CDF note 8850

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

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: September 14, 2007)

A combined mass and dE/dx fit is used to obtain the first observation of the decay $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ and measure the relative branching fraction of $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ to $\overline{B}_s^0 \to D_s^{+} \pi^{-}$. This analysis uses 1.2 fb^{-1} of $D_s(\phi\pi)X$ data collected with the CDF II detector in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron collider. We measure $\mathcal{B}(\overline{B}_s^0 \to D_s^{\pm} K^{\mp})/\mathcal{B}(\overline{B}_s^0 \to D_s^{+} \pi^{-}) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$. The statistical significance of the $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ signal is 7.90σ.

Figure 1: Diagrams contributing to $B_s^0 \to D_s^{\mp} K^{\pm}$. The relative weak phase between the diagrams is $\arg\left(-V_{cs}V_{ub}^*/V_{us}V_{cb}^*\right) = \gamma$.

I. INTRODUCTION

The decay mode $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ is interesting for several reasons. Because the $\overline{B}_s^0$ can decay both to $D_s^+ K^-$ and to $D_s^- K^+$, and because these decay amplitudes can interfere through B_s mixing, the relative branching fraction $\mathcal{B}(\overline{B}_s^0 \to D_s^{\pm} K^{\mp})/\mathcal{B}(\overline{B}_s^0 \to D_s^+ \pi^-)$ can be significantly larger or smaller than $\mathcal{B}(\overline{B}^0 \to D^+ K^-)/\mathcal{B}(\overline{B}^0 \to D^+ \pi^-)$. Furthermore, the relative weak phase between $\overline{B}_s^0 \to D_s^+ K^-$ and $\overline{B}_s^0 \to D_s^- K^+$ is the CKM angle γ, as shown in Figure 1. This means that a time-dependent, flavor-tagged analysis of $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ can provide a measurement of γ that is theoretically very clean [1]. An untagged analysis that relies on a sizeable lifetime difference between the *CP* eigenstates of the B_s [2] may also be possible.

In this note we report the first observation of $\overline{B}_s^0 \to D_s^{\pm} K^{\mp}$ decays and the first measurement of $\mathcal{B}(\overline{B}_s^0 \to D_s^{\pm} K^{\mp})/\mathcal{B}(\overline{B}_s^0 \to D_s^+ \pi^-)$ using $D_s^+ \to \phi\pi^+$ decays. This measurement sheds light on the size of the interference between $\overline{B}_s^0 \to D_s^+ K^-$ and $\overline{B}_s^0 \to D_s^- K^+$. It also proves that CDF can disentangle the Cabibbo-suppressed $D_s K$ mass peak from that of the Cabibbo-allowed $D_s \pi$, which is a prerequisite for the measurement of γ.

II. DATA SAMPLE AND EVENT SELECTION

We use 1.2 fb^{-1} of $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV recorded by the CDF II detector at the Fermilab Tevatron collider. The components of the detector relevant to this analysis are briefly described below. A more detailed description can be found elsewhere [3]. Charged particle tracks are reconstructed in the pseudorapidity range $|\eta| \lesssim 1$, where η is defined as $-\ln\tan(\theta/2)$ [4], from hits in a silicon microstrip vertex detector and a cylindrical drift chamber. Both detectors are immersed in a 1.4 T solenoidal magnetic field. The vertex detector is composed of L00 (a single layer of axial strip detectors placed close to the beampipe), SVX II (five cylindrical layers of double-sided sensors providing five axial, three 90° stereo and two small-angle stereo measurements) and ISL (a double-sided axial and small-angle stereo layer) [5]. Surrounding the vertex detector is the COT, an open-cell drift chamber with 96 layers of axial and small-angle stereo sense wires [6]. The specific energy loss (dE/dx) of charged particles in the COT can be measured from the collected charge, which is encoded in the output pulse width of each wire.

A sample rich in bottom hadrons is selected by a three-level trigger system. At level 1, tracks are reconstructed in the COT by the trigger track processor (XFT) [7]. The level 2 silicon vertex trigger (SVT) [8] associates SVX II r–ϕ position measurements with XFT tracks, providing a precise measurement of the track impact parameter d_0. Decays of heavy flavor particles are identified by requiring two tracks with 0.12 mm $< d_0 <$ 1 mm and making additional cuts on the two-track combination. The level 3 trigger [9] performs a full event reconstruction and applies the level 1 and 2 selections.

We use three data samples in this analysis: two control samples to measure the relative branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}K^-)/\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)$ and $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$ and the signal sample to measure $\mathcal{B}(\overline{B}_s^0 \to D_s^{\pm}K^{\mp})/\mathcal{B}(\overline{B}_s^0 \to D_s^+\pi^-)$. The samples are reconstructed in the $D^{*+}\left(D^0(K^-\pi^+)\pi^+\right)\pi^-$, $D^+(K^-\pi^+\pi^+)\pi^-$ or $D_s^+(\phi\pi^+)\pi^-$ mass hypothesis [10]. No particle identification is used in the reconstruction, and tracks are assumed to be either a pion or a kaon to match the reconstruction hypothesis. Tracks are combined to form ϕ, D^0, D^+, D_s^+, D^{*+}, B^0 and B_s^0 candidates. The track combinations must be consistent with forming a vertex. Combinations falling outside a window around the invariant mass of the respective meson (and a window in $\Delta m = m(D^{*+}) - m(D^0)$ in the case of the D^{*+}) are rejected.

Further selection cuts are applied to reduce combinatorial background. The discriminating variables are the transverse (d_0) and longitudinal (z_0) impact parameter of the B candidate with respect to the primary event vertex; the transverse momentum (p_T) of the B candidate; the isolation of the B candidate (defined as the fraction of the scalar p_T sum of tracks within a cone of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 1$ around the B direction that is carried by the B candidate); the opening angle between the D candidate and the track recoiling against the D; and the projection of the B and D decay length along the transverse momentum of the B candidate (L_{xy}). In addition, we reconfirm the trigger requirements on the B candidate. Because the dE/dx calibrations were tuned on trigger tracks, we require that the track recoiling against the D be a trigger track.

Monte Carlo simulation is used to model signal and background and to determine trigger and reconstruction efficiencies. We generate single B hadrons with `Bgenerator` [11, 12] and simulate their decays with `EvtGen` [13]. A detailed detector and trigger simulation is then performed. Additional corrections to the Monte Carlo resolution and mass scale are applied. These corrections are determined by comparing the peak mean and width in a clean $\overline{B}^0 \to D^{*+}\pi^-$ data sample with Monte Carlo. For each sample used in this analysis, the corrections are cross-checked *in situ* by scanning the maximum likelihood as a function of resolution and mass scale.

III. ANALYSIS METHOD

The central challenge in this analysis is to disentangle the various components of the $D_s^+(\phi\pi^+)\pi^-$ data sample [14]. Apart from the $\overline{B}_s^0 \to D_s^\pm K^\mp$ and $\overline{B}_s^0 \to D_s^+\pi^-$ signals, the sample contains various other partially reconstructed $\overline{B}_s^0 \to D_s^+ X$ modes, decays of other bottom hadrons and combinatorial background.

A. Mass and dE/dx fit

To separate the components and determine the number of candidates of each component type, we perform a maximum-likelihood fit. The fit variables are the invariant mass m of the candidate in the $D_s^+\pi^-$ mass hypothesis and the particle-identification (PID) variable

$$Z = \log\left(\frac{dE/dx(\text{measured})}{dE/dx(\text{expected for } \pi)}\right) \tag{1}$$

for the track recoiling against the D_s^+ (the "B daughter track"). (We choose Z because we expect it to be normally distributed instead of dE/dx, which follows a log-normal distribution.) The likelihood function is

$$L(f_1, \ldots, f_{M-1}) = \prod_{i=1}^{N}\sum_{j=1}^{M} f_j\, p_j(m_i)\, q_j(Z_i), \quad \text{where } f_M = 1 - \sum_{j=1}^{M-1} f_j. \tag{2}$$

$\prod_i$ is a product over all N candidates, and $\sum_j$ is a sum over the M components; f_j is the fraction of candidates in the jth component. The mass p.d.f.'s $p_j(m)$ and Z p.d.f.'s $q_j(Z)$ (also referred to as dE/dx p.d.f.'s) will be described shortly.

The fit components used in this analysis can be characterized as coming from three possible sources:

- D-plus-track combinations where the D candidate and the track come from a single B hadron (B^0, B^+, B_s^0, Λ_b^0). We will call the p.d.f.'s from such samples "single-B templates". This category includes D_s^+-plus-track combinations but also reflections of other charmed hadrons (D^+, Λ_c^+). The mass p.d.f.'s for these components are determined from Monte Carlo simulation. Because we know for each Monte Carlo sample the species of the track that is combined with the D, the dE/dx p.d.f.'s are known.

- Cases where the D candidate does not come from a real D. We will call the p.d.f.'s for this case "fake-D templates." Both the mass and dE/dx p.d.f.'s for this component are measured in the data using the D_s sidebands.

- D plus track combinations where a real D is combined with a track coming from fragmentation, the underlying event or the other B. We will call the p.d.f.'s from such samples "real-D background templates".

