Measurement of \( CP \) Violation in \( B^0 \to D^+ D^- \) Decays

R. Aaij et al. *(LHCb Collaboration)*

(Received 24 August 2016; revised manuscript received 14 October 2016; published 23 December 2016)

The \( CP \) violation observables \( S \) and \( C \) in the decay channel \( B^0 \to D^+ D^- \) are determined from a sample of proton-proton collisions at center-of-mass energies of 7 and 8 TeV, collected by the LHCb experiment and corresponding to an integrated luminosity of 3 fb\(^{-1}\). The observable \( S \) describes \( CP \) violation in the interference between mixing and the decay amplitude, and \( C \) parametrizes direct \( CP \) violation in the decay. The following values are obtained from a flavor-tagged, decay-time-dependent analysis: 

\[
S = -0.54_{-0.17}^{+0.17} \text{(stat)} \pm 0.05 \text{(syst)}, \quad C = 0.26_{-0.17}^{+0.18} \text{(stat)} \pm 0.02 \text{(syst)}.
\]

These values provide evidence for \( CP \) violation at a significance level of 4.0 standard deviations. The phase shift due to higher-order standard model corrections is constrained to a small value of \( \Delta \phi = -0.16_{-0.21}^{+0.19} \text{ rad} \).

DOI: 10.1103/PhysRevLett.117.261801

Studies of beauty hadron decays into pairs of charm hadrons give access to a multitude of observables that probe the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [1,2] of the standard model (SM). Comparisons of these observables with each other and with similar observables from beauty hadron decays to charmonia allow higher-order SM contributions, like loop diagrams, to be separated from effects caused by physics beyond the SM [3–7]. For example, under the assumption that flavor symmetries hold to a good approximation, higher-order corrections in the measurement of \( \phi_d \) in \( B^0 \to D^0 \bar{D}^0 \) [8] can be constrained. (The inclusion of charge-conjugate processes is implied throughout the Letter, unless otherwise noted.)

In the \( B^0 \) meson system, \( CP \) violation in the mixing is negligible, as is the decay width difference \( \Delta \Gamma \) of the mass eigenstates [9]. In contrast, sizable \( CP \) violation from the interference between the direct (unmixed) decay into the \( CP \)-even final state \( D^+ D^- \) and the decay to the same final state after \( B^0 - \bar{B}^0 \) mixing, or from the interference of different decay processes, leads to a decay-time-dependent decay rate of

\[
d\Gamma(t, d) = \alpha e^{-i \beta} \left[ 1 - dS \sin(\Delta \tau) + dC \cos(\Delta \tau) \right],
\]

where \( t \) is the proper decay time, \( d \) represents the \( B^0 \) flavor at production and takes a value of +1 for mesons whose initial flavor is \( B^0 \) and −1 for \( \bar{B}^0 \), \( \tau \) is the mean lifetime, and \( \Delta \tau \) is the mass difference between the physical \( B^0 \) meson eigenstates. The \( CP \) observables \( S \) and \( C \) are related to the \( B^0 \) mixing phase \( \phi_d \) and a phase shift \( \Delta \phi \) from the decay amplitudes via \( S/\sqrt{1 - C^2} = -\sin(\phi_d + \Delta \phi) \) [10]. In the SM, \( \phi_d = 2\beta \), where \( \beta \equiv \arg[\langle V_{cd}V_{cb}^\ast \rangle/(V_{ud}V_{ub}^\ast)] \) is an angle of one of the CKM unitary triangles and \( V_{qq'} \) are elements of the CKM matrix. If the \( B^0 \to D^+ D^- \) decay amplitude can be described by a dominant tree-level \( b \to c \bar{c} d \) transition, the phase shift \( \Delta \phi \) vanishes and the \( CP \) observables are given by \( C = 0 \) and \( S = -\sin \beta \). The value of the latter has been measured to be \( \sin \beta = +0.679 \pm 0.020 \) [9] in \( b \to c \bar{c} s \) decays such as \( B^0 \to J/\psi K_S^0 \), in which the contribution from loop processes in the decay can be constrained to high precision [11]. In contrast, previous measurements of the \( CP \) observables in the decay \( B^0 \to D^+ D^- \) by the BABAR and Belle collaborations [12,13] give world average values of \( S = -0.98 \pm 0.17 \) and \( C = -0.31 \pm 0.14 \) [9]. The values are at the edge of the physically allowed region of \( S^2 + C^2 \leq 1 \), which leaves room for a large value of \( \Delta \phi \).

