Search for the decay B[0 over s] D[^*][^±]
Decays of $B_s^0$ mesons to final states such as $D^+D^-$, $D^0\bar{D}^0$ [1] and $\pi^+\pi^-$ [2] have been recently observed by LHCb. Such decays can proceed, at short distances, by two types of amplitudes, referred to as weak exchange and rescattering processes, and therefore a measurement of the branching fraction will help us to understand the mechanism behind related decays such as $B_s^0 \rightarrow \pi^+\pi^-$ and $B_s^0 \rightarrow D\bar{D}$. Systematic uncertainties are minimized by using $B^0 \rightarrow D^{*\pm}\pi^{\mp}$ as a normalization channel. We find no evidence for a signal, and set an upper limit on the branching fraction of $B(B_s^0 \rightarrow D^{*\pm}\pi^{\mp}) < 6.1(7.8) \times 10^{-6}$ at 90% (95%) confidence level.

DOI: 10.1103/PhysRevD.87.071101

PACS numbers: 13.25.Hw

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
FIG. 1. Decay diagrams for (a) \( B^0 \rightarrow D^{(*)+}D^{(*)-} \) via weak exchange, (b) \( B^0 \rightarrow D^{(*)+}D^{(*)-} \) via penguin annihilation, and (c) \( B^0 \rightarrow D^{(*)+}\pi^- \) via weak exchange.

signatures of \( b \)-hadron decays: at least one track, with \( p_T > 1.7 \text{ GeV}/c \) and \( \chi^2_{IP} \) with respect to any primary interaction vertex (PV) greater than 16, that subsequently forms a two-, three- or four-track secondary vertex with a high sum of the \( p_T \) of the tracks and significant displacement from the PV. The \( \chi^2_{IP} \) is the difference between the \( \chi^2 \) of the PV reconstruction with and without the considered track. In the offline analysis, the software trigger decision is required to be due to the candidate signal decay.

Candidates that are consistent with the decay chain \( B^0_{(s)} \rightarrow D^{(*)+}\pi^- \), \( D^{(*)-} \rightarrow D^{(*)-}\pi^+ \), \( D^{(*)-} \rightarrow K^+\pi^- \) are selected. The \( D^0 \) and \( D^{(*)-} \) candidate invariant mass values are required to satisfy \( 1814 < m_{D^{(*)-}} < 1914 \text{ MeV}/c^2 \) and \( 2008.78 < m_{D^0p} < 2011.78 \text{ MeV}/c^2 \), respectively, where a \( D^0 \) mass constraint is applied in the evaluation of \( m_{D^0p} \). The bachelor pion, from the \( B^0_{(s)} \) decay, is required to be consistent with the pion mass hypothesis, based on particle identification (PID) information from the ring-imaging Cherenkov detectors [11]. All other selection criteria were tuned on the \( B^0 \rightarrow D^{(*)-}\pi^- \) control channel in a similar manner to that used in another recent LHCb publication [12]. The large yield in the normalization sample allows the selection to be based on data, though the efficiencies are determined using Monte Carlo simulated events in which \( pp \) collisions are generated using PYTHIA 6.4 [13] with a specific LHCb configuration [14]. Decays of hadronic particles are described by EVTGEN [15]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [16] as described in Ref. [17].

The selection requirements include criteria on the quality of the tracks forming the signal candidate, their \( p \), \( p_T \) and inconsistency with the hypothesis of originating from the PV \( (\chi^2_{IP}) \). Requirements are also placed on the corresponding variables for candidate composite particles \( (\tilde{D}^0, B^0_{(s)}) \) together with restrictions of the decay fit \( (\chi^2_{vertex}, \chi^2_{flight}) \), the flight distance \( (\chi^2_{flight}) \), and the cosine of the angle between the momentum vector and the line joining the PV to the \( B^0_{(s)} \) vertex \( (\cos \theta_{dir}) \) [18].

Further discrimination between signal and background categories is achieved by calculating weights for the remaining \( B^0 \) candidates [19]. The weights are based on a simplified fit to the \( B \) candidate invariant mass distribution, where the \( B^0 \) region is neither examined nor included in the fit. The weights are used to train a neural network [20] in order to maximize the separation between categories.

To retain sufficient background events for the network training, the requirement on \( m_{D^0p} \) is not applied. A total of 15 variables are used as input to the network. They include the \( \chi^2_{IP} \) of the four candidate tracks, the \( \chi^2_{vertex}, \chi^2_{flight} \) and \( \cos \theta_{dir} \) of the \( \tilde{D}^0 \) and \( B^0_{(s)} \) candidates, and the \( B^0_{(s)} \) candidate \( p_T \). The \( p_T \) asymmetry and track multiplicity in a cone with half-angle of 1.5 units in the plane of pseudorapidity and azimuthal angle (measured in radians) [21] around the \( B^0_{(s)} \) candidate flight direction are also used. The input quantities to the neural network depend only weakly on the kinematics of the \( B^0_{(s)} \) decay. A requirement on the network output is imposed that reduces the combinatorial background by an order of magnitude while retaining about 75% of the signal. Potential biases from this data-driven method are investigated by training the neural network with different fractions of the data sample. The same results are obtained using a neural network trained on 30%, 40%, 50%, 60%, and 70% of the total data sample.

After all selection requirements are applied, approximately 50,000 candidates are selected in the invariant mass range \( 5150 < m_{D^0p} < 5600 \text{ MeV}/c^2 \). About 1% of events with at least one candidate also contain a second candidate. Such multiple candidates are retained and treated the same as other candidates.

In addition to combinatorial background, candidates may be formed from misidentified or partially reconstructed \( B^0_{(s)} \) decays. Contributions from partially reconstructed decays are reduced by requiring the invariant mass of the \( B^0_{(s)} \) candidate to be above 5150 MeV/c^2. The contribution from \( B^0_{(s)} \) decays to identical final states but without intermediate charmed mesons is negligible due to the requirement on the \( D^{(*)-} \) candidate invariant mass.

A small but significant number of background events are expected from \( B^0 \rightarrow D^-K^+ \) decays with the \( K^+ \) misidentified as a pion. The branching fractions of \( B_0^+ \rightarrow D^{(*)+}K^- \) and \( B_0^- \rightarrow D^{(*)-}p \) are expected to be small due to Cabibbo-Kobayashi-Maskawa supression, so that these potential backgrounds are negligible.

Since the \( B^0 \) decay mode is several orders of magnitude more abundant than the \( B^0_s \) decay, it is critical to understand precisely the shape of the \( B^0 \) signal peak. The dependence of the width of the peak on different kinematic variables of the \( B^0 \) decay was investigated. The strongest correlation was found to be with the angle between the momenta of the \( D^{(*)-} \) candidate and the bachelor \( \pi^+ \) in the lab frame,
denoted as $\theta_{\text{bach}}$. Simulated pseudo-experiments were used to find an optimal number of $\theta_{\text{bach}}$ bins to be used in a simultaneous fit. The outcome is that five bins are used, with ranges 0–0.046, 0.046–0.067, 0.067–0.092, 0.092–0.128, and 0.128–0.4 rad, chosen to have approximately equal numbers of $B^0$ decays in each. The peak width in the highest bin is approximately 60% of that in the lowest bin. The pseudo-experiments show that the simultaneous fit in bins of $\theta_{\text{bach}}$ is approximately 20% more sensitive to a potential $B^0_s$ signal than the fit without binning.

The signal yields are obtained from a maximum likelihood fit to the $D^{*-}\pi^+$ invariant mass distribution in the range 5150–5600 MeV/c$^2$. The fit is performed simultaneously in the five $\theta_{\text{bach}}$ bins. The fit includes double Gaussian shapes, where the two Gaussian functions share a common mean, for $B^0$ and $B^0_s$ signals, together with an exponential component for the partially reconstructed background, a linear component for the combinatorial background and a nonparametric function, derived from simulation, for $B^0 \rightarrow D^{*-}K^+$ decays. The probability density function (PDF) for the $B^0 \rightarrow D^{*-}K^+$ background is shifted by the mass difference between data and simulation for each bin of $\theta_{\text{bach}}$.

The parameters of the double Gaussian shapes are constrained to be identical for $B^0$ and $B^0_s$ signals, with
an offset in their mean values fixed to the known $B^0_s$-$B^0$ mass difference [8]. Additionally, the relative normalization of the two Gaussian functions and the ratio of their widths are constrained within uncertainties to the value obtained in simulation. A total of 33 parameters are allowed to vary in the fit: the ratio of yields $N(B^0_s)/N(B^0)$, the linear slope of the combinatorial background and the exponential parameter of the partially reconstructed background, plus separate parameters in each of the $\theta_{\text{bch}}$ bins to describe the peak position and core Gaussian width of the signal PDF, and the yields of the $B^0$ peak, the combinatorial background, the partially reconstructed background, and the background from $B^0 \rightarrow D^{*-}K^+$.

The results of the fit are shown in Fig. 2. The total number of $B^0 \rightarrow D^{*-}\pi^\pm$ decays is found to be $29400 \pm 400$, and the ratio of yields is determined to be $N(B^0_s)/N(B^0) = (1.4 \pm 3.5) \times 10^{-4}$, where the uncertainty is statistical only. The number of $B^0 \rightarrow D^{*-}K^+$ decays found is $1200 \pm 200$, with a correlation of 7% to the ratio of signal yields.

The ratio of yields is converted to a branching fraction following

$$B(B^0_s \rightarrow D^{*-}\pi^\pm) = \frac{N(B^0_s)}{N(B^0)} \times \frac{\epsilon(B^0)}{\epsilon(B^0_s)} \times \frac{f_d}{f_s} \times \frac{f_s}{f_d} \times B(B^0 \rightarrow D^{*-}\pi^\pm),$$

where $\epsilon(B^0)$ and $\epsilon(B^0_s)$ are the efficiencies for the $B^0$ and $B^0_s$ decay modes respectively, while $f_d (f_s)$ is the probability that a $b$ quark produced in the acceptance results in a $B^0$ ($B^0_s$) meson. Their ratio has been determined to be $f_d/f_s = 0.256 \pm 0.020$ [22].

The total efficiencies are $(0.165 \pm 0.002)\%$ and $(0.162 \pm 0.002)\%$ for the $B^0$ and $B^0_s$ decay modes, respectively, including contributions from detector acceptance, selection criteria, PID and trigger effects. The ratio is consistent with unity, as expected. The PID efficiency is measured using a control sample of $D^{*-} \rightarrow \bar{D}^0\pi^-$, $\bar{D}^0 \rightarrow K^+\pi^-$ decays to obtain background-subtracted efficiency tables for kaons and pions as functions of their $p$ and $p_T$ [2]. The kinematic properties of the tracks in signal decays are obtained from simulation, allowing the PID efficiency for each event to be obtained from the tables. Note that this calibration sample is dominated by promptly produced $D^*$ mesons. The remaining contributions to the total efficiency are determined from simulation and validated using data.

Systematic uncertainties on $B(B^0_s \rightarrow D^{*-}\pi^\pm)$ are assigned due to the following sources, given in units of $1 \times 10^{-6}$, summarized in Table I. Event selection efficiencies for both modes are found to be consistent in simulation to within 2%, yielding a systematic uncertainty of 0.02. The fit model is varied by replacing the double Gaussian signal shapes with double Crystal Ball [23] functions (with both upper and lower tails), changing the linear combinatorial background shape to quadratic and including a possible contribution from $B^0_s \rightarrow D^{*-}K^+$. The nonparametric function for the $B^0 \rightarrow D^{*-}K^+$ background was scaled in each bin to account for the change in the width of the $B^0$ signal. Combined in quadrature these sources contribute 1.44 to the systematic uncertainty. Possible biases in the determination of the fit parameters are investigated by simulated pseudo-experiments, leading to an uncertainty of 0.12. Events with multiple candidates are investigated by performing a fit, having chosen one candidate at random. This fit is performed 100 times, with different seeds, and the spread of the results, 0.22, is taken as the systematic uncertainty. The uncertainty on the quantity $f_s/f_d$ contributes 0.12, while that on $B(B^0 \rightarrow D^{*-}\pi^\pm)$ gives 0.08. Combining all sources in quadrature, the total absolute systematic uncertainty is $1.47 \times 10^{-6}$, and the $B^0_s$ branching fraction is determined to be $B(B^0_s \rightarrow D^{*-}\pi^\pm) = (1.5 \pm 3.8 \pm 1.5) \times 10^{-6}$, where the first uncertainty is statistical and the second is systematic.

A number of cross-checks are performed to test the stability of the result. Candidates are divided based upon the hardware trigger decision into three groups: events in which a particle from the signal decay created a large enough cluster in the calorimeter to fire the trigger, events that were triggered independently of the signal decay and those events that were triggered by both the signal decay and the rest of the event. The neural network and PID requirements are tightened and loosened. The nonparametric function used to describe the background from $B^0 \rightarrow D^{*-}K^+$ decays is smoothed to reduce potential statistical fluctuations. All cross-checks give consistent results.

Since no significant signal is observed, upper limits are set, at both 90% and 95% confidence level (CL), using a Bayesian approach. The statistical likelihood curve from the fit is convolved with a Gaussian function of width given by the systematic uncertainty, and the upper limits are taken as the values containing 90% (95%) of the integral of the likelihood in the physical region. The obtained limits are

$$B(B^0_s \rightarrow D^{*-}\pi^\pm) < 6.1(7.8) \times 10^{-6} \text{ at 90\% (95\%) CL}.$$
of a detectable signal indicates that rescattering effects may make significant contributions to other hadronic decays, such as $B^0_s \rightarrow \pi^+ \pi^-$ and $B^0 \rightarrow DD$, as recently suggested [5].

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCSR (Poland); ANCS/IFA (Romania); MinES, Rosatom, RFBR and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (USA). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centers are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), and GridPP (United Kingdom). We are thankful for the computing resources put at our disposal by Yandex LLC (Russia), as well as to the communities behind the multiple open source software packages on which we depend.


SEARCH FOR THE DECAY $B^0_s \rightarrow D^{*+} \pi^-$

\[ \text{PHYSICAL REVIEW D 87, 071101(R) (2013)} \]