B. Mass templates

The single-B templates are divided into several components corresponding to different decay modes. Many mass p.d.f.'s are not easily described by a simple parametric shape; we therefore choose to represent the mass p.d.f.'s for the single-B fit components as histograms. We extract mass templates from large Monte Carlo samples of $B \to DX$ decays, where the B decay is inclusive and the D decay is forced to be one of the decay modes ($D_s^+ \to \phi\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $D^0 \to K^-\pi^+$) used in this analysis. These samples are then subdivided by B decay mode. The exact list of modes used varies between the three analysis samples, but all samples use templates for the Cabibbo-allowed $B \to D\pi$ and Cabibbo-suppressed $B \to DK$ fully reconstructed decays, and for the partially reconstructed

modes of significance in the DK mass region ($B \to D\rho$ and $B \to D^*\pi$). Modes that are not separated into one of the aforementioned categories are collected in a catch-all "remainder" template. If large reflections exist that fall into the mass window of the charmed meson, templates for these reflections are also used. The details for each sample are described in the respective fit configuration sections below.

One place where special care has been taken is in the treatment of the radiative tail in the Cabibbo-allowed decay $B \to D\pi$. This radiative tail is important for two reasons. First, proper treatment of the tail is important for properly determining the number of events in the Cabibbo-allowed mode. Second, this tail puts additional $D\pi$ events into the region of the Cabibbo-suppressed decay $B \to DK$. The radiative tail on the Cabibbo-allowed mode is modeled in `EvtGen` by using the `PHOTOS` algorithm for radiative corrections [15]. This model is expected to describe the shape of the tail well, but its reliability for determining the rate is not known. We therefore choose to leave the size of the radiative tail as a free parameter in our fit. Monte Carlo events in the Cabibbo-allowed mode are separated into a non-radiative subsample and a radiative one, and separate mass templates are created for either subsample. The `PHOTOS` cut-off in `EvtGen` is 10 MeV.

In contrast to the single-B components, the combinatorial background mass shape is determined from samples of limited size. We therefore describe the combinatorial background mass template using a smooth function:

$$p_{\rm bg}(m) = \frac{\alpha(1-\beta)\exp(-\alpha m)}{\exp(-\alpha m_{\rm min}) - \exp(-\alpha m_{\rm max})} + \frac{\beta}{m_{\rm max} - m_{\rm min}} \tag{3}$$

where $m_{\rm min}$ and $m_{\rm max}$ are the limits of the fit range. In the case of the fake-D + track background, the parameters of the function are determined by fitting the $D\pi$ mass distribution for events in the D sideband to a fake-D component of the form (3). Care must be taken because of real D candidates far in the tails of the D mass distribution which leak into the D sideband region. We add a "poorly measured real-D" component to the $D\pi$ fit to model this signal leakage. The shape (but not the normalization) of the real-D part of this fit is a histogram taken from the D sidebands in our single-B Monte Carlo samples.

The real-D combinatorial background is modeled using the wrong-sign data. We plot the invariant mass of the wrong-sign candidates in the $D\pi$ mass assumption and sideband-subtract in the D mass. Then we perform a binned fit of this background-subtracted mass plot using a function of the form (3).

C. Z templates

We determine the Z templates for pions and kaons from $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. The flavor of the daughter tracks of the D^0 in the decay $D^{*+} \to D^0(K^-\pi^+)\pi^+$ is tagged by the charge of the soft pion from the D^{*+} decay. Taken together with the large signal-to-background ratio of the Δm peak, this yields a very clean sample of pions and kaons. In addition, we sideband-subtract in Δm.

The Z calculated for the track includes hit-based and track-based corrections. We correct for two additional effects. The first effect is a momentum dependence of the Z of kaon tracks even after correction, which persists since the corrections are performed under an incorrect mass hypothesis. The second effect is that the measured dE/dx of a track depends on the tracking chamber activity in the vicinity of the track. It is necessary to reweight the templates to account for higher activity in the template samples. Several variables are useful proxies for the chamber activity. The default variable is the track *secance*, or number of r–ϕ intersections with other tracks between the COT inner and outer radius. We choose this variable because it results in the best agreement between templates derived from non-prompt and prompt D^* decays. One of the other plausible proxy variables, the total number of tracks in the event, is considered in a systematics study (Section IV D).

Because the $B \to DX$ templates contain semileptonic decays, we need to parameterize the dE/dx response for muons and electrons. For muons, we use the same templates as for pions. (Because these semileptonic components are small in the mass region of interest, systematic uncertainties associated with this choice are negligible.) For electrons, we generate a template from parametric Monte Carlo tuned on a $J/\psi \to e^+e^-$ sample.

The Z templates for the combinatorial backgrounds are taken directly from the data. A Z distribution of the appropriate species composition for the signal leakage (see the description in Section III B) is subtracted from the Z distribution of the D-sideband candidates; this process is referred to as "signal subtraction".

D. Efficiency corrections

We fit for the number of candidates in each sample component. To calculate ratios of branching fractions from the fit parameters, we proceed as follows. Let f_1 and f_2 be the fit fractions of the components whose ratio of branching

ratios $\mathcal{B}_1/\mathcal{B}_2$ we want to measure. The relationship between these quantities is

$$\frac{\mathcal{B}_1}{\mathcal{B}_2} = \frac{f_1/(\varepsilon_1\phi_1\eta_1)}{f_2/(\varepsilon_2\phi_2\eta_2)}, \tag{4}$$

where ε_i is the kinematic efficiency for a channel i, ϕ_i is the fraction of candidates in the fit mass range and η_i is the efficiency for the track species recoiling against the D in channel i to be reconstructed in the tracking algorithm for the level 1 trigger. Monte Carlo simulation can be used to determine ε_i and ϕ_i. The relative level 1 trigger efficiency η_K/η_π between K and π tracks is determined from a prompt D^+ sample.

IV. CONTROL SAMPLES

As a crosscheck of our method, we measure the independently known $\mathcal{B}(\overline{B}^0 \to D^{*+}K^-)/\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)$ and $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$. The measurement on the signal sample was not performed until all procedures, including the estimate of the systematic uncertainties, had been established on the control samples.

A. $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$ fit configuration

The fit to the $\overline{B}^0 \to D^+X$ control sample is a maximum-likelihood fit. The fit variables are the B candidate mass in the $D^+\pi^-$ mass hypothesis with no mass constraint on the D^+ candidate; and the PID variable Z defined in (1). In place of the likelihood (2) we use the extended likelihood

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

where the candidate yields μ_j in each component are fit parameters.

1. *Single-B templates*

The mass templates marked "Monte Carlo" are selected from single-B Monte Carlo according to the decay channels shown in Table I. With a few exceptions, the partially reconstructed modes are collected in the "remainder" $\overline{B}^0/B^- \to D^+X$ template. $D^+\rho^-$ and $D^{*+}\pi^-$ are split out because these components leak into the D^+K^- mass region. The branching ratio $\mathcal{B}(B^- \to D_2^*\pi^-)$ used in our Monte Carlo generation was estimated based on an upper limit. The branching ratio has since been measured [16]. The $D^{(*)}K^{(*)}$ modes are the largest group of modes in which the track recoiling against the D^+ is a kaon; we split this group from the remainder template so we can give it a kaon dE/dx template. Likewise, we split out the $D^+e^\pm X$ decays and give them an electron dE/dx template. The $D^{(*)}K^{(*)}$ and $D^+e^\pm X$ components that are "split out" of the remainder template do not have floating normalizations but are instead fixed to the remainder template with ratios determined from the Monte Carlo. Lastly, a $B_s \to D_s^-(\phi\pi^-)X$ reflection template is also included in the fit. The "parameter" column of the table summarizes the connections between the various templates.

2. *Fake-D templates*

The fake-D mass template is obtained as follows. The invariant mass under the $D^+\pi^-$ mass assumption is calculated in the sidebands of the D mass: 1.825 GeV $< m(D^+) <$ 1.84 GeV or 1.9 GeV $< m(D^+) <$ 1.915 GeV. This data is used in an unbinned mass-only fit with two components. The first component is a function of the form (3). The second component is a histogram obtained from $B^0 \to D^-X$ MC selecting on the D sideband. Figure 3 shows the results of this mass fit. The shape parameters obtained from this auxiliary fit are fixed in the $\overline{B}^0 \to D^+\pi^-$ fit.

3. *Real-D templates*

The real-D template is obtained from a binned fit of the $D\pi$ mass of wrong-sign candidates selected with the same cuts as the right-sign candidates but sideband-subtracted in the D mass. The fit function is of the form (3). The shape parameters obtained from this auxiliary fit are fixed in the $\overline{B}^0 \to D^+\pi^-$ fit; the normalization is left floating.

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

Table I: Fit components for the $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ fits. See the text for a description of the "mass p.d.f." and "dE/dx p.d.f." categories.

4. *Z templates*

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

The Z template for the fake-D component is taken directly from the D sideband data. The Z template for the real-D component is taken from the wrong-sign D data, selecting on the D peak. We divide both combinatorial-background templates into two B-mass regions (4.85 GeV $< m(B) <$ 5.4 GeV and 5.4 GeV $< m(B) <$ 6.45 GeV) to account for their B-mass dependence. In addition, signal subtraction is performed on the fake-D template to account for the leakage of poorly measured real D into the D sidebands. Because of limited statistics, sideband subtraction is not performed on the real-D Z template.

B. $\mathcal{B}(\overline{B}^0 \to D^{*+}K^-)/\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-)$ fit configuration

The fit is a non-extended maximum-likelihood fit. The $D^*\pi$ mass is calculated with the D^0 mass constrained to the PDG value. Table I lists the fit components. The mass templates marked "Monte Carlo" are selected from single-B Monte Carlo according to the decay channels shown in the table.

