# Study of orbitally excited ($L = 1$) $B$ mesons

**Text for the blessed web page – CDF note 8945**

The CDF Collaboration
*URL http://www-cdf.fnal.gov*
(Dated: August 21, 2008)

We study neutral orbitally excited $B$ mesons in decays to $B^{(*)+}\pi^-$. The $B^+$ is reconstructed in two independent data samples in the decay modes $B^+ \to \mathrm{J}/\psi\mathrm{K}^+$, with $\mathrm{J}/\psi \to \mu^+\mu^-$, $B^+ \to \bar{D}^0\pi^+$, with $\bar{D}^0 \to K^+\pi^-$, and $B^+ \to \bar{D}^0\pi^+\pi^+\pi^-$, with $\bar{D}^0 \to K^+\pi^-$. From an unbinned maximum likelihood fit to the combined mass difference $Q = m(B\pi) - m(B) - m_\pi$, we extract the masses and width of the narrow $B_1^0$ and $B_2^{*0}$ states:

$$m(B_1^0) = 5725.3^{+1.6}_{-2.2}\ \text{(stat.)}\ ^{+1.4}_{-1.5}\ \text{(syst.)}\ \text{MeV/c}^2$$

$$m(B_2^{*0}) = 5740.2^{+1.7}_{-1.8}\ \text{(stat.)}\ ^{+0.9}_{-0.8}\ \text{(syst.)}\ \text{MeV/c}^2$$

$$\Gamma(B_2^{*0}) = 22.7^{+3.8}_{-3.2}\ \text{(stat.)}\ ^{+3.2}_{-10.2}\ \text{(syst.)}\ \text{MeV/c}^2$$

This is currently the most precise measurement of the narrow $B^{**0}$ masses, and the first measurement of the $B_2^{*0}$ width.

*Preliminary Results for Summer 2007 Conferences*

## I. INTRODUCTION

This note describes a measurement of the first orbitally excited $B$ mesons in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF II detector at the Fermilab Tevatron. The purpose of this analysis is to more precisely measure the properties of the excited $B$ mesons, the spectroscopy of which has not been well-studied. There exist many theoretical predictions for the excited $B$ states using different models of the heavy-light quark interactions, including heavy quark effective theories. Experimental confirmation of these predictions will lead to a better understanding of how heavy quarks and light quarks interact in bound states and help the development of non-perturbative QCD.

The $B$ meson consists of a heavy $b$ quark bound to a light $u$ or $d$ quark. In the limit of infinite $b$ quark mass, the angular momentum and flavor of the light quark become good quantum numbers. The $b$ quark has a spin angular momentum $J_Q = \frac{1}{2}$, while the light quark has a total angular momentum

$$J_l = \left| L \pm \frac{1}{2} \right| \tag{1}$$

where $L$ is the orbital angular momentum of the light quark. The resulting physical hadron state has a total angular momentum of

$$\mathbf{J} = \left| J_l \pm \frac{1}{2} \right| \tag{2}$$

In the ground state of $L = 0$, $J_l = \frac{1}{2}$ and $\mathbf{J} = 0$ ($B$) or $\mathbf{J} = 1$ ($B^*$).

The first excited state of the $B$ meson occurs when the light quark has an orbital angular momentum of $L = 1$. This leads to two isodoublets of excited states, one with $J_l = \frac{1}{2}$, $\mathbf{J} = 0$ or 1, and another with $J_l = \frac{3}{2}$, $\mathbf{J} = 1$ or 2. These four states are collectively referred to as $B^{**}$. The states belonging to the first doublet are called $B_0^*$ and $B_1^*$, and they decay to $B^{(*)}\pi$ via a $S$-wave transition. Consequently, these states are expected to be very broad and have not yet been observed. The states belonging to the second doublet are called $B_1$ and $B_2^*$, and they decay to $B^{(*)}\pi$ via a $D$-wave transition; therefore these two states are much narrower than the $J_l = \frac{1}{2}$ states. The decay $B_1 \to B\pi$ is forbidden by angular momentum and parity conservation, while both $B_2^* \to B\pi$ and $B_2^* \to B^*\pi$ are allowed. Tab. I summarizes the $L = 1$ states and their decays. The $B$ spectrum is depicted in Fig. 1. A similar spectrum exists for the $B_s^{**0}$, which consists of the $b$ and $s$ quarks. The narrow $B_{s1}^0$ and $B_{s2}^{*0}$ have been measured in decays to $B^{(*)+}K^-$ [1]. When the kaon is misreconstructed as a pion, the $B_s^{**}$ states will reflect into the $B^{**}$ mass distribution.

TABLE I: Properties of the four orbitally excited ($L = 1$) $B^{**}$ states.

| State | $J_l$ | $\mathbf{J}^P$ | Width | Decay |
|---|---|---|---|---|
| $B_0^*$ | $\frac{1}{2}$ | $0^+$ | broad | $(B\pi)$ |
| $B_1^*$ | $\frac{1}{2}$ | $1^+$ | broad | $(B^*\pi)$ |
| $B_1$ | $\frac{3}{2}$ | $1^+$ | narrow | $(B^*\pi)$ |
| $B_2^*$ | $\frac{3}{2}$ | $2^+$ | narrow | $(B\pi, B^*\pi)$ |

Decays to $B^*\pi$ are immediately followed by the decay of $B^*$ to $B$ by emission of a photon with energy $E(\gamma) = 45.78 \pm 0.35$ MeV/c$^2$ [2] virtually 100% of the time. These low energy photons cannot be separated from the large amount of other electromagnetic background sources in the CDF II detector; consequently, the reconstructed mass of the $B^{**}$ states is decreased by the energy of this photon. This adds extra complication to the mass spectrum, as the $B_2^* \to B^*\pi$ mass peak will be displaced from the $B_2^* \to B\pi$ mass peak by the photon energy. The theoretical relative branching ratio of the two $B_2^*$ decay modes is based on observations of the charm sector, and found to be [2]

$$\frac{BR(B_2^* \to B\pi)}{BR(B_2^* \to B^*\pi)} = 1.1 \pm 0.3 \tag{3}$$

The four orbitally excited $B^{**}$ states exist for both $B^\pm$ and $B^0$, denoted respectively by $B^{**\pm}$ and $B^{**0}$. This note presents the reconstruction of $B^{**0}$ states decaying to $B^{(*)\pm}\pi^\mp$. For ease of reference, $B^{**}$ will be used in place of $B^{**0}$ and $B_1$ and $B_2^*$ in place of $B_1^0$ and $B_2^{*0}$ from this point on. The narrow $B_1^0$ and $B_2^{*0}$ states were first observed by the LEP experiments [3–5]. More recently, the narrow $B^{**0}$ states have been precisely measured by the DELPHI [6] and DØ experiments [7]. A previous CDF II analysis also measured these states in $370 \pm 20$ pb$^{-1}$ of data [8]. Thus far, no experiment has measured the widths of the $B^{**0}$ mesons. One theorectical prediction for the narrow state widths is shown in Tab. II, and will be used in this analysis to constrain the $B_1$ width.