This Letter reports a measurement of \( CP \) violation in \( B^0 \to D^+ D^- \) decays with the LHCb experiment. The measurement is based on samples of \( pp \) collision data corresponding to integrated luminosities of 1 and 2 fb\(^{-1}\) at center-of-mass energies of 7 and 8 TeV, respectively, recorded by the LHCb experiment. The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range \( 2 < \eta < 5 \), designed for the study of particles containing \( b \) or \( c \) quarks, and is described in detail in Refs. [14,15]. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. Simulated events are produced with the software described in Refs. [16–21].

Candidate \( B^0 \to D^+ D^- \) decays are reconstructed through the subsequent decays \( D^+ \to K^- \pi^+ \pi^+ \) and

---

\*Full author list given at the end of the article.

*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.*
$D^+ \rightarrow K^- K^+ \pi^+$, with combinations of two $D \rightarrow K K \pi$ candidates omitted due to the low branching fraction. The kaon and pion candidates, which must fulfill loose particle identification (PID) criteria, are required to have transverse momentum $p_T > 100$ MeV/c, to have a good track quality, and to be inconsistent with originating from a primary vertex (PV). The three hadron tracks must form a good common vertex and their combined invariant mass has to be in the range $\pm 25$ MeV/c$^2$ around the known $D^+$ mass [22]. The scalar sum of the $p_T$ of the three hadrons has to exceed 1800 MeV/c and the $D^+$ vertex has to be significantly displaced from all PVs. Defining $\cos \theta_D$ as the angle between the momentum vector of a particle $X$ and the displacement vector from the best-matched PV to the $X$ decay vertex, $\cos \theta_D$ is required to be positive.

To suppress contributions from misreconstructed $D^+_s \rightarrow K^- K^+ \pi^+$ decays, which proceed predominantly through $D^+_s \rightarrow \phi \pi^+$, $D^+ \rightarrow K^- \pi^+ \pi^+$ candidates are rejected if, after assigning the kaon mass hypothesis to the $\pi^+$ with the higher $p_T$, the invariant mass $m(K^- K^+)$ is within 10 MeV/c$^2$ of the known $\phi$ meson mass. Furthermore, if the invariant mass $m(K^- K^+)$ is within 25 MeV/c$^2$ of the known $D^+_s$ meson mass, the requirement on the PID information of the higher-$p_T$ pion to be consistent with the pion hypothesis is tightened. Similarly, protons can be misidentified as pions, resulting in background contributions from $\Lambda_c^+ \rightarrow K^- p \pi^+$. To suppress these processes, the candidate with the higher $p_T$ of $D^+ \rightarrow K^- \pi^+ \pi^+$ is required to be well identified as a pion if $|m(K^- p \pi^+) - m_{\Lambda_c^+}| < 25$ MeV/c$^2$.