Because the combinatorial background under the D^{*+} is small, we cannot extract the background shape parameters from the data. We instead use the shape parameters of the fake-D background in the D^+ fit. These parameters are fixed in the $\overline{B}^0 \to D^{*+}\pi^-$ fit. Since the background is very small ($< 4\%$), the fit results are insensitive to the background.

The Z templates marked "prompt D^*" are described in Section III C. Corrections as described in Section III C are applied.

Decay modes	Fit result
$\overline{B}^0 \to D^+\pi^-$ fit	
$\overline{B}^0 \to D^+\pi^-$	13336.0 ± 145.4
$\overline{B}^0 \to D^+\pi^-(n\gamma)$	734.5 ± 119.8
$\overline{B}^0 \to D^+K^-$	1230.0 ± 57.6
$\overline{B}^0 \to D^+\rho^-$	8799.9 ± 519.2
$\overline{B}^0 \to D^{*+}\pi^-$	4118.4 ± 248.5
$D_2^*\pi^-$	fixed to $\overline{B}^0/B^- \to D^+X$
$D^{(*)}K^{(*)}$	fixed to $\overline{B}^0/B^- \to D^+X$
$D^{(*)}eX$	fixed to $\overline{B}^0/B^- \to D^+X$
$\overline{B}^0/B^- \to D^+X$	7262.7 ± 296.6
$\overline{B}_s^0 \to D_s^+X$	2156.9 ± 318.6
D^+(sb) (fake-D bg)	7361.2 ± 321.4
D^+X^+ (real-D bg)	1498.5 ± 180.1
$\overline{B}^0 \to D^{*+}\pi^-$ fit	
$D^{*+}\pi^-$	0.359 ± 0.005
$D^{*+}K^-$	0.032 ± 0.003
$D^{*+}\rho^-$	0.287 ± 0.011
$D^{*+}\pi^-(n\gamma)$	0.026 ± 0.005
remainder ($D^{*+}X$)	0.257 ± 0.011
combinatorial bg	norm $= 1 - \sum_i f_i$

Table II: Fit parameters determined for the $\overline{B}^0 \to D^+X$ and $\overline{B}^0 \to D^{*+}X$ control samples by the combined mass and Z fit. For the $\overline{B}^0 \to D^+X$ sample, the fit results are the candidate yields μ_j as in (5); for the $\overline{B}^0 \to D^{*+}X$ sample, the fit results are the fractions f_j as in (2).

Sample	$D^+K^-/D^+\pi^-$ (%)	$D^+\rho^-/D^+\pi^-$	$D^{*+}\pi^-/D^+\pi^-$
PDG 2006	6.80 ± 1.66	2.21 ± 0.68	0.81 ± 0.22
our measurement	9.21 ± 0.53 (1.38σ)	2.49 ± 0.15 (0.40σ)	1.11 ± 0.07 (1.26σ)
	$D^{*+}K^-/D^{*+}\pi^-$ (%)	$D^{*+}\rho^-/D^{*+}\pi^-$	
PDG 2006	7.76 ± 0.45	2.46 ± 0.38	
our measurement	8.39 ± 0.77 (0.71σ)	3.13 ± 0.16 (1.62σ)	

Table III: Comparison with PDG 2006 for the $\overline{B}^0 \to D^+\pi^-$ and $\overline{B}^0 \to D^{*+}\pi^-$ fits. The numbers in parentheses indicate the deviation from the PDG value in standard deviations. The errors quoted here for our measurement are statistical only; for the PDG value, the statistical and systematic uncertainties are added in quadrature.

C. Control sample fit results and comparison with world averages

Results of the likelihood fit are given in Table II. The mass projection is shown in Figures 4 and 7. The Z projection is shown in Figures 5 and 8. The four mass regions used in the Z projections are the same as described above for the D_sX template.

Table III shows our measurement of relative branching fractions for several modes after all corrections described in Section III D have been applied. We calculate the errors taking into account the correlations amongst the fit parameters. The results for the Cabibbo-suppressed modes are $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-) = 0.092 \pm 0.005(\text{stat}) \pm 0.007(\text{sys})$ and $\mathcal{B}(\overline{B}^0 \to D^{*+}K^-)/\mathcal{B}(\overline{B}^0 \to D^{*+}\pi^-) = 0.084 \pm 0.008(\text{stat})$. The determination of the systematic uncertainty on $\mathcal{B}(\overline{B}^0 \to D^+K^-)/\mathcal{B}(\overline{B}^0 \to D^+\pi^-)$ is described in Section IV D below.

In all cases, the agreement between our measurement and the world average is satisfactory. This provides validation for our method, allowing us to proceed to the signal sample.

D. Systematic uncertainties

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

1. *Combinatorial background mass shape*

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

2. *Monte Carlo smearing and scaling*

As explained in Section II, it is necessary to scale and smear the Monte Carlo mass templates to achieve agreement with the data. We assess the systematic uncertainty associated with the errors on the scale factor by varying the scale factor up and down within the error. We use the greater of the changes in fit value resulting from an upward or downward variation as the systematic uncertainty due to the scale factor. The same method is employed to determine the systematic uncertainty due to the smearing.

3. *D_sX and fake-D species composition*

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

4. *Choice of dE/dx reweighting variable*

As described in Section III C, the reweighting of the templates can be performed in several variables, each of which is a plausible proxy variable for the local hit density. The default fit uses secance, which is the most "local" of the variables we have studied: it attempts to measure the chamber activity in the vicinity of the track. A variable which behaves almost as well in reducing the difference between prompt and non-prompt templates is the total number of tracks $n_{\rm tr}$ in the event; this is also the most "global" proxy conceivable for the chamber activity. We assess a systematic uncertainty by fitting with an alternate model in which dE/dx reweighting in $n_{\rm tr}$ is used.

5. *Z residuals in the $D^+\pi^-$-dominated region*

Even after reweighting to account for the local chamber activity, residual discrepancies between the template and physics Z distributions persist, as can be seen in Figure 5(c). Our second dE/dx systematic uncertainty assessment quantifies the effect of the discrepancy between data and fit Z distributions in the $D^+\pi^-$-dominated region. We add the difference between data and fit in the region 5.26 GeV $< m(B) <$ 5.4 GeV to the Z templates for $D^+\pi^-$ and $D^+\pi^-(n\gamma)$ and use these modified templates as an alternate fit model.

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

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

V. SIGNAL SAMPLE

Having validated the method on two control samples, we are ready to proceed to the measurement of $\mathcal{B}(\overline{B}^0_s \to D_s^\pm K^\mp)/\mathcal{B}(\overline{B}^0_s \to D_s^+\pi^-)$.

A. Fit configuration

The fit is a non-extended maximum-likelihood fit for the fractions f_j of (2). The $D_s\pi$ mass is calculated without a mass constraint on the D_s.

1. *Mass templates*

The mass p.d.f.'s for the single-B modes are represented as histograms. These modes are selected from single-B Monte Carlo according to the decay channels shown in Table V. With a few exceptions, the partially reconstructed modes are collected in the "remainder" $B_s \to D_s^{(*)}X$ template. $D_s^+\rho^-$ and $D_s^{*+}\pi^-$ are split out because these components leak into the D_sK mass region; the normalization of both modes floats in the fit. The $D_s^{(*)}K^{(*)}$ modes are the largest group of "remainder" modes in which the track recoiling against the D_s^+ is a kaon; we split this group from the remainder template so we can give it a kaon dE/dx template. Likewise, we split out the $D_s^+e^\pm X$ decays and give them an electron dE/dx template. The $D_s^{(*)}K^{(*)}$ and $D_s^+e^\pm X$ components that are "split out" of the remainder template do not have floating normalizations but are instead fixed to the remainder template with ratios determined from the MC. The fully reconstructed decays $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^-K^+$ can give localized peaks in the mass plot. Since a measurement of the relative branching ratios of $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ has recently been published [17], we include templates for these decays. The templates are split according to the species recoiling against the $D_s^{(*)}$ with ratios fixed to the values in [17]. A template for $\overline{B}^0/B^0 \to D^+(K^-\pi^+\pi^+)X$ and $B^-/B^+ \to D^+(K^-\pi^+\pi^+)X$ reflections is also included, as is a template for $\Lambda_b \to \Lambda_c(pK^-\pi^+)X$ reflections. The "parameter" column of Table V summarizes the connections between the various templates.

2. *Fake-D templates*

The fake-D_s mass template is obtained as follows. The invariant mass under the $D_s^+\pi^-$ mass assumption is calculated in the sidebands of the D_s mass: 1.924 GeV $< m(D_s^+) <$ 1.939 GeV or 1.999 GeV $< m(D_s^+) <$ 2.014 GeV. This data is used in an unbinned mass-only fit with two components. The first component is a function of the form (3). The second component is a histogram obtained from the $\overline{B}^0_s \to D_s^+(\phi\pi^+)X$ MC selecting on the D_s sideband. Figure 10 shows the results of this fit. The shape parameters obtained from this auxiliary fit are fixed in the main $\overline{B}^0_s \to D_s^+(\phi\pi^+)X$ fit. Because the fake-D_s fit gives better agreement with data when performed on candidates that are not mass-constrained, we use unconstrained candidates for all fit components.

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

Table V: Fit components for the $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$ fit. See the text for a description of the "mass p.d.f." and "dE/dx p.d.f." categories.