TABLE II: Theoretical predictions for the intrinsic widths of the two narrow $B^{**}$ states [9].

| Name | Prediction |
|---|---|
| $\Gamma(B_2^*)$ | $16 \pm 6$ MeV/c$^2$ |
| $\Gamma(B_1)/\Gamma(B_2^*)$ | 0.9 (for pure $D$-wave) |

## II. DATA SAMPLE & EVENT SELECTION

The current analysis uses data collected with the CDF II detector by the $J/\psi \to \mu\mu$ and two displaced track triggers. The $J/\psi$ trigger starts from two tracks which have information in the muon chambers and $p_{\mathrm{T}}$ larger than 1.5 (2.0) GeV/c$^2$ for different subdetectors and/or data taking periods. For the final online selection, the two tracks are required to have an opposite charge, an opening angle smaller than 135° and an invariant mass around the world average $J/\psi$ mass. The two track trigger selects events based on the large impact parameter of the tracks coming from $b$ hadron decays. It requires two tracks with an impact parameter in the range from 120 $\mu$m to 1 mm together with a minimal transverse momentum of each track and minimal scalar sum of the transverse momenta of the two tracks. This analysis utilizes data corresponding to an integrated luminosity of 1.7 fb$^{-1}$ for both triggered datasets.

The offline reconstruction starts with reconstructing $B^+$ candidates in the $J/\psi K^+$ and $\overline{D}^0(3)\pi$ decay modes with $J/\psi \to \mu^+\mu^-$ and $\overline{D}^0 \to K^+\pi^-$ [11]. The $B^+ \to J/\psi K^+$ is reconstructed from the $J/\psi$ trigger data while the $B^+ \to \overline{D}^0(3)\pi$ is from the two track trigger data.

Reconstructed candidates are preselected using separate neural networks for each of the three channels. The neural networks are based on the NeuroBayes [10] package. They combine topological, kinematic, and particle identification quantities of the $B^+$ and its daughters. The neural networks for the $B^+ \to J/\psi K^+$ and $B^+ \to \overline{D}^0\pi^+$ are taken from the $B_s^{**}$ analysis [1]. The network for $B^+ \to J/\psi K^+$ is trained on Monte Carlo events with the full CDF detector simulation for the signal patterns and data from the $B^+$ mass sidebands for the background patterns. For the $B^+ \to \overline{D}^0\pi^+$ channel we use the possibility to train the Neural Network with weights and use only data from both the $B^+$ mass signal and sideband regions, and subtract background statistically during neural network training. For selecting the $B^+ \to \overline{D}^0\pi^+\pi^+\pi^-$ we set up a further neural network which was trained using Monte Carlo events for the signal pattern and data from the upper $B^+$ mass sideband for the background pattern. The preselection is done by cutting on the neural network output, which is chosen to keep as much signal as possible, while removing a large part of the background. The invariant mass distributions of the $B^+$ candidates in the three channels are shown in Fig. 2. In total, 51500 $B^+$ signal events are selected in the $J/\psi K^+$ decay channel, 40100 in the $\overline{D}^0\pi^+$ channel, and 11000 in the $\overline{D}^0 3\pi$ channel.

The $B^{**}$ candidates are constructed by combining $B^+$ candidates with a track which is assumed to be a pion. For the selection, three neural networks are trained on a combination of Monte Carlo events for the signal pattern and data for the background pattern. The data for the background pattern are taken from the entire $Q$ range of 0 to 1.0 GeV/c$^2$, where the $Q$ value is defined as $Q = m\,(B^+\pi^-) - m\,(B^+) - m_\pi^-$, with $m$ denoting the invariant mass of the $B^+\pi^-$ pair, the $B^+$ candidate, and the pion respectively. Over the entire $Q$ range, the $B^{**}$ signal contribution in the background samples is marginal and can therefore be neglected. To avoid biasing the network training, the Monte Carlo events are generated with the same $Q$ distribution as the data. The distributions of the neural network outputs can be found in Fig. 3. The plots show a good signal and background separation which is true for well-trained neural networks.

For the final selection we cut on the number of candidates per event and on the output of the neural networks. The cut on the number of candidates is fixed to be the same for all three $B^+$ decay channels, and requires fewer than six $B^{**}$ candidates in an event. The cut on the neural network output is chosen to maximize $N_{MC}/\sqrt{N_{data}}$, which is proportional to the significance $S/\sqrt{S+B}$. The optimization is done by counting the number of Monte Carlo events $N_{MC}$ and the number of data events $N_{data}$ in the $Q$ value range from 0.2 GeV/c$^2$ to 0.4 GeV/c$^2$ for a given cut on the network output. The best cut is found to be 0.5 for the decay $B^{**0} \to B^+\pi^- \to (J/\psi K^+)\pi^-$, 0.7 for $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+)\pi^-$, and 0.75 for $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+\pi^+\pi^-)\pi^-$. Fig. 4 shows the $Q$ distributions of the $B^{**}$ candidates for these network cuts.

In view of the combined fit performed for the analysis, we also optimized the selection using the combined significance, which is a function of all three network outputs. In this case, the cuts on the network outputs are optimized simultaneously when the combined significance is maximized. The best cuts on the network outputs for the combined significance are 0.85 for the decay $B^{**0} \to B^+\pi^- \to (J/\psi K^+)\pi^-$, 0.7 for $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+)\pi^-$, and 0.75 for $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+\pi^+\pi^-)\pi^-$. Fig. 5 shows the $Q$ distributions of the $B^{**}$ candidates at these network cuts.

