



Observation of orbitally excited ($L=1$) B_s mesons

The CDF Collaboration
URL <http://www-cdf.fnal.gov>
(Dated: September 6, 2006)

We observe the two narrow states of the orbitally excited ($L = 1$) B_s mesons using fully reconstructed $B^+ \rightarrow J/\psi K^+$, $J/\psi \rightarrow \mu\mu$ and $B^+ \rightarrow D^0 \pi^+$, $D^0 \rightarrow K^+ \pi^-$ decays. In addition to the previously observed B_{s2}^* state also the B_{s1} state is observed here. A fit to the Q distribution yields $Q(B_{s1}) = 10.73 \pm 0.21$ (stat) ± 0.14 (syst) MeV/c^2 and $Q(B_{s2}^*) = 66.96 \pm 0.39$ (stat) ± 0.14 (syst) MeV/c^2 . The obtained result allows us for the first time a unique interpretation and mass assignment of the two narrow B_s^{**} states, $m(B_{s1}) = 5829.41 \pm 0.21$ (stat) ± 0.14 (syst) ± 0.6 (PDG) MeV/c^2 and $m(B_{s2}^*) = 5839.64 \pm 0.39$ (stat) ± 0.14 (syst) ± 0.5 (PDG) MeV/c^2 .

Preliminary Results for Summer 2006 Conferences

I. INTRODUCTION

Studies of excited states of hadrons play a vital role in understanding Quantum Chromodynamics, the theory of the strong interaction. The excited heavy mesons play a similar role as the hydrogen atom in Quantum electrodynamics. In this analogy, the heavy quark plays the role of the atomic nucleus, while the light quark takes the role of the electron in the hydrogen atom. Few predictions for masses and widths are available from theory [1–4]. They are summarised in Table I. Table II summarises the basic properties of the B_s^{**} mesons.

state/reference	Ebert et al. [1]		Godfrey et al. [2]		Eichten et al. [3]		Falk et al.[4]	
	M [MeV/ c^2]	Γ [MeV]	M [MeV/ c^2]	Γ [MeV]	M [MeV/ c^2]	Γ [MeV]	M [MeV/ c^2]	Γ [MeV]
$B_{s0}^*(1/2)$	5841	-	5830	170	-	-	-	-
$B_{s1}(1/2)$	5859	-	5860	-	-	-	-	-
$B_{s1}(3/2)$	5831	-	5860	-	5834	< 1.0	5886	2.8 ± 1.2
$B_{s2}^*(3/2)$	5844	-	5880	2.6	5846	1.0	5899	7 ± 3

TABLE I: Theory predictions for masses and widths of the orbitally excited ($L = 1$) B_s mesons.

j_q	J^P	B_s^{**} state	decay mode	width
1/2	0^+	B_{s0}^*	BK	broad (S-wave)
1/2	1^+	B_{s1}	B^*K	broad (S-wave)
3/2	1^+	B_{s1}	B^*K	narrow (D-wave)
3/2	2^+	B_{s2}^*	BK, B^*K	narrow (D-wave)

TABLE II: Properties of the orbitally excited ($L = 1$) B_s mesons. J is the total orbital momentum of the state, P its parity and j_q is the total orbital momentum of the light quark.

A lot of experimental results have been obtained in recent years for the orbital excitations with $L = 1$ of D , D_s and B mesons. On the contrary for orbital excitations of the B_s meson, only one of the four states was observed up to now, first by the OPAL Collaboration [5] and later confirmed by the DELPHI Collaboration [6]. Recently also the DØ experiment reported the observation of the same state [7]. In all three experiments, only one narrow state was observed and it is a priori unknown which one was observed. The DELPHI Collaboration interpreted the observed signal to stem more probably from the B_{s2}^* than from the B_{s1} . This interpretation is based on the width of the observed state, which was more consistent with the B_{s2}^* .

In this note we report on the search and observation of the narrow doublet of orbital excitations with $L = 1$ of the B_s mesons (commonly named as B_s^{**}) using the CDF II detector, which is described in [8]. The observed pattern of two narrow states allows us for the first time to uniquely determine the masses of the narrow B_s^{**} states.

II. DATA SAMPLE & EVENT SELECTION

This analysis is based on an integrated luminosity of 1 fb^{-1} collected with the CDF II detector between March 2002 and February 2006. Data are collected using the $J/\psi \rightarrow \mu\mu$ and the two track trigger. The J/ψ trigger starts from two tracks, which have information in the muon chambers and p_T larger than 1.5 (2.0) GeV/ c^2 for different subdetectors and/or data taking periods. For the final online selection, two tracks are required to have an opposite charge, an opening angle smaller than 135° and an invariant mass around the world average J/ψ mass. The two track trigger selects events based on the large impact parameter of the tracks coming from the B-hadron decays. It requires two tracks with an impact parameter in the range from $120 \mu\text{m}$ to 1 mm together with minimal transverse momentum of each track and minimal scalar sum of the transverse momenta of the two tracks.