3. *Z templates*

The Z templates marked "prompt D^*" are described in Section III C. Two-dimensional reweighting in secance and momentum is applied. The DeX template uses the parametric dE/dx simulation described in Section III C. The D^+ and Λ_c reflection templates are composed from e, μ (using π as proxy), π and K templates according to species fractions determined from the Monte Carlo samples. The reflection templates are divided into four mass regions: the region below the D_sK (4.85 GeV $< m(B_s) <$ 5.26 GeV); the D_sK region (5.26 GeV $< m(B_s) <$ 5.35 GeV); the $D_s\pi$ region (5.35 GeV $< m(B_s) <$ 5.5 GeV); the region above the $D_s\pi$ (5.5 GeV $< m(B_s) <$ 6.45 GeV). The Z template for the fake-D_s component is taken directly from the D_s sideband data. Signal subtraction is then performed on the fake-D_s template to account for the leakage of poorly measured real D_s into the D_s sidebands.

B. Fit results

Results of the likelihood fit are given in Table VI. The mass projection is shown in Figure 11; the Z projections are shown in Figure 12. The four mass regions used in the Z projections are the same as described above for the D^+ and Λ_c reflection templates.

With all corrections applied, we find $\mathcal{B}(\overline{B}_s^0 \to D_s^\pm K^\mp)/\mathcal{B}(\overline{B}_s^0 \to D_s^+\pi^-) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$. We calculate the statistical error taking into account the correlations amongst the fit parameters. The determination of the systematic uncertainty will be described in Section V C below. Within the uncertainties, this result is compatible with the analogous B^0 branching fractions.

The $\overline{B}_s^0 \to D_s^\pm K^\mp$ event yield is 109 ± 19 (statistical error only). To determine the statistical significance of the observation of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ signal, we perform a fit in which the $\overline{B}_s^0 \to D_s^\pm K^\mp$ fraction is fixed to zero. This fit is worse than the default fit by $\Delta \log L = -31.21$; the corresponding statistical significance of the $\overline{B}_s^0 \to D_s^\pm K^\mp$ signal is 7.90σ.

C. Systematic uncertainties

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

1. *Combinatorial background model*

From the high-statistics $\overline{B}^0 \to D^+X$ fit, we know that a combinatorial background model comprising two sources of background (fake D + track, real D + track) is a better model than a single-source background; because we are limited

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

Table VI: Fit parameters determined for $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$ by the combined mass and Z fit. The fit results are the fractions f_j as in (2).

in the $D_s \to \phi\pi$ case by the wrong-sign ($D_s^-\pi^-$) sample statistics, we fit without a real-D background component. We assess the systematic uncertainty associated with the background model by comparing different background models. The default model uses a single background of the form (3) with shape parameters determined from the fit of the D_s sidebands described in Section V A 2. We compare this configuration to a two-background configuration in which the real-D background shape parameters and Z distribution from the $\overline{B}^0 \to D^+X$ fit are also used.

2. *Monte Carlo smearing and scaling*

As explained in Section II, it is necessary to scale and smear the Monte Carlo mass templates to achieve agreement with the data. We assess the systematic uncertainty associated with the errors on the scale factor by varying the scale factor up and down within the error. We use the greater of the changes in fit value resulting from an upward or downward variation as the systematic uncertainty due to the scale factor. The same method is employed to determine the systematic uncertainty due to the smearing.

3. *Λ_cX, D^+X and fake-D_s species composition*

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

4. *Choice of dE/dx reweighting variable*

As described in Section III C, the reweighting of the templates can be performed in several variables, each of which is a plausible proxy variable for the local hit density. The default fit uses secance, which is the most "local" of the variables we have studied: it attempts to measure the chamber activity in the vicinity of the track. A variable which behaves almost as well in reducing the difference between prompt and non-prompt templates is the total number of tracks $n_{\rm tr}$ in the event; this is also the most "global" proxy conceivable for the chamber activity. We assess a systematic uncertainty by fitting with an alternate model in which dE/dx reweighting in $n_{\rm tr}$ is used.

5. *Z residuals in the $D_s^+\pi^-$-dominated region*

Finally we assess whether the residual Z-distribution discrepancies observed in the $\overline{B}^0 \to D^+X$ sample exist in the $\overline{B}_s^0 \to D_s^+X$ fit. We use the pion Z templates modified according to Section IV D 5: the discrepancy between $\overline{B}^0 \to D^+X$ data and fit in the $D^+\pi^-$-dominated region is added to the Z templates for $D_s^+\pi^-$ and $D_s^+\pi^-(n\gamma)$.

Effect	$D_s^\pm K^\mp / D_s^+ \pi^-$ uncertainty (%)
Z reweighting in $n_{\rm tr} : p$	0.45
Residual discrepancy in π Z template	0.38
Signal subtraction in fake-D Z template	0.29
Composition of D^+X Z template	0.21
Combinatorial bg model	0.18
MC mass scale factor	0.18
MC mass smearing	0.13
Composition of $\Lambda_c X$ Z template	0.11
Sum in quadrature	0.76

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

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

VI. CONCLUSIONS

We have presented the first observation of the decay $\overline{B}_s^0 \to D_s^\pm K^\mp$; the statistical significance of the observation is 7.90σ. The $\overline{B}_s^0 \to D_s^\pm K^\mp$ event yield is 109 ± 19 (statistical error only). We use this sample to measure $\mathcal{B}(\overline{B}_s^0 \to D_s^\pm K^\mp)/\mathcal{B}(\overline{B}_s^0 \to D_s^+ \pi^-) = 0.107 \pm 0.019(\text{stat}) \pm 0.008(\text{sys})$. This result is statistically consistent with the analogous relative B^0 branching fractions.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; the European Community's Human Potential Programme; the Slovak R&D Agency; and the Academy of Finland.

[1] R. Aleksan, I. Dunietz and B. Kayser, Z. Phys. **C54**, 653 (1992).
[2] I. Dunietz, Phys. Rev. **D52**, 3048 (1995), hep-ph/9501287.
[3] D. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005), hep-ex/0412071.
[4] CDF II uses a cylindrical coordinate system in which θ and ϕ are the polar and azimuthal angle, r is the radius from the symmetry axis, y points up, and z points in the proton direction, with the origin at the center of the detector. The transverse plane is the plane perpendicular to the z axis.
[5] A. Sill, Nucl. Instrum. Meth. **A447**, 1 (2000).
[6] A. A. Affolder *et al.*, Nucl. Instrum. Meth. **A526**, 249 (2004).
[7] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).
[8] B. Ashmanskas *et al.*, Nucl. Instrum. Meth. **A518**, 532 (2004), physics/0306169.
[9] G. Gomez-Ceballos *et al.*, Nucl. Instrum. Meth. **A518**, 522 (2004).
[10] Reference to the charge-conjugate decays is implied here and throughout the text.
[11] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).
[12] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. **B327**, 49 (1989).

[13] D. J. Lange, Nucl. Instrum. Meth. **A462**, 152 (2001).
[14] We describe the fit used in the D_s analysis; the control-sample fits are analogous, with the appropriate particles replacing the D_s.
[15] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).
[16] K. Abe *et al.*, Phys. Rev. **D69**, 112002 (2004), hep-ex/0307021.
[17] B. Aubert *et al.*, Phys. Rev. Lett. **98**, 081801 (2007), hep-ex/0604012.

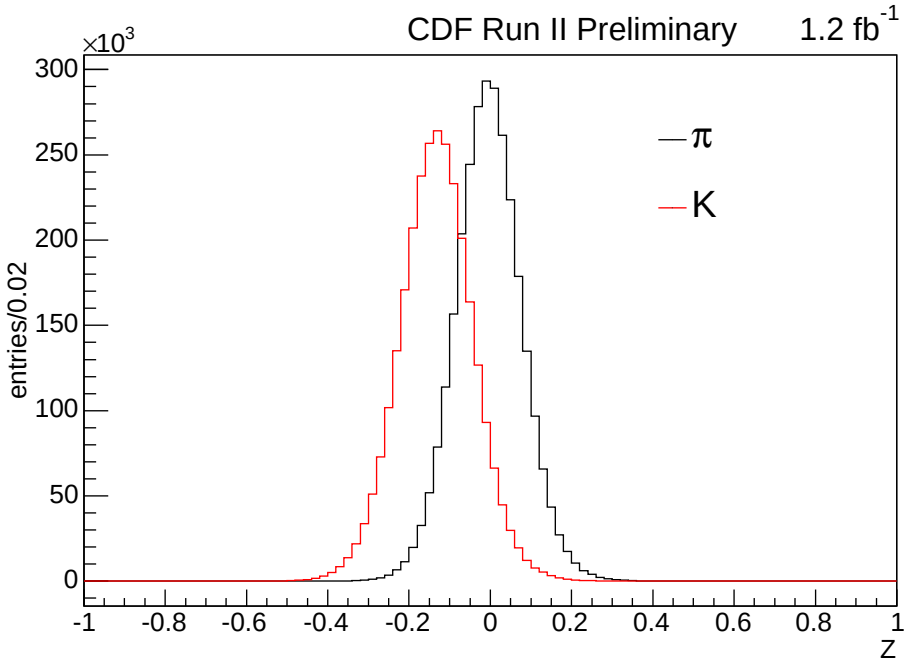

Figure 2: Z templates for K and π tracks reweighted for $\overline{B}^0 \to D^+X$