## III. $B^{**}$ FIT DESCRIPTION

To reduce the effects of detector and mass resolution, we fit the $B^{**}$ mass difference $Q$. The expected $B^{**}$ signal structure was described in Sec. I. We fit for three peaks, one for each of the decays $B_2^* \to B\pi$, $B_2^* \to B^*\pi$, and $B_1 \to B^*\pi$. Each peak is modeled as a non-relativistic Breit-Wigner convoluted with a double Gaussian detector resolution model. Theoretical predictions enter into the fit as Gaussian constraints to the likelihood. These predictions and their Gaussian uncertainties are:

- Ratio of $B_1$ and $B_2^*$ widths: $\frac{\Gamma(B_1)}{\Gamma(B_2^*)} = 0.9 \pm 0.2$
- Energy of $B^*$ photon: $E(\gamma) = 45.78 \pm 0.35$ MeV/c$^2$
- Ratio of $B_2^*$ branching fractions: $\frac{BR(B_2^* \to B\pi)}{BR(B_2^* \to B^*\pi)} = 1.1 \pm 0.3$
- Number of misreconstructed $B_{s1}^0$ events in combined data: $24 \pm 12$
- Number of misreconstructed $B_{s2}^{*0}$ events in combined data: $62 \pm 31$

We derive the shape of misreconstructed $B_s^{**}$ events from Monte Carlo simulations, and fix these shapes in the fit to data. The number of misreconstructed $B_s^{**}$ events is constrained to expectations from the $B_s^{**}$ analysis [1]. The background is modeled by a smooth function, a power law times an exponential. There is a small fixed component to the background at high mass, but the background shape under the signal region is left floating in the fit to data. Tests of the fit show that it is stable, with negligible fit bias on the $B^{**}$ signal parameters.

To cross-check that the $B^{**}$ signal is consistent in each $B^+$ decay channel, we first fit each of the individually optimized $B^+$ samples separately. This cross-check uses only the first 1.3 fb$^{-1}$ of data in the two $D\pi$ decay modes, as the remaining data was not available at the time. The fit results to data are shown in Fig. 6. The $B^{**}$ signal parameter values are given in Tab. III. The signal location is consistent among the three channels within statistical uncertainties.

TABLE III: $B^{**}$ signal parameters from fits to each of the $B^+$ decay channels separately. Overall the signal location is consistent among the decay channels. These fits are only used as a cross-check of the final $B^{**}$ fit to the combined events.

| Parameter | $J/\psi K$ channel | $\bar{D}^0\pi$ channel | $\bar{D}^0 3\pi$ channel |
|---|---|---|---|
| $B_2^*$ $Q$ (MeV/c$^2$) | $321 \pm 3$ | $320 \pm 3$ | $324 \pm 3$ |
| $m(B_2^*) - m(B_1)$ (MeV/c$^2$) | $13 \pm 4$ | $14 \pm 3$ | $17 \pm 3$ |
| $\Gamma(B_2^*)$ (MeV/c$^2$) | $35 \pm 9$ | $20 \pm 6$ | $21 \pm 6$ |
| Number of $B_1 \to B^*\pi$ events | $260 \pm 102$ | $208 \pm 57$ | $145 \pm 49$ |
| Number of $B_2^* \to B\pi$ events | $454 \pm 78$ | $127 \pm 31$ | $112 \pm 24$ |

## IV. RESULTS

For the final $B^{**}$ fit, we combine the events in all $B^+$ decay channels after optimizing the signal samples for the combined significance. The result of this fit is shown in Fig. 7. The $\chi^2$ probablity of the fit is 73% in the range $Q \in [0., 0.5]$ GeV/c$^2$, and 54% over the full range $Q \in [0., 1.0]$ GeV/c$^2$. The fit results for the $B^{**}$ signal parameters are shown in Tab. IV. The data is consistent with having only two signal peaks of widely different widths, so the interpretation of three signal peaks is motivated by theoretical predictions. The signal is consistent with theoretical predictions, including those entered as Gaussian constraints in the likelihood. We do not measure the yield of $B$ mesons from $B^{**}$ decay, so for the number of events in each signal peak we quote only statistical uncertainties.

## V. SYSTEMATIC UNCERTAINTIES

Sources of systematic uncertainty on the mass-difference and width measurements include: mass scale, mass-dependent efficiency for the signal, fit model bias, assumptions entered as Gaussian constraints in the fit, choice of background and resolution models, and location and amount of the $B^{**0}$ broad states. The effect of each systematic is summarized in Table V.

TABLE IV: Final $B^{**}$ signal parameters from the fit to the combined events. We use asymmetric uncertainties as the likelihood is asymmetric for some parameters.

| Parameter | Value | Stat. uncertainties |
|---|---|---|
| $B_2^*$ $Q$ (MeV/c$^2$) | 321.5 | $(+1.7, -1.8)$ |
| $\Gamma(B_2^*)$ (MeV/c$^2$) | 22.7 | $(+3.8, -3.2)$ |
| $m(B_2^*) - m(B_1)$ $Q$ (MeV/c$^2$) | 14.9 | $(+2.2, -2.5)$ |
| Number of $B_1 \to B^*\pi$ events | 503 | $(+75, -68)$ |
| Number of $B_2^* \to B\pi$ events | 384 | $(+48, -45)$ |
| Number of $B_2^* \to B^*\pi$ events | 351 | $(+48, -45)$ |

TABLE V: Systematic uncertainties of the $B^{**0}$ parameters. Each row corresponds to one source of systematic uncertainty. The columns show the resulting uncertainties for the three $B^{**0}$ signal parameters. The final row shows the total systematic uncertainty on each parameter. Uncertainties are in units of MeV/c$^2$.

| Source | $Q(B_2^{*0})$ | $\Gamma(B_2^{*0})$ | $m(B_2^{*0}) - m(B_1^0)$ |
|---|---|---|---|
| Mass scale | $\pm 0.20$ | – | $< 0.01$ |
| Efficiency | $^{+0}_{-0.03}$ | $^{+0.40}_{-0}$ | $^{+0}_{-0.30}$ |
| Fit constraints | $^{+0.35}_{-0.29}$ | $^{+2.09}_{-1.45}$ | $^{+0.65}_{-0.94}$ |
| Fit bias | $^{+0}_{-0.12}$ | $^{+0.44}_{-0}$ | $^{+0}_{-0.21}$ |
| Resolution | $< 0.01$ | $^{+0}_{-0.40}$ | $< 0.01$ |
| Background | $^{+0.24}_{-0}$ | $^{+0}_{-1.63}$ | $^{+0.16}_{-0}$ |
| Broad states | $^{+0.65}_{-0.51}$ | $^{+2.27}_{-9.92}$ | $^{+0.89}_{-0.95}$ |
| Total | $^{+0.81}_{-0.63}$ | $^{+3.14}_{-10.16}$ | $^{+1.11}_{-1.39}$ |

To determine the mass scale uncertainty, we reconstruct $\psi(2S) \to J/\psi\pi^+\pi^-$, $J/\psi \to \mu^+\mu^-$, which has a similar $Q$ value as the $B^{**0}$ decays. We compare our measured $Q$ value to the world average $Q$ value [2] and take the difference as the mass scale uncertainty. To evaluate the effect of changing signal efficiency with $Q$ value, we generate pseudo-experiments that include the efficiency and then model these experiments with the default fit.

Tests of the fit model on generated pseudo-experiments show a small fit bias on the $B^{**0}$ signal parameters, which is included as a systematic uncertainty. Signal parameters entered as Gaussian constraints in the fit model contribute to the fit uncertainty. To determine their systematic contribution, we refit the data with these constrained parameters fixed. This fit returns the statistical fit uncertainties, which we subtract in quadrature from the total fit uncertainties to obtain the systematic contribution.

To estimate the uncertainties due to the choice of background and resolution models, we generate pseudo-experiments using an alternate background parameterization and increased resolution width. The background is also well-modeled by the sum of a broad Breit-Wigner and the product of a power law and an exponential function. From comparisons of the detector resolution in data and Monte Carlo for the $\psi(2S)$ sample, we expect the Monte Carlo to underestimate the resolution by no more than 20%. The pseudo-experiments are fit and the results compared with those from pseudo-experiments generated assuming the default model; the distribution of the differences between fit results is modeled by a Gaussian whose mean is assigned as systematic uncertainty.

To account for the possible effect of the yet-unobserved $j_q = \frac{1}{2}$ $B^{**0}$ states on our measurement of the $j_q = \frac{3}{2}$ $B^{**0}$ states, we add two broad Breit-Wigner shapes to our background model. We give these states the same width and vary it between 100 and 200 MeV/c$^2$. The $Q$ values of the states are independently varied between the range 240 to 360 MeV/c$^2$, the region around the narrow $B^{**0}$ peaks. We refit the data for various masses and widths of the broad states, with the normalizations of the broad Breit-Wigners as additional free parameters in the fit model. We then take the largest variation in the narrow $B^{**0}$ parameters from any configuration of broad states as the systematic uncertainty due to the $B^{**0}$ broad states.

## VI. SUMMARY

Using three fully reconstructed decay modes, $B^{**0} \to B^+\pi^- \to (J/\psi K^+)\pi^-$, $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+)\pi^-$, and $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+\pi^+\pi^-)\pi^-$, we observe the two narrow $B^{**0}$ states in agreement with previous measurements and theoretical predictions. The results of this study show:

- $m(B_2^{*0}) - m(B) - m_\pi = 321.5^{+1.7}_{-1.8}$ (stat.) $^{+0.9}_{-0.7}$ (syst.) MeV/c$^2$
- $m(B_2^{*0}) - m(B_1^0) = 14.9^{+2.2}_{-2.5}$ (stat.) $^{+1.2}_{-1.4}$ (syst.) MeV/c$^2$
- $\Gamma(B_2^*) = 22.7^{+3.8}_{-3.2}$ (stat.) $^{+3.2}_{-10.2}$ (syst.) MeV/c$^2$

The ratio of the widths, $\frac{\Gamma(B_1)}{\Gamma(B_2^*)}$, while not measured in this analysis, is consistent with a theoretical prediction of 0.9.

The $B_2^*$ $Q$ value is converted into an absolute mass by adding the $B^+$ and charged pion masses to the $Q$ value. There is a small uncertainty on the $B^+$ world average mass, $B^+ = 5279.1 \pm 0.4$ MeV/c$^2$ [2], which is added in quadrature to the systematic uncertainty of the absolute $B_2^*$ mass. The $B_1$ absolute mass may be obtained by subtracting $m(B_2^{*0}) - m(B_1^0)$ from the $B_2^*$ $Q$ value, and then adding the $B^+$ and charged pion masses. This results in absolute mass values of:

- $m(B_2^{*0}) = 5740.2^{+1.7}_{-1.8}$ (stat.) $^{+0.9}_{-0.8}$ (syst.) MeV/c$^2$
- $m(B_1^0) = 5725.3^{+1.6}_{-2.2}$ (stat.) $^{+1.4}_{-1.5}$ (syst.) MeV/c$^2$

Because we are not measuring $B^{**}$ production, we quote the number of events with only statistical uncertainties. The number of events in each peak is found to be:

- $N(B_2^{*0} \to B^+\pi^-) = 385^{+48}_{-45}$ (stat.)
- $N(B_2^{*0} \to B^{*+}\pi^-) = 351^{+48}_{-45}$ (stat.)
- $N(B_1^0 \to B^*\pi) = 503^{+75}_{-68}$ (stat.)

The ratio $\frac{BR(B_2^* \to B\pi)}{BR(B_2^* \to B^*\pi)}$, while not measured in this analysis, is consistent with the theoretical prediction of 1.1.

This is the most precise measurement of the narrow $B^{**0}$ states to date. This is also the first measurement of the $B_2^{*0}$ width.

## 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 under contract HPRN-CT-2002-00292; and the Academy of Finland.

[1] CDF Collaboration, "Observation of orbitally excited $L = 1$ $B_s$ mesons", CDF Note 8468 (2006).
[2] W. M. Yao *et al.* [Particle Data Group], "Review of particle physics," J. Phys. G **33**, 1 (2006).
[3] ALEPH Collaboration, "Resonant Structure and Flavor-tagging in the $B\pi^\pm$ System Using Fully Reconstructed $B$ Decays," contribution to the International Europhysics Conference on High Energy Physics, Brussels, Belgium (1995).
[4] DELPHI Collaboration, "Observation of Orbitally Excited $B$ Mesons," CERN-PPE/94-210 (1994); DELPHI Collaboration, "Observation of Orbitally Excited $B$ and $B_s$ Mesons," DELPHI 95-105 PHYS 540 (1995).
[5] R. Akers *et al.* [OPAL Collaboration], "Observations of pi - B charge - flavor correlations and resonant B pi and B K production," Z. Phys. C **66**, 19 (1995).
[6] Delphi Collab., Z. Albrecht et al., DELPHI 2004-025 CONF 700; Z. Albrecht, *Analysis of Excited B-Mesons*, PhD-Thesis, IEKP-KA/03-16, Universität Karlsruhe, 2003; M. Moch, *Study of $B^+$-Mesons and of Excited b-Hadron Properties*, PhD-Thesis, IEKP-KA/2004-16, Universität Karlsruhe, 2004. Both theses available at `http://www-ekp.physik.uni-karlsruhe.de/pub/web/thesis`.
[7] V. M. Abazov *et al.*, "Observation and Properties of L=1 $B_1$ and $B_2*$ Mesons," arXiv:0705.3229 [hep-ex].

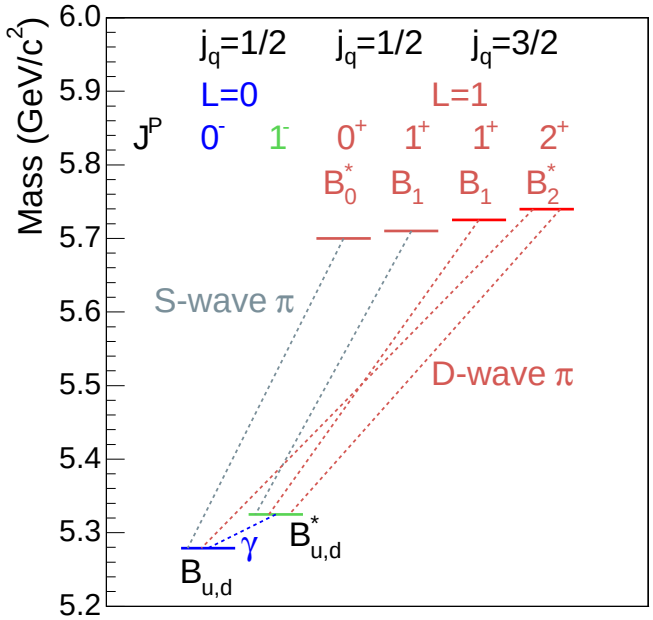


FIG. 1: Predicted spectrum and dominant decays of the lowest lying $B$ meson states.


[8] CDF Collaboration, "Observation of orbitally excited ($L = 1$) $B$ mesons", CDF note 7938 (2005).

[9] A. F. Falk and T. Mehen, "Excited heavy mesons beyond leading order in the heavy quark expansion," Phys. Rev. D **53**, 231 (1996) [arXiv:hep-ph/9507311].

[10] M. Feindt, A Neural Bayesian Estimator for Conditional Probability Densities, physics/0402093, 2004; M. Feindt and U. Kerzel, Nucl. Instrum. Methods A **559**, 190 (2006).

[11] Unless otherwise noted, any reference to a specific charge state implies the charge conjugate state as well.

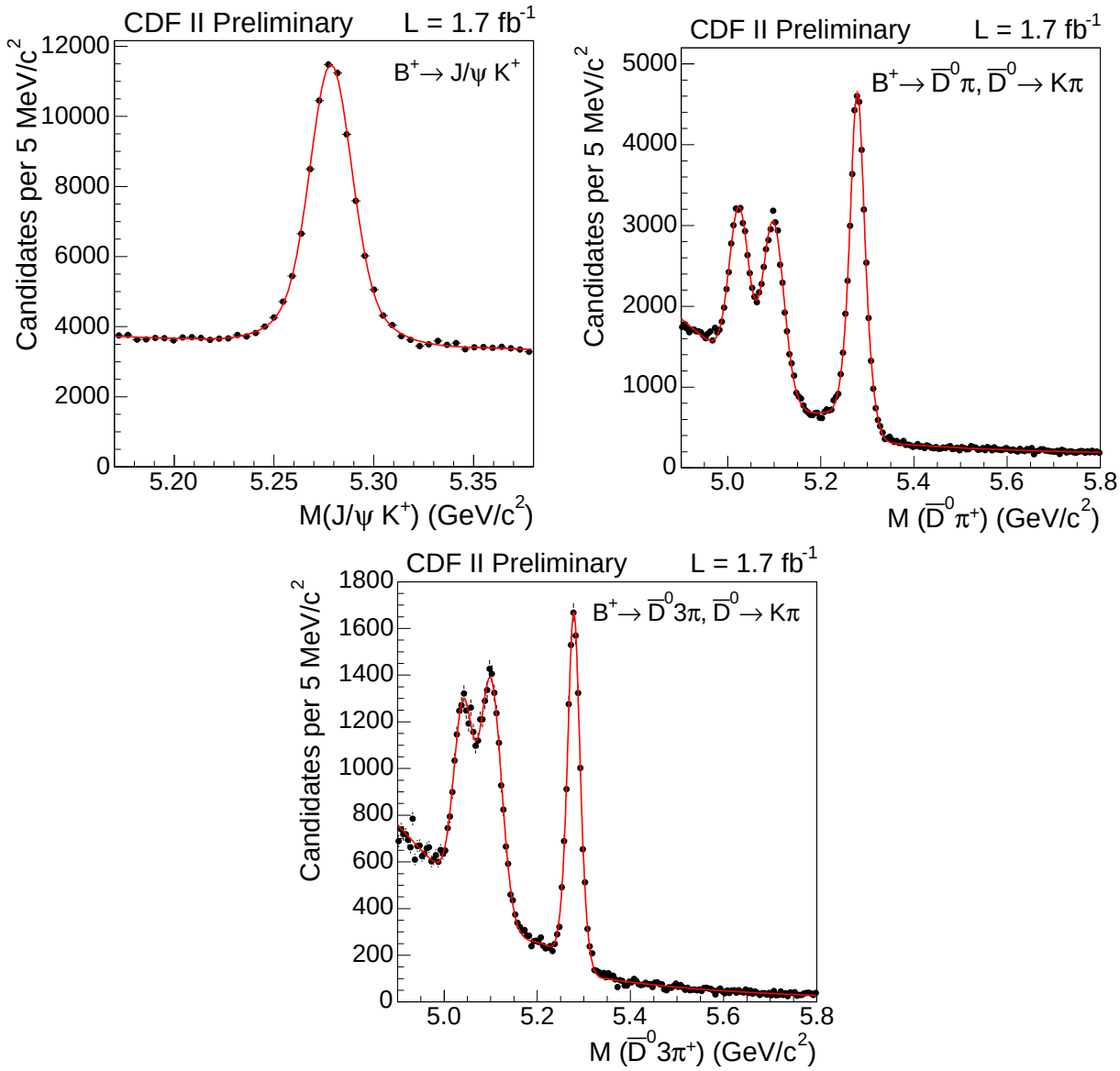


FIG. 2: The invariant mass distributions of the $B^+$ candidates at the best neural network cuts.

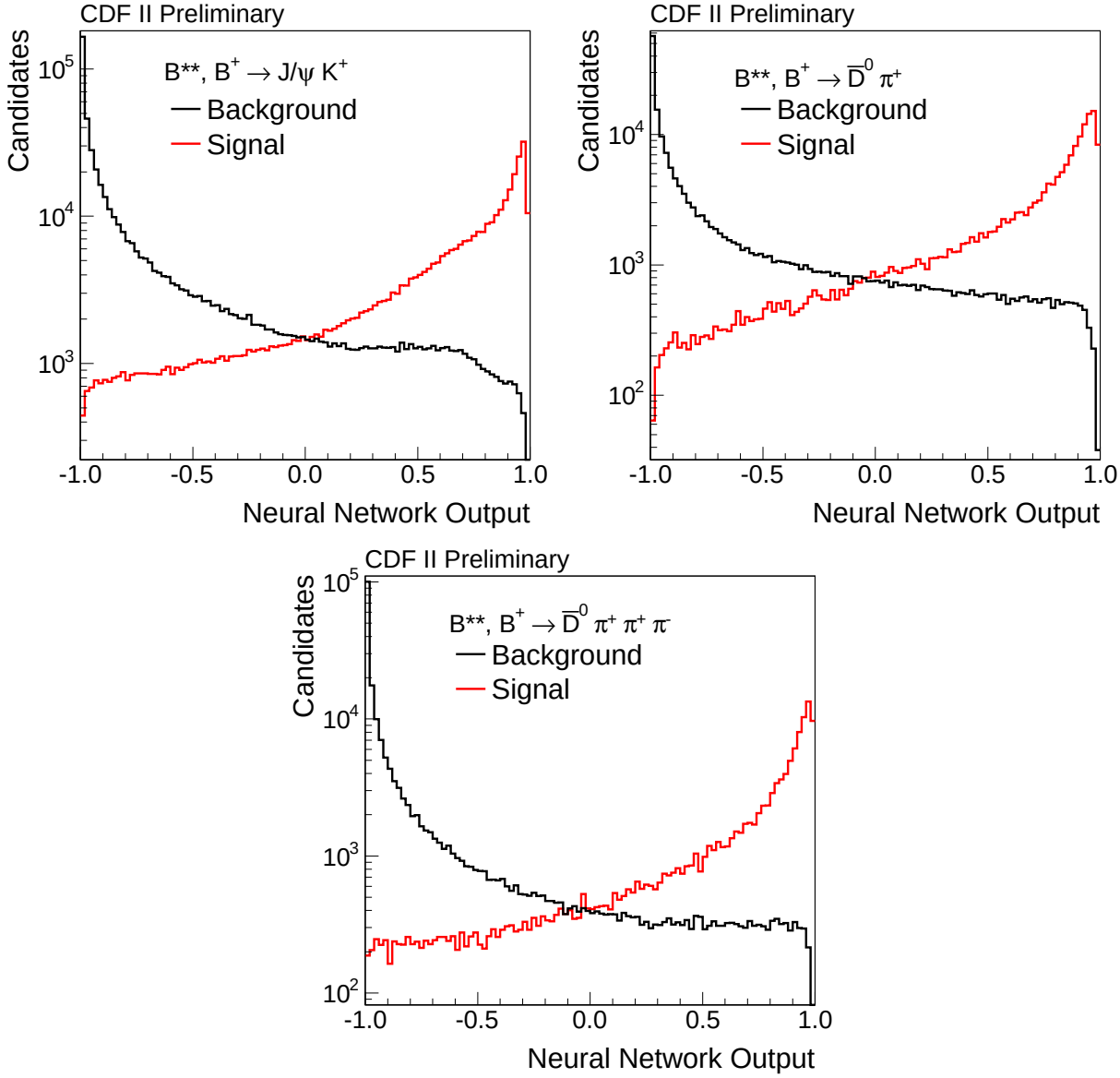


FIG. 3: The distributions of the neural network outputs for signal (red) and background (black) events using the training samples.

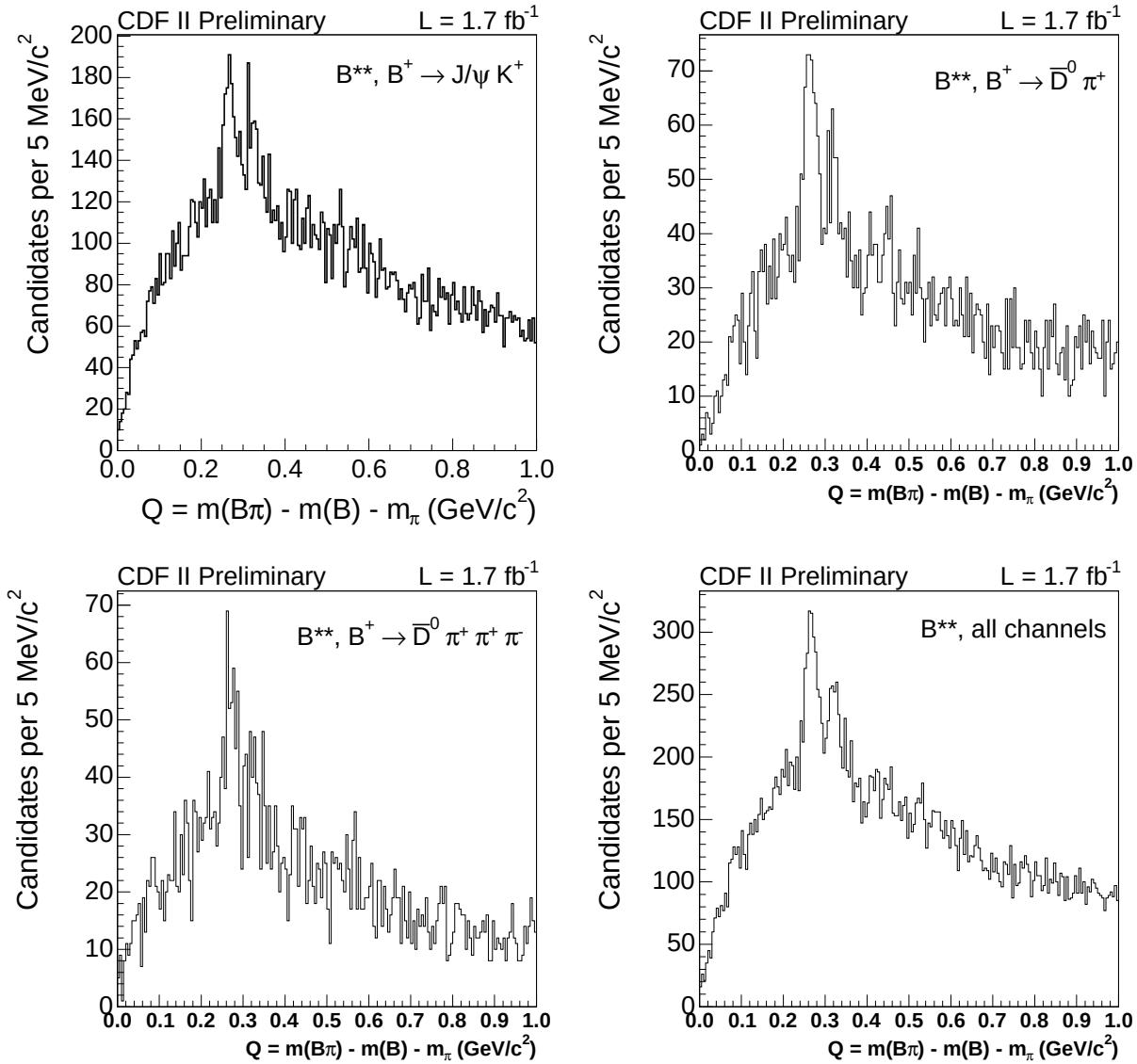


FIG. 4: The $Q$ distributions of the $B^{**}$ candidates at the best neural network cuts when channels are optimized separately. The final plot shows the events of all three channels combined.

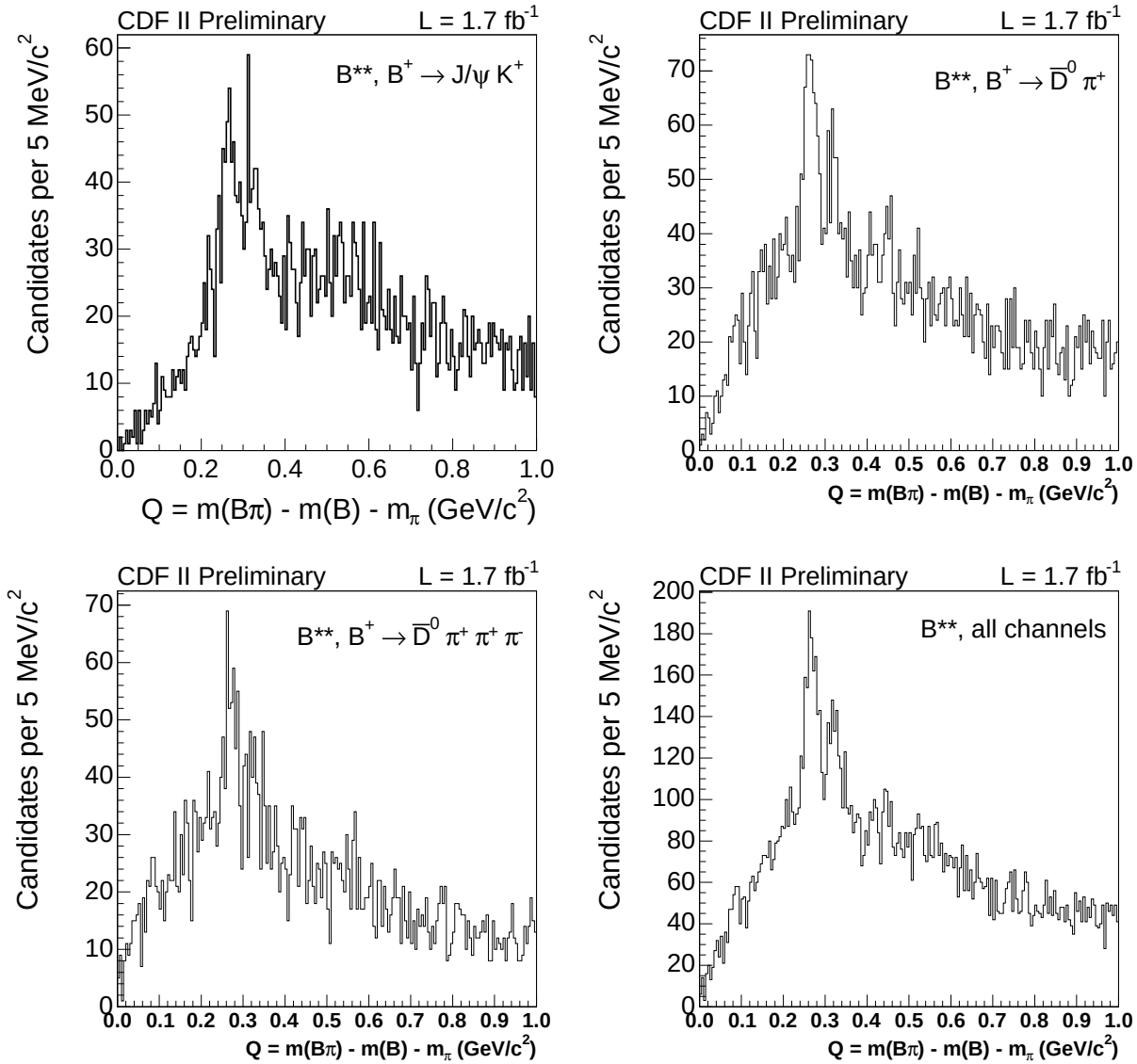


FIG. 5: The $Q$ distributions of the $B^{**}$ candidates at the best neural network cuts when channels are optimized simultaneously. The final plot shows the events of all three channels combined.

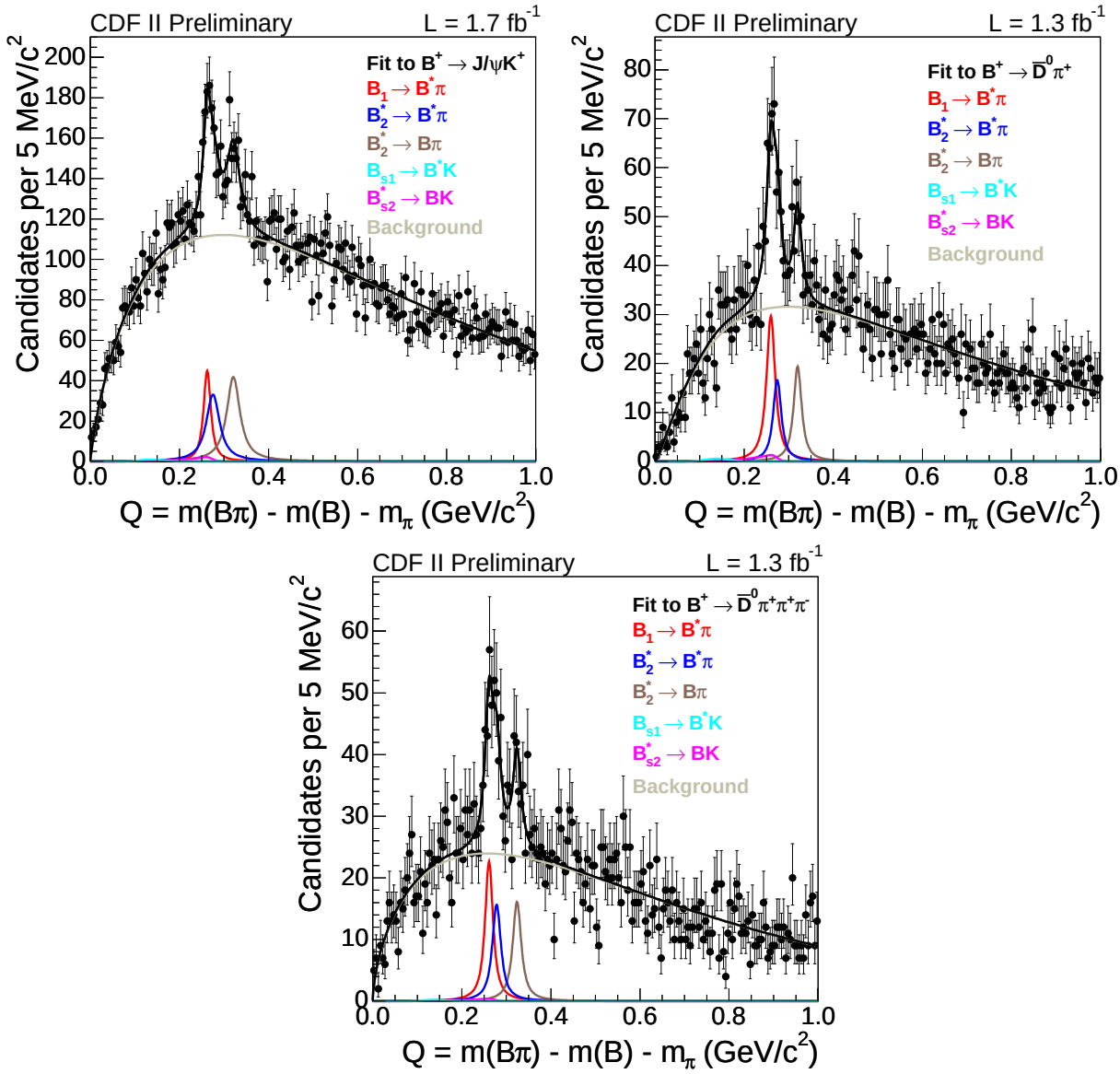


FIG. 6: Individual $B^{**}$ fits to the three $B^+$ samples. From left to right: $B^{**0} \to B^+\pi^- \to (J/\psi K^+)\pi^-$, $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+)\pi^-$, and $B^{**0} \to B^+\pi^- \to (\bar{D}^0\pi^+\pi^+\pi^-)\pi^-$. These fits are used as cross-checks of the final combined fit, to check that the $B^{**}$ signal is consistent in all three channels separately.

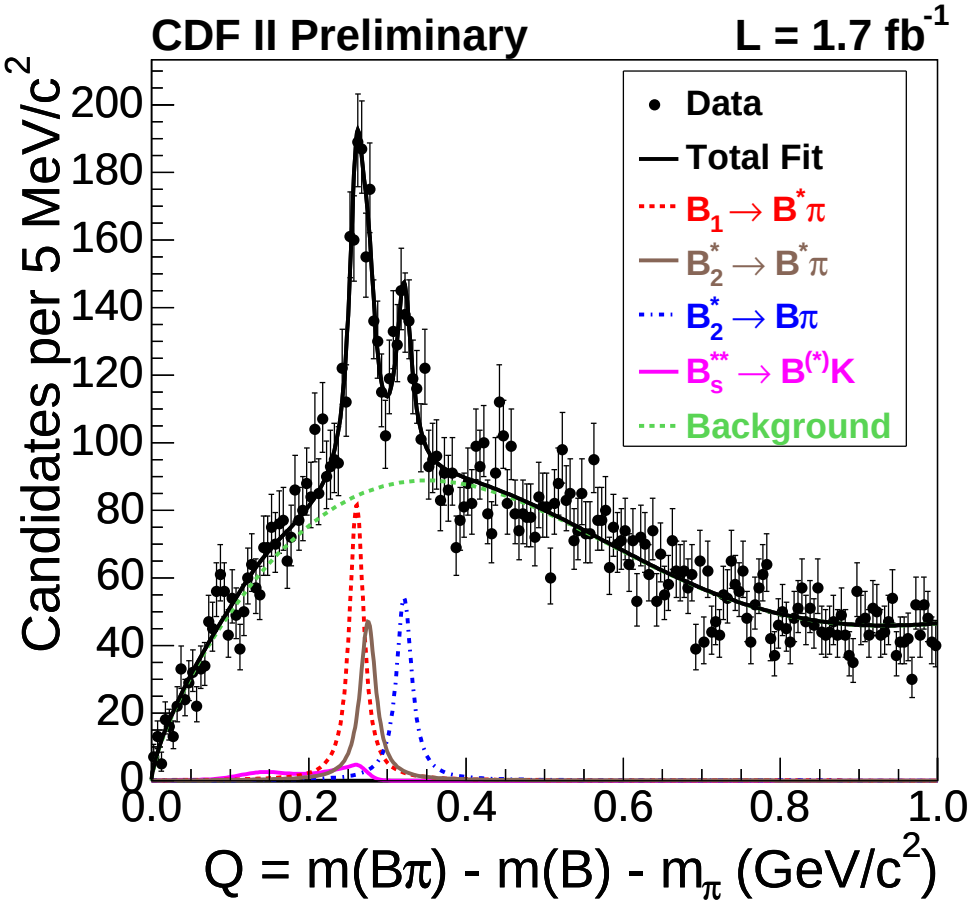


FIG. 7: $B^{**}$ fit to the combined data sample. The three $B^{**}$ signal peaks are shown separately.