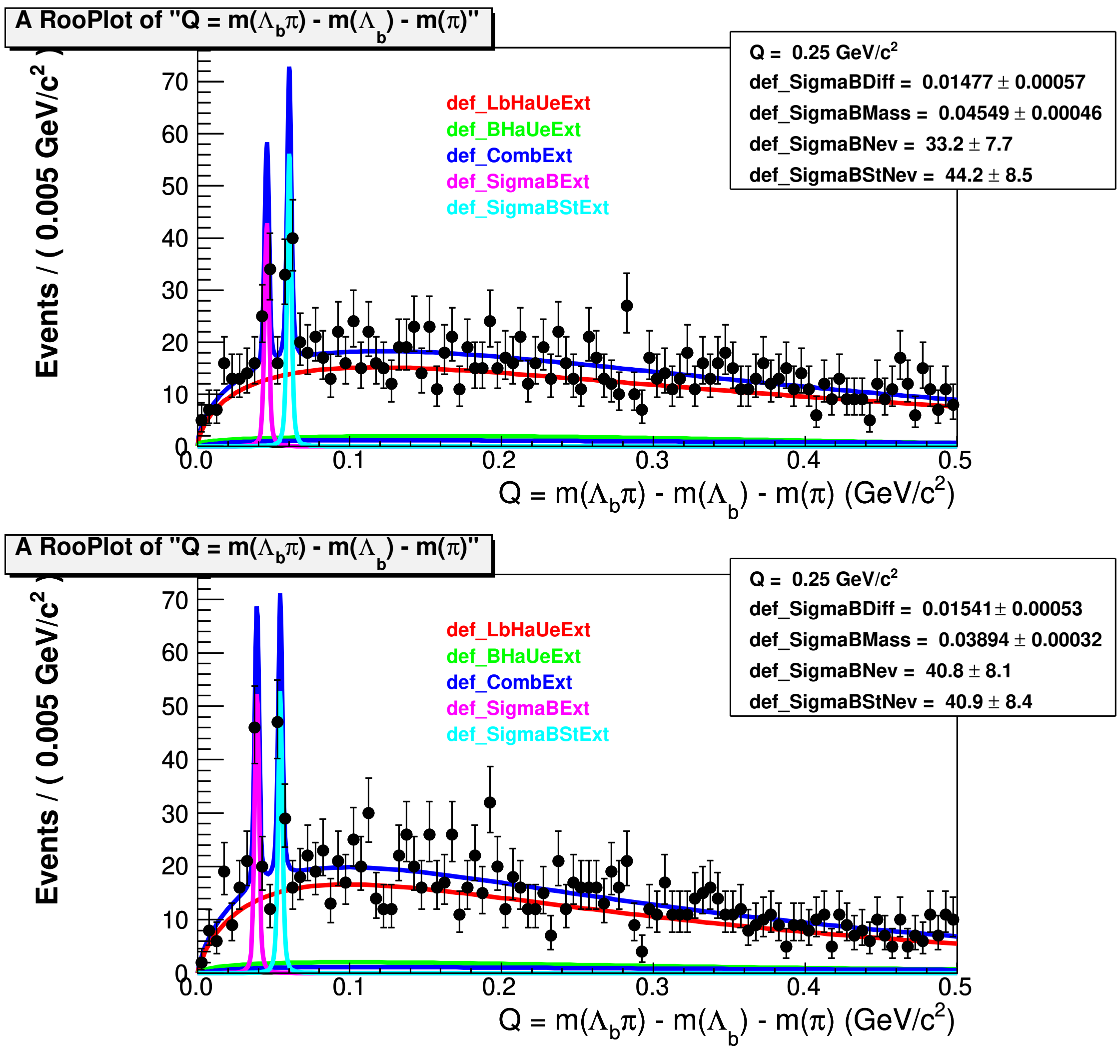


Figure 85: Plots of a Toy MC sample generated with the input signal parameters given in Table 43. The unbinned fit to the Toy MC is also shown.

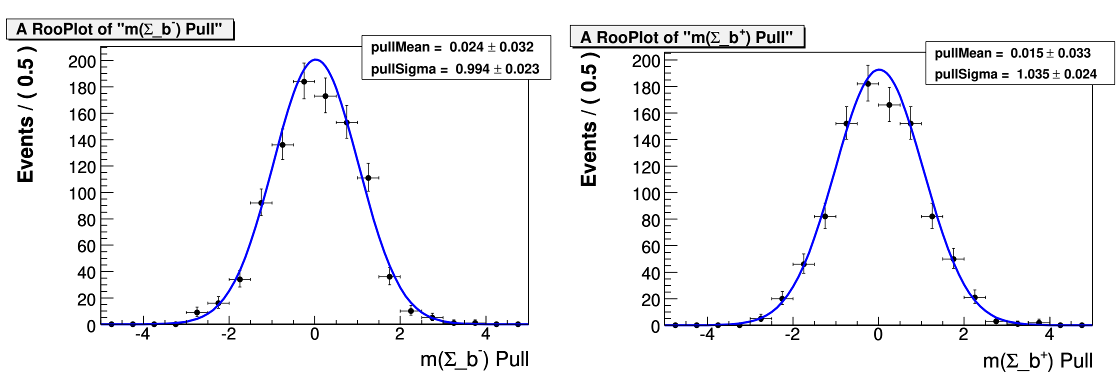


Figure 86: Plot of the Toy MC pull for the Σ_b mass, for both $\Lambda_b^0\pi^-$ (left) and $\Lambda_b^0\pi^+$ (right).

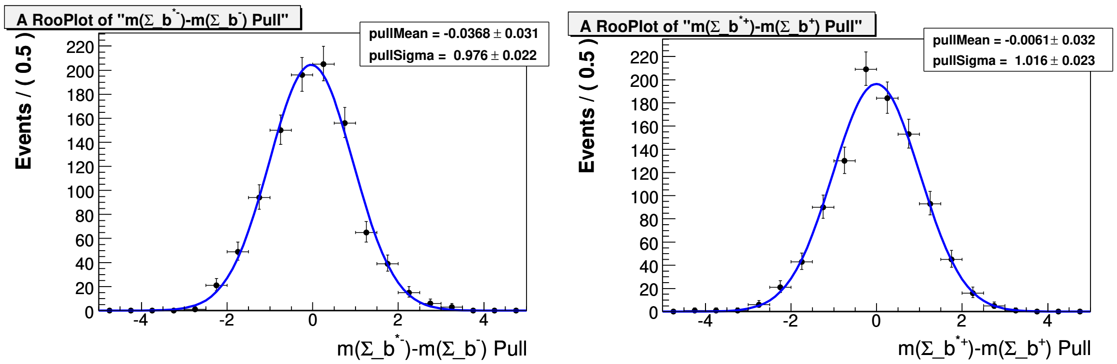


Figure 87: Plot of the Toy MC pull for the $\Sigma_b^* - \Sigma_b$ mass difference, for both $\Lambda_b^0\pi^-$ (left) and $\Lambda_b^0\pi^+$ (right).

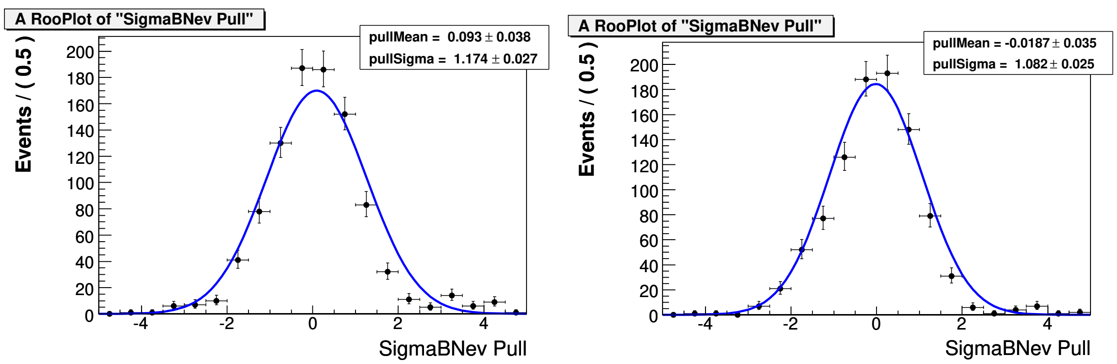


Figure 88: Plot of the Toy MC pull for the number of Σ_b events, for both $\Lambda_b^0\pi^-$ (left) and $\Lambda_b^0\pi^+$ (right).

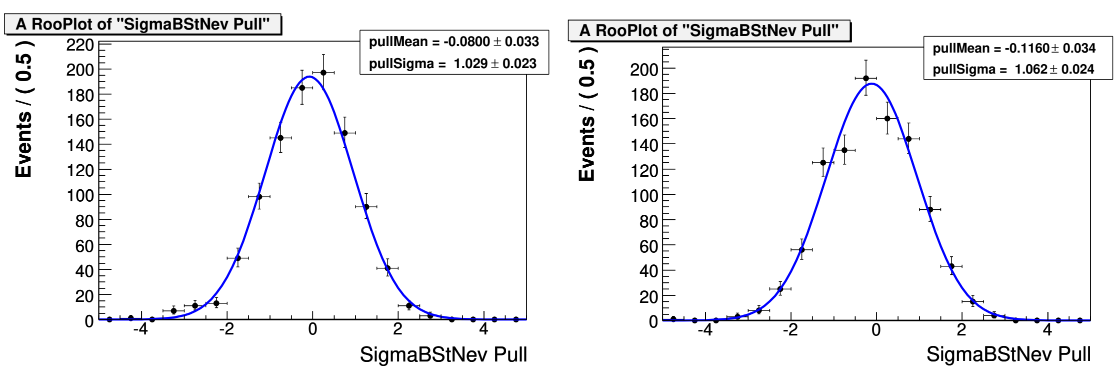


Figure 89: Plot of the Toy MC pull for the number of Σ_b^* events, for both $\Lambda_b^0\pi^-$ (left) and $\Lambda_b^0\pi^+$ (right).