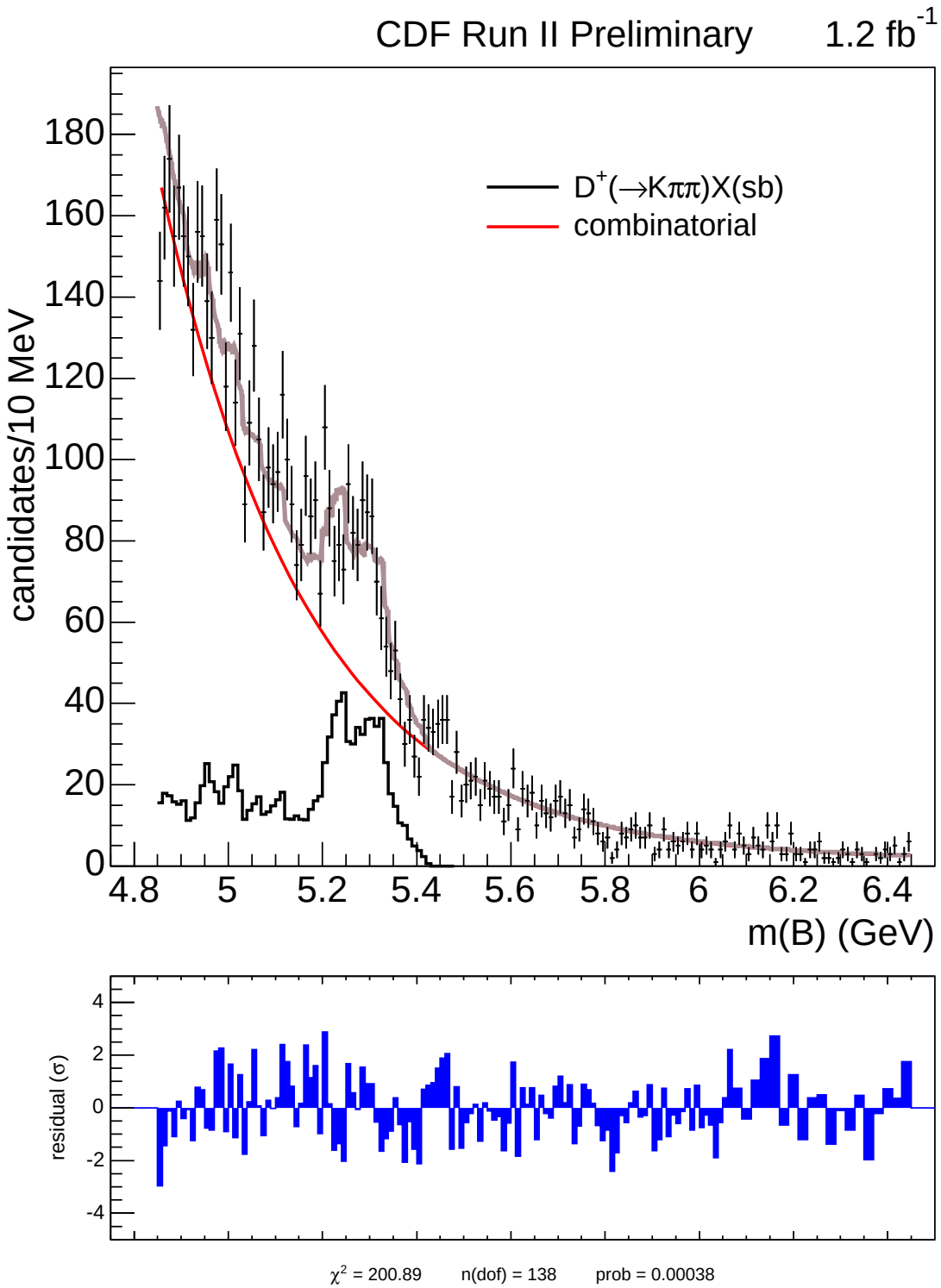

Figure 3: Fake-D^+ combinatorial template. The histogram is obtained using the D^+ sidebands. The combinatorial background shape is an exponential plus a flat background of the form (3).

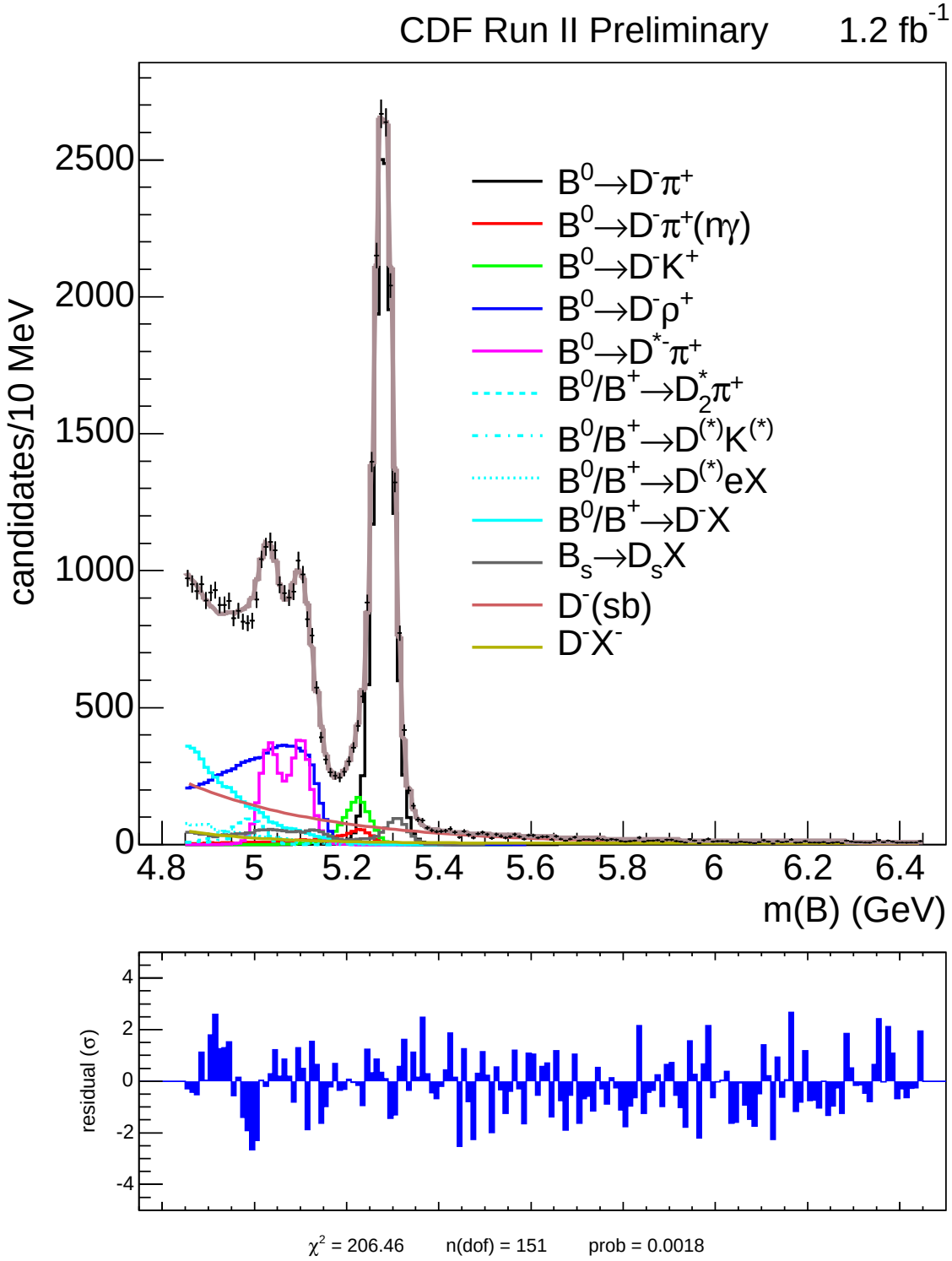

Figure 4: Mass projection of the likelihood fit results for $\overline{B}^0 \to D^+\pi^-$, $D^+ \to K^-\pi^+\pi^+$ candidates. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

