



## Observation of new suppressed $B_s$ decays and measurement of their branching ratios

The CDF Collaboration  
URL <http://www-cdf.fnal.gov>  
(Dated: August 26, 2010)

We present the observation of new suppressed  $B_s$  decays,  $B_s \rightarrow J/\psi K^*(892)^0$  and  $B_s \rightarrow J/\psi K_S$ , and the measurement of their branching ratios. This measurement is based on an integrated luminosity of  $5.9 \text{ fb}^{-1}$  of CDF data collected by a dedicated di-muon trigger. A cut based optimization is carried out for the observation of  $B_s \rightarrow J/\psi K^*$ , while a Neural Network is used for the  $B_s \rightarrow J/\psi K_S$ . In addition to the observation of the new decay modes, we measured the following quantities:

$$\begin{aligned} f_s Br(B_s \rightarrow J/\psi K_S) / f_d Br(B^0 \rightarrow J/\psi K_S) &= 0.0109 \pm 0.0019(\text{stat.}) \pm 0.0011(\text{sys.}) \\ f_s Br(B_s \rightarrow J/\psi K^*) / f_d Br(B^0 \rightarrow J/\psi K^*) &= 0.0168 \pm 0.0024(\text{stat.}) \pm 0.0068(\text{sys.}) \end{aligned}$$

Using the CDF measurement of  $f_s/f_d$ , the ratio of branching fractions to the reference  $B^0$  decays are:

$$\begin{aligned} Br(B_s \rightarrow J/\psi K_S) / Br(B^0 \rightarrow J/\psi K_S) &= 0.041 \pm 0.007(\text{stat.}) \pm 0.004(\text{sys.}) \pm 0.005(\text{frag.}) \\ Br(B_s \rightarrow J/\psi K^*) / Br(B^0 \rightarrow J/\psi K^*) &= 0.062 \pm 0.009(\text{stat.}) \pm 0.025(\text{sys.}) \pm 0.008(\text{frag.}) \end{aligned}$$

Using PDG values for  $Br(B^0 \rightarrow J/\psi K^*)$  and  $Br(B^0 \rightarrow J/\psi K^0)$ , the absolute branching fractions are calculated:

$$\begin{aligned} Br(B_s \rightarrow J/\psi K^0) &= (3.5 \pm 0.6(\text{stat.}) \pm 0.4(\text{sys.}) \pm 0.4(\text{frag.}) \pm 0.1(\text{PDG})) \cdot 10^{-5} \\ Br(B_s \rightarrow J/\psi K^*) &= (8.3 \pm 1.2(\text{stat.}) \pm 3.3(\text{sys.}) \pm 1.0(\text{frag.}) \pm 0.4(\text{PDG})) \cdot 10^{-5} \end{aligned}$$

## I. MOTIVATION

While  $B^0$  decays have been extensively studied at the B factory experiments, much less is about  $B_s$  decays. We report about studies for two specific  $B_s$  decays:  $B_s \rightarrow J/\psi K_S$ , with  $J/\psi \rightarrow \mu^+\mu^-$  and  $K_S \rightarrow \pi^+\pi^-$ , and  $B_s \rightarrow J/\psi K^*(892)^0$ , with  $J/\psi \rightarrow \mu^+\mu^-$  and  $K^*(892)^0 \rightarrow K\pi$ .  $B_s \rightarrow J/\psi K_S$  is a CP eigenstate and has never been observed. Measurement of its lifetime directly probes the lifetime of the heavy mass eigenstate,  $\tau_{B_{s,H}}$ . Additionally, large samples of  $B_s \rightarrow J/\psi K_S$  can be used to extract the angle  $\gamma$  of the unitary triangle [1]. The  $B_s \rightarrow J/\psi K^*$  decay is yet another unobserved mode which contains an admixture of CP final states. An angular analysis of a significantly large enough sample of  $B_s \rightarrow J/\psi K^*$  can be carried out to extract  $\sin(2\beta_s)$  as a compliment to  $B_s \rightarrow J/\psi\phi$  [2].

In addition to the first observation of these two decays, the purpose of this analysis is to measure the ratio of branching ratios of  $B_s \rightarrow J/\psi K_S$  to  $B^0 \rightarrow J/\psi K_S$  and  $B_s \rightarrow J/\psi K^*$  to  $B^0 \rightarrow J/\psi K^*$  using the relation

$$\frac{Br(B_s \rightarrow J/\psi K)}{Br(B^0 \rightarrow J/\psi K)} = A_{rel} \frac{f_d N(B_s \rightarrow J/\psi K)}{f_s N(B^0 \rightarrow J/\psi K)},$$

where  $K$  represents  $K_S$  or  $K^*$ . By measuring the ratio of the number of decays,  $N(B_s \rightarrow J/\psi K)/N(B^0 \rightarrow J/\psi K)$ , from data and the relative acceptance,  $A_{rel}$ , between the  $B^0$  and  $B_s$  from Monte Carlo simulation (MC), the value  $Br(B_s \rightarrow J/\psi K)/Br(B^0 \rightarrow J/\psi K)$  can be extracted by inputting the ratio of fragmentation fractions  $f_s/f_d$ .

## II. DATA SAMPLE & EVENT SELECTION

The data used in these analyses are selected from a  $J/\psi$  dataset, collected from March 2002 to February 2010 by the CDF Run II detector [3]. It corresponds to an integrated luminosity of  $5.9 \text{ fb}^{-1}$ . The  $J/\psi$  dataset contains events with at least one reconstructed  $J/\psi$  selected by dedicated di-muon triggers. The muon identification begins with hits in the muon chambers reconstructed into stubs, and then matched with a reconstructed track in the open-cell wire drift chamber (COT). In addition to the selected  $J/\psi$ , two tracks are found to get  $B^0 \rightarrow J/\psi K_S$  and  $B^0 \rightarrow J/\psi K^*$  candidates. In the  $B^0 \rightarrow J/\psi K_S$  analysis, the two tracks are reconstructed as pions and combined to define a  $K_S$  candidate. To reduce backgrounds from prompt sources, a displacement between the  $K_S$  candidate vertex and the B candidate vertex is required in the event selection. The  $K^*$  candidate for the  $B^0 \rightarrow J/\psi K^*$  decay is reconstructed from the combination of a  $\pi$  and a  $K$ . If two candidates are reconstructed with the same tracks, with the only difference that the kaon and pion hypotheses are swapped, the  $K^*$  candidate closest to the PDG value of  $896 \text{ MeV}/c^2$  is selected. All event selection cuts for  $B^0 \rightarrow J/\psi K_S$  candidates and  $B^0 \rightarrow J/\psi K^*$  candidates are listed in Table I and II, respectively. With the purpose of removing some background, specific optimization selection is applied in each analysis.

### Optimization of candidates selection in the $B_s \rightarrow J/\psi K_S$ analysis

A Neural Network is used in this analysis to remove further combinatorial background. In order to train the Neural Network, simulated  $B_s$  MC events are used as signal. Data from the upper side band in the  $B^0$  invariant mass distribution, well separated from the signal region, are used as a background data sample. Twenty-two variables are chosen as inputs for the Neural Network training, the list can be found in Table III. After the training, the Neural Network achieves strong discrimination between signal and background as shown in Figure 1. The selection is optimized by maximizing  $S/(1.5 + \sqrt{B})$ . This quantity is well accepted for signal discovery as described in [4]. The events selected as signal,  $S$ , are in the reconstructed mass range  $5.35 \text{ GeV}/c^2 < M_B < 5.4 \text{ GeV}/c^2$ . For the background sample,  $B$ , the range is  $5.43 \text{ GeV}/c^2 < M_B < 5.48 \text{ GeV}/c^2$ . The figure of merit suggests a cut in the Neural Network response of 0.88 as shown in Figure 1.

### Optimization of candidates selection in the $B_s \rightarrow J/\psi K^*$ analysis

Likewise the figure of merit chosen for optimization is  $S/(1.5 + \sqrt{B})$ . For the signal sample,  $S$ , a  $B_s \rightarrow J/\psi K^*$  MC is used. For the background sample,  $B$ , an upper sideband in the  $B^0 \rightarrow J/\psi K^*$  reconstructed invariant mass plot is chosen,  $5.6 \text{ GeV}/c^2 < M_B < 5.8 \text{ GeV}/c^2$ . A simultaneous four-dimensional optimization is carried out over  $\pi p_T$ ,

$K$   $p_T$ , transverse decay length  $L_{xy}(B_s)$  and  $B_s$  vertex fit probability. The final cuts used are  $p_T(K, \pi) > 1.5$  GeV/ $c$ ,  $L_{xy}(B) > 300$   $\mu\text{m}$  and fit probability  $> 10^{-5}$ . Figure 2 shows the  $S/(1.5 + \sqrt{B})$  variation as a function of the four optimization variables.

### III. BINNED LIKELIHOOD FIT

For the purpose of extracting the yields of  $B^0 \rightarrow J/\psi K_S$ ,  $B_s \rightarrow J/\psi K_S$ ,  $B^0 \rightarrow J/\psi K^*$  and  $B_s \rightarrow J/\psi K^*$  signals in the invariant mass distributions, an accurate modeling for signals and backgrounds is needed prior to the fit.

#### Signals

In both analyses, the signal contributions are modeled with three Gaussians template obtained from fits to  $B^0$  MC. The relative contributions, means and widths from each Gaussian are fixed in the final fit. For both analyses, the  $B_s$  templates used in the final fit are identical to  $B^0$  templates, except for a shift of 86.8 MeV/ $c^2$  in the mean value of the three Gaussians. This value corresponds to the PDG [5] mass difference between  $B_s$  and  $B^0$ . It is important to note that the MC generally underestimates the widths of the mass distribution. Therefore, the Gaussian widths of the two narrowest Gaussians are multiplied by a scale factor, which is allowed to float in the final fit. The scale factor is not applied to the third Gaussian since it is not expected to be governed by detector resolution effects as the other two. Moreover, a mass shift is added to the means of all Gaussians templates to account for a possible mass mismodeling in the MC.

#### Common backgrounds

Both analyses have two common background contributions: the combinatorial background and the partially reconstructed background. The first one is resulting from different sources, for example a real  $J/\psi$  plus two random tracks, where the  $J/\psi$  could be a prompt  $J/\psi$  or coming from a B decay. Other sources that could contribute to it are fake  $J/\psi$  reconstructed with prompt fake muons or fake muons coming from heavy flavor. The combinatorial background is modeled in the final fit with an exponential function, where the fraction of combinatorial background events and the decay constant are allowed to float. The other background that emerges is partially reconstructed B-hadrons where a five-body decay occurs where a  $\pi$ ,  $K$ , or  $\gamma$  is not reconstructed. This background is fitted with an ARGUS function [6]. All the ARGUS function parameters are allowed to float in the final fit.

#### Specific backgrounds in the $B_s \rightarrow J/\psi K_S$ analysis

$\Lambda_b \rightarrow J/\psi \Lambda$ , where  $\Lambda \rightarrow p\pi$ , is a background in  $B_s \rightarrow J/\psi K_S$  analysis when the  $p$  is reconstructed as a  $\pi$ . In order to suppress the  $\Lambda_b$  contribution, a cut in the angular variable  $\cos(\theta_{K_S, \pi_2})$ , where  $\pi_2$  is the  $\pi$  with lower  $p_T$ , is applied. The angle  $\theta$  is defined as the angle between the  $K_S$  and the  $\pi_2$  in the  $K_S$  center of mass frame. The cut at  $\cos(\theta_{K_S, \pi_2}) > -0.75$  decreases the acceptance for  $\Lambda_b$  by a factor of 99.8%, but only 14.2% for  $B_s$ .

#### Specific backgrounds in the $B_s \rightarrow J/\psi K^*$ analysis

With the use of the  $J/\psi$  di-muon trigger a background that must be considered is  $B_s \rightarrow J/\psi \phi$ . A template, consisting in two Gaussians, extracted from simulation is used to model this background. The widths, means and relative contributions from each Gaussian are fixed in the final fit. The constant width of the narrowest Gaussian is multiplied by the same scale factor used in the signals templates. The  $B_s \rightarrow J/\psi \phi$  contribution is constrained using data, basically by measuring the  $B_s \rightarrow J/\psi \phi$  in the data, and then using simulation to calculate the fraction of those  $J/\psi \phi$  events that would show up in  $J/\psi K^*$ .

Another potential background contribution in this analysis is  $B_s \rightarrow J/\psi f_0$ . However, a specific study has estimated the size of this contribution is negligible.

## Fit Result

A binned log likelihood fit is performed to the invariant mass distributions using the templates for signals and the functions described above. The mass distributions in data for  $B^0 \rightarrow J/\psi K_S$  and  $B^0 \rightarrow J/\psi K^*$ , the final fit, and the residuals appear in Figure 3 to 6. The yields of the  $B^0 \rightarrow J/\psi K_S$  and  $B_s \rightarrow J/\psi K_S$  signal are determined to be  $5954 \pm 79$  and  $64 \pm 14$ , respectively. The value of  $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$  extracted from the fit is  $0.0108 \pm 0.0019$  (stat.). The yields for  $B^0 \rightarrow J/\psi K^*$  and  $B_s \rightarrow J/\psi K^*$  signal are  $9530 \pm 110$  and  $151 \pm 25$ , respectively. In this case, the ratio  $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$  is  $0.0159 \pm 0.0022$  (stat.).

## IV. SIGNAL SIGNIFICANCE

The statistical significances of the  $B_s \rightarrow J/\psi K_S$  and  $B_s \rightarrow J/\psi K^*$  signals are determined by fitting the mass distribution without these  $B_s$  contributions. For the  $B_s \rightarrow J/\psi K_S$  analysis, being this  $B_s$  signal the only degree of freedom separating the two hypothesis, a  $\Delta\chi^2$  distribution gives a  $p$ -value of  $3.85 \times 10^{-13}$  or  $7.2\sigma$ . Similarly, for  $B_s \rightarrow J/\psi K^*$ , we determined a  $p$ -value of  $8.9 \times 10^{-16}$  or  $8.0\sigma$ . These  $p$ -values are with respect to the background hypothesis.

## V. SYSTEMATICS UNCERTAINTIES IN THE RATIO OF YIELDS

Different sources of systematic uncertainties, which can influence the measured ratio of  $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$  and  $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$ , are discussed below and summarized in Table IV.

### Signal Modeling

The modeling of the  $B^0$  and  $B_s$  signal peaks can influence the measurement of the ratio. To determine the size of the effect that mismodeling has, the fit is repeated using two Gaussians template for the signal in the  $B_s \rightarrow J/\psi K^*$  analysis. The fit value of  $N(B_s)/N(B^0)$  is shifted by  $7 \times 10^{-4}$ . Another method is used in the  $B_s \rightarrow J/\psi K_S$  analysis because the two Gaussians template does not describe the signal shape. In this case, the widths of the three Gaussians are allowed to float in the final fit. This results in a shift of  $5 \times 10^{-4}$  in the ratio of yields.

### Mass difference between $B_s$ and $B^0$

The PDG mass difference between  $B_s$  and  $B^0$ ,  $86.8 \text{ MeV}/c^2$ , is used in the  $B_s$  templates. This value has a  $\pm 0.7 \text{ MeV}/c^2$  uncertainty. Therefore, the mass shift is varied within its uncertainty leading to an average change in the  $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$  value of  $1.3 \times 10^{-5}$ . Alternatively, the ratio  $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$  is shifted on average by  $2 \times 10^{-5}$ . This systematic uncertainty does not have a significant effect on the measured ratio.

### Combinatorial Background Modeling

The shape of the combinatorial background is another source of systematic uncertainty. In this case, a polynomial is used instead of an exponential. Additional systematic uncertainties of  $\pm 6 \times 10^{-4}$  for the  $K_s$  channel and  $\pm 2 \times 10^{-4}$  for the  $K^*$  channel are included in the final measurements to take into account this effect.

### Combinatorial Background Contribution

In the likelihood fits, the combinatorial background contribution is allowed to float. A study was done to evaluate how the ratio of yields change if this contribution is fixed in the final fit. In both analyses, the upper sideband in the invariant mass distribution is used to obtain the combinatorial background contribution before the final fit. A systematic uncertainty is included in both analyses to account for the difference in the ratio of yields between this method and the final one. Therefore an additional systematic uncertainty of  $\pm 0.0006$  is included

in  $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$  value. Likewise, a systematic uncertainty of  $\pm 0.0050$  is added in the  $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$  ratio.

### $B_s \rightarrow J/\psi\phi$ Contribution

In order to study the uncertainty in the  $B_s \rightarrow J/\psi\phi$  contribution, the fraction of candidates that are  $B_s \rightarrow J/\psi\phi$  is doubled from its normal contribution and the fit is performed again. A shift of  $2 \times 10^{-4}$  is assigned to cover the size of the uncertainty in the  $B_s \rightarrow J/\psi\phi$  contribution.

The different contributions are added in quadrature resulting in a total systematic uncertainty of  $\pm 0.0010$  for  $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$  and  $\pm 0.0050$  for  $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$ .

## VI. ACCEPTANCE CALCULATION

To determine the  $Br(B_s \rightarrow J/\psi K)/Br(B^0 \rightarrow J/\psi K)$ , where  $K$  represents  $K_S$  or  $K^*$ , the relative acceptance of  $B^0 \rightarrow J/\psi K_S$  to  $B_s \rightarrow J/\psi K_S$  and  $B^0 \rightarrow J/\psi K^*$  to  $B_s \rightarrow J/\psi K^*$  need to be determined. MC samples of  $B^0 \rightarrow J/\psi K_S$ ,  $B_s \rightarrow J/\psi K_S$ ,  $B^0 \rightarrow J/\psi K^*$ , and  $B_s \rightarrow J/\psi K^*$  are used to extract  $A_{rel}$  as follows:

$$A_{rel} = \frac{N(B^0 \rightarrow J/\psi K \text{ passed})/N(B^0 \rightarrow J/\psi K \text{ generated})}{N(B_s \rightarrow J/\psi K \text{ passed})/N(B_s \rightarrow J/\psi K \text{ generated})}, \quad (1)$$

where  $K$  represents  $K_S$  or  $K^*$ . The number of *passed* candidates is simply the number that passed the event selection criteria and the number of *generated* is number of candidates generated by the MC.

The value for  $A_{rel}$  is determined to be  $A_{rel} = 1.012 \pm 0.010$  for the  $K_S$  channel and  $1.057 \pm 0.010$  for the  $K^*$  channel. The statistical uncertainty on the acceptances for  $B^0$  and  $B_s$  are determined assuming binomial statistics. These uncertainties are then propagated through using Gaussian uncertainties for  $A_{rel}$  and added in as a systematic uncertainty for the branching ratio calculations. Different systematics uncertainties for  $A_{rel}$  have being evaluated.

### Lifetime for $B^0$ and $B_s$

The  $B_s$  and  $B^0$  lifetimes play a role in how well  $A_{rel}$  is known. For  $B_s \rightarrow J/\psi K_S$  analysis, different MC samples have been generated modifying the lifetime  $1\sigma$  up and down from their PDG values. The maximum deviation is 0.028 and this value has been taken as a systematic uncertainty. For  $B_s \rightarrow J/\psi K^*$  analysis, the procedure to evaluate the systematic uncertainty is slightly different. It is likely that  $B_s \rightarrow J/\psi K^*$  is not equally CP even or odd. The PDG currently gives  $\Delta\Gamma_{B_s}/\Gamma_{B_s} = 0.092^{+0.051}_{-0.054}$  for  $\Gamma_{B_s} = \frac{1}{2}(\Gamma_{B_{sH}} + \Gamma_{B_{sL}})$  where  $\Gamma_{B_{sH}}$  and  $\Gamma_{B_{sL}}$  are the widths of the heavy and light mass eigenstates, respectively. If the  $B_s$  was either all  $B_{sH}$  or  $B_{sL}$ , the most the lifetime would change is 5%. Therefore, the lifetime is shifted up and down by 5% in order to evaluate the systematic uncertainty due to the  $B_s$  lifetime. Additionally,  $B^0$  MC samples are generated with lifetime varied by  $\pm 1\sigma$ . All of this leads to a systematic uncertainty on  $A_{rel}$  of  $\pm 0.009$ .

### B hadron $p_T$ spectrum

The default  $B_s$  and  $B^0$  samples are generated using the NDE NLO calculation [7]. Additional samples are produced using the  $p_T$  spectrum measured in the  $B \rightarrow J/\psi X$  analysis [8]. For  $B_s \rightarrow J/\psi K_S$  analysis, the value of  $A_{rel}$  varies by 0.032 using these additional samples and this is added as systematic uncertainty. Likewise, for  $B_s \rightarrow J/\psi K^*$  analysis, the change in  $A_{rel}$  is 0.029.

### Polarization in $B_s \rightarrow J/\psi K^*$ decay

To compute a systematic arising from the polarization in  $B_s \rightarrow J/\psi K^*$  decay, the most conservative approach possible is taken. With the assumption that the polarization can take on any possible value, different MC samples

are generated. The maximum variation from any of these polarizations leads to a systematic uncertainty on  $A_{rel}$  of 0.261.

Table V shows a summary of the systematic uncertainties on  $A_{rel}$  for both measurements. The different contributions are added in quadrature in the total systematic uncertainties. The values of  $A_{rel}$ , including all uncertainties, are  $A_{rel} = 1.012 \pm 0.010$  (stat.)  $\pm 0.042$  (sys.) for the  $K_S$  analysis and  $A_{rel} = 1.057 \pm 0.010$  (stat.)  $\pm 0.263$  (sys.) for the  $K^*$  analysis.

## VII. RESULTS

With the values of  $A_{rel}$ , the measurements of the  $f_s Br(B_s \rightarrow J/\psi K_S)/f_d Br(B^0 \rightarrow J/\psi K_S)$  and  $f_s Br(B_s \rightarrow J/\psi K^*)/f_d Br(B^0 \rightarrow J/\psi K^*)$  are made to be:

$$f_s Br(B_s \rightarrow J/\psi K_S)/f_d Br(B^0 \rightarrow J/\psi K_S) = 0.0109 \pm 0.0019(stat.) \pm 0.0011(sys.)$$

$$f_s Br(B_s \rightarrow J/\psi K^*)/f_d Br(B^0 \rightarrow J/\psi K^*) = 0.0168 \pm 0.0024(stat.) \pm 0.0068(sys.).$$

To determine  $f_s/f_d$ , the most recent CDF measurement [9] of  $f_s/(f_u + f_d) \times Br(D_s \rightarrow \phi\pi)$  is combined with the actual PDG value [5] for  $Br(D_s \rightarrow \phi\pi)$ . With the input of  $f_s/f_d = 0.269 \pm 0.033$ , the ratio of branching fractions to the reference  $B^0$  decays are:

$$Br(B_s \rightarrow J/\psi K^*)/Br(B^0 \rightarrow J/\psi K^*) = 0.062 \pm 0.009(stat.) \pm 0.025(sys.) \pm 0.008(frag.)$$

$$Br(B_s \rightarrow J/\psi K_S)/Br(B^0 \rightarrow J/\psi K_S) = 0.041 \pm 0.007(stat.) \pm 0.004(sys.) \pm 0.005(frag.).$$

The PDG values for  $Br(B^0 \rightarrow J/\psi K^*)$  and  $Br(B^0 \rightarrow J/\psi K^0)$  are used to calculate the absolute branching fractions:

$$Br(B_s \rightarrow J/\psi K^*) = (8.3 \pm 1.2(stat.) \pm 3.3(sys.) \pm 1.0(frag.) \pm 0.4(PDG)) \times 10^{-5}$$

$$Br(B_s \rightarrow J/\psi K^0) = (3.5 \pm 0.6(stat.) \pm 0.4(sys.) \pm 0.4(frag.) \pm 0.1(PDG)) \times 10^{-5}.$$

## VIII. 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 Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

- 
- [1] R. Fleischer, Eur. Phys. J. C 10, 299-306, (1999).  
[2] S. Faller, R. Fleischer and T. Mannel, *Phys. Rev. D* 79, 014005, (2009).  
[3] R. Blair et. al., FERMILAB-PUB-96/390-E, (1996).  
[4] G. Punzi, *arXiv:physics/0308063*.

- [5] C. Amsler et. al., *Phys. Lett. B* **667**, 1, (2008).
- [6] ARGUS Collaboration, H. Albrecht, *Phys. Lett. B* **241**, 278, (1990)
- [7] P. Nason, S. Dawson and R.K. Ellis, *Nucl. Phys. B* **303**, 607, (1988).
- [8] D. Acosta et. al., *Phys. Rev. D* **71**, 032001, (2005).
- [9] T. Aaltonen et. al., *Phys. Rev. D* **77**, 072003, (2008).

Selection Cuts for $B^0 \rightarrow J/\psi K_S$		
Particle	Variable	Cut
$B^0$	mass	$3 \text{ GeV}/c^2 < \text{Mass}(B^0) < 6.7 \text{ GeV}/c^2$
	4 track vertex fit $\chi^2$	$< 50$
	4 track vertex fit Probability	$> 10^{-5}$
	transverse decay length significance, $L_{xy}/\sigma$	$> 2$
	transverse momentum, $p_T$	$> 4 \text{ GeV}/c$
$J/\psi$	mass	$2.8 \text{ GeV}/c^2 < \text{Mass}(J/\psi) < 3.3 \text{ GeV}/c^2$
	2 track vertex fit $\chi^2$	$< 30$
$K_S$	mass	$0.45 \text{ GeV}/c^2 < \text{Mass}(K_S) < 0.55 \text{ GeV}/c^2$
	2 track vertex fit $\chi^2$	$< 20$
	transverse decay length, $L_{xy}$	$> 0.5 \text{ cm}$
$\mu$	transverse momentum, $p_T$	$> 1.5 \text{ GeV}/c$
	$\Delta\phi$ between the two muons	$< 2.25 \text{ radians}$
	Charge( $\mu_1$ )*Charge( $\mu_2$ )	$< 0$
	$\Delta z(\text{vertex})$ between the two muons	$< 5 \text{ cm}$
$\pi$	transverse momentum, $p_T$	$> 0.5 \text{ GeV}/c$

TABLE I: Selection cuts for  $B^0 \rightarrow J/\psi K_S$  candidates.

Selection Cuts for $B^0 \rightarrow J/\psi K^*$		
Particle	Variable	Cut
$B^0$	mass	$3 \text{ GeV}/c^2 < \text{Mass}(B^0) < 6.7 \text{ GeV}/c^2$
	4 track vertex fit $\chi^2$	$< 50$
	transverse momentum, $p_T$	$> 6 \text{ GeV}/c$
	impact parameter	$< 50 \mu\text{m}$
$J/\psi$	mass	$3.05 \text{ GeV}/c^2 < \text{Mass}(J/\psi) < 3.15 \text{ GeV}/c^2$
	2 track vertex fit $\chi^2$	$< 30$
$K^*$	mass	$0.846 \text{ GeV}/c^2 < \text{Mass}(K^*) < 0.946 \text{ GeV}/c^2$
	2 track vertex fit $\chi^2$	$< 30$
$\mu$	transverse momentum, $p_T$	$> 1.5 \text{ GeV}/c$
	$\Delta\phi$ between the two muons	$< 2.25 \text{ radians}$
	Charge( $\mu_1$ )*Charge( $\mu_2$ )	$< 0$
	$\Delta z(\text{vertex})$ between the two muons	$< 5 \text{ cm}$

TABLE II: Selection cuts for  $B^0 \rightarrow J/\psi K^*$  candidates.

Input variables in the Neural Network	
Particle	Variable
$B^0$	transverse momentum, $p_T$
	4 track vertex fit $\chi^2$
	proper decay length, ct
	impact parameter, $d_{xy}$
$J/\psi$	$p_T$
	mass
	proper decay length, ct
	impact parameter, $d_{xy}$
$K_S$	transverse momentum, $p_T$
	mass
	proper decay length, ct
	impact parameter, $d_{xy}$
both $\pi$ 's	transverse momentum, $p_T$
	impact parameter, $d_{xy}$
both $\mu$ 's	transverse momentum, $p_T$
	impact parameter, $d_{xy}$
	$\cos(\theta_{B,\mu_i})$

TABLE III: Variables used as input in the Neural Network training.

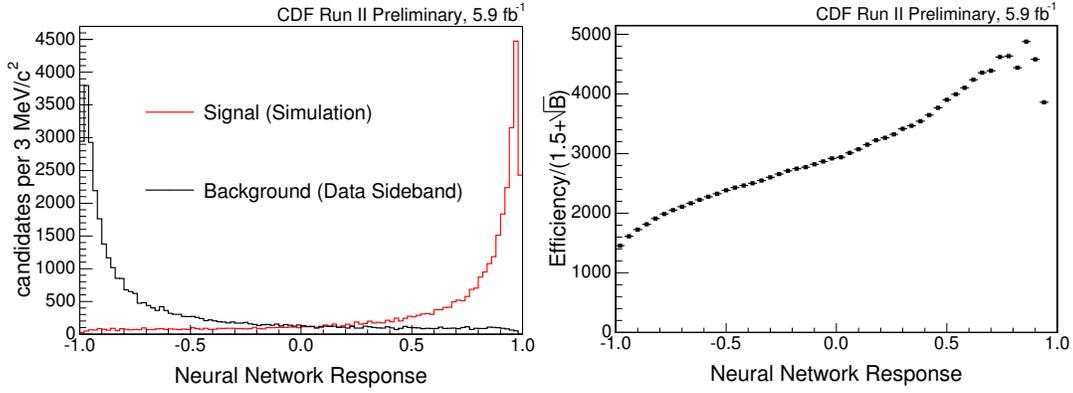


FIG. 1: On the left, Neural Network response where the red histogram is signal MC and the black one is sideband data. On the right, figure of merit  $S/(1.5 + \sqrt{B})$  as a function of Neural Network response.

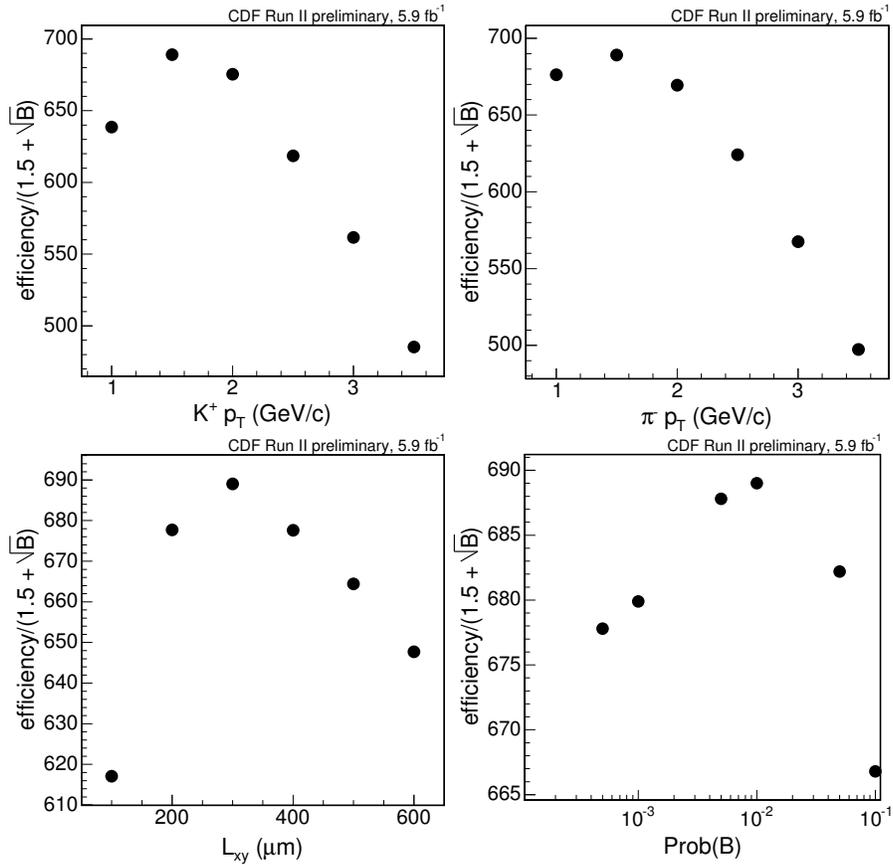


FIG. 2:  $S/(1.5 + \sqrt{B})$  as a function of  $K p_T$  (top left),  $\pi p_T$  (top right),  $L_{xy}(B_s)$  (bottom left) and  $B_s$  vertex fit probability (bottom right).

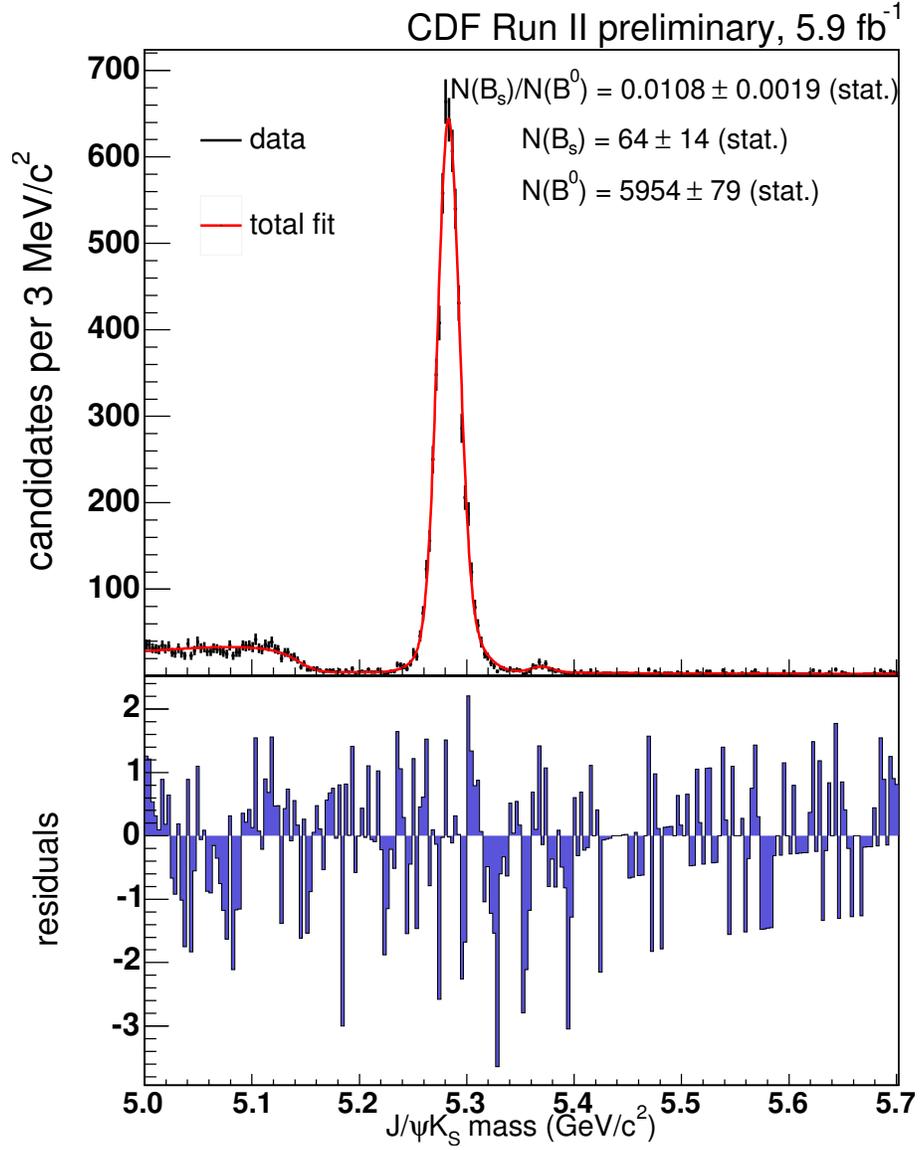


FIG. 3: Invariant mass distribution for  $J/\psi K_S$  overlaid with the binned likelihood fit and the residuals.

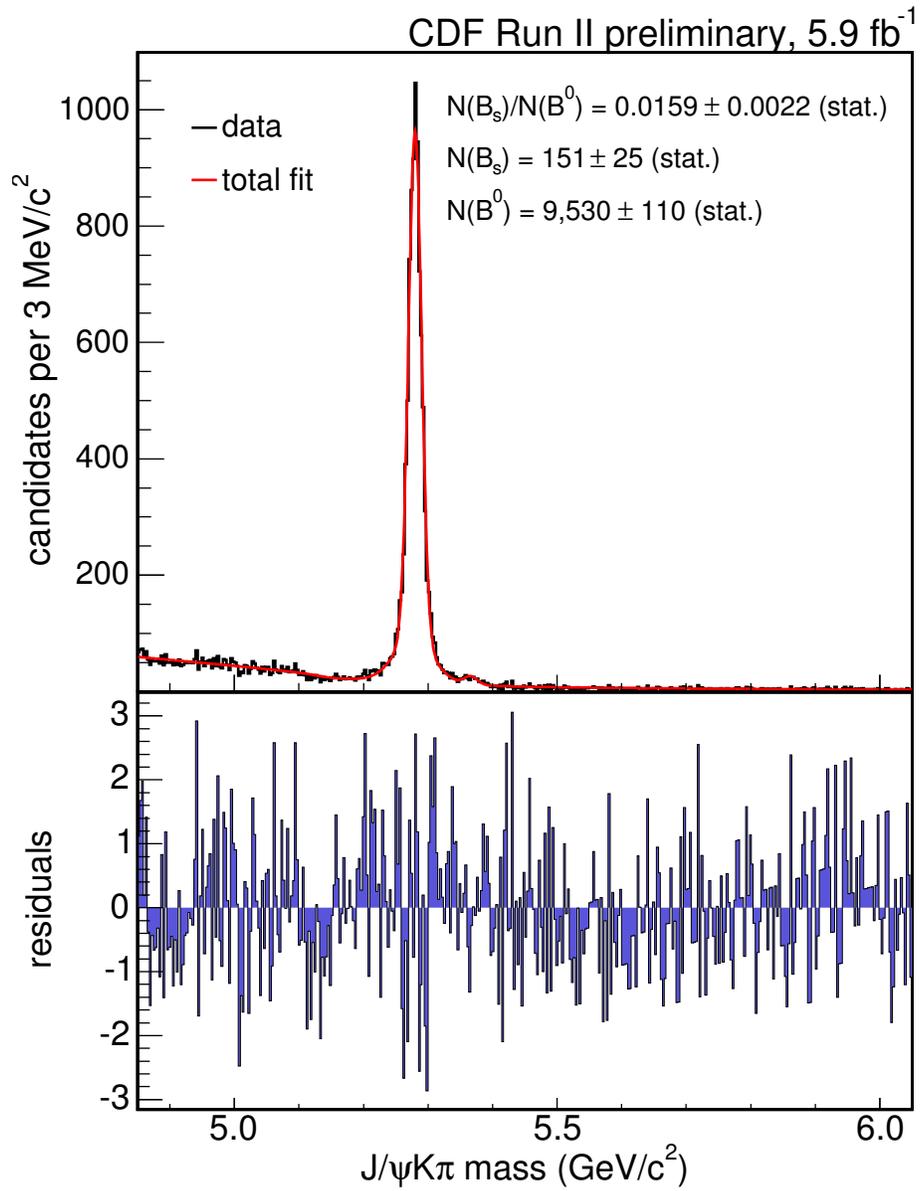


FIG. 4: Invariant mass distribution for  $J/\psi K^*$  overlaid with the binned likelihood fit and the residuals.

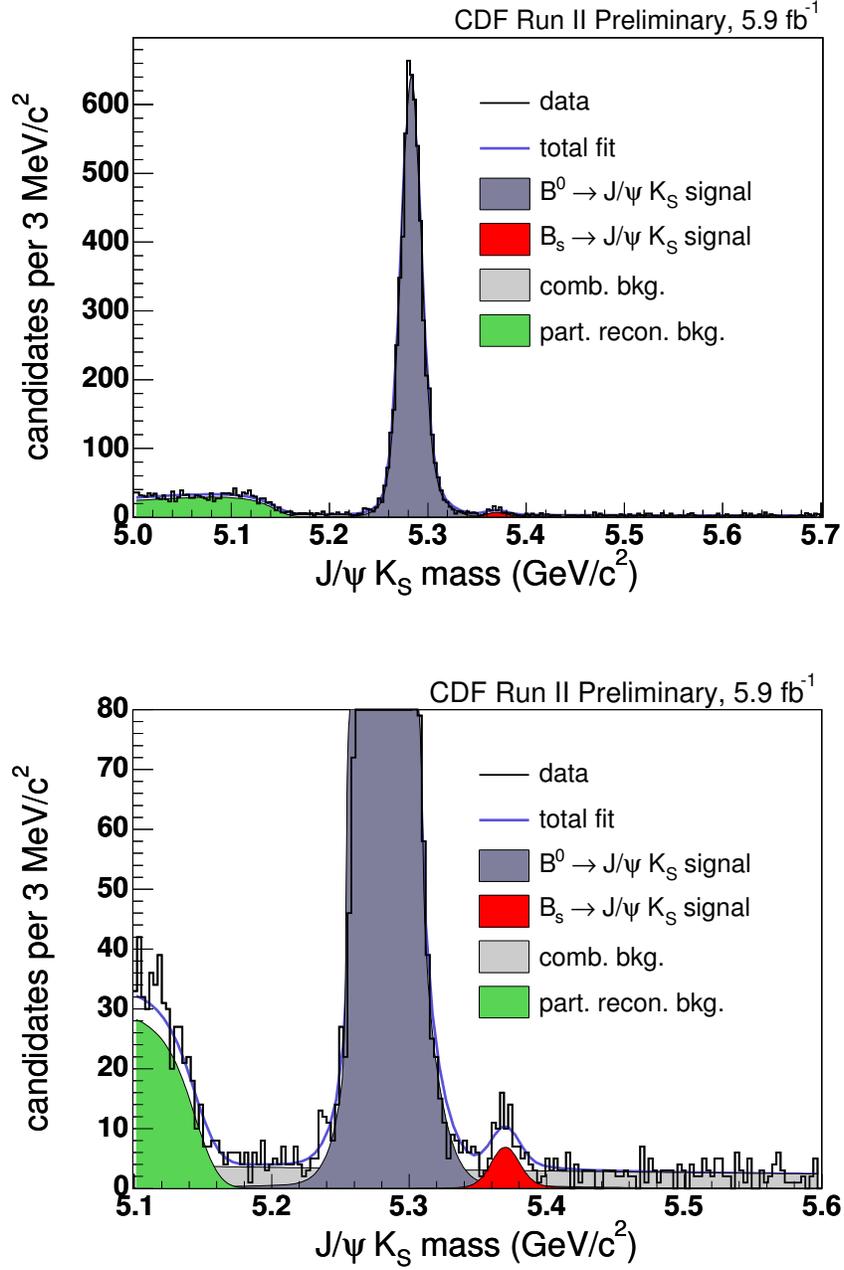


FIG. 5: Invariant mass distribution for  $J/\psi K_S$  and fit including the different contributions (top). The distribution is enlarged in the signal region for more detail (bottom).

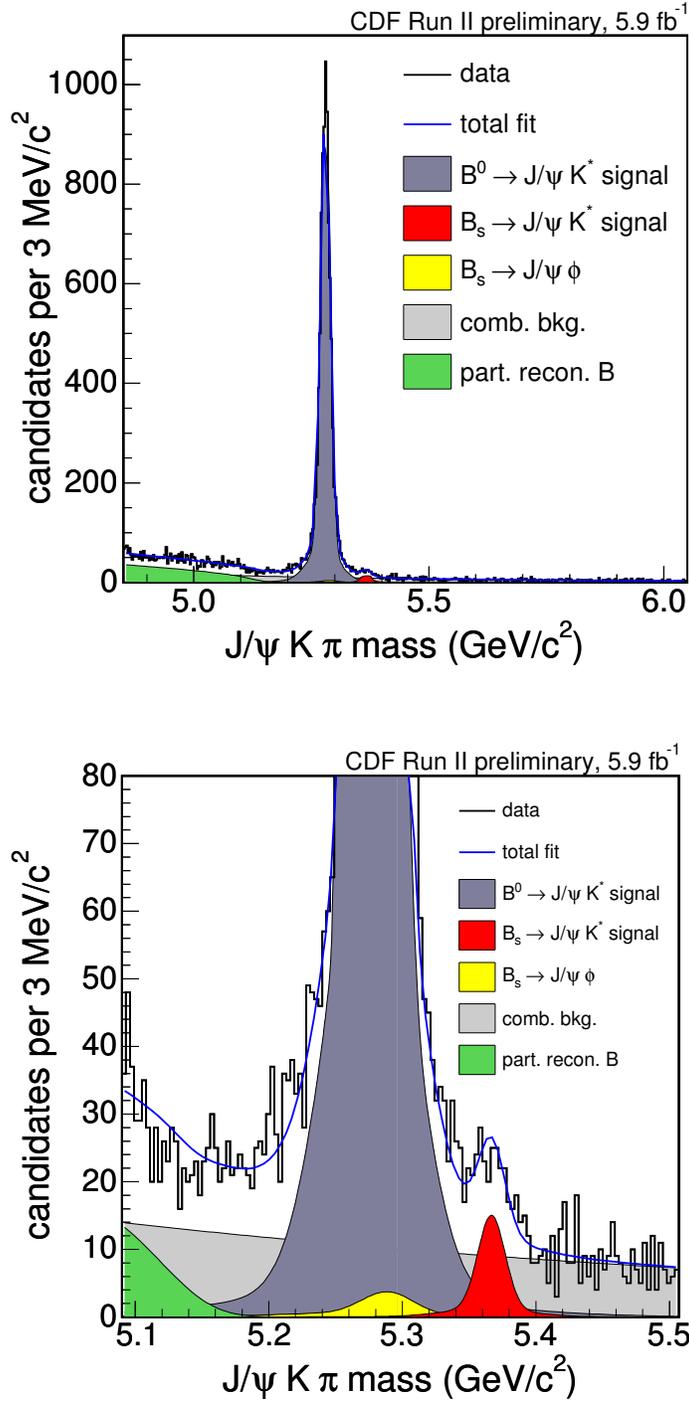


FIG. 6: Invariant mass distribution for  $J/\psi K^*$  and fit including the different contributions (top). The distribution is enlarged in the signal region for more detail (bottom).

Source of Systematic Uncertainties	Relative Uncertainty for $N(B_s \rightarrow J/\psi K_S)/N(B^0 \rightarrow J/\psi K_S)$	Relative Uncertainty for $N(B_s \rightarrow J/\psi K^*)/N(B^0 \rightarrow J/\psi K^*)$
Signal Modeling	4.6 %	4.4 %
Mass difference between $B^0$ and $B_s$	0.1 %	0.1 %
Combinatorial Background Modeling	5.6 %	1.3 %
Combinatorial Background Contribution	5.6 %	31.4 %
$B_s \rightarrow J/\psi \phi$ Contribution	-	1.3 %

TABLE IV: Systematic uncertainties for the ratio of yields.

Source of Systematic Uncertainties	Relative Uncertainty for $A_{rel}$ in $B_s \rightarrow J/\psi K_S$	Relative Uncertainty for $A_{rel}$ in $B_s \rightarrow J/\psi K^*$
Lifetime for $B^0$ and $B_s$	2.8 %	0.9 %
B hadron $p_T$ spectrum	3.2 %	2.7 %
Polarization	-	24.7 %

TABLE V: Systematic uncertainties for the relative acceptances.