The offline reconstruction starts with reconstructing B^+ candidates in the $J/\psi K^+$ and $\bar{D}^0 \pi^+$ decay modes with $J/\psi \rightarrow \mu^+ \mu^-$ and $\bar{D}^0 \rightarrow K^+ \pi^-$ [11]. The $B^+ \rightarrow J/\psi K^+$ is reconstructed from the J/ψ trigger data while the $B^+ \rightarrow \bar{D}^0 \pi^+$ is from the two track trigger data. Reconstructed candidates are preselected using separate Neural Networks for each of the two channels. To construct a Neural Network the NeuroBayes[©] [9] package is used. Both Neural Networks combine topological, kinematic and particle identification quantities of the B^+ and its daughters. The Neural Network for the $B^+ \rightarrow J/\psi K^+$ is trained on the Monte Carlo events with the full CDF detector simulation for signal patterns and data from sidebands for background patterns. For the $B^+ \rightarrow \bar{D}^0 \pi^+$ channel we use the possibility to train the Neural Network with weights and use only data from both signal and sideband regions and subtract background statistically during Neural Network training. The preselection is done by cutting on the Neural Network

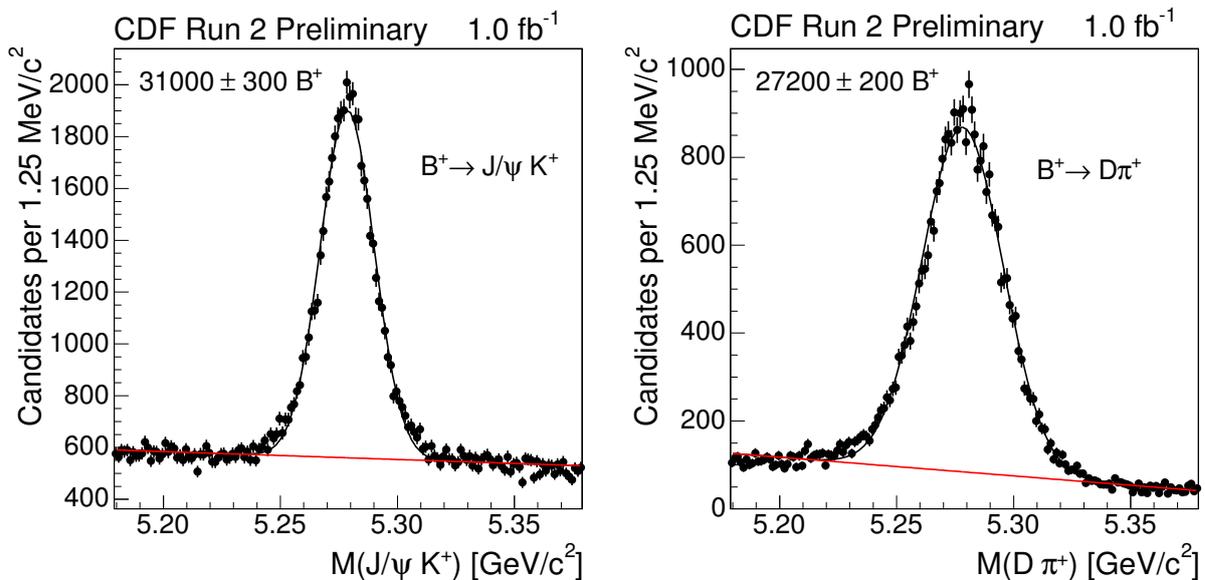


FIG. 1: Invariant mass distribution of the $B^+ \rightarrow J/\psi K^+$ (left) and $B^+ \rightarrow \bar{D}^0 \pi^+$ (right) candidates after preselection.

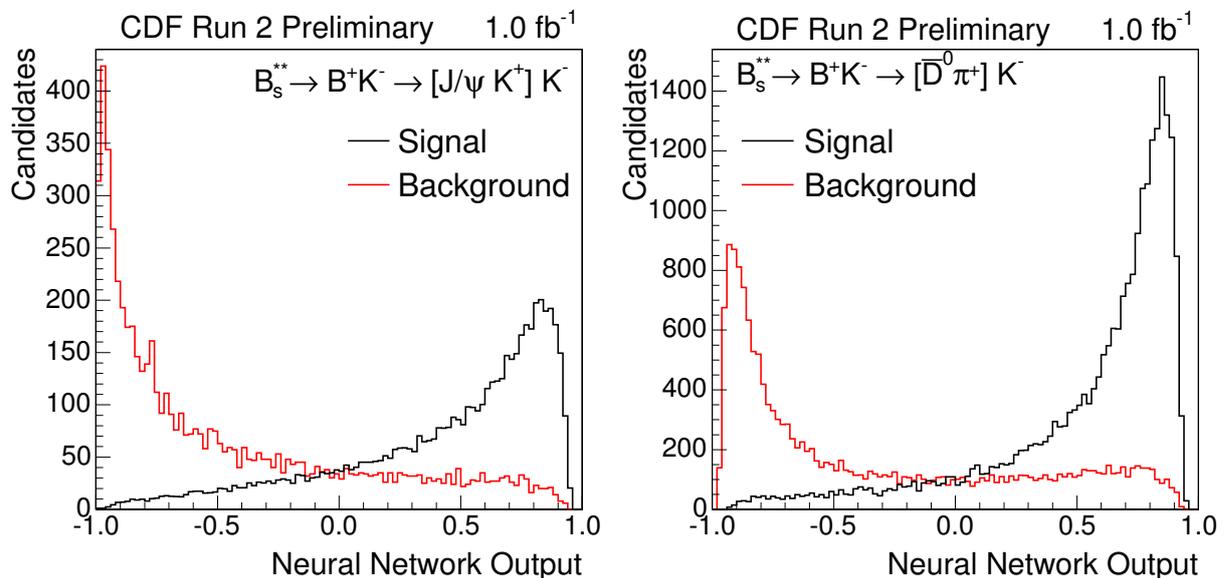


FIG. 2: Distribution of the Neural Network output for the signal and background events used in the training of the Neural Network for $B_s^{**} \rightarrow B^+ K^- \rightarrow [J/\psi K^+] K^-$ (left) and $B_s^{**} \rightarrow B^+ K^- \rightarrow [\bar{D}^0 \pi^+] K^-$ (right) channel.

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 two channel are shown in Figure 1. In total, 31000 B^+ signal events are selected in the $J/\psi K^+$ decay channel and 27200 in the $\bar{D}^0 \pi^+$ channel.

The B_s^{**} candidates are constructed by combining B^+ candidates with a track which is assumed to be a kaon. For the selection two Neural Networks are trained on a combination of Monte Carlo events for signal patterns and data for background patterns. The data for background patterns are taken from the Q value range from 0 to 200 MeV/c^2 . Q is defined as $M(B^+ K^-) - M(B^+) - M(K^-)$ where $M(B^+ K^-)$ is the invariant mass of the $B^+ K^-$ pair, $M(B^+)$ is the invariant mass of the B^+ candidate and $M(K^-)$ is the mass of the kaon. In order to avoid possible biases of the Neural Network in the mass, the signal patterns from Monte Carlo are reweighted to have the same Q distribution as data. Neural Networks for both channels combine topological, kinematical and particle identification quantities of the various particles in the decay chain. The distribution of the Neural Network output for the two Neural Networks for the signal and the background used in the training are shown in Figure 2.

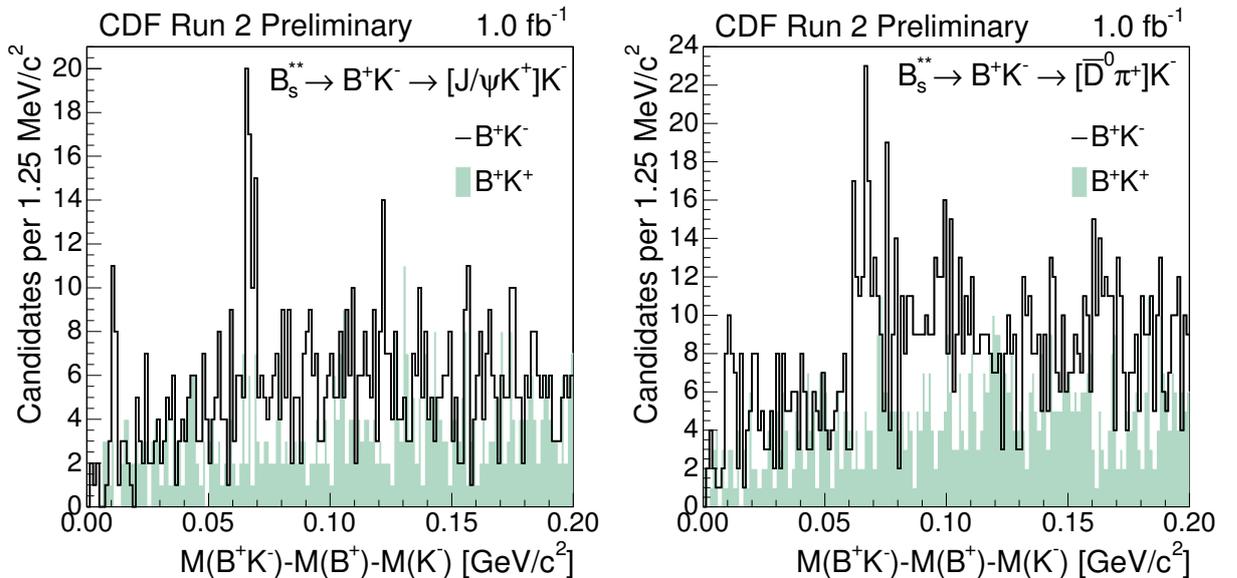


FIG. 3: $Q = M(B^+K^-) - M(B^+) - M(K^-)$ distribution of the B_s^{**} candidates (full line) with the distribution of wrong sign candidates (green area) for $B_s^{**} \rightarrow B^+K^- \rightarrow [J/\psi K^+]K^-$ (left) and $B_s^{**} \rightarrow B^+K^- \rightarrow [\bar{D}^0 \pi^+]K^-$ (right) channel.

For the final selection we cut on two quantities, which are the number of candidates in the event and the output of the Neural Network. The number of candidates is used separately and not in the Neural Network because a correct simulation of this quantity is nontrivial and difficult. Therefore we also select the cut value rather arbitrarily just from the fact, that more candidates in the event means a smaller signal to background fraction. Before extracting any information from data we fix the cut on the number of candidates to be less than four candidates in the event. The value of the cut on the Neural Network output is chosen to optimise the significance $S/\sqrt{S+B}$ in the Q window from 60 to 70 MeV/c^2 . The optimisation is done on a combination of a Monte Carlo simulation which is used to obtain a nominator and data from which the denominator is extracted at each tested cut value. The best cut is found to be 0.5 for the $B_s^{**} \rightarrow B^+K^- \rightarrow [J/\psi K^+]K^-$ channel and 0.3 for $B_s^{**} \rightarrow B^+K^- \rightarrow [\bar{D}^0 \pi^+]K^-$ channel.