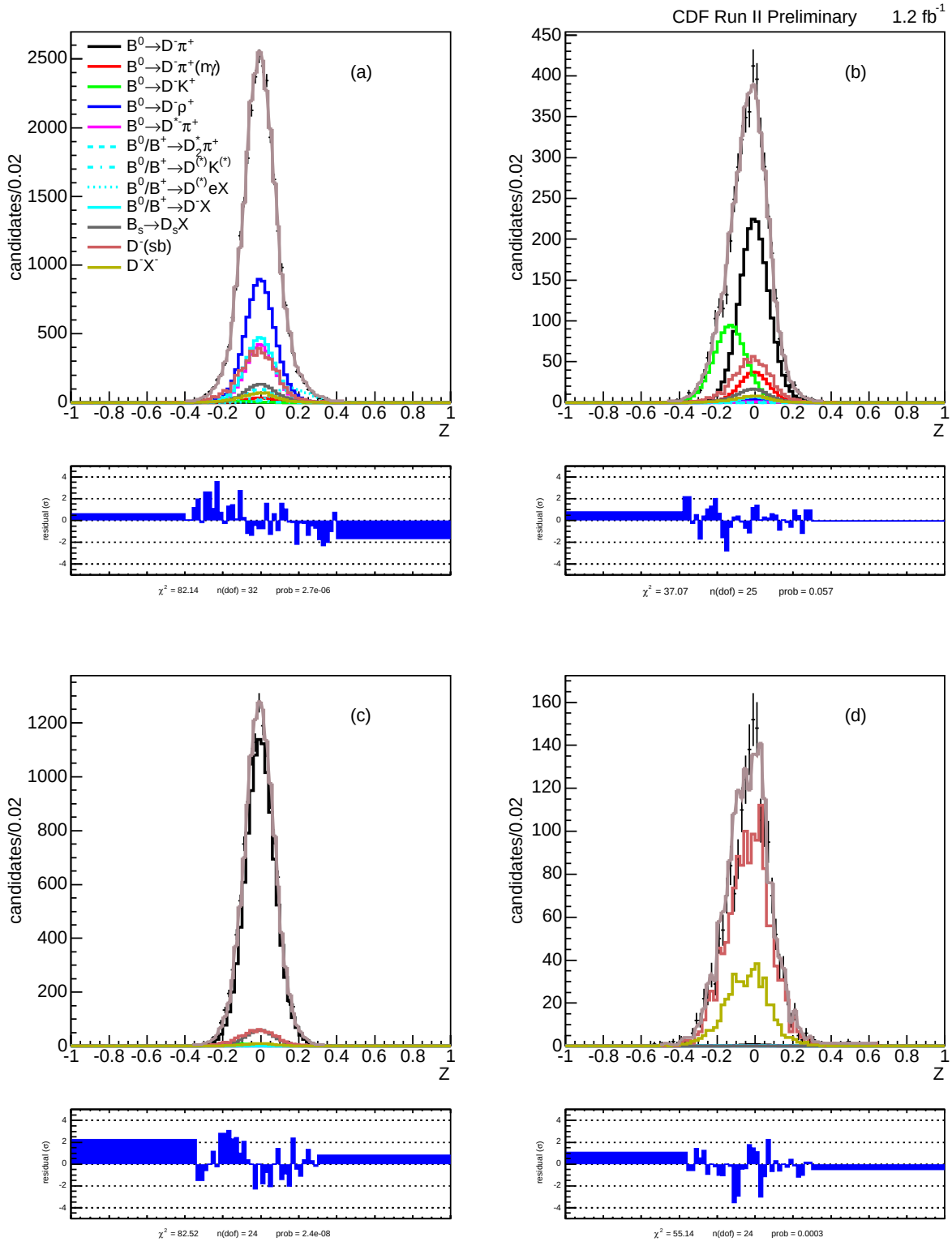

Figure 5: Z projection of the likelihood fit results for $\overline{B}^0 \to D^+\pi^-$, $D^+ \to K^-\pi^+\pi^+$ candidates. The projections are shown in four mass regions: (a) the region below the D^+K^- (4.85 GeV $< m(B) <$ 5.17 GeV); (b) the D^+K^- region (5.17 GeV $< m(B) <$ 5.26 GeV); (c) the $D^+\pi^-$ region (5.26 GeV $< m(B) <$ 5.4 GeV); and (d) the region above the $D^+\pi^-$ (5.4 GeV $< m(B) <$ 6.45 GeV). The residual plots show the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

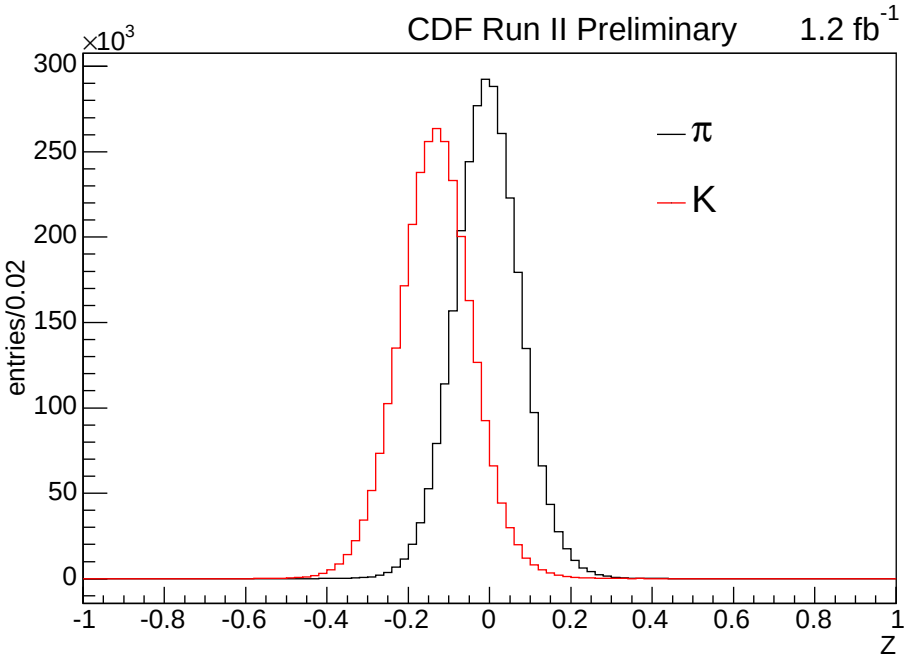

Figure 6: Z templates for K and π tracks in reweighted for $\overline{B}^0 \to D^{*+}X$

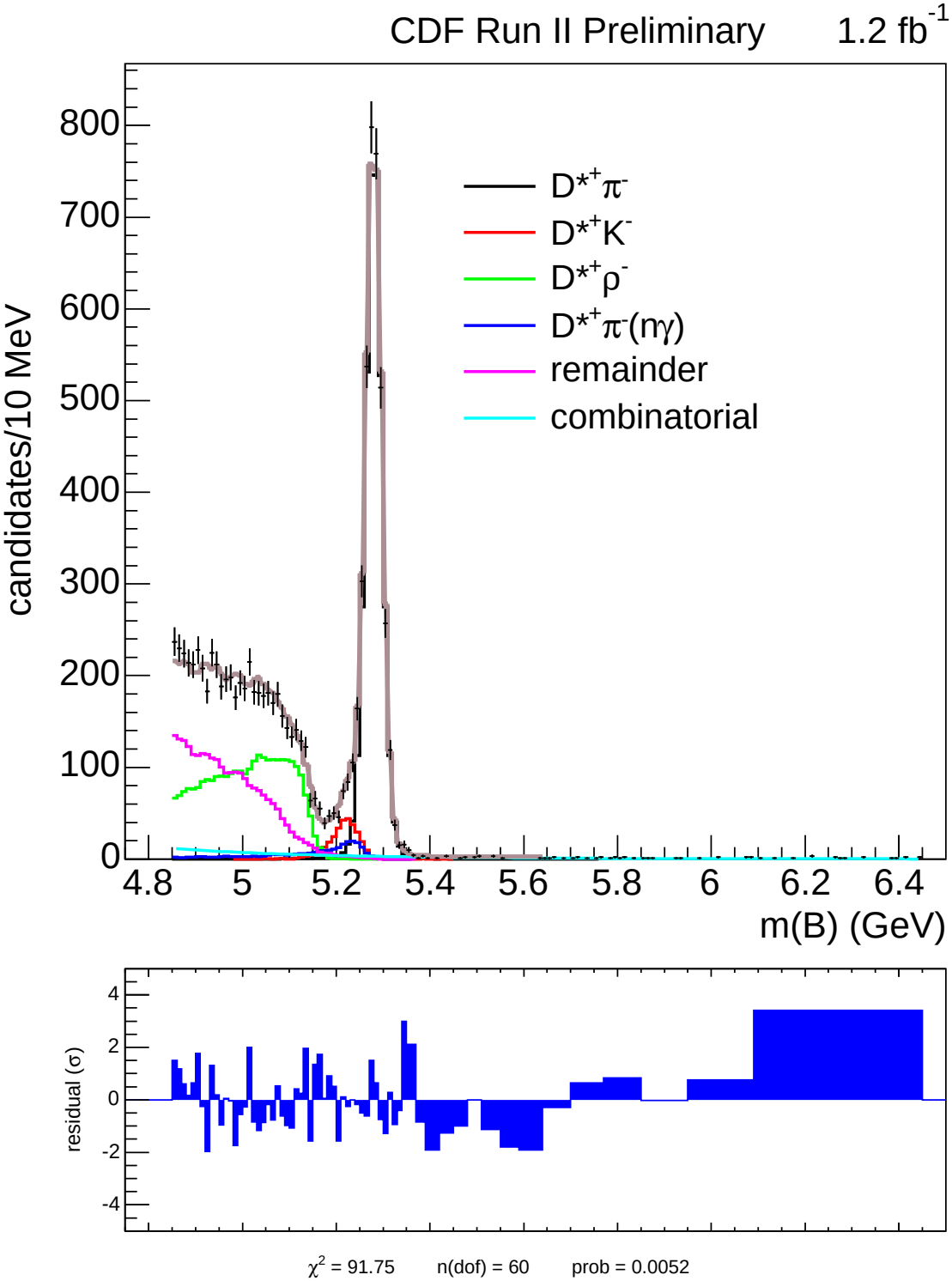

Figure 7: Mass projection of the likelihood fit results for $\overline{B}^0 \to D^{*+}\pi^-$, $D^{*+} \to D^0\pi^+$, $D^0 \to K\pi$ candidates. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

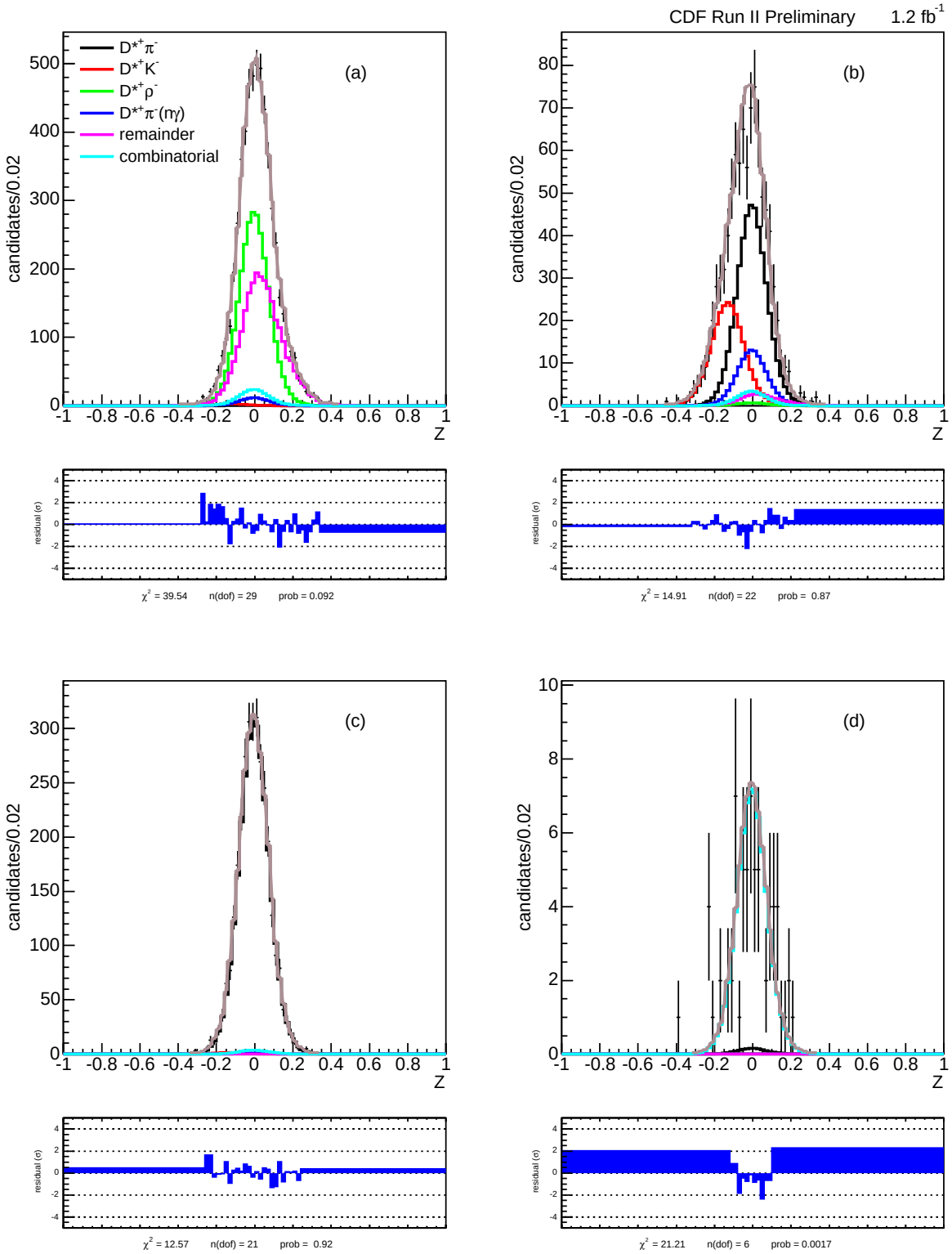

Figure 8: Z projection of the likelihood fit results for $\overline{B}^0 \to D^{*+}\pi^-$, $D^{*+} \to D^0\pi^+$, $D^0 \to K\pi$ candidates. The projections are shown in four mass regions: (a) the region below the $D^{*+}K^-$ (4.85 GeV $< m(B) <$ 5.17 GeV); (b) the $D^{*+}K^-$ region (5.17 GeV $< m(B) <$ 5.26 GeV); (c) the $D^{*+}\pi^-$ region (5.26 GeV $< m(B) <$ 5.4 GeV); and (d) the region above the $D^{*+}\pi^-$ (5.4 GeV $< m(B) <$ 6.45 GeV). The residual plots show the number of σ discrepancy (data minus fit). In the calculation of the residuals, neighboring bins are combined until the predicted number of events is > 5.

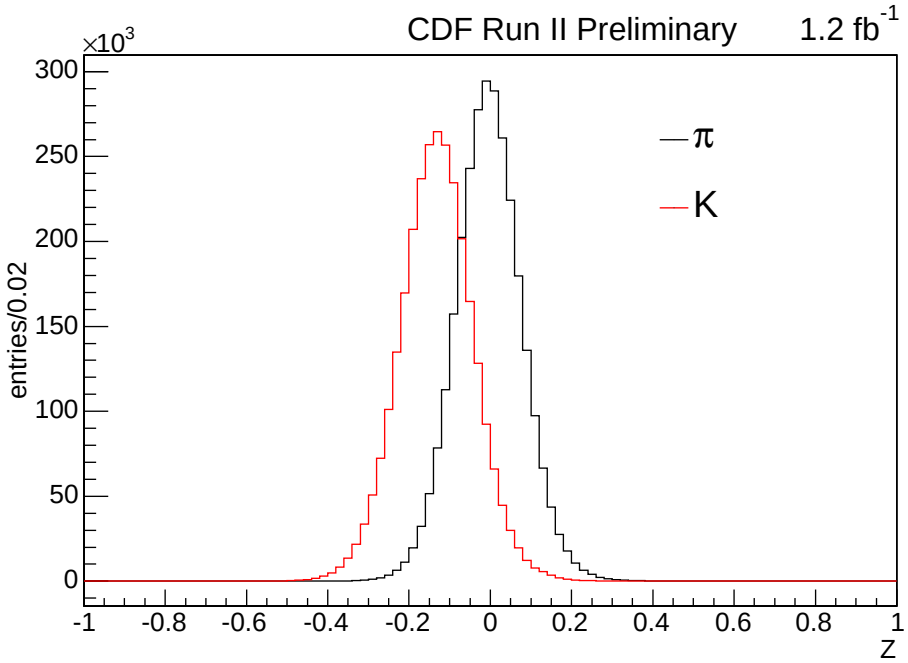

Figure 9: Z templates for K and π tracks reweighted for $\overline{B}_s^0 \to D_s^+(\phi\pi^+)X$

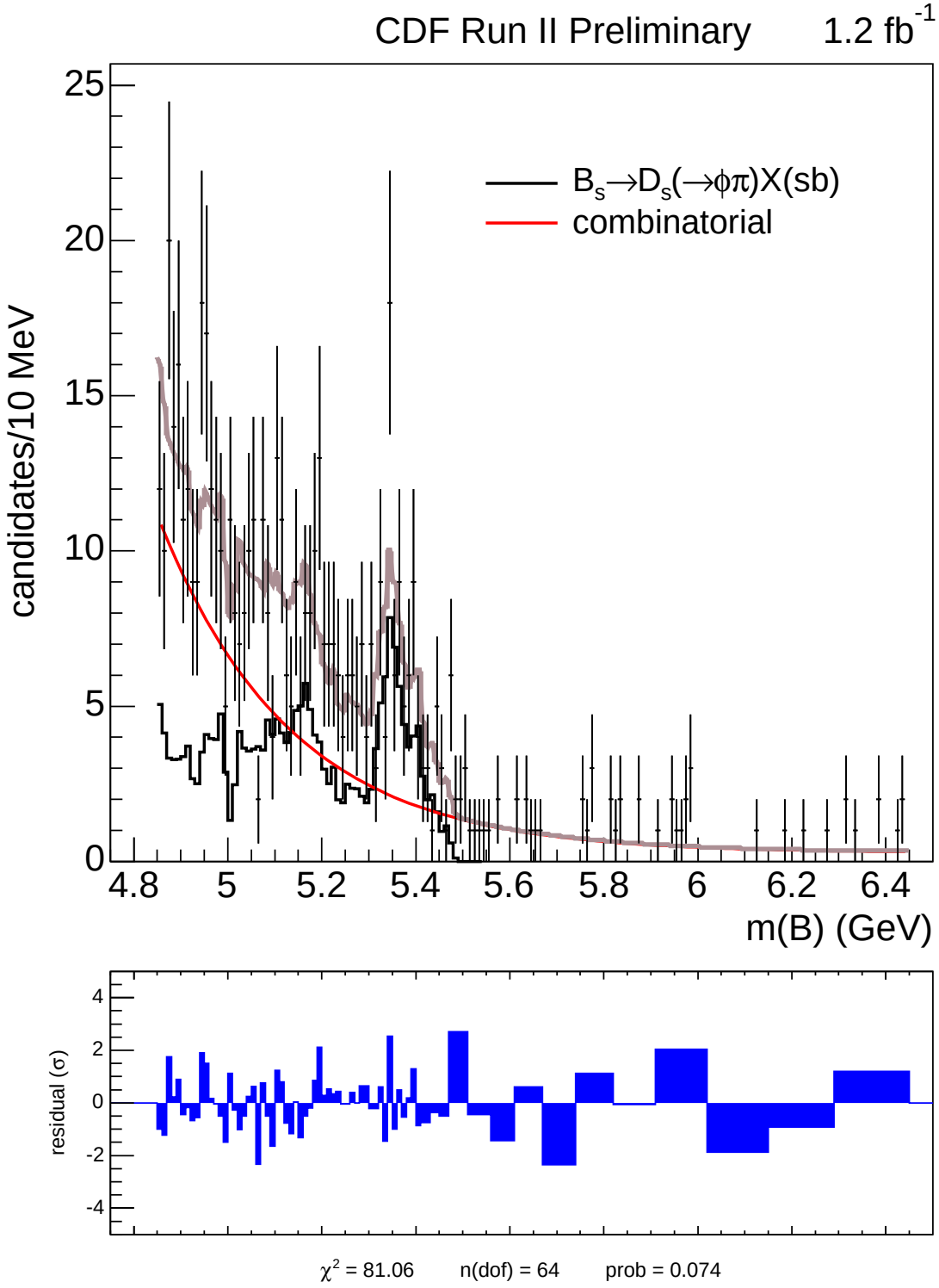

Figure 10: Fake-D_s^+ combinatorial template. The histogram is obtained using the D_s^+ sidebands. The combinatorial background shape is an exponential plus a flat background of the form (3).

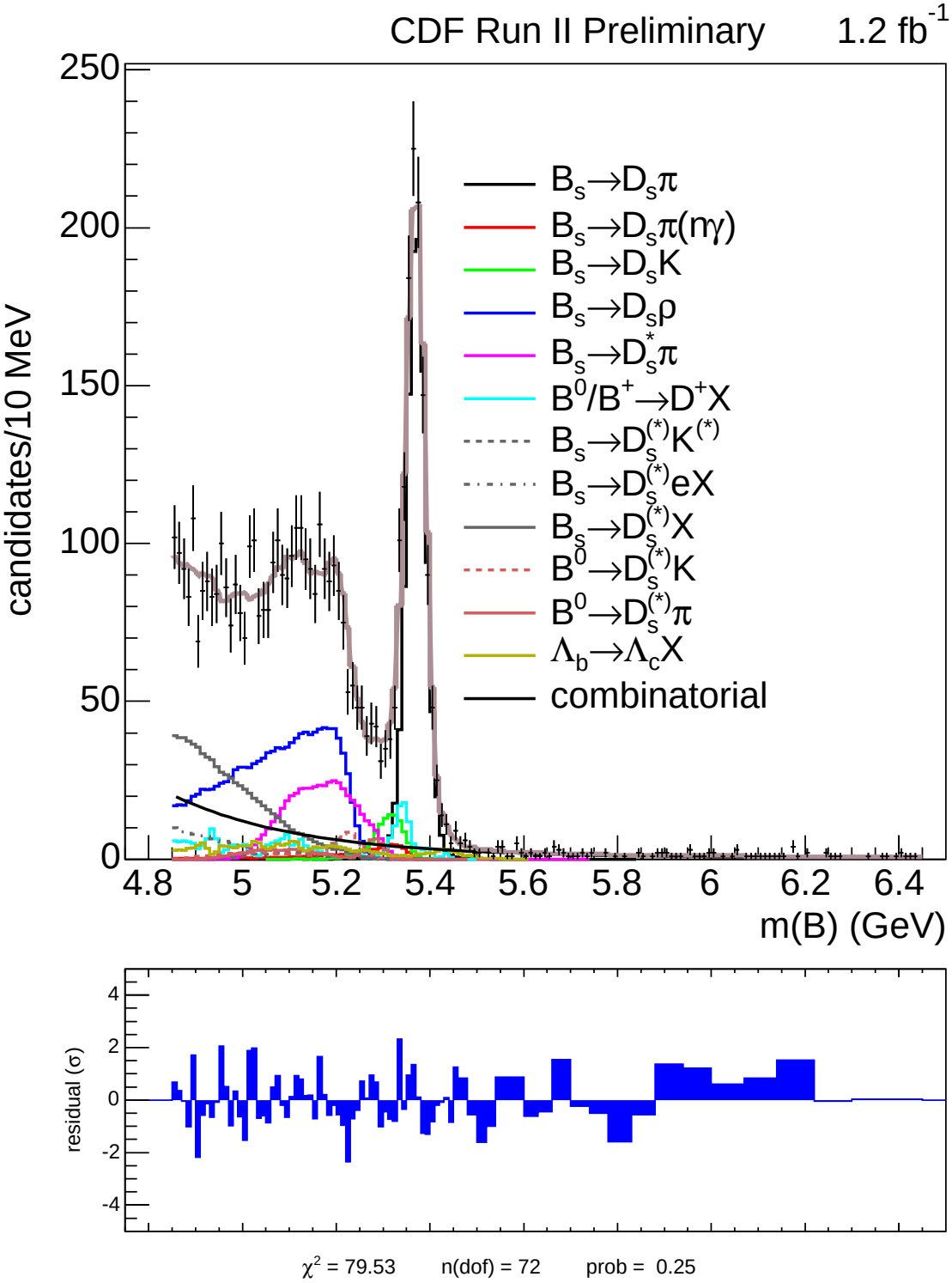

Figure 11: Mass projection of the likelihood fit results for $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$ candidates. The residual plot at the bottom shows the number of σ discrepancy (data minus fit). For the bins with low statistics, neighboring bins are combined until the predicted number of events is > 5.

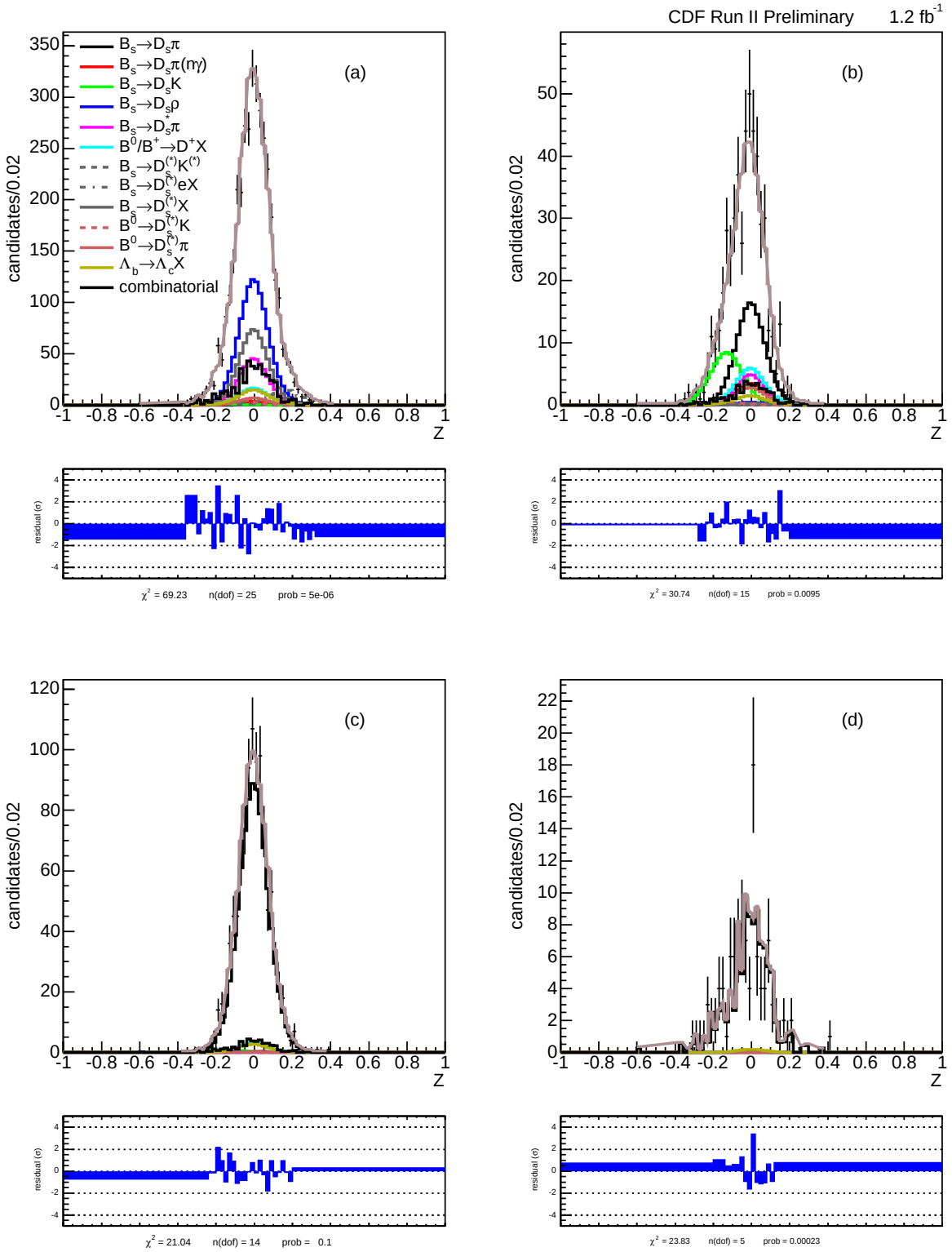

Figure 12: Z projection of the likelihood fit results for $\overline{B}_s^0 \to D_s^+(\phi\pi^+)\pi^-$ candidates. The projections are shown in four mass regions: (a) the region below the D_sK (4.85 GeV $< m(B_s) <$ 5.26 GeV); (b) the D_sK region (5.26 GeV $< m(B_s) <$ 5.35 GeV); (c) the $D_s\pi$ region (5.35 GeV $< m(B_s) <$ 5.5 GeV); and (d) the region above the $D_s\pi$ (5.5 GeV $< m(B_s) <$ 6.45 GeV). The residual plots at the bottom show the number of σ discrepancy (data minus fit). For the bins with low statistics, neighboring bins are combined until the predicted number of events is > 5.