References

[1] Particle Data Group, W-M Yao *et al.* 2006 J. Phys. G: Nucl. P art. Phys.**33** 1

[2] DELPHI Collaboration (M. Feindt *et al.*), CERN PPE/95-139

[3] M. Feindt, private communications

[4] F. Bedeschi *et al.*, “Observation of orbitally excited (L=1) B mesons”, CDF/ANAL/BOTTOM/CDFR/7820.
F. Bedeschi *et al.*, “Study of orbitally excited (L=1) B mesons”, CDF/ANAL/BOTTOM/CDFR/7471

[5] J. L. Rosner, Phys. Rev. D **57** (1998) 4310-4317 [arXiv:hep-ph/9707473].

[6] J. L. Rosner, private communications.

[7] D. P. Stanley and D. Robson, Phys. Rev. Lett. **45**, 235 (1980).

[8] D. P. Stanley and D. Robson, Phys. Rev. D **21**, 3180 (1980).

[9] D. Izatt, C. DeTar and M. Stephenson, Nucl. Phys. B **199**, 269 (1982).

[10] D. B. Lichtenberg, R. Roncaglia, J. G. Wills and E. Predazzi, Z. Phys. C **47**, 83 (1990).

[11] A. Martin and J. M. Richard, Phys. Lett. B **185**, 426 (1987).

[12] J. M. Richard and P. Taxil, Phys. Lett. B **128**, 453 (1983).

[13] J. L. Basdevant and S. Boukraa, Z. Phys. C **30**, 103 (1986).

[14] W. Kwong, J. L. Rosner and C. Quigg, Ann. Rev. Nucl. Part. Sci. **37**, 325 (1987).

[15] K. C. Bowler *et al.* [UKQCD Collaboration], Phys. Rev. D **54**, 3619 (1996) [arXiv:hep-lat/9601022].

[16] E. Jenkins, Phys. Rev. D **54**, 4515 (1996) [arXiv:hep-ph/9603449].

[17] E. Jenkins, Phys. Rev. D **55** (1997) 10 [arXiv:hep-ph/9609404].

[18] M. Karliner and H. J. Lipkin, hep-ph/0307243.

[19] N. Mathur, R. Lewis and R. M. Woloshyn, Phys. Rev. D **66**, 014502 (2002) [arXiv:hep-ph/0203253].

[20] C. Albertus, J. E. Amaro, E. Hernandez and J. Nieves, Nucl. Phys. A **740**, 333 (2004) [arXiv:nucl-th/0311100].

[21] D. Ebert, R. N. Faustov and V. O. Galkin, Phys. Rev. D **72**, 034026 (2005) [arXiv:hep-ph/0504112].

[22] A. F. Falk, Int. J. Mod. Phys. A **12**, 4079 (1997) [arXiv:hep-ph/9609208].

[23] J. G. Korner, M. Kramer and D. Pirjol, Prog. Part. Nucl. Phys. **33**, 787 (1994) [arXiv:hep-ph/9406359].

[24] W-Y. P. Hwang and D. B. Lichtenberg, Phys. Rev. D **35**, 3526 (1987)

[25] S. Capstick, Phys. Rev. D **36**, 2800 (1987)

[26] I. Gorelov *et al.*, "Λ_b^0 Data Sample and Mass Fit with Two Displaced Track Trigger Using 1 fb^{-1}," CDF/PHYS/BOTTOM/CDFR/8395

[27] *Good Run List Version 13*: http://www-cdf.fnal.gov/internal/dqm/goodrun/v13/goodv13.html

[28] *Sin2BetaMods*: http://cdfcodebrowser.fnal.gov/CdfCode/source/Sin2BetaMods

[29] *JpsiXMods*: http://cdfcodebrowser.fnal.gov/CdfCode/source/JpsiXMods

[30] *BStntuples*: http://www-cdf.fnal.gov/htbin/twiki/bin/view/BStntuples/DataSetsSpring2006

[31] S.S. Yu, R.J. Tesarek, D. Litvintsev, "User Guide of the Λ_b^0 p_{T} Spectrum" CDF/ANAL/BOTTOM/CDFR/8111

[32] The default track quality cuts have not changed since CDF software version 5.1.0. For software versions 5.1.0 up to 6.1.4 (used in this analysis), defTracks selection criteria are listed on the Tracking Group webpage: http://www-cdf.fnal.gov/internal/upgrades/computing/projects/reconstruction/tracking/user-docs/padtrack510.html

[33] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," Nucl. Instrum. Meth. A **389**, 81 (1997).

[34] K. Cranmer, "Kernel Estimation in High-Energy Physics," Comput. Phys. Commun. 136, 198-207 (2001), hep-ex/0011057.

[35] A. Korn, Ch. Paus, "Measurement of b hadron masses in exclusive J/ψ decays with the CDF detector," CDF/PHYS/BOTTOM/CDFR/7691.