To check for possible systematic effects we examine also wrong sign combinations (B^+K^+). They are selected using the same selection criteria as the signal right sign candidates. The cut on the Neural Network output and a cut on the number of candidates which has the same value, but a different definition. While for the right sign combinations the number of right sign candidates is used, for the wrong sign combinations the number wrong sign candidates are counted to follow the logic of the selection for the signal.

III. RESULTS

In Figure 3 we show the Q value distributions for the two studied channels. In the $B_s^{**} \rightarrow B^+K^- \rightarrow [J/\psi K^+]K^-$ channel, two clear peaks are visible at around 67 MeV/c^2 and 10 MeV/c^2 . In the channel $B_s^{**} \rightarrow B^+K^- \rightarrow [\bar{D}^0 \pi^+]K^-$ the peaks are not as clear but there is still evidence of the peaks in the same positions. We also varied the selection cuts and the binning and the two peaks showed consistent behaviour. In contrast to right sign combinations, the wrong sign combinations don't show any significant and consistent structure. Figure 4 shows the Q distribution of the two channels added together.

To extract the Q values from the data we use an unbinned maximum likelihood fit. The two peaks are described by two Gaussians. We use a phenomenological function able to describe the background without any attempt to distinguish different types of backgrounds. The basic properties of the function are that it should go to zero at $Q = 0$ and should have a maximum followed by a slow decrease. The functional form is

$$B(Q) = \alpha[Q(\beta - Q)]^\gamma \exp[-\gamma(Q - \delta)] \quad (1)$$

where α is normalisation constant and β , γ and δ are free parameters. The fit has 3 parameters for each of the Gaussians and 4 parameters for the background. All parameters are free during the fit. The fitting code was validated using Toy Monte Carlo experiments of same statistics and sample composition as observed in the data. The fit projections

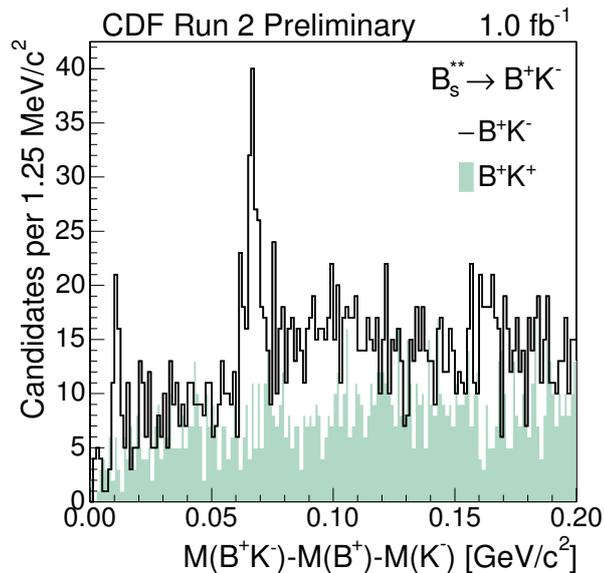


FIG. 4: $Q = M(B^+K^-) - M(B^+) - M(K^-)$ distribution of the B_s^{**} candidates (full line) with distribution of wrong sign candidates (green area) for two channels added together.

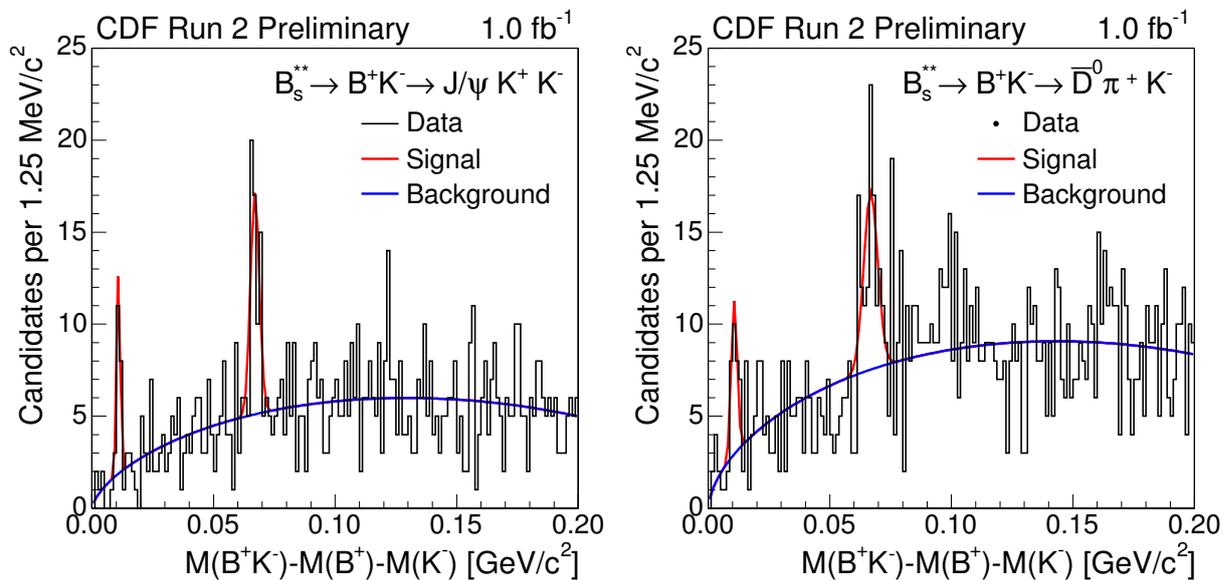


FIG. 5: Fit projections of the unbinned maximum likelihood fits done on the two decay channels separately. On the left side $B_s^{**} \rightarrow B^+K^- \rightarrow [J/\psi K^+]K^-$ is shown while on the right we show the $B_s^{**} \rightarrow B^+K^- \rightarrow [\bar{D}^0 \pi^+]K^-$ channel.

to the two subsamples from the different decay channels are shown in Figure 5. The parameters for the two signals are listed in Table III. The two samples are consistent with each other. The only small difference is in the resolution, where the channel $B_s^{**} \rightarrow B^+K^- \rightarrow [\bar{D}^0 \pi^+]K^-$ has slightly worse resolution than $B_s^{**} \rightarrow B^+K^- \rightarrow [J/\psi K^+]K^-$. In addition we also compare the background shapes of the two channels. They are shown in Figure III. The functions are scaled to have the same total area as the function from the fit to the full sample to facilitate a shape comparison. Both channels are consistent in the signal parameters and background shapes and therefore we add both channels together to do the final fit. The resulting fit projection is shown in Figure III. The quantities extracted from this fit

	B_{s1}		B_{s2}^*	
	$B^+ \rightarrow J/\psi K^+$	$B^+ \rightarrow \bar{D}^0 \pi^+$	$B^+ \rightarrow J/\psi K^+$	$B^+ \rightarrow \bar{D}^0 \pi^+$
Q [MeV/c ²]	10.87 ± 0.19	10.68 ± 0.46	67.03 ± 0.44	66.85 ± 0.76
σ [MeV/c ²]	0.64 ± 0.25	1.18 ± 0.56	1.79 ± 0.42	2.88 ± 0.75
N	16.98 ± 5.14	20.66 ± 7.12	44.15 ± 13.36	55.74 ± 19.20

TABLE III: Comparison of the unbinned maximum likelihood fits of the two data subsets coming from different B^+ decays. Those two subsamples are independent and have roughly the same B^+ yield. The errors are statistical only.

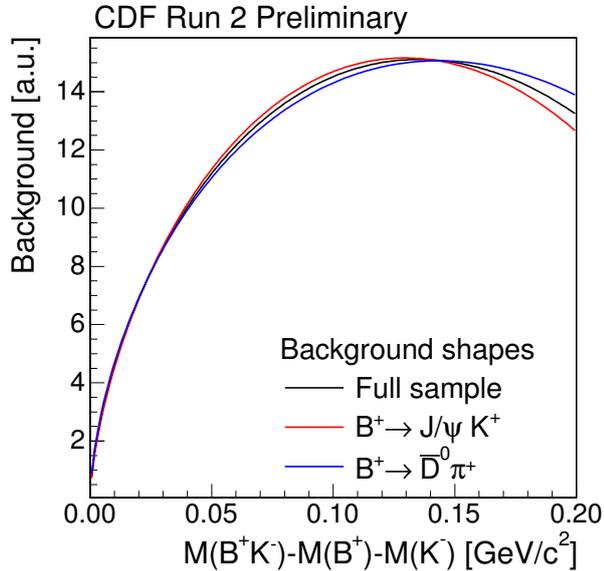


FIG. 6: Comparison of the fitted background shapes between the $B_s^{**} \rightarrow B^+ K^- \rightarrow [J/\psi K^+] K^-$ and the $B_s^{**} \rightarrow B^+ K^- \rightarrow [\bar{D}^0 \pi^+] K^-$ channel. The two functions are scaled to the same area to facilitate a shape comparison.

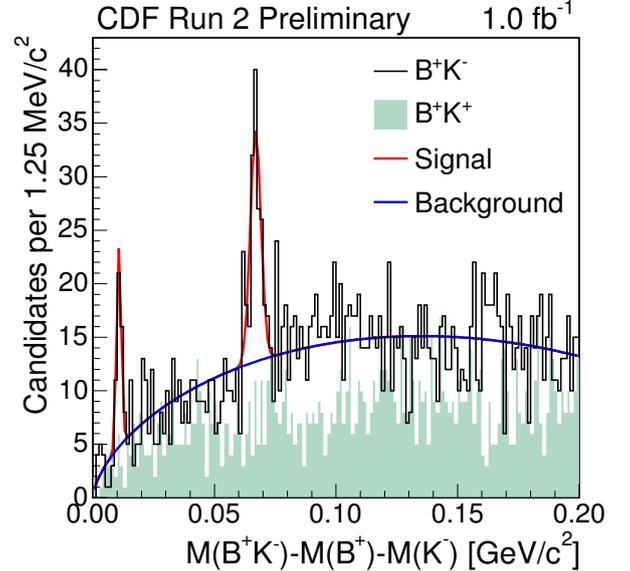


FIG. 7: Fit projection of the final fit using both channels added together.

are Q values for the two peaks and number of signal events. The two Q values are

$$\begin{aligned} Q(B_{s1}) &= 10.73 \pm 0.21 \text{ (stat) MeV/c}^2 \\ Q(B_{s2}^*) &= 66.96 \pm 0.39 \text{ (stat) MeV/c}^2 \end{aligned} \quad (2)$$

with 36.4 ± 9.0 events in the peak at 10.73 MeV/c^2 and 94.8 ± 23.4 events in the peak at 66.96 MeV/c^2 . In both cases, the likelihood function has a well defined minimum. They are shown in Figure 8 for the peak at 10.73 MeV/c^2 and in Figure 9 for peak at 66.96 MeV/c^2 .

Systematic effects due to the tracking and fitting procedure are studied for the obtained Q values. For the tracking, there are two main sources of uncertainty and those are the COT error matrix scaling and the calibration of the material and magnetic field inside the tracking volume. A detailed study of the effect was done for the measurement of mass and width of the orbitally excited charm D^{**} states [10]. The combined uncertainty of the two tracking sources is found to be 0.14 MeV/c^2 . The fitting procedure can contribute from the fact that unknown shape of the background, which can be wrongly modelled, and the simplification, where a single Gaussian is used in the fit, while two decay channels can have different resolutions and therefore two Gaussians would be more appropriate. All those effects are studied using large statistics of Toy Monte Carlo experiments, which are generated with the studied effects and fitted with the original fit. In all three cases, the pulls of the Q values were consistent with a Gaussian with mean zero and unit width. Therefore we don't assign any systematic uncertainty coming from the fitting procedure.

To estimate the statistical significance of the two peaks, we repeat the fit without the studied peak. From the value of the likelihood functions of the original fit \mathcal{L} and the fit without the given peak \mathcal{L}_0 we can form $-2 \ln \mathcal{L}_0 / \mathcal{L}$. This quantity asymptotically behaves as a χ^2 distribution with degrees of freedom equal to the difference in the number of free parameters in the two fits. This allows us to convert the $-2 \ln \mathcal{L}_0 / \mathcal{L}$ to a probability and therefore also a statistical significance. From this procedure we obtain a significance of 6.3σ for the peak at 10.73 MeV/c^2 and 7.7σ

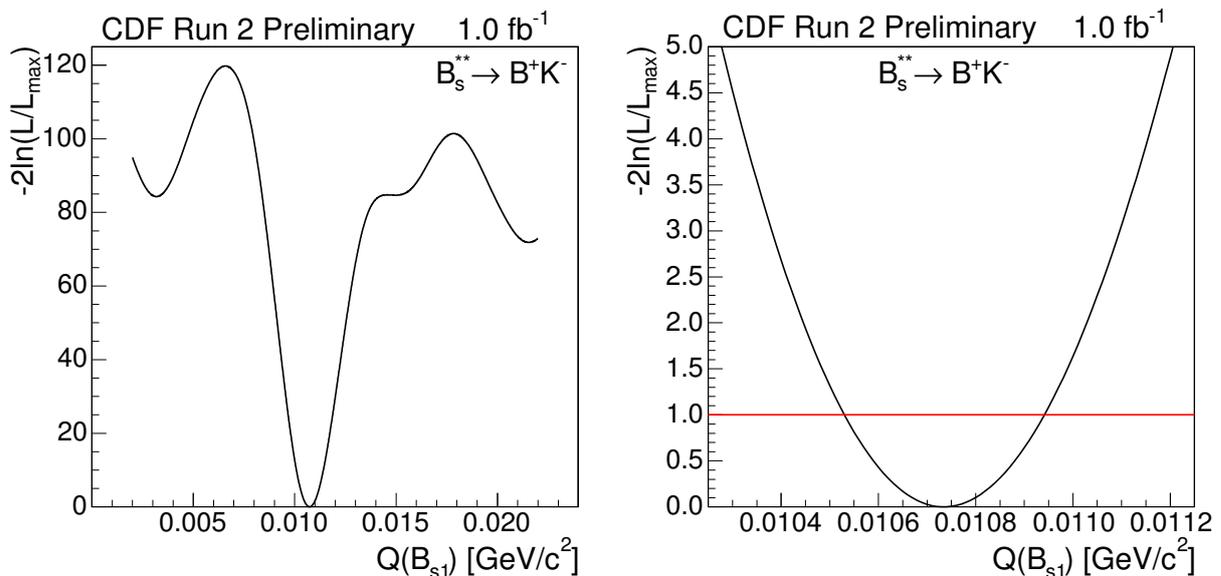


FIG. 8: Dependence of the likelihood function on the Q value for the peak at 10.73 MeV/c². The left plot shows the function in a wider range around the minimum, while the right plot zooms into the region around the minimum.

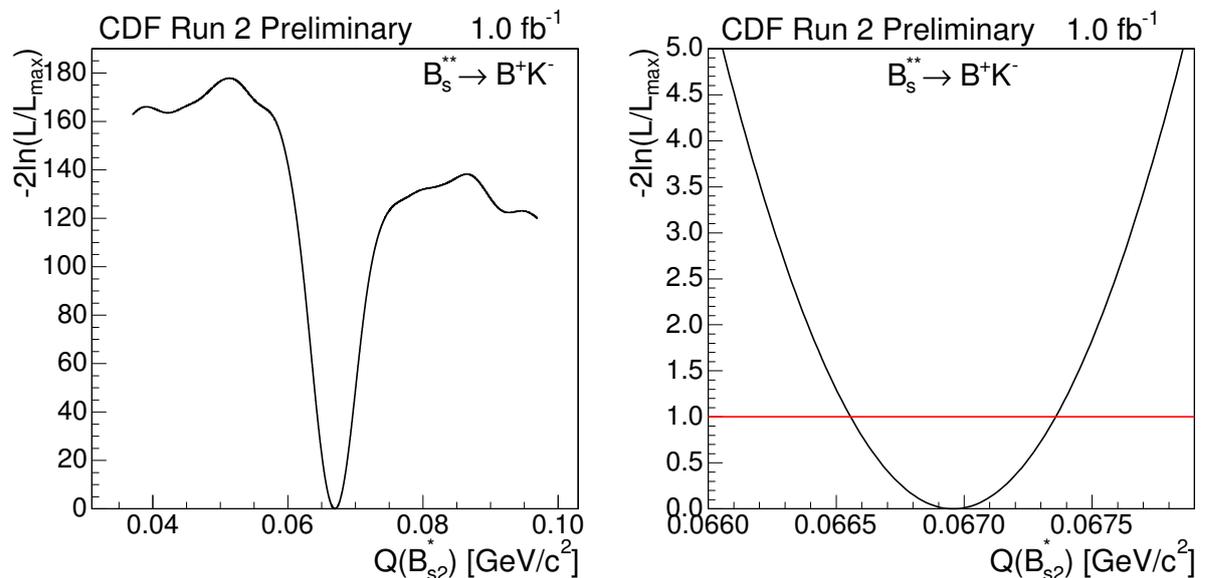


FIG. 9: Dependence of the likelihood function on the Q value for the peak at 66.96 MeV/c². The left plot shows the function in a wider range around the minimum, while the right plot zooms into the region around the minimum.

for the peak at 66.96 MeV/c².

As the peak at the 10.73 MeV/c² is observed here for the first time additional a study of the significance is done. We generate a large sample of Toy experiments with background according to our data and the signal at 66.96 MeV/c². Each of the Toy experiments is fitted in the same way as data and $-2\ln \mathcal{L}_0/\mathcal{L}$ is evaluated. In Figure 10 we show the distribution of $-2\ln \mathcal{L}_0/\mathcal{L}$ from the Toy experiments where the fits converged and the Q value for the peak with lower Q -value was between 0 and 50 MeV/c². This distribution is then converted to the dependence of the p-Value on $-2\ln \mathcal{L}_0/\mathcal{L}$, which is shown in Figure 10 in right panel. The value observed in the data is larger than any seen in the Toy experiments. Last non-zero p-Value occurs around $-2\ln \mathcal{L}_0/\mathcal{L} \approx 35.15$ with p-Value $2.13 \cdot 10^{-7}$. This corresponds to the significance slightly above 5 sigma and gives additional support for the observation claim.

The two peaks seen in the data can be interpreted as the two states of the narrow doublet of orbitally excited B_s mesons. The natural interpretation is that the peak at $Q = 66.96$ MeV/c² stems from the $B_{s2}^* \rightarrow B^+ K^-$ decay while

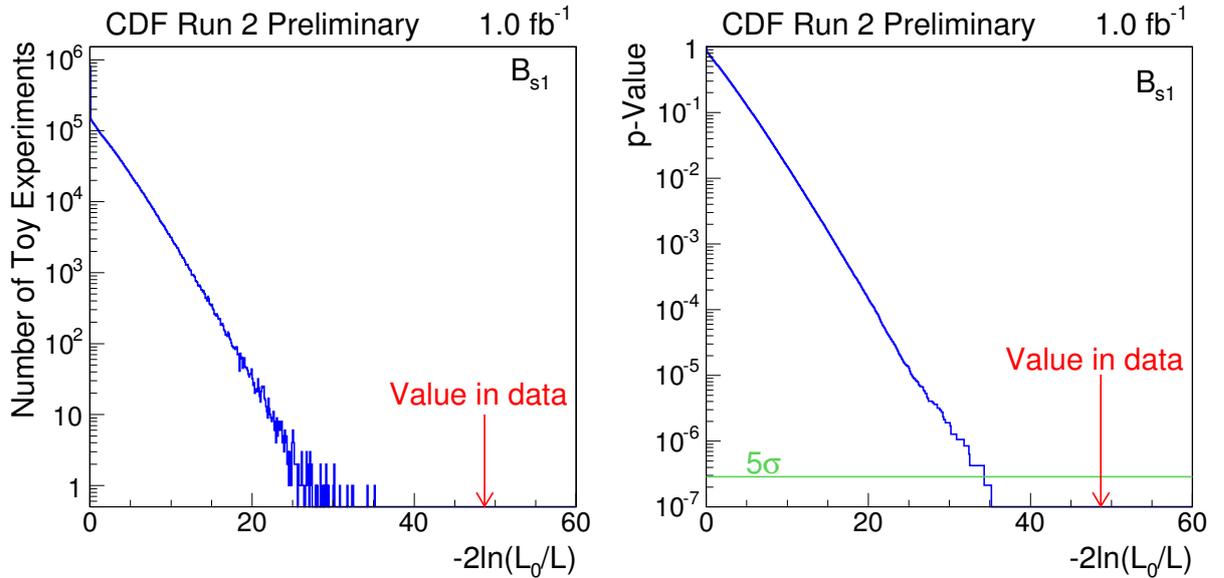


FIG. 10: (Left) Distribution of the $-2\ln\mathcal{L}_0/\mathcal{L}$ for evaluating the p-Value for the peak at $10.73\text{ MeV}/c^2$. Toy Monte Carlo experiments without the corresponding peak were generated and fitted with and without the peak. All experiments where the fit converged in the Q -range from 0 to $50\text{ MeV}/c^2$ are used here. (Right) P-Value as a function of $-2\ln\mathcal{L}_0/\mathcal{L}$ for the peak at $10.73\text{ MeV}/c^2$.

the peak at $Q = 10.73\text{ MeV}/c^2$ would stem from the decay $B_{s1} \rightarrow B^{*+}K^-$. Another peak from decay $B_{s2}^* \rightarrow B^{*+}K^-$ is expected at $Q \approx 21\text{ MeV}/c^2$. This peak however will be much smaller as the predicted ratio of the branching fraction for the two B_{s2}^* decays is

$$\frac{\mathcal{B}(B_{s2}^* \rightarrow BK)}{\mathcal{B}(B_{s2}^* \rightarrow B^*K)} = 12.0 \pm 3.5 \quad (3)$$

Our data indeed show few bins above the fitted background at the corresponding Q value of around $21\text{ MeV}/c^2$. This slight excess is consistent with the expectation for the $B_{s2}^* \rightarrow B^{*+}K^-$ decay, but is far from being statistically significant. The swapped assignment of the two peaks at 10.73 and $66.96\text{ MeV}/c^2$, would result in a very large mass difference between the B_{s2}^* and B_{s1} which would be rather unnatural for the hyperfine splitting within the doublet. Therefore we assign to the peak at $Q = 10.73\text{ MeV}/c^2$ the decay $B_{s1} \rightarrow B^{*+}K^-$ and to the peak at $Q = 66.96\text{ MeV}/c^2$ the decay $B_{s2}^* \rightarrow B^{*+}K^-$.

Assigning the two observed peaks to the decays of the two narrow states we can transform the measured Q values to the masses of the two narrow B_s^{**} states. The corresponding masses are

$$\begin{aligned} m(B_{s1}) &= 5829.41 \pm 0.21(\text{stat}) \pm 0.14(\text{syst}) \pm 0.6(\text{PDG})\text{ MeV}/c^2 \\ m(B_{s2}^*) &= 5839.64 \pm 0.39(\text{stat}) \pm 0.14(\text{syst}) \pm 0.5(\text{PDG})\text{ MeV}/c^2 \end{aligned}$$

where the PDG error is due to the error on the masses of B^+ and the mass and B^{*+} . Finally the mass difference of the two narrow B_s^{**} states is

$$\Delta m(B_{s2}^*, B_{s1}) = 10.20 \pm 0.44(\text{stat}) \pm 0.35(\text{PDG})\text{ MeV}/c^2$$

with PDG error coming from the mass difference between B^+ and B^{*+} . The measured value is close to the model expectations.

IV. SUMMARY

In this note we report on the observation of the two narrow states of the orbitally excited B_s mesons. For the first time we observe two narrow peaks in the B^+K^- distribution. The observed pattern of the two peaks allows us for the first time to make unambiguous assignments of the peaks to the states which translate to the mass measurements

of the two narrow B_s^{**} . By this analysis we confirm, that the signal seen by previous experiments really stems from the decay $B_{s2}^* \rightarrow B^+ K^-$ as was interpreted by some of the experiments. In addition, for the first time we observe a signal coming from the decay $B_{s1} \rightarrow B^{*+} K^-$. The measured Q values translate to the following masses of the two states

$$\begin{aligned} m(B_{s1}) &= 5829.41 \pm 0.21 (\text{stat}) \pm 0.14 (\text{syst}) \pm 0.6 (\text{PDG}) \text{ MeV}/c^2 \\ m(B_{s2}^*) &= 5839.64 \pm 0.39 (\text{stat}) \pm 0.14 (\text{syst}) \pm 0.5 (\text{PDG}) \text{ MeV}/c^2 \end{aligned}$$

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 fuer 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 Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

-
- [1] Ebert, Garkin, Faustov, Phys. Rev. **D57**, 5663; Ebert, Garkin, Faustov, hep-ph/9712318
 - [2] Godfrey, Kokoshi, Phys. Rev. **D43**, 1679
 - [3] Eichten, Hill, Quigg, Phys. Rev. Lett. **71**, 4116
 - [4] Falk, Mehen, Phys. Rev. **D53**, 231; hep-ph/9507311
 - [5] OPAL Collab., R. Akers et al., Z. Phys. **C66**, 19
 - [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 thesis available at <http://www-ekp.physik.uni-karlsruhe.de/pub/web/thesis>.
 - [7] DØ Collab., *First Direct Observation of B_{s2}^{*0} meson*, DØnote 5027-Conf
 - [8] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, et al., Nucl. Instrum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, et al., Phys. Rev. D **52**, 4784 (1995); P. Azzi, et al., Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E
 - [9] M. Feindt, A Neural Bayesian Estimator for Conditional Probability Densities, physics/0402093, 2004
 - [10] CDF Collaboration, Measurement of mass and width of orbitally-excited charm mesons D_1 and D_2^* , CDF Note 7191, <http://www-cdf.fnal.gov/physics/new/bottom/040805.blesses-dss/>; CDF Collaboration, Phys. Rev. D **73**, 051104 (2006)
 - [11] Charge conjugates are used implicitly