Observation of the $B^0_s \rightarrow D^{*\pm} D^{\mp}$ decay

LHCb collaboration†

Abstract

A search for the $B^0_s \rightarrow D^{*\pm} D^{\mp}$ decay is performed using proton-proton collision data at centre-of-mass energies of 7, 8 and 13 TeV collected by the LHCb experiment, corresponding to an integrated luminosity of $9 \text{ fb}^{-1}$. The decay is observed with a high significance and its branching fraction relative to the $B^0 \rightarrow D^{*\pm} D^{\mp}$ decay is measured to be

$$\frac{B(B^0_s \rightarrow D^{*\pm} D^{\mp})}{B(B^0 \rightarrow D^{*\pm} D^{\mp})} = 0.137 \pm 0.017 \pm 0.002 \pm 0.006,$$

where the first uncertainty is statistical, the second systematic and the third is due to the uncertainty on the ratio of the $B^0_s$ and $B^0$ hadronisation fractions.

Submitted to JHEP

© 2020 CERN for the benefit of the LHCb collaboration. CC BY 4.0 licence

†Authors are listed at the end of this paper.
1 Introduction

The family of $B$-meson decays into a pair of open-charm mesons are sensitive to elements of the Cabibbo–Kobayashi–Maskawa matrix [1,2]. While $B^0 \to D^{(*)+}D^{(*)-}$ decays can be used to measure the $B^0\bar{B}^0$ mixing phase, $\sin(2\beta)$ [3,7], $B^0_s \to D^{(*)+}_sD^{(*)-}_s$ decays provide access to the $B^0_s\bar{B}^0_s$ mixing phase, $\phi_s$ [8]. Information on additional decays, such as $B^0_s \to D^{(*)+}_sD^{(*)-}_s$, can be exploited to constrain loop and non-factorisable contributions [9–14], which can be notably prominent [15].

Both $B^0 \to D^{(*)+}D^{(*)-}$ and $B^0_s \to D^{(*)+}_sD^{(*)-}_s$ decays occur predominantly through tree or penguin transitions, as shown in Fig. 1. Subleading contributions are expected from $W$-exchange and penguin-annihilation transitions, illustrated in Fig. 2. In contrast, the $B^0_s \to D^{(*)+}D^{(*)-}$ decay is forbidden at tree level and its dominant contributions originate from $W$-exchange and penguin-annihilation diagrams shown in Fig. 2 or from rescattering of intermediate states [16]. Thus, the $B^0_s \to D^{(*)+}D^{(*)-}$ decay can be used to estimate the subleading contributions of the $B^0 \to D^{(*)+}D^{(*)-}$ decay mode.

The $B^0_s \to D^{(*)+}D^{(*)-}$ decay has not been previously observed, but an excess of possible $B^0_s \to D^{(*)+}D^{(*)-}$ candidates was seen in a recent measurement of $CP$ violation in $B^0 \to D^{(*)+}D^{(*)-}$ decays by the LHCb experiment [7]. Assuming prominent contributions from rescattering of e.g. $D^{(*)+}_sD^{(*)-}_s$ states, the branching fraction is predicted to be $(6.1 \pm 3.6) \times 10^{-5}$ [16]. A perturbative QCD approach predicts a much larger branching fraction of $(3.6 \pm 0.6) \times 10^{-3}$ [17].

This paper presents the first observation of the $B^0_s \to D^{(*)+}D^{(*)-}$ and $B^0_s \to D^{(*)+}_sD^{(*)-}_s$ decays, which have indistinguishable final states. Throughout this paper these decays are treated together and charge conjugation is applied. The branching fraction of the $B^0_s \to D^{(*)+}D^{(*)-}$ decay is measured relative to the $B^0 \to D^{(*)+}D^{(*)-}$ decay. Since both decay channels have the same final state, the experimental systematic uncertainties on the ratio of branching fractions are expected to be small. The measurement uses proton-proton ($pp$) collision data collected with the LHCb detector in the years 2011, 2012, and 2015–2018 at centre-of-mass energies of 7, 8, and 13 TeV, respectively, corresponding to an integrated luminosity of 9 fb$^{-1}$.

Figure 1: (Left) Tree-level and (right) penguin diagrams contributing to $B^0 \to D^{(*)+}_sD^{(*)-}_s$ and $B^0 \to D^{(*)+}D^{(*)-}$ decays.
2 Detector and simulation

The LHCb detector \cite{18,19} is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of the momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a primary $pp$ collision vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T) \mu$m, where $p_T$ is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

Simulated data samples are used to train a multivariate algorithm, model the shapes of mass distributions and calculate efficiencies. In the simulation, $pp$ collisions are generated using PYTHIA \cite{20} with a specific LHCb configuration \cite{21}. Decays of unstable particles are described by EVTGEN \cite{22}, in which final-state radiation is generated using PHOTOS \cite{23}. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit \cite{24} as described in Ref. \cite{25}.

The distributions of particle identification (PID) variables do not match perfectly between simulation and data. This difference is corrected using an approach where functions are constructed that transform the simulated PID response to match calibration samples of recorded data. This is based on a four-dimensional kernel density estimation for distributions in PID value, $p_T$ and $\eta$ of the track and the event multiplicity \cite{26}.

3 Candidate selection

Due to varying data-taking conditions, the data samples for the three periods 2011–2012, 2015–2016 and 2017–2018 are treated differently. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter.

Figure 2: (Left) $W$-exchange and (right) penguin-annihilation diagrams contributing to $B_d(s) \rightarrow D_s^{(*)} D^{(*)} - D^{(*)}$ decays.
and muon systems, followed by a software stage, which applies a full event reconstruction. At the hardware trigger stage, events are required to have a muon with high $p_T$ or a hadron, photon or electron with high transverse energy in the calorimeters. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from any PV. At least one charged particle must have a large transverse momentum and be inconsistent with originating from any PV. A multivariate algorithm \cite{27,28} is used for the identification of secondary vertices consistent with the decay of a $b$ hadron.

The $B^0_{(s)} \to D^{*\pm} D^\mp$ candidates are reconstructed through the decays $D^{*+} \to D^0 \pi^+$ with $D^0 \to K^- \pi^+$ and $D^- \to K^+ \pi^- \pi^-$. The tracks of the final-state particles are required to have a good quality, fulfil loose PID criteria, and have a high $\chi^2_{IP}$ value with respect to any PV, where $\chi^2_{IP}$ is defined as the difference in the vertex-fit $\chi^2$ of a given PV reconstructed with and without the particle being considered. The probability of a candidate being a duplicate track is required to be small. Additionally, the distance of closest approach between all possible combinations of tracks is required to be small. The reconstructed masses of the $D^{*+}$, $D^0$ and $D^-$ candidates are required to lie inside a mass window of $\pm 50$ MeV/$c^2$ around their known values \cite{29}, and the difference of the reconstructed masses between the $D^{*+}$ and $D^0$ candidates is required to be smaller than 150 MeV/$c^2$. The ratio of the $D^-$ decay time and its uncertainty, $t/\sigma_t$, is required to be larger than $-1$. The $B^0_{(s)}$ candidate is reconstructed by combining the $D^{*\pm}$ and $D^\mp$ candidates to form a common vertex. In case multiple PVs are reconstructed in the same event, the PV for which the $B^0_{(s)}$ candidate has the lowest $\chi^2_{IP}$ is assigned as the associated PV. The sum of the transverse momenta of the decay products of the $B^0_{(s)}$ candidate is required to be larger than 5 GeV/$c$ and the $\chi^2_{IP}$ of the $B^0_{(s)}$ candidate for the associated PV is required to be small. Furthermore, the decay time of the $B^0_{(s)}$ candidate is required to be larger than 0.2 ps. Candidates are retained if the mass of the $D^{*\pm}D^\mp$ system, $m_{D^{*\pm}D^\mp}$, is in the range $5000 \text{ MeV}/c^2 < m_{D^{*\pm}D^\mp} < 5600 \text{ MeV}/c^2$.

Background contributions to $D^+$ candidates arise when kaons or protons stemming from hadronic decays of $D_s^+$ and $A_c^+$ hadrons are misidentified as pions. A combination of mass and PID requirements is used to suppress contributions from $B^0 \to D^+ D_s^-$ ($A_c^0 \to D^+ A_c^-$) decays with $D_s^+ \to K^- K^+ \pi^+$ ($A_c^+ \to K^- p \pi^+$) to a negligible level. The mass of the $K^- \pi^+ \pi^+$ system from the $D^+$ candidate is recalculated using a kaon (proton) mass hypothesis for either of the pions. The candidate is rejected if the pion has a high probability to be identified as a kaon (proton) and the recomputed mass is compatible with the known $D_s^+$ ($A_c^+$) mass \cite{29}. Background contributions that arise from $\phi(1020) \to K^- K^+$ transitions in the $D_s^+$ decay chain are further suppressed by rejecting candidates if the pion has a high probability to be identified as a kaon and the mass of the $K^- \pi^+$ system, where the kaon mass is assigned to either of the pions from the $D^+$ decay, is compatible with the known $\phi(1020)$ meson mass \cite{29}. Decays of $B^0_{(s)}$ mesons of the form $B^0_{(s)} \to D^{*\pm} h^- h^+ h^+$ are suppressed by ensuring that the $B^0_{(s)}$ and $D^+$ decay vertices are well separated. Partially reconstructed decays, \textit{i.e.} decays where one or more final-state particles are not reconstructed, contribute to the lower-mass sideband and are accounted for in the fit to the data.

To suppress combinatorial background from random combinations of final-state tracks, a boosted decision tree (BDT) classifier \cite{30,31}, implemented in the TMVA toolkit \cite{32}, utilising the AdaBoost method is used. The BDT classifier is trained using simulated $B^0 \to D^{*\pm} D^\mp$ decays as signal proxy and the upper-mass sideband
mass \( m_{D^{±}D^{∓}} < 6000 \text{ MeV}/c^2 \) as background proxy to avoid contributions from signal and partially reconstructed decays. For each data-taking period a \( k \)-folding technique with \( k = 5 \) is adopted. The following variables are used in the training of the BDT classifier: the mass difference of the \( D^{∗+} \) and \( D^{0} \) candidates; PID variables of the final-state particles of the \( D^{-} \) candidate decay, the kaon coming from the \( D^{0} \) candidate decay and the pion coming from the \( D^{∗+} \) decay; the transverse momenta of the \( B^{0}_{(s)} \) candidate and the kaon from the \( D^{-} \) decay; \( t/σ_t \) of the \( D^{-} \) candidate; the \( \chi^2_{IP} \) of the \( B^{0}_{(s)} \) and \( D^{-} \) candidates; the \( \chi^2 \) of the flight distance of the \( D^{-} \) and \( D^{0} \) candidates and the \( \chi^2 \) of a kinematic fit to the whole decay chain.

The optimal requirement on the BDT response (also referred to as working point) is determined by maximising the figure-of-merit \( ε/(a/2 + \sqrt{N_B}) \) [34]. The efficiency of signal decays, \( ε \), for a specific working point is determined by fits to the data around the known \( B^{0} \) mass [29] before and after the application of the BDT requirement. The number of background candidates in the \( B^{0} \) signal region, \( N_B \), is estimated with the upper-mass sideband, and the targeting significance in numbers of the standard deviation, \( a \), is set to three. A three-dimensional scan of the figure-of-merit in the three data-taking periods is conducted, resulting in slightly different working points.

Afterwards, a single candidate is selected randomly from each event containing multiple candidates. The total selection efficiencies of the \( B^{0}_{(s)} \) \( → D^{∗±}D^{∓} \) and \( B^{0} \rightarrow D^{∗±}D^{∓} \) decays are needed for the calculation of the branching fraction ratio and are calculated using simulated samples.

### 4 Candidate mass fit

To improve the \( B^{0}_{(s)} \) mass resolution, a kinematic fit is applied to the decay chain, where the masses of the \( D^{∗+} \), \( D^{0} \) and \( D^{-} \) candidates are constrained to their known values [29]. An unbinned maximum-likelihood fit to the mass distribution of the \( D^{∗±}D^{∓} \) system is performed separately for each data-taking period to determine the number of signal candidates. To determine the significance of the observation of the \( B^{0}_{s} \rightarrow D^{∗±}D^{∓} \) decay, the three likelihoods are added together. The fit model consists of the signal \( B^{0} \rightarrow D^{∗±}D^{∓} \) and \( B^{0}_{s} \rightarrow D^{∗±}D^{∓} \) decays, a contribution from combinatorial background and components for partially reconstructed \( B^{0} \rightarrow D^{∗±}D^{∓} \) and \( B^{0}_{s} \rightarrow D^{∗±}D^{∓} \) decays, where one of the \( D^{∗} \) mesons decays into a charged \( D \) meson and an unreconstructed \( π^0 \) meson or photon. The \( B^{0} \rightarrow D^{∗±}D^{∓} \) component is modelled by the sum of two Crystal Ball functions [35], with the same mean but different widths and tail parameters. The \( B^{0}_{s} \rightarrow D^{∗±}D^{∓} \) component is described by the same model but with the mean shifted by the difference of the known \( B^{0} \) and \( B^{0}_{s} \) masses [29]. The parameters of the Crystal Ball functions are determined using fits to simulated \( B^{0} \rightarrow D^{∗±}D^{∓} \) decays, apart from their mean and a single scale factor, which corrects the widths for inaccuracies in simulation. The combinatorial background component is described by an exponential function. The functional forms of the \( B^{0} \rightarrow D^{∗±}D^{∗∓} \) and \( B^{0}_{s} \rightarrow D^{∗±}D^{∗∓} \) contributions depend on the polarisation of the \( D^{∗±} \) mesons and are modelled using simulated decays with a combination of functions corresponding to pure longitudinal and transverse polarisations. For a longitudinally polarised \( D^{∗±} \) meson the shape is a double peak, in contrast to a single broad peak for the case of a transversely polarised \( D^{∗±} \) system. The free parameters in the fit are the mass of the \( B^{0} \rightarrow D^{∗±}D^{∓} \) peak, the scaling factor, the slope of the exponential function,
the relative fractions between longitudinally and transversely polarised $D^\pm$ mesons in $B^0 \to D^{\ast\pm}D^{\mp}$ and $B^0_s \to D^{\ast\pm}\bar{D}^{\ast\mp}$ decays, and the yields of all shapes. Pseudoexperiments are used to validate that the model provides unbiased results.

The resulting yields of $B^0 \to D^{\ast\pm}D^{\mp}$ and $B^0_s \to D^{\ast\pm}\bar{D}^{\ast\mp}$ decays are $466 \pm 22$ and $12 \pm 4$ in 2011–2012, $780 \pm 29$ and $34 \pm 7$ in 2015–2016, and $1263 \pm 36$ and $49 \pm 8$ in 2017–2018, respectively, where the quoted uncertainties are statistical only. The mass distributions and fit projections are shown in Fig. 3 for the three data-taking periods.

5 Systematic uncertainties

The measurement of the ratio of branching fractions, $\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})/\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})$, relies on input from the measurement of the ratio of the $b$ quark hadronisation fractions to $B^0_s$ and $B^0$ mesons, $f_s/f_d$. The precision on $f_s/f_d$ results in the dominant source of
systematic uncertainty. The values are taken from Ref. [36] for 2011–2012 and from Ref. [37] for 2015–2016 and 2017–2018 data-taking periods. Both measurements share sources of systematic uncertainty and thus are treated as partially correlated.

Two sources of systematic uncertainty on the efficiency ratio are considered. The first is caused by the finite size of the simulated data samples. The second originates in the use of PID variables, whose distributions do not match perfectly between data and simulation. This uncertainty is determined by choosing a different kernel density estimation in the transformation of the simulated PID response and calculating the difference of the resulting efficiency ratio.

The systematic uncertainties on the signal yields due to the fit models are evaluated using pseudoexperiments. A systematic uncertainty due to the choice of the signal model and the assumption that the $B^0$ and $B^0_s$ distributions have identical shape in the mass fit is evaluated. Candidates are generated with a mass distribution described by a Hypatia function [38] with different parameters for the $B^0$ and $B^0_s$ models. The values of the parameters of the Hypatia function are determined by a fit to simulated $B^0 \to D^{\ast\pm} D^\mp$ and $B^0_s \to D^{\ast\pm} D^\mp$ data, respectively. All yields and background parameters are set to the values found in the default fit to the data. The generated candidates are then fitted with the default model and the result for the branching fraction ratio is calculated for each fit. The mean of all experiments and its residual are calculated for the three periods separately. The residual and its uncertainty are summed in quadrature and the square root is assigned as the systematic uncertainty.

In addition, a systematic uncertainty due to the model of the combinatorial background is evaluated. Pseudoexperiments are used with parameters of the signal and partially reconstructed background models set to the values found in the default result. The slope of the exponential function is extracted by a fit to the data, where a looser BDT requirement is applied to enhance the contribution of the combinatorial background. The generated candidates are fitted with the default model and the systematic uncertainty is calculated in the same way as the systematic uncertainty for the signal model.

To determine the systematic uncertainty on the combined result, the systematic uncertainties related to the PID variables, signal model and background model are assumed to be fully correlated between the data-taking periods when calculating the weighted average. The systematic uncertainties due to the finite size of the simulated data are assumed to be uncorrelated. All systematic uncertainties are added in quadrature to obtain the total systematic uncertainty per data-taking period, and are listed together with their contributions in Table 1.

6 Results
The $B^0_s \to D^{\ast\pm} D^\mp$ decay is observed with a high significance, which is calculated using Wilk’s theorem [39]. The relative branching ratio is calculated using the expression

$$\frac{B(B^0 \to D^{\ast\pm} D^\mp)}{B(B^0 \to D^{\ast\pm} D^\mp)} = \frac{N_{B^0_s} \varepsilon_{B^0} \varepsilon_{d}}{N_{B^0} \varepsilon_{B^0_s} \varepsilon_{s}},$$

where the $B^0$ and $B^0_s$ yields, $N_{B^0}$ and $N_{B^0_s}$, are determined from the fit to the $D^{\ast\pm} D^\mp$ mass distribution. The ratios of the $B^0_s \to D^{\ast\pm} D^\mp$ and $B^0 \to D^{\ast\pm} D^\mp$ selection efficiencies, $\varepsilon_{B^0_s}/\varepsilon_{B^0}$, calculated using simulation samples for the data-taking periods 2011–2012,
Table 1: Systematic uncertainties on $\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp})/\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})$. The systematic uncertainty is given relative to the measured value.

| Source                     | 2011–2012 [%] | 2015–2016 [%] | 2017–2018 [%] | Combined [%] |
|----------------------------|---------------|---------------|---------------|--------------|
| $f_s/f_d$                  | 36            | 37            |               |              |
| Simulated data size        | 0.8           | 1.2           | 0.8           | 0.6          |
| PID                        | 0.7           | 0.7           | 0.8           | 0.7          |
| Signal model               | 0.1           | 0.1           | 0.5           | 0.3          |
| Background model           | 1.7           | 1.3           | 0.8           | 1.1          |
| Total without $f_s/f_d$    | 2.0           | 1.9           | 1.5           | 1.5          |
| Total                      | 6.1           | 5.3           | 5.1           | 4.8          |

2015–2016 and 2017–2018 are $1.063 \pm 0.008$, $1.062 \pm 0.013$ and $1.074 \pm 0.009$, respectively, where the uncertainties are statistical. The ratios of the hadronisation fractions are taken as $0.259 \pm 0.015$ and $0.244 \pm 0.012$ for the 2011–2012 and 2015–2018 data-taking periods, respectively.

The ratio of branching fractions are found to be

$$
\frac{\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp})}{\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})} = 0.939 \pm 0.032 \pm 0.002 \pm 0.005,
$$

$$
\frac{\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp})}{\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})}_{\text{2011–2012}} = 0.168 \pm 0.034 \pm 0.003 \pm 0.008,
$$

$$
\frac{\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp})}{\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})}_{\text{2015–2016}} = 0.149 \pm 0.024 \pm 0.002 \pm 0.007,
$$

where the first uncertainty is statistical, the second systematic and the third is due to the uncertainty of the fragmentation fraction ratio $f_s/f_d$. Using the quadratic sums of the uncertainties as weights and including the correlation of the systematic uncertainties, the average of these measurements is

$$
\frac{\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp})}{\mathcal{B}(B^0 \to D^{\ast\pm}D^{\mp})} = 0.137 \pm 0.017 \pm 0.002 \pm 0.006.
$$

Using the measured value of the $B^0 \to D^{\ast\pm}D^{\mp}$ branching fraction from Ref. [5], the $B_s^0 \to D^{\ast\pm}D^{\mp}$ branching fraction is determined to be

$$
\mathcal{B}(B_s^0 \to D^{\ast\pm}D^{\mp}) = (8.41 \pm 1.02 \pm 0.12 \pm 0.39 \pm 0.79) \times 10^{-5},
$$

where the fourth uncertainty is due to the $B^0 \to D^{\ast\pm}D^{\mp}$ branching fraction.

This result assumes an average $B_s^0$ lifetime for the $B_s^0 \to D^{\ast\pm}D^{\mp}$ decay. The heavy and light eigenstates of the $B_s^0$ meson have significantly different lifetimes. As the selection efficiency depends on the lifetime, correction factors for the efficiency are calculated following the procedure outlined in Ref. [40] that considers either a purely heavy or a purely light $B_s^0$ eigenstate. The correction factors are found to be compatible for all data-taking periods. The integrated correction factors are 1.042 (0.949) for a purely heavy (light) $B_s^0$ eigenstate. The equivalent effect in the selection efficiency of the $B^0$ decay is negligible due to the small value of $\Delta \Gamma_d$ [29].
7 Conclusion

This paper presents the first observation of the $B^0_s \to D^{\ast\pm} D^\mp$ decay along with the measurement of its branching fraction relative to the $B^0 \to D^{\ast\pm} D^\mp$ decay. The analysis is performed with data collected by the LHCb experiment in the years 2011, 2012, and 2015 to 2018, corresponding to an integrated luminosity of 9 fb$^{-1}$. The combined ratio of branching fractions for all data-taking periods is determined to be

$$\frac{\mathcal{B}(B^0_s \to D^{\ast\pm} D^\mp)}{\mathcal{B}(B^0 \to D^{\ast\pm} D^\mp)} = 0.137 \pm 0.017 \pm 0.002 \pm 0.006,$$

where the first uncertainty is statistical, the second systematic and the third is due to the uncertainty of the fragmentation fraction ratio $f_s/f_d$. The $B^0_s \to D^{\ast\pm} D^\mp$ branching fraction is determined to be

$$\mathcal{B}(B^0_s \to D^{\ast\pm} D^\mp) = (8.41 \pm 1.02 \pm 0.12 \pm 0.39 \pm 0.79) \times 10^{-5},$$

where the fourth uncertainty is due to the $B^0 \to D^{\ast\pm} D^\mp$ branching fraction [29]. The result is in agreement with predictions from other $B$-meson decays [16] and disagrees with predictions from a perturbative QCD approach [17]. It can be used to constrain subleading contributions in the measurement of the $CP$-violating parameter $\sin(2\beta)$ with $B^0 \to D^{\ast\pm} D^\mp$ decays.

Acknowledgements

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); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MSHE (Russia); MICINN (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); DOE NP and NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); A*MIDEX, ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, CAS CCEPP, Fundamental Research Funds for Central Universities, and Sci. & Tech. Program of Guangzhou (China); RFBR, RSF and Yandex LLC (Russia); GVA, XuntaGal and GENCAT (Spain); the Royal Society and the Leverhulme Trust (United Kingdom).

References

[1] N. Cabibbo, Unitary symmetry and leptonic decays, Phys. Rev. Lett. 10 (1963) 531
[2] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. 49 (1973) 652.

[3] BaBar collaboration, B. Aubert *et al.*, Measurements of time-dependent CP asymmetries in $B^0 \rightarrow D^{(*)+}D^{(*)-}$ decays, Phys. Rev. D79 (2009) 032002, arXiv:0808.1866.

[4] Belle collaboration, S. Fratina *et al.*, Evidence for CP violation in $B^0 \rightarrow D^+D^-$ decays, Phys. Rev. D79 (2009) 032002, arXiv:hep-ex/0702031.

[5] Belle collaboration, M. Röhrken *et al.*, Measurements of branching fractions and time-dependent CP violating asymmetries in $B^0 \rightarrow D^{(*)+}D^{(*)-}$ decays, Phys. Rev. D85 (2012) 091106, arXiv:1203.6647.

[6] LHCb collaboration, R. Aaij *et al.*, Measurement of $CP$ violation in $B \rightarrow D^{(*)+}D^{(*)-}$ decays, Phys. Rev. Lett. 117 (2016) 261801, arXiv:1608.06620.

[7] LHCb collaboration, R. Aaij *et al.*, Measurement of $CP$ violation in $B \rightarrow D^{(*)+}D^{(*)-}$ decays, JHEP 03 (2020) 147, arXiv:1912.03723.

[8] LHCb collaboration, R. Aaij *et al.*, Measurement of the $CP$-violating phase $\phi_s$ in $B^0_s \rightarrow D^{(*)+}D^{(*)-}$ decays, Phys. Rev. Lett. 113 (2014) 211801, arXiv:1409.4619.

[9] R. Aleksan *et al.*, The decay $B \rightarrow DD^*$ in the heavy quark limit and tests of $CP$ violation, Phys. Lett. B317 (1993) 173.

[10] A. I. Sanda and Z.-z. Xing, Towards determining $\phi_1$ with $B \rightarrow D^{(*)}D^{(*)}$, Phys. Rev. D56 (1997) 341, arXiv:hep-ph/9702297.

[11] Z.-z. Xing, Measuring CP violation and testing factorization in $B_d \rightarrow D^+D^-$ and $B_s \rightarrow D^{(*)+}D^{(*)-}$ decays, Phys. Lett. B443 (1998) 365, arXiv:hep-ph/9809496.

[12] Z.-z. Xing, CP violation in $B_d \rightarrow D^+D^-$, $D^{(*)+}D^{(*)-}$, $D^+D^{*-}$ and $D^{(*)+}D^{(*)-}$ decays, Phys. Rev. D61 (2000) 014010, arXiv:hep-ph/9907455.

[13] X.-Y. Pham and Z.-z. Xing, $CP$ asymmetries in $B_d \rightarrow D^{(*)+}D^{(*)-}$ and $B_s \rightarrow D^{(*)+}D^{(*)-}$ decays: $P$ wave dilution, penguin and rescattering effects, Phys. Lett. B458 (1999) 375, arXiv:hep-ph/9904360.

[14] M. Jung and S. Schacht, Standard model predictions and new physics sensitivity in $B \rightarrow DD$ decays, Phys. Rev. D91 (2015) 034027, arXiv:1410.8396.

[15] L. Bel *et al.*, Anatomy of $B \rightarrow D\bar{D}$ decays, JHEP 07 (2015) 108, arXiv:1505.01361.

[16] M. Gronau, D. London, and J. L. Rosner, Rescattering contributions to rare $b$-meson decays, Phys. Rev. D87 (2013) 036008, arXiv:1211.5785.

[17] Y. Li, C.-D. Lu, and Z.-J. Xiao, Rare decays $B^0 \rightarrow D^{(*)+}D_s^{(*)-}$ and $B^0 \rightarrow D^{(*)+}D^{(*)-}$ in perturbative QCD approach, J. Phys. G31 (2005) 273, arXiv:hep-ph/0308243.

[18] LHCb collaboration, A. A. Alves Jr. *et al.*, The LHCb detector at the LHC, JINST 3 (2008) S08005.
[19] LHCb collaboration, R. Aaij et al., LHCb detector performance, Int. J. Mod. Phys. A30 (2015) 1530022, arXiv:1412.6352

[20] T. Sjöstrand, S. Mrenna, and P. Skands, A brief introduction to PYTHIA 8.1, Comput. Phys. Commun. 178 (2008) 852, arXiv:0710.3820; T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026, arXiv:hep-ph/0603175.

[21] I. Belyaev et al., Handling of the generation of primary events in Gauss, the LHCb simulation framework, J. Phys. Conf. Ser. 331 (2011) 032047.

[22] D. J. Lange, The EvtGen particle decay simulation package, Nucl. Instrum. Meth. A462 (2001) 152.

[23] P. Golonka and Z. Was, PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays, Eur. Phys. J. C45 (2006) 97, arXiv:hep-ph/0506026.

[24] Geant4 collaboration, J. Allison et al., Geant4 developments and applications, IEEE Trans. Nucl. Sci. 53 (2006) 270; Geant4 collaboration, S. Agostinelli et al., Geant4: A simulation toolkit, Nucl. Instrum. Meth. A506 (2003) 250.

[25] M. Clemencic et al., The LHCb simulation application, Gauss: Design, evolution and experience, J. Phys. Conf. Ser. 331 (2011) 032023.

[26] A. Poluektov, Kernel density estimation of a multidimensional efficiency profile, JINST 10 (2015) P02011, arXiv:1411.5528.

[27] V. V. Gligorov and M. Williams, Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree, JINST 8 (2013) P02013, arXiv:1210.6861.

[28] T. Likhomanenko et al., LHCb topological trigger reoptimization, J. Phys. Conf. Ser. 664 (2015) 082025.

[29] Particle Data Group, P. A. Zyla et al., Review of particle physics, to be published in Prog. Theor. Exp. Phys. 6 (2020) 083C01.

[30] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and regression trees, Wadsworth international group, Belmont, California, USA, 1984.

[31] Y. Freund and R. E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1997) 119.

[32] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, TMVA - Toolkit for Multivariate Data Analysis with ROOT, PoS ACAT (2007) 040; A. Hoecker et al., TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide., arXiv:physics/0703039.

[33] A. Blum, A. Kalai, and J. Langford, Beating the hold-out: bounds for k-fold and progressive cross-validation, in Proceedings of the Twelfth Annual Conference on Computational Learning Theory, COLT ’99, (New York, NY, USA), 203–208, ACM 1999.
[34] G. Punzi, *Sensitivity of searches for new signals and its optimization*, eConf C030908 (2003) MODT002, arXiv:physics/0308063

[35] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[36] LHCb collaboration, R. Aaij et al., *Measurement of the fragmentation fraction ratio $f_s/f_d$ and its dependence on $B$ meson kinematics*, JHEP 04 (2013) 001, arXiv:1301.5286, $f_s/f_d$ value updated in LHCb-CONF-2013-011.

[37] LHCb collaboration, R. Aaij et al., *Measurement of $b$-hadron fractions in 13 TeV pp collisions*, Phys. Rev. D100 (2019) 031102(R), arXiv:1902.06794.

[38] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. A764 (2014) 150, arXiv:1312.5000.

[39] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. 9 (1938) 60.

[40] K. De Bruyn et al., *Branching ratio measurements of $B_s$ decays*, Phys. Rev. D86 (2012) 014027, arXiv:1204.1735.
R. Vazquez Gomez, P. Vazquez Regueiro, C. Vázquez Sierra, S. Vecchi, J.J. Velthuis, M. Veltri, A. Venkateswaran, M. Veronesi, M. Vesterinen, D. Vieira, M. Vieites Diaz, H. Vienemann, X. Vilasis-Cardons, E. Villela Figueras, P. Vincent, G. Vitali, A. Vollhardt, D. Vom Bruch, A. Vorobyev, V. Vorobyev, N. Voropaev, R. Waldi, J. Walsh, J. Wang, J. Wang, J. Wang, M. Wang, R. Wang, Y. Wang, Z. Wang, H.M. Wark, N.K. Watson, S.G. Weber, D. Websdale, C. Weiss, B.D.C. Westhenry, D.J. White, M. Whitehead, D. Wiedner, G. Wilkinson, M. Wilkinson, I. Williams, M. Williams, M.R.J. Williams, F.F. Wilson, W. Wislicki, M. Witkowska, L. Witola, G. Wornser, S.A. Wotton, H. Wu, K. Wyllic, Z. Xiang, D. Xiao, Y. Xie, A. Xu, J. Xu, L. Xu, M. Xu, Q. Xu, Z. Xu, Z. Xu, D. Yang, Y. Yang, Z. Yang, Z. Yang, Y. Yao, L.E. Yeomans, H. Yin, J. Yu, X. Yuan, O. Yushchenko, E. Zaffaroni, K.A. Zarebski, M. Zavertyaev, M. Zdyba, O. Zenaiev, S. Zucchelli.

1 Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
2 Universidad Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
3 Center for High Energy Physics, Tsinghua University, Beijing, China
4 School of Physics State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
5 University of Chinese Academy of Sciences, Beijing, China
6 Institute of High Energy Physics (IHEP), Beijing, China
7 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China
8 Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IN2P3-LAPP, Annecy, France
9 Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
10 Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
11 Université Paris-Saclay, CNRS/IN2P3, IJCLab, Orsay, France
12 LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
13 I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany
14 Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
15 Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
16 Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
17 School of Physics, University College Dublin, Dublin, Ireland
18 INFN Sezione di Bari, Bari, Italy
19 INFN Sezione di Bologna, Bologna, Italy
20 INFN Sezione di Ferrara, Ferrara, Italy
21 INFN Sezione di Firenze, Firenze, Italy
22 INFN Laboratori Nazionali di Frascati, Frascati, Italy
23 INFN Sezione di Genova, Genova, Italy
24 INFN Sezione di Milano-Bicocca, Milano, Italy
25 INFN Sezione di Milano, Milano, Italy
26 INFN Sezione di Cagliari, Monserrato, Italy
27 Universita degli Studi di Padova, Universita e INFN, Padova, Padova, Italy
28 INFN Sezione di Pisa, Pisa, Italy
29 INFN Sezione di Roma Tor Vergata, Roma, Italy
30 INFN Sezione di Roma La Sapienza, Roma, Italy
31 Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
32 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, Netherlands
33 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
34 AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
35 National Center for Nuclear Research (NCBJ), Warsaw, Poland
36 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
Università di Bologna, Bologna, Italy
Università di Cagliari, Cagliari, Italy
Università di Ferrara, Ferrara, Italy
Università di Firenze, Firenze, Italy
Università di Genova, Genova, Italy
Università di Milano Bicocca, Milano, Italy
Università di Roma Tor Vergata, Roma, Italy
AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
Università di Padova, Padova, Italy
Università di Pisa, Pisa, Italy
Università degli Studi di Milano, Milano, Italy
Università di Urbino, Urbino, Italy
Università della Basilicata, Potenza, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Modena e Reggio Emilia, Modena, Italy
Università di Siena, Siena, Italy
MSU - Iligan Institute of Technology (MSU-IIT), Iligan, Philippines
Novosibirsk State University, Novosibirsk, Russia