Candidate $B^0$ mesons are reconstructed from pairs of oppositely charged $D^\pm$ candidates that form a common vertex. The scalar sum of the $p_T$ of the $D^\pm$ mesons must exceed 5 GeV/c. The decay time significance of each $D^\pm$ meson, defined as its decay time divided by its estimated uncertainty, is required to be greater than zero, or greater than 3 if one of the $D^\pm$ mesons is reconstructed in the $K^- K^+ \pi^+$ final state. This reduces the contamination of $B^0 \rightarrow D^+ K^- K^+ \pi^+$ decays. The $B^0$ candidate is required to have momentum $p > 10$ GeV/c, $\cos \theta_{B^0} > 0.999$, and to not originate from the associated PV. A fit to the full decay chain, in which the $B^0$ production vertex is constrained to the position of the associated PV, is performed to determine the reconstructed decay time $t'$ of the $B^0$ candidate, which differs from the true time $t$. Only candidates with decay times in the range 0.25–10.25 ps are kept. The invariant mass $m_{B^0}$ of the $B^0$ candidate is calculated from a similar fit to the full decay chain, while additionally constraining the invariant masses of $K^- \pi^+ \pi^+$ and $K^- K^+ \pi^+$ to the known $D^+$ mass, and is required to be in the range 5150–5500 MeV/c$^2$.

Two boosted decision trees [23,24], for $B^0$ final states with two and three kaons, are used to suppress the combinatorial background. Both are trained on simulated signal samples and on background samples formed from $B^0$ candidates at high invariant masses (> 5500 MeV/c$^2$), and exploit observables related to the kinematics of the decay, PID information, and track and vertex quality. The requirements on the boosted decision tree classifier outputs are chosen to optimize the precision of both $CP$ observables $S$ and $C$.

To separate the remaining background from the signal a fit to the $D^+ D^-$ invariant mass distribution is performed to calculate signal candidate weights via the $sPlot$ technique [25]. The mass fit is performed simultaneously in four categories, split by the data-taking period (7, 8 TeV) and the number of kaons in the final state. The probability density function (PDF) used to parametrize the mass distribution consists of four contributions: signal, $B^0 \rightarrow D^+ D^-$, combinatorial background, and a component that includes both $B^0 \rightarrow D_s^+ D^-$ and $B^0 \rightarrow D_s^- D^+$ decays. The signal is modeled by the sum of three Crystal Ball functions [26] with a common mean. The parameters of the tails (two towards lower and one towards higher mass) and the three widths are determined from simulated samples. To account for differences in the mass resolution in simulation and data, the width parameters are multiplied by a common scale factor, which is free to vary in the fit to data. The $B^0 \rightarrow D^+ D^-$ component shares all shape parameters with the signal PDF except for the peak position, which is constrained by the known value of the difference between the $B^0$ and the $B^0_s$ masses [22]. Each peak in the $B^0 \rightarrow D_s^+ D^-$ and $B^0 \rightarrow D_s^- D^+$ component is described by the sum of two Crystal Ball functions (one with a tail towards lower and one with a tail towards higher masses) whose parameters are taken from simulation. The widths and the $B^0$ peak position are free to vary in the fit while the $B^0_s$ peak offset is constrained in the same way as that of the $B^0_s \rightarrow D^+ D^-$ component. The combinatorial background is parametrized with an exponential function, with separate exponents used for the final states with two or three kaons. Partially reconstructed $B^0 \rightarrow D^{*+} D^-$ decays with $D^{*+} \rightarrow D^+ \pi^0$, where the neutral pion is missed, lie outside the mass range used for the fit. The equivalent $B^0 \rightarrow D^{*+} D^- \pi^0$ decays and decay modes with only one or no charm meson, such as $B^0 \rightarrow D^- K^- K^+ \pi^+$, are also neglected in the mass fit. The influence of their omission on the $CP$ measurement is treated as a systematic uncertainty. The mass distribution is shown in Fig. 1(a). The combined $B^0 \rightarrow D^+ D^-$ signal yield is $1610 \pm 50$, of which $1347 \pm 45$ are in the Cabibbo-favored final state with two $D^+ \rightarrow K^- \pi^+ \pi^+$ decays.

The measurement of decay-time-dependent $CP$ violation requires knowledge of the initial flavor of each reconstructed $B^0$ meson. Flavor-tagging algorithms deliver a measured tag decision $d'$ for the flavor of the $B^0$ meson, which takes the value $+1$ for a $B^0$, $-1$ for a $B^0$ initial state, and zero if no decision is possible, and an estimate $\eta$ of the probability for the tag decision to be incorrect. The latter is
referred to as the mistag probability. Two classes of flavor-tagging algorithms are used: opposite-side (OS) and same-side (SS) taggers [27–29]. In $b\bar{b}$ pair production, the dominant source of $b$ hadrons at LHCb, the signal $B^0$ meson is accompanied by a second $b$ hadron. The OS taggers determine the flavor of the signal by examining the decay products of this second $b$ hadron. The information from the decay products consists of the charge of muons or electrons produced in semileptonic decays, the charge of kaons from $b\to c\to s$ transitions, the charge of charm hadrons from $b\to c$ transitions, and the net charge of all decay products. The SS taggers analyze pions and protons related to the hadronization process of the $B^0$ meson. This is the first analysis to use the LHCb SS proton and OS charm taggers, and the first to use the new SS pion tagger.

The outputs of all OS algorithms are combined into an overall OS tagging decision and mistag estimate, and the same is done for the SS algorithms. The mistag estimates $\eta \in \{\eta_{OS}, \eta_{SS}\}$ are calibrated using linear functions $\omega(\eta)d_0$.

so that $\eta$ on average matches the true mistag probability $\omega$, which depends on the true production flavor $d$ of the $B^0$ meson. The calibration studies are performed with a sample of $B^0 \to D^{(*)}_s D^-$ decays, for which the final state determines the flavor of the $B^0$ at decay. Since the calibration and signal channels are kinematically very similar, the calibration can be applied to the signal channel without further corrections. To ensure that the same calibration is valid for both, the same selection is used as for the signal decay with one $D^{*-} \to K^-\pi^+$, apart from requiring that the $K^-\pi^+$ invariant mass lie within 25 MeV/c$^2$ of the known $D_s^+$ mass [22] and dropping the vetoes against misidentified backgrounds. Background is subtracted from the calibration sample via the splot technique [25]. The tagging calibration parameters are determined from a fit to the decay time and tag distributions of $B^0 \to D^{(*)}_s D^-$ candidates, in which the detection asymmetry, the production asymmetry of the $B^0$ mesons, and the flavor-specific semileptonic asymmetry $a^d_{l\nu}$ are taken into account. Here, the detection asymmetry describes the difference in reconstruction efficiency between the $D^{(*)}_s D^-$ and $D^{(*)}_s D^+$ final states, and $A_P \equiv [\sigma(B^0) - \sigma(B^0)]/[\sigma(B^0) + \sigma(B^0)]$, where $\sigma$ denotes the production cross section inside the LHCb acceptance. The values of all these parameters are fixed according to the latest LHCb measurements [30,31], and their uncertainties are treated as sources of systematic uncertainty on the calibration parameters. Further systematic uncertainties are assigned due to the calibration method, the dependence of the efficiency on decay time, the decay time resolution, and the background subtraction. More details on the calibration studies are given in Ref. [32].

In the $B^0 \to D^{(*)} D^-$ signal data sample, the correlation between the OS and the SS mistag estimates is found to be negligible. A small correlation of the mistag probability with decay time is seen; this is neglected in the main fit but considered as a source of systematic uncertainty.

The effective tagging efficiency is the product of the probability for reaching a tagging decision, $\epsilon_{\text{tag}} = (87.6 \pm 0.8)\%$, and the square of the effective dilution $D = 1 - 2\omega = (30.3 \pm 1.1)\%$. Its value is $\epsilon_{\text{tag}}D^2 = (8.1 \pm 0.6)\%$, the highest effective tagging efficiency to date in tagged $CP$ violation measurements at LHCb thanks to the improved flavor-tagging algorithms and the kinematic properties of the selected $B^0 \to D^{(*)} D^-$ decays.

The $CP$ violation observables $S$ and $C$ are determined from a multidimensional fit to the background-subtracted tag and decay time distributions of the tagged $B^0 \to D^{(*)} D^-$ candidates; a projection of the decay time distribution summed over the nonzero tag decisions is shown in Fig. 1(b). The conditional PDF describing the reconstructed decay time $t'$ and tag decisions $d' = (d'_{\text{OS}}, d'_{\text{SS}})$, given a per-event decay time resolution $\sigma_t$ and per-event mistag probability estimates $\bar{\eta} = (\bar{\eta}_{\text{OS}}, \bar{\eta}_{\text{SS}})$, is
\[ P(t', \bar{d}[\sigma_{\tau}, \bar{n}]) \propto e(t')(\mathcal{P}(t, \bar{d}[\bar{n}]) \otimes \mathcal{R}(t' - t|\sigma_{\tau})), \]  

(2)

where

\[ \mathcal{P}(t, \bar{d}[\bar{n}]) \propto \sum_{d} \mathcal{P}(d', \bar{d}[\bar{n}])[1 - dA_{P}]e^{-t'/t} 
= \{1 - dS \sin(\Delta m t) + dC \cos(\Delta m t)\}, \]  

(3)

and where \( t \) is the true decay time, \( d \) is the true production flavor, \( A_{P} \) is the production asymmetry, and \( \mathcal{P}(d', \bar{d}[\bar{n}] \) is a two-dimensional binomial PDF describing the distribution of tagging decisions given \( \bar{n} \) and \( d \). Normalization factors are omitted for brevity. In the fit, the mass difference \( \Delta m \) and the lifetime \( \tau \) are constrained to their known values within uncertainties [22]. The production asymmetry \( A_{P} \) is constrained separately for the 7 and 8 TeV samples to the values obtained from weighting the results from the constrained separately for the 7 and 8 TeV samples to the values obtained from weighting the results from the measurements in Ref. [31] according to the kinematic distribution of the \( B^{0} \) signal candidates. The decay time resolution model \( \mathcal{R} \) is the sum of three Gaussian functions, two of which have event-dependent widths proportional to \( \sigma_{\tau} \), and one of which has a global width that describes the effect of candidates matched to a wrong PV; all three share a common mean. All parameters of the resolution model are determined from simulation. The average decay time resolution in data is 49 fs. The function \( e(t') \) describes the efficiency for all reconstruction and selection steps as a function of the reconstructed decay time. It is represented by cubic splines [33], with the spline coefficients left unconstrained in the fit.

The statistical uncertainties are estimated using the bootstrap method [34]. Individual bootstrap samples are drawn from the candidates in data that pass the full selection; the analysis procedure described above, consisting of the mass fit, background subtraction, and decay time fit, is then applied to obtain the values of the \( CP \) observables for each such sample. Two-sided 68\% confidence intervals, with equal tail probabilities on either side, are obtained from the distributions of fitted parameters in the bootstrapped samples. To account for the uncertainties of the flavor-tagging calibration parameters, which are fixed in the likelihood fit, further pseudoexperiments are generated in which these flavor-tagging calibration parameters are varied within their combined statistical and systematic uncertainties. The results are then used to correct the uncertainties from the bootstrapping procedure.

The \( CP \) observables are measured to be \( S = -0.54^{+0.17}_{-0.16} \) and \( C = 0.26^{+0.18}_{-0.17} \) with a correlation coefficient of \( \rho = 0.48 \). The decay-time-dependent signal yield asymmetry \( (N_{B^{0}} - N_{\bar{B}^{0}})/(N_{B^{0}} + N_{\bar{B}^{0}}) \), where \( N_{B^{0}} \) is the number of \( B^{0} \rightarrow D^{+}D^{-} \) decays with a \( B^{0} \) flavor tag, and \( N_{\bar{B}^{0}} \) the number with a \( B^{0} \) tag, is shown in Fig. 2.

Several sources of systematic uncertainties on the \( CP \) observables are studied with pseudoexperiments. The largest systematic uncertainty arises from neglecting backgrounds in which the final state contains only one charm meson, such as \( B^{0} \rightarrow D^{-}K^{-}K^{+}\pi^{+} \). The yield of these backgrounds is estimated to be about 2\% of the signal yield and their impact is assessed by assuming that they maximally violate \( CP \) symmetry and have the eigenvalue opposite to the signal mode. This leads to a systematic uncertainty of \( \pm 0.05 \) on \( S \) and \( \pm 0.013 \) on \( C \). Further systematic uncertainties on \( S \) are related to the assumption \( \Delta \Gamma = 0 \) (\( \pm 0.014 \)), and to the modeling of the dependence of the efficiency on decay time (\( \pm 0.007 \)). For \( C \) the second largest systematic uncertainty of \( \pm 0.007 \) is due to neglecting the correlation between the invariant mass and the decay time. Additional systematic uncertainties arise from the decay time resolution, from the uncertainty on the knowledge of the length scale, from the parametrization of the mass model, and from uncertainties on the \( B^{0} \) production asymmetry and mass difference \( \Delta m \). The total systematic uncertainty, calculated as the sum in quadrature of all contributions, is \( \pm 0.05 \) for \( S \) and \( \pm 0.02 \) for \( C \), with a correlation coefficient of \( \rho = -0.69 \).

In conclusion, a measurement of the \( CP \) observables \( S \) and \( C \) in the decay channel \( B^{0} \rightarrow D^{+}D^{-} \) is performed. Using the full data sample collected by the LHCb experiment during Run 1, which corresponds to a total integrated luminosity of 3 fb\(^{-1} \), they are determined to be

\[ S = -0.54^{+0.17}_{-0.16} \text{(stat)} \pm 0.05 \text{(syst)}, \]
\[ C = 0.26^{+0.18}_{-0.17} \text{(stat)} \pm 0.02 \text{(syst)} \]

with a statistical correlation coefficient of \( \rho = 0.48 \). This result excludes the conservation of \( CP \) symmetry by 4.0 standard deviations. It is compatible with the previous measurement by the BABAR experiment of \( S = -0.63 \pm 0.36 \pm 0.05 \) and \( C = -0.07 \pm 0.23 \pm 0.03 \) [12] while being significantly more precise. A proper evaluation of the compatibility with the result from the Belle experiment [13] could not be performed due to its non-Gaussian
uncertainties. The result presented here corresponds to
\sin(\phi_D + \Delta \phi) = 0.56^{\pm 0.18}_{-0.17}, which constrains the phase shift
to the world’s most precise value of \Delta \phi = -0.16^{\pm 0.19}_{-0.21} \text{rad,}
and thus implies only a small contribution from higher-order standard model corrections.

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 (France); BMBF, DFG, and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.). 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), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

[3] R. Fleischer, Extracting \gamma from \B_{u(B)} \to J/\psi K_S and \B_{d(s)} \to D^{\pm}_{s(B)} D_{s(B)}^{-}, Eur. Phys. J. C 10, 299 (1999).
[5] R. Fleischer, Exploring CP violation and penguin effects through \B^0 \to D^+ D^- and \B^0_l \to D_{s(K)}^0 D_{s(K)}^{*}, Eur. Phys. J. C 51, 849 (2007).
[10] K. De Bruyn and R. Fleischer, A Roadmap to Control Penguin Effects in \B^0 \to J/\psi K_S^0 and \B^0_l \to J/\psi \phi, J. High Energy Phys. 03 (2015) 145.


(LHCb Collaboration)
1Also at Università di Pisa, Pisa, Italy
2Also at Università di Roma La Sapienza, Roma, Italy
3Also at Università della Basilicata, Potenza, Italy
4Also at Università di Urbino, Urbino, Italy
5Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia