Measurement of
\[ \mathcal{B}(B_s^0 \rightarrow \psi(2S)\phi)/\mathcal{B}(B_s^0 \rightarrow J/\psi\phi) \]

The LHCb Collaboration

Abstract

We describe a measurement of the ratio of the branching fractions for \( B_s^0 \rightarrow \psi(2S)\phi \) and \( B_s^0 \rightarrow J/\psi\phi \) with approximately 36 pb\(^{-1}\) of data collected at \( \sqrt{s} = 7 \) TeV during the 2010 run at LHCb. The result is \( 0.68 \pm 0.10 \) (stat) \( \pm 0.09 \) (syst) \( \pm 0.07 \) (B), which is compatible with previous measurements from experiments at the Tevatron.

\(^1\)Conference report prepared for the 13th International Conference on B-Physics at Hadron Machines (BEAUTY), Amsterdam, 4–8 April 2011; contact author: Paul Schaack
1 Introduction

We report a measurement of the ratio of branching fractions for the two decays $B_0^s \rightarrow \psi(2S)\phi$ and $B_0^s \rightarrow J/\psi \phi$. The data set for this analysis corresponds to 36 pb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 7$ TeV recorded with the LHCb detector. The LHCb detector is described in detail elsewhere [1]. We consider the decay modes where $\psi(2S) \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$, such that the branching fraction ratio is given by the formula:

$$\frac{\mathcal{B}(B_0^s \rightarrow \psi(2S)\phi)}{\mathcal{B}(B_0^s \rightarrow J/\psi \phi)} = \frac{N_{\psi(2S)\phi}}{N_{J/\psi \phi}} \times \frac{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}{\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)} \times \frac{\epsilon_{J/\psi \phi}}{\epsilon_{\psi(2S)\phi}} \quad (1)$$

where $N_{\psi(2S)\phi}/N_{J/\psi \phi}$ is the ratio of the numbers of signal events and $\epsilon_{J/\psi \phi}/\epsilon_{\psi(2S)\phi}$ is the ratio of the combined reconstruction, selection and trigger efficiencies. Many of the techniques employed in this analysis are applicable to a measurement of the differential branching fraction of the rare radiative decay $B_0^s \rightarrow \phi\mu^+\mu^-$, that we expect to be able to make with the data collected in 2011.

2 Event selection

Four charged tracks ($\mu^+\mu^-K^+K^-$) are combined to form a $B_0^s$ candidate, accepting the entire dimuon mass range ($2M_{\mu^+\mu^-} < M_{\mu^+\mu^-} < M_{B_0^s} - M_{\phi}$). The $B_0^s$ decay vertex is on average displaced by $\mathcal{O}(10\,\text{mm})$ from the primary vertex (PV), where the $pp$ collision takes place in LHCb. We take advantage of the $B_0^s$ lifetime to reject tracks coming directly from the PV. This is achieved by requiring that the $\chi^2_{IP}$ of each track be greater than 9, where the $\chi^2_{IP}$ is formed by the hypothesis that the track’s impact parameter with respect to the PV is equal to zero. Furthermore, we require the $K^+$ and $K^-$ candidates to be identified as kaons and the $\mu^+$ and $\mu^-$ candidates to be identified as muons by the LHCb particle identification system. For the kaons we require the difference in log-likelihood ($\Delta \log L$) between the kaon and pion hypotheses based on information from the ring imaging Cherenkov detectors to be greater than 0. This requirement is $> 95\%$ efficient for kaons and typically rejects about 95\% of the pions. Additionally we require a good track quality ($\chi^2/\text{n.d.o.f.} < 5$) and the transverse momentum with respect to the beam axis, $p_T$, to be greater than 300 MeV/$c$ for each of the four daughters. All duplicate tracks are removed by only selecting tracks with a Kullback-Liebler (KL) distance greater than 5000 [2]. The requirements on the $B_0^s$ candidate are that its proper time ($\tau$) is greater than 0.2 ps, its vertex (formed by the four daughter tracks) is of good quality ($\chi^2_{VX} < 40$) and its $\chi^2_{IP}$ with respect to the PV is less than 9. We require the dikaon mass, $M_{KK}$, to be within $\pm 10\,$MeV of the nominal $\phi$ mass [3]. A summary of the selection requirements for this analysis can be found in Table 1. All candidates passing the LHCb trigger system are considered.

The selection requirements have been optimised with Monte Carlo simulated data (MC) in order to keep signal and reduce background events in real data. For the optimisation we used data events in the upper $B_0^s$ mass sideband (described below) as background and MC as signal. The MC samples used in this study have been tuned to resemble closely the detector performance.
during the 2010 data taking. We use a \( B_0^s \to J/\psi \phi \) sample as well as a \( B_0^s \to \phi \mu \mu \) sample in order to model all regions of the \( M_{\mu \mu} \) spectrum with sufficient statistics.

Table 1: Variables used to select \( B_0^s \to \psi(2S)\phi \) and \( B_0^s \to J/\psi \phi \) events for this analysis. The column on the right shows the values of the selection variables.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Selection value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K^\pm, \mu^\pm, \chi^2_{IP} )</td>
<td>&gt; 9</td>
</tr>
<tr>
<td>( K^\pm, \Delta \log \mathcal{L} )</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>( K^\pm, \mu^\pm ) track ( \chi^2/\text{n.d.o.f.} )</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>( K^\pm, \mu^\pm p_T )</td>
<td>&gt; 300 MeV/c</td>
</tr>
<tr>
<td>( K^\pm, \mu^\pm ) KL distance</td>
<td>&gt; 5000</td>
</tr>
<tr>
<td>( B_0 \tau )</td>
<td>&gt; 0.2 ps</td>
</tr>
<tr>
<td>( B_0 \chi^2_{VX} )</td>
<td>&lt; 40</td>
</tr>
<tr>
<td>( B_0 \chi^2_{IP} )</td>
<td>&lt; 9</td>
</tr>
<tr>
<td>( M_{KK} )</td>
<td>∈ [1009 : 1029] MeV/c^2</td>
</tr>
</tbody>
</table>

The reconstructed mass distribution of all \( B_0^s \) candidates in the data set is shown in Figure 1. We perform an extended unbinned maximum likelihood fit to the reconstructed mass distribution, with the signal and background parametrised by a Gaussian and a linear function, respectively. From the fit, we obtain the number of signal events (820 ± 34) and the Gaussian parameters (\( M_{J/\psi \phi} = 5363.0 \pm 0.7 \) MeV/c^2 and \( \sigma = 18.4 \pm 0.7 \) MeV/c^2), where the uncertainties are statistical only. From this we define a signal window of \( M_{B_0^s} = [5300 : 5430] \) MeV/c^2, which corresponds to a width of about ±3\( \sigma \). We define lower and upper background windows of \( M_{B_0^s} = [5200 : 5300] \) MeV/c^2 and \( M_{B_0^s} = [5430 : 5700] \) MeV/c^2, respectively. Furthermore, we define the dimuon mass windows for the \( J/\psi \) and the \( \psi(2S) \) to be \( M_{\mu \mu} = [3040 : 3150] \) MeV/c^2 and \( M_{\mu \mu} = [3630 : 3740] \) MeV/c^2, which also correspond to widths of about ±3\( \sigma \).

3 Measurement of \( N_{\psi(2S)\phi}/N_{J/\psi \phi} \)

By selecting events from the data set that lie inside the dimuon mass windows as defined in Section 2 we obtain two samples of potential \( B_0^s \to J/\psi \phi \) and \( B_0^s \to \psi(2S)\phi \) events. For each we perform an extended unbinned maximum likelihood fit to the reconstructed \( B_0^s \) mass distribution. The signal is parametrised by a Gaussian on top of a linear background. The two fitted functions are shown in Figure 2. From the fitted functions, we obtain the Gaussian parameters (\( M_{J/\psi \phi} = 5363.0 \pm 0.7 \) MeV/c^2 with \( \sigma_{J/\psi \phi} = 18.4 \pm 0.7 \) MeV/c^2 and \( M_{\psi(2S)\phi} = 5364.0 \pm 3.3 \) MeV/c^2 with \( \sigma_{\psi(2S)\phi} = 21.2 \pm 2.9 \) MeV/c^2) and the numbers of signal events (\( N_{J/\psi \phi} = 760 \pm 29 \) and \( N_{\psi(2S)\phi} = 62 \pm 9 \)), where the uncertainties are statistical only. We obtain:

\[
\frac{N_{\psi(2S)\phi}}{N_{J/\psi \phi}} = 0.082 \pm 0.012(\text{stat}) \pm 0.010(\text{syst})
\]  

where the first uncertainty is statistical and the second is systematic. There are two main contribu-
Figure 1: Fitted $B_s^0$ mass ($M_{B_s}$) distribution for all candidates in the data set. The points show the data, the dashed line (red) is the linear background component, the thin solid line (blue) is the signal component of a Gaussian curve and the thick solid line (blue) is the combined fit to the data.

The overall efficiency is the product of the geometrical acceptance of LHCb ($\varepsilon_{\text{geo}}$), the detection, reconstruction and selection efficiencies ($\varepsilon_{\text{rec&sel}}$) and the trigger efficiency ($\varepsilon_{\text{trigger}}$). For our analysis we need to evaluate the ratio of these efficiencies between $B_s^0 \rightarrow J/\psi \phi$ and $B_s^0 \rightarrow \psi(2S)\phi$ separately to obtain the overall efficiency ratio, given by:

$$\frac{\varepsilon_{J/\psi \phi}}{\varepsilon_{\psi(2S)\phi}} = \frac{\varepsilon_{\text{geo}}}{\varepsilon_{\psi(2S)\phi}} \times \frac{\varepsilon_{\text{rec&sel}}}{\varepsilon_{\psi(2S)\phi}} \times \frac{\varepsilon_{\text{trigger}}}{\varepsilon_{\psi(2S)\phi}}$$

where the first and second terms need to be evaluated from MC. For the ratio of trigger efficiencies we make the assumption that the ratio is equal to unity and validate it with a method that uses a trigger-unbiased subset of the data. In the following, the errors on each efficiency ratio
term include both statistical and systematic uncertainties, and are combined and accounted as systematic error in the final result.

The ratio of $\varepsilon_{J/\psi \phi}^{\text{geo}} / \varepsilon_{\psi(2S)\phi}^{\text{geo}}$ is estimated using MC. We find:

$$\frac{\varepsilon_{J/\psi \phi}^{\text{geo}}}{\varepsilon_{\psi(2S)\phi}^{\text{geo}}} = 0.996 \pm 0.004$$  \hspace{1cm} (4)

where the error is obtained by varying the parameters of the decay model for $B_s^0 \rightarrow \psi(2S)\phi$. The error corresponds to the difference in the central values for the largest possible variation of the $\psi(2S)\phi$ polarisation (either CP-even or CP-odd).

The selection and reconstruction efficiency, $\varepsilon_{J/\psi \phi}^{\text{rec&sel}}$, is also obtained from MC. We find:

$$\frac{\varepsilon_{J/\psi \phi}^{\text{rec&sel}}}{\varepsilon_{\psi(2S)\phi}^{\text{rec&sel}}} = 1.086 \pm 0.013$$  \hspace{1cm} (5)

where the error accounts for small discrepancies between data and MC. The value of the ratio is significantly different from unity. The reason is that in the MC sample used $\varepsilon_{\text{rec&sel}}$ is a function of $M_{\mu\mu}$ due to one of the applied selection requirements. By tightening each selection requirement in turn we identify that the requirement on $K^\pm \chi_{IP}^2$ generates the efficiency difference between the decays. In regions of large $M_{\mu\mu}$ the $\phi$ has a smaller momentum and transverse momentum ($p_T$) component than in other regions. Therefore the produced kaons have less $p_T$ available and their momentum vector is more likely to point to the primary vertex.
The LHCb trigger is highly efficient in selecting B meson decays with two muons in the final state. For the low level trigger, L0, candidates are typically selected by requiring one or two high-\(p_T\) muons in the final state. In the first stage of LHCb’s high level trigger, HLT, these candidates are confirmed using a form of muon identification similar to that used offline. The events may also be selected if they contain a single high-\(p_T\) and large-IP track. In the second stage of the HLT candidates are selected based on their topology by requiring two-to-three tracks to be reconstructed with a large mass and a vertex that is displaced from the PV. At no stage do we explicitly require all four tracks to be found. One benefit of this ‘inclusive’ approach is that we expect trigger efficiencies to be very close for \(B^0_s \rightarrow \phi \mu \mu\), \(J/\psi \phi\) and \(\psi(2S)\phi\). We assume the trigger efficiency ratio to be unity and validate this assumption with data and MC.

We can calculate the trigger efficiency by splitting the data into three groups of events: events triggered only by the daughters of the \(B^0_s\) candidate (A), events explicitly not triggered by the daughters of the \(B^0_s\) candidate (B) and events triggered for both reasons (C). The trigger efficiency is given by
\[
\epsilon_{\text{trigger}} = \frac{N_{\text{observed}} \times (N^C + N^B + N^A + \frac{N^B \times N^A}{N^C})}{1}.
\]
Using this technique we can extract the trigger efficiency for the combination of trigger levels or for each trigger level separately (with the condition that the global trigger decision is true). The described technique to measure these efficiencies has been used in data and MC. The results are in agreement between data and MC and are compatible with the ratio being equal to unity, given large statistical errors in data. The uncertainty on our assumption is evaluated using MC as the statistical precision in data is not sufficient. We use:
\[
\frac{\epsilon_{\psi(2S)\phi}}{\epsilon_{J/\psi \phi}} = 1.00 \pm 0.04.
\]

## 5 Results

Using Equation 1 and combining the results of Section 3 and 4 together with the PDG values for \(\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) = (5.93 \pm 0.06) \times 10^{-2}\) and \(\mathcal{B}(\psi(2S) \rightarrow \mu^+ \mu^-) = (7.7 \pm 0.8) \times 10^{-3}\), we derive:
\[
\frac{\mathcal{B}(B^0_s \rightarrow \psi(2S)\phi)}{\mathcal{B}(B^0_s \rightarrow J/\psi \phi)} = 0.68 \pm 0.10(\text{stat}) \pm 0.09(\text{syst}) \pm 0.07(\text{B})
\]

where the first error is statistical, the second error is systematic and the third error is due to the uncertainty on the branching fractions of the \(J/\psi\) and \(\psi(2S)\) decays. A breakdown of the systematic error can be seen in Table 2.

## 6 Conclusions

Our result of the ratio of branching fractions for \(B^0_s \rightarrow \psi(2S)\phi\) and \(B^0_s \rightarrow J/\psi \phi\) is compatible with measurements from CDF and D0, which are shown together with the ratios of similar decays for comparison in Table 3. The statistical and systematic components of the error on our measured ratio are of equal order. When the LHCb detector collects more data, a reduction of the systematic uncertainty will be possible using data-driven methods.
Table 2: The relative contribution of each term to the total systematic error of the final result.

<table>
<thead>
<tr>
<th>Term</th>
<th>syst</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{N_{\psi^\prime}/N_{J/\psi}}{\epsilon_{\psi^\prime}/\epsilon_{\psi}} )</td>
<td>0.08</td>
</tr>
<tr>
<td>( \epsilon_{J/\psi}/\epsilon_{\psi} )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \epsilon_{\text{rec&amp;sel}}/\epsilon_{\psi} )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \epsilon_{\text{trigger}}/\epsilon_{\psi} )</td>
<td>0.03</td>
</tr>
<tr>
<td>Total</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 3: Various branching fraction ratios from BABAR, CDF, D0 and the PDG as comparison.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{B(B_s^0 \to \psi(2S)\phi)}{B(B_s^0 \to J/\psi \phi)} ) (CDF)</td>
<td>0.52 ( \pm ) 0.13(stat) ( \pm ) 0.04(syst) ( \pm ) 0.06(B) [4]</td>
</tr>
<tr>
<td>( \frac{B(B_s^0 \to \psi(2S)\phi)}{B(B_s^0 \to J/\psi \phi)} ) (D0)</td>
<td>0.53 ( \pm ) 0.10(stat) ( \pm ) 0.07(syst) ( \pm ) 0.06(B) [5]</td>
</tr>
<tr>
<td>( \frac{B(B_s^0 \to \psi(2S)K^+)}{B(B_s^0 \to J/\psi K^+)} )</td>
<td>0.82 ( \pm ) 0.12(stat) ( \pm ) 0.13(syst) [6]</td>
</tr>
<tr>
<td>( \frac{B(B^+ \to \psi(2S)K^+)}{B(B^+ \to J/\psi K^+)} )</td>
<td>0.47 ( \pm ) 0.10 [3]</td>
</tr>
<tr>
<td>( \frac{B(B^0 \to \psi(2S)K^{<em>-})}{B(B^0 \to J/\psi K^{</em>-})} )</td>
<td>0.46 ( \pm ) 0.04 [3]</td>
</tr>
<tr>
<td>( \frac{B(B^+ \to \psi(2S)K^{<em>-})}{B(B^+ \to J/\psi K^{</em>-})} )</td>
<td>0.63 ( \pm ) 0.05(stat) ( \pm ) 0.08(syst) [5]</td>
</tr>
</tbody>
</table>

References


[4] A. Abulencia, et al., CDF Collaboration, Observation of \( B_s \to \psi(2S)\phi \) and measurement of ratio of branching fractions \( B(B_s \to \psi(2S)\phi) / B(B_s \to J/\psi \phi) \), Phys. Rev. Lett. 96 (2006) 